AcademyU® Your Otolaryngology Education Source | www.entnet.org/hsc

Laryngology, Voice Disorders, and Bronchoesophagology

Hsc
Home Study Course

Home Study Course

Section 6
November 2016

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

© 2016 American Academy of Otolaryngology—Head and Neck Surgery Foundation
Empowering otolaryngologist-head and neck surgeons to deliver the best patient care

THE
HOME STUDY COURSE
IN
OTOLARYNGOLOGY — HEAD AND NECK SURGERY

SECTION 6

Laryngology, Voice Disorders and Bronchoesophagology

November 2016

SECTION FACULTY:

Thomas L. Carroll, MD**
Brad W. deSilva, MD**
Priya Krishna, MD, MS**
C. Kwang Sung, MD, MS**
Lee M. Akst, MD
Maggie A. Kuhn, MD
Ashli K. O'Rourke, MD, MS
David E. Rosow, MD

American Academy of Otolaryngology - Head and Neck Surgery Foundation
1650 Diagonal Road, Alexandria, VA 22314

Section 6 suggested exam deadline: January 3, 2017
Expiration Date: August 4, 2017; CME credit not available after that date

Introduction
The Home Study Course is designed to provide relevant and timely clinical information for physicians in training and current practitioners in otolaryngology - head and neck surgery. The course, spanning four sections, allows participants the opportunity to explore current and cutting edge perspectives within each of the core specialty areas of otolaryngology.

The **Selected Recent Material** represents primary fundamentals, evidence-based research, and state of the art technologies in Laryngology, Voice Disorders and Bronchoesophagology. The scientific literature included in this activity forms the basis of the assessment examination.

The number and length of articles selected are limited by editorial production schedules and copyright permission issues, and should not be considered an exhaustive compilation of knowledge Laryngology, Voice Disorders and Bronchoesophagology.

The **Additional Reference Material** is provided as an educational supplement to guide individual learning. This material is not included in the course examination and reprints are not provided.

Needs Assessment
AAO-HNSF's education activities are designed to improve healthcare provider competence through lifelong learning. The Foundation focuses its education activities on the needs of providers within the specialized scope of practice of otolaryngologists. Emphasis is placed on practice gaps and education needs identified within eight subspecialties. The *Home Study Course* selects content that addresses these gaps and needs within all subspecialties.

Target Audience
The primary audience for this activity is physicians and physicians-in-training who specialize in otolaryngology-head and neck surgery.

Outcomes Objectives

1. Evaluate the utility of computed tomography in the evaluation of patients with idiopathic vocal fold paresis.
2. Communicate the current practice patterns for otolaryngologists in diagnosing unilateral vocal fold paresis.
3. Discuss the various etiologies of unilateral vocal fold paralysis and concurrent dysphagia findings that these patients exhibit.
4. Evaluate the utility of stroboscopy in evaluating patients with laryngeal dysplasia, vocal fold paresis, and vocal fold paralysis.
5. Articulate the effectiveness of photoangiolytic laser treatment for Reinke's edema and expected voice outcomes following treatment.
6. Determine the benefits of office-based biopsy of laryngeal lesions versus surgical intervention for these pathologies.
7. Describe the anatomic changes that occur with vocal fold atrophy in the aging larynx and expected voice outcomes following voice therapy for this disorder.
8. Evaluate the utility of impedance testing in patients with extraesophageal reflux symptoms.
9. Measure body mass index and communicate how it predicts tracheal size.
10. Recognize the outcomes of tracheal stenosis and other airway complications in trauma patients that undergo percutaneous versus open tracheostomy.
11. Measure pulmonary function testing in patients that have undergone endoscopic treatment for subglottic stenosis.

Medium Used
The Home Study Course is available in electronic or print format. The activity includes a review of outcomes objectives, selected scientific literature, and a self-assessment examination.

Method of Physician Participation in the Learning Process
The physician learner will read the selected scientific literature, reflect on what they have read, and complete the self-assessment exam. After completing this section, participants should have a greater understanding of Laryngology, Voice Disorders and Bronchoesophagology as they affect the head and neck area, as well as useful information for clinical application.

Estimated time to complete this activity: 40.0 hours

Accreditation Statement
The American Academy of Otolaryngology—Head and Neck Surgery Foundation (AAO-HNSF) is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

Credit Designation
The AAO-HNSF designates this enduring material for a maximum of 40.0 *AMA PRA Category 1 Credit(s)*™. Physicians should claim credit commensurate with the extent of their participation in the activity.

ALL PARTICIPANTS must achieve a post-test score of 70% or higher for a passing completions to be recorded and a transcript to be produced. Residents' results will be provided to the Training Program Director.

PHYSICIANS ONLY: In order to receive *Credit* for this activity **a post-test score of 70% or higher is required**. Two retest opportunities will automatically be available if a minimum of 70% is not achieved.

Disclosure
The American Academy of Otolaryngology Head and Neck Surgery/Foundation (AAO-HNS/F) supports fair and unbiased participation of our volunteers in Academy/Foundation activities. All individuals who may be in a position to control an activity's content must disclose all relevant financial relationships or disclose that no relevant financial relationships exist. All relevant financial relationships with commercial interests[1] that directly impact and/or might conflict with Academy/Foundation activities must be disclosed. Any real or potential conflicts of interest[2] must be identified, managed, and disclosed to the learners. In addition, disclosure must be made of presentations on drugs or devices, or uses of drugs or devices that have not been approved by the Food and Drug Administration. This policy is intended to openly identify any potential conflict so that participants in an activity are able to form their own judgments about the presentation.

[1] A "Commercial interest" is any entity producing, marketing, re-selling, or distributing health care goods or services consumed by, or used on, patients.

[2] "Conflict of interest" is defined as any real or potential situation that has competing professional or personal interests that would make it difficult to be unbiased. Conflicts of interest occur when an individual has an opportunity to affect education content about products or services of a commercial interest with which they have a financial relationship. **A conflict of interest depends on the situation and not on the character of the individual.**

2016 SECTION 6 LARYNGOLOGY, VOICE DISORDERS AND BRONCHOESOPHAGOLOGY FACULTY

Co-Chairs:

Thomas L. Carroll, MD, Director, The Voice Program at Brigham and Women's Hospital, Instructor, Department of Otolaryngology, Harvard Medical School, Boston, Massachusetts.
Disclosure: Consultant: Pentax Medical.

Brad W. deSilva, MD, Assistant Professor, Residency Program Director, Voice and Swallow Disorders Clinic. Ohio State University Wexner Medical Center, Columbus, Ohio.
Disclosure: No relationships to disclose.

Priya Krishna, MD, MS, Co-Director, Loma Linda University Center for Voice and Swallowing, Assistant Professor, Department of Otolaryngology, Loma Linda University, Loma Linda, California.
Disclosure: Consultant: Merz North America, Olympus Medical.

C. Kwang Sung, MD, MS, Assistant Professor**,** Department of Otolaryngology-Head and Neck Surgery, Stanford University School of Medicine, Stanford, California.
Disclosure: Consultant: Olympus Medical; Expert Witness: Girardi – Keese.

Faculty:

Lee M. Akst, MD, Director, Division of Laryngology and Assistant Professor, Department of Otolaryngology, Johns Hopkins Department of Otolaryngology-Head and Neck Surgery, Baltimore, Maryland.
Disclosure: Consultant: Merz Pharmaceuticals and Olympus.

Maggie A. Kuhn, MD, Assistant Professor, University of California, Davis, Department of Otolaryngology – Head & Neck Surgery, Sacramento, California
Disclosure: No relationships to disclose.

Ashli K. O'Rourke, MD, MS, Assistant Professor, Evelyn Trammell Institute for Voice and Swallowing, Department Otolaryngology – Head and Neck Surgery, Medical University of South Carolina, Charleston, South Carolina.
Disclosure: Other: Carolina Speech Services; Consultant: Medtronic.

David E. Rosow, MD, Director, Division of Laryngology and Voice, Assistant Professor, Department of Otolaryngology, University of Miami Miller School of Medicine, Miami, Florida. Assistant Professor, Department of Vocal Performance, University of Miami Frost School of Music, Coral Gables, Florida.
Disclosure: No relationships to disclose.

Planner(s):

Linda Lee, AAO─HNSF Education Program Manager	No relationships to disclose
Stephanie Wilson, Stephanie Wilson Consulting, LLC; Production Manager	No relationships to disclose
Richard V. Smith, MD, chair, AAO-HNSF Education Steering Committee	Expert Witness: various legal firms
Catherine R. Lintzenich, MD, chair, AAO-HNSF Laryngology & Bronchoesophagology Education Committee	No relationships to disclose

This 2016 Section 6 Home Study Course does not include discussion of off-label uses of drugs or devices which have not been approved by the United States Food and Drug Administration:

Disclaimer

The information contained in this activity represents the views of those who created it and does not necessarily represent the official view or recommendations of the American Academy of Otolaryngology – Head and Neck Surgery Foundation.

January 3, 2017: Suggested Section 6 Exam submission deadline; course closed August 4, 2017.

EVIDENCE BASED MEDICINE

The AAO-HNSF Education Advisory Committee approved the assignment of the appropriate level of evidence to support each clinical and/or scientific journal reference used to authenticate a continuing medical education activity. Noted at the end of each reference, the level of evidence is displayed in this format: **[EBM Level 3]**.

Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001)	
Level 1	Randomized[1] controlled trials[2] or a systematic review[3] (meta-analysis[4]) of randomized controlled trials[5].
Level 2	Prospective (cohort[6] or outcomes) study[7] with an internal control group or a systematic review of prospective, controlled trials.
Level 3	Retrospective (case-control[8]) study[9] with an internal control group or a systematic review of retrospective, controlled trials.
Level 4	Case series[10] without an internal control group (retrospective reviews; uncontrolled cohort or outcome studies).
Level 5	Expert opinion without explicit critical appraisal, or recommendation based on physiology/bench research.

Two *additional ratings* to be used for articles that do not fall into the above scale. Articles that are informational only can be rated N/A , and articles that are a review of an article can be rated as Review. All definitions adapted from Glossary of Terms, Evidence Based Emergency Medicine at New York Academy of Medicine at www.ebem.org.

[1] A technique which gives every patient an equal chance of being assigned to any particular arm of a controlled clinical trial.
[2] Any study which compares two groups by virtue of different therapies or exposures fulfills this definition.
[3] A formal review of a focused clinical question based on a comprehensive search strategy and structure critical appraisal.
[4] A review of a focused clinical question following rigorous methodological criteria and employing statistical techniques to combine data from independently performed studies on that question.
[5] A controlled clinical trial in which the study groups are created through randomizations.
[6] This design follows a group of patients, called a "cohort", over time to determine general outcomes as well as outcomes of different subgroups.
[7] Any study done forward in time. This is particularly important in studies on therapy, prognosis or harm, where retrospective studies make hidden biases very likely.
[8] This might be considered a randomized controlled trial played backwards. People who get sick or have a bad outcome are identified and "matched" with people who did better. Then, the effects of the therapy or harmful exposure which might have been administered at the start of the trial are evaluated.
[9] Any study in which the outcomes have already occurred before the study has begun.
[10] This includes single case reports and published case series.

OUTLINE

NOVEMBER 2016 SECTION 6

LARYNGOLOGY, VOICE DISORDERS AND BRONCHOESOPHAGOLOGY

I. LARYNGOLOGY
   - A. Neurolaryngology: Injection and Diagnostics
   - B. Stroboscopy
   - C. Office-Based Procedures
   - D. Aging Larynx

II. BRONCHOESOPHAGOLOGY
   - A. Esophageal Dysphagia
   - B. Esophagopharyngeal Reflux
   - C. Tracheobronchial Disorders

TABLE OF CONTENTS
Selected Recent Materials - Reproduced in this Study Guide

NOVEMBER 2016 SECTION 6:
LARYNGOLOGY, VOICE DISORDERS, AND BRONCHOESOPHAGOLOGY

ADDITIONAL REFERENCE MATERIAL…….....i-ii

I. LARYNGOLOGY

A. Neurolaryngology: Injection and Diagnostics

Domer AS, Leonard R, Belafsky PC. Pharyngeal weakness and upper esophageal sphincter opening in patients with unilateral vocal fold immobility. *Laryngoscope*. 2014; 124(10):2371-2374. EBM level 4.....1-4

Summary: Individuals with unilateral vocal fold immobility of iatrogenic and idiopathic etiologies with subjective dysphagia demonstrate objective evidence of pharyngeal weakness. The increased prevalence of aspiration in this population may not solely be the result of impaired airway protection.

Li M, Chen S, Wang W, et al. Effect of duration of denervation on outcomes of ansa-recurrent laryngeal nerve reinnervation. *Laryngoscope*. 2014; 124(8):1900-1905. EBM level 4.....5-10

Summary: Although delayed laryngeal reinnervation is proved valid for unilateral vocal fold paralysis, surgical outcome is better if the procedure is performed within 2 years after nerve injury than after 2 years.

Paddle PM, Mansor MB, Song PC, Franco RA Jr. Diagnostic yield of computed tomography in the evaluation of idiopathic vocal fold paresis. *Otolaryngol Head Neck Surg*. 2015; 153(3):414-419. EBM level 4.....11-16

Summary: This study reviewed charts of 174 patients with a diagnosis of idiopathic unilateral vocal fold paresis (IUVFP) who underwent CT scan from skull base to mediastinum in a tertiary laryngology practice over a 10-year period. Of the 174 patients, 5 patients had a cause for their paresis identified on CT. This equated to a diagnostic yield of 2.9% (95% confidence interval, 0.94% to 6.6%). In addition, 48 patients had other incidental lesions identified that required further follow-up, investigation, or treatment. This equated to an incidental yield of 27.6% (95% confidence interval, 21.1% to 34.9%). This review demonstrates a low diagnostic yield and a high incidental yield. These findings suggest that the routine use of CT in the evaluation of idiopathic vocal fold paresis should be given careful consideration and that a tailored approach to investigation with good otolaryngologic follow-up is warranted.

Spataro EA, Grindler DJ, Paniello RC. Etiology and time to presentation of unilateral vocal fold paralysis. *Otolaryngol Head Neck Surg*. 2014; 151(2):286-293. EBM level 4.....17-24

Summary: Records were reviewed for 10 years at a single institution for unilateral vocal fold paralysis (UVFP); 938 patients met inclusion criteria. Of this group, 522 patients (55.6%) had UVFP due to surgery; 158 patients (16.8%) had UVFP associated with thyroid/parathyroid surgery, while in 364 patients (38.8%), UVFP was due to non-thyroid surgery. Of the total group, 416 patients (44.4%) had nonsurgical etiologies, 124 patients (13.2%) had idiopathic UVFP, and 621 patients (66.2%) had left-sided UVFP. Thyroidectomy remains the leading cause of surgery-related UVFP. Patients typically are seen within 3 to 4 months of onset.

Wu AP, Sulica L. Diagnosis of vocal fold paresis: current opinion and practice. *Laryngoscope*. 2015; 125(4):904-908. EBM level 5.....25-29

Summary: Wu and Sulica surveyed expert laryngologists who diagnosed vocal fold paresis predominantly on stroboscopic examination. Gross motion abnormalities had the highest positive predictive value. Laryngeal electromyography was infrequently used to assess for vocal fold paresis.

B. Stroboscopy

Djukic V, Milovanovic J, Jotic AD, Vukasinovic M. Stroboscopy in detection of laryngeal dysplasia: effectiveness and limitations. *J Voice*. 2014; 28(2):262.e13-262.e21. EBM level 3a.....30-38

Summary: This is a large study examining pre- and posttreatment stroboscopic findings in a prospectively collected group of 112 patients with dysplasia. There were fairly stringent exclusion criteria. The main finding was that abnormal amplitude of vocal fold vibration was significantly associated with recurrence. The type of cordectomy performed for the dysplasia and involvement of the vibratory segment was also associated with recurrence of dysplasia. Most recurrence occurred in a moderate dysplasia group. One limitation was that there were smaller numbers of patients in each dysplasia category; however, the study did highlight that caution should be exercised in the posttreatment follow-up period, and that stroboscopy has to be used in combination with other methods for an accurate diagnosis but can be helpful in predicting recurrence.

Reiter R, Pickhard A, Sander S, Brosch S. Prognostic relevance of mucosal waves in patients with unilateral vocal fold paralysis. *Ann Otol Rhinol Laryngol*. 2016; 125(4):331-335. EBM level 4.....39-43

Summary: This is a retrospective review of 100 patients with unilateral vocal fold paralysis looking at stroboscopic characteristics and relationship to eventual outcome. The majority of patients (75%) recovered function, and in all patients who recovered, there was an existing mucosal wave. In patients who did not recover function, only 10% had a mucosal wave.

Rosow DE, Sulica L. Laryngoscopy of vocal fold paralysis: evaluation of consistency of clinical findings. *Laryngoscope*. 2010; 120(7):1376-1382. EBM level 2.....44-50

Summary: This study sent videostroboscopy examination results from patients with unilateral vocal fold paralysis (VFP) to 22 blinded laryngologists and asked them to rate the results on twelve different criteria. The interrater reliability for each criterion was then calculated. The criteria with the best interrater agreement were glottic insufficiency, vocal fold bowing, and salivary pooling, which showed moderate agreement. All other criteria showed fair or poor agreement. The authors concluded that while it would be ideal to have a standardized rating scale for evaluation of VFP, the lack of interrater agreement across a wide range of laryngologists with different training and different backgrounds suggests that this may be very difficult to achieve.

Simpson CB, May LS, Green JK, et al. Vibratory asymmetry in mobile vocal folds: is it predictive of vocal fold paresis? *Ann Otol Rhinol Laryngol*. 2011; 120(4):239-242. EBM level 4.....51-54

Summary: This study is a retrospective review of 23 patients with symptoms suggestive of glottic insufficiency and stroboscopy examinations showing normal vocal fold mobility and vibratory asymmetry. All patients underwent laryngeal electromyography (LEMG) to determine presence of paresis. A total of 19 patients (83%) had evidence of paresis on LEMG. Blinded reviewers evaluated stroboscopy examinations for presence of paresis, but their ability to predict the distribution (sidedness) of the paresis was 37% or worse. The authors note that their findings suggest that all clinical and stroboscopic diagnoses of vocal fold paresis should be followed up with LEMG as the gold standard for diagnosis.

C. Office-Based Procedures

Croake DJ, Stemple JC, Uhl T, et al. Reliability of clinical office-based laryngeal electromyography in vocally healthy adults. *Ann Otol Rhinol Laryngol*. 2014; 123(4):271-278. EBM level 3.....55-62

Summary: Using a quantitative analysis protocol to inform an essentially qualitative technique, the study results indicated that there was generally poor to fair reliability in the laryngeal electromyography (LEMG) signal over testing sessions. Vocal intensity was an important variable that affected LEMG signal reliability. Standardization of LEMG protocols using vocal control parameters and quantitative analyses may help improve LEMG reliability in clinical settings.

Koszewski IJ, Hoffman MR, Young WG, et al. Office-based photoangiolytic laser treatment of Reinke's edema: safety and voice outcomes. *Otolaryngol Head Neck Surg*. 2015; 152(6):1075-1081. EBM level 4.....63-69

Summary: This study provides a retrospective analysis of patients undergoing office-based laser treatment of endoscopically proven Reinke's edema. Nineteen patients met criteria for the study inclusion. Five procedures were truncated due to patient intolerance. Phonatory frequency range increased (N = 12, $p = 0.003$), while percent jitter decreased (N = 12, $p = 0.004$). Phonation threshold pressure decreased after treatment (N = 4, $p = 0.049$). The Voice Handicap Index also decreased (N = 14, $p = 0.001$).

Richards AL, Sugumaran M, Aviv JE, et al. The utility of office-based biopsy for laryngopharyngeal lesions: comparison with surgical evaluation. *Laryngoscope*. 2015; 125(4):909-912. EBM level 4.....70-73

Summary: Office biopsy for laryngopharyngeal lesions may offer early detection and avoid operative intervention in some cases; however, for suspected dysplastic or malignant lesions, direct microlaryngoscopy should be the standard of care to ensure adequate full-thickness sampling and staging. For benign pathology, office biopsy is a safe and viable alternative to direct microlaryngoscopy and biopsy/excision.

Verma SP, Dailey SH. Office-based injection laryngoplasty for the management of unilateral vocal fold paralysis. *J Voice*. 2014; 28(3):382-386. EBM level 4.....74-78

Summary: This study is a retrospective chart review of 82 consecutive office-based injection laryngoplasty (OBIL) attempts on 57 patients. The most common route of access was transoral (85.6%). All OBILs were able to be completed. Injectates used were hyaluronic acid derivatives (57.3%), calcium hydroxyapatite (16%), and Cymmetra (16.5%). Three complications (3.7%) occurred. Thirty percent of patients ultimately elected for thyroplasty or ansa reinnervation, 22% found their condition to self-resolve, 14% died, and 25% were lost to follow-up.

D. Aging Larynx

Branco A, Todorovic Fabro A, Gonçalves TM, Garcia Martins RH. Alterations in extracellular matrix composition in the aging larynx. *Otolaryngol Head Neck Surg*. 2015; 152(2):302-307. EBM level 4.....79-84

Summary: This cadaver study proposes to further characterize extracellular matrix composition (ECM) changes in the aged vocal fold. Through immunohistochemistry, an overall increase in ECM mediated by increased collagen as well as decreased elastin were demonstrated. This work further highlights the histologic changes responsible for age-related voice changes.

Lee YC, Lee JS, Kim SW, et al. Influence of age on treatment with proton pump inhibitors in patients with laryngopharyngeal reflux disease: a prospective multicenter study. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(12):1291-1295. EBM level 4.....85-89

Summary: The authors aim to evaluate differences in laryngopharyngeal reflux (LPR) symptom severity among different age cohorts as well as response to treatment. The authors demonstrate that patients over 60 years of age experience greater symptoms and impact on quality of life from LPR; however, they seem to achieve less benefit from proton pump inhibitor therapy.

Yamauchi A, Yokonishi H, Imagawa H, et al. Vocal fold vibration in vocal fold atrophy: quantitative analysis with high-speed digital imaging. *J Voice*. 2015; 29(6):755-762. EBM level 2.....90-97

Summary: This study shows that high-speed digital imaging gives more insight into characterizing atrophic vocal folds.

Zeigler A, Verdolini Abbott K, Johns M, et al. Preliminary data on two voice therapy interventions in the treatment of presbyphonia. *Laryngoscope*. 2014; 124(8):1869-1876. EBM level 2b.....98-105

Summary: Zeigler et al demonstrate that vocal function exercise and PhoRTE voice therapy techniques appear to be effective in atrophy patients.

II. BRONCHOESOPHAGOLOGY

A. Esophageal Dysphagia

Kocdor P, Siegel ER, Tulunay-Ugur OE. Cricopharyngeal dysfunction: a systematic review comparing outcomes of dilatation, botulinum toxin injection, and myotomy. *Laryngoscope*. 2016; 126(1):135-141. EBM level 2a.....106-112

Summary: This systematic review of cohort studies evaluated the outcomes between different interventions for cricopharyngeal (CP) dysfunction, including CP dilation, botulinum toxin injections, and myotomy. The authors found that there was a significant increase in the odds of success and decreased complication rates with endoscopic myotomy versus open myotomy. They also found that myotomy was more effective than botulinum toxin injections.

Miles A, McMillan J, Ward K, Allen J. Esophageal visualization as an adjunct to the videofluoroscopic study of swallowing. *Otolaryngol Head Neck Surg*. 2015; 152(3):488-493. EBM level 4.....113-118

Summary: Miles et al hypothesize that esophageal disorders are the cause for dysphagia in many patients and propose that studying the esophagus as part of videofluoroscopic study of swallowing will yield greater diagnosis of abnormalities. Their findings suggest that esophageal disease is common and sometimes is the only abnormality in patients with cervical dysphagia. Furthermore, esophageal abnormalities frequently coexist with oral and pharyngeal disorders.

Moawad FJ, Veerappan GR, Dias JA, et al. Randomized controlled trial comparing aerosolized swallowed fluticasone to esomeprazole for esophageal eosinophilia. *Am J Gastroenterol*. 2013; 108(3):366-372. EBM level 1b.....119-125

Summary: This single-blinded, randomized controlled trial compared the efficacy of swallowed aerosolized steroids (fluticasone) to a proton pump inhibitor (omeprazole) for the treatment of eosinophilic esophagitis (EOE)–an important cause of esophageal dysphagia. Gastroesophageal reflux disease (GERD) patients were stratified equally into each arm. GERD patients with eosinophilia had improvement in their dysphagia symptoms and eosinophilia in biopsy specimens with PPI treatment alone. Current guidelines recommend failed trial of PPI prior to formal diagnosis of EOE due to this "newer" entity of PPI-responsive EOE.

Peng KA, Feinstein AJ, Salinas JB, Chhetri DK. Utility of the transnasal esophagoscope in the management of chemoradiation-induced esophageal stenosis. *Ann Otol Rhinol Laryngol*. 2015; 124(3):221-226. EBM level 4.....126-131

Summary: This article establishes the applications and safety of the transnasal esophagoscope (TNE) for chemoradiation-induced pharyngoesophageal swallowing dysfunction. Through use of a modified dysphagia score, the Functional Outcome Swallowing Scale (FOSS), the authors also suggest efficacy of TNE-based procedures in the population.

B. Esophagopharyngeal Reflux

Jetté ME, Gaumnitz EA, Birchall MA, et al. Correlation between reflux and multichannel intraluminal impedance pH monitoring in untreated volunteers. *Laryngoscope*. 2014; 124(10):2345-2351. EBM level 1b.....132-138

Summary: This study shows that the Reflux Finding Score (RFS) is not specific to detect laryngopharyngeal reflux in healthy volunteers, suggesting that other things can cause laryngeal inflammation.

Kavitt RT, Yuksel ES, Slaughter JC, et al. The role of impedance monitoring in patients with extraesophageal symptoms. *Laryngoscope*. 2013; 123(10):2463-2468. EBM level 2b.....139-144

Summary: Kavitt et al demonstrate that typical gastroesophageal reflux disease (GERD) testing parameters cannot predict the same conclusions when impedance is performed while the patient is on proton pump inhibitor therapy and recommend caution on over-interpreting impedance data in laryngopharyngeal reflux.

Reichel O, Dressel H, Wiederänders K, Issing WJ. Double-blind, placebo-controlled trial with esomeprazole for symptoms and signs associated with laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2008; 139(3):414-420. EBM level 1.....145-151

Summary: This study is a prospective, double-blinded randomized controlled trial examining esomeprazole versus placebo in managing the symptoms and signs of laryngopharyngeal reflux (LPR). A total of 62 patients with LPR were enrolled, and ultimately 30 patients were in the esomeprazole group, with 28 in the control group. Study subjects were given either esomeprazole 20 mg twice daily or identical placebo. They underwent laryngoscopic examinations and completed a Reflux Symptom Index (RSI) at 6 weeks and 12 weeks. There was minimal difference between the two groups at 6 weeks, but at 12 weeks, there was a significant difference in RSI as well as the Reflux Finding Score on laryngoscopy. The authors also found a high percentage of placebo patients (42%) experienced complete relief of symptoms at the end of the 3-month trial, though the percentage of the esomeprazole group was significantly higher (over 78%). One possible limitation is that pH monitoring was not used to diagnose LPR.

Youssef TF, Ahmed MR. Treatment of clinically diagnosed laryngopharyngeal reflux disease. *Arch Otolaryngol Head Neck Surg*. 2010; 136(11):1089-1092. EBM level 1.....152-155

Summary: This randomized controlled trial looked at the difference between monotherapy with esomeprazole versus triple therapy with esomeprazole, amoxicillin, and clarithromycin in treating patients with laryngopharyngeal reflux disease (LPRD) that have been found to be positive for *H. pylori* stool antigen (HPSA). The authors first determined that there is no statistical difference in symptoms between HPSA-positive and HPSA-negative patients with LPRD. Next, patients with clinically diagnosis of LPRD confirmed on pH testing were divided into two treatment groups based on HPSA status. HPSA-negative patients received esomeprazole 40 mg once daily for 4 weeks, while HPSA-positive patients were divided into monotherapy and triple therapy groups. The monotherapy treatment was identical to the HPSA-negative treatment, while the triple therapy group received esomeprazole 40 mg daily, 1 g amoxicillin daily, and 500 mg clarithromycin daily for 4 weeks. The HPSA-negative group showed improvement in 97% of patients, while in the HPSA-positive group, patients on monotherapy showed improvement in 40%, and 90% of those on triple therapy showed improvement. A major weakness of the study is that "improvement" is based on self-reporting by patients and by blinded laryngoscopic evaluation by the senior author. While these are no doubt important factors to take into consideration, they do not make use of objective testing, such as repeat 24-hour pH probe or validated questionnaires.

C. Tracheobronchial Disorders

D'Anza B, Knight J, Greene JS. Does body mass index predict tracheal airway size? *Laryngoscope*. 2015; 125(5):1093-1097. EBM level 4.....156-160

Summary: This study reviewed information on 123 patients who underwent tracheotomy over a 4-year period who also had CT imaging of the trachea in the 3 months preceding tracheostomy. The size of the endotracheal tube at time of the tracheotomy was also noted. Measurements were taken at the level of the first tracheal ring, as this was the most likely area for cuff-related injury of the airway. Important findings from the study were that airway area was correlated with height, and body mass index was inversely related to tracheal width after controlling for gender and age.

Gelbard A, Francis DO, Sandulache VC, et al. Causes and consequences of adult laryngotracheal stenosis. *Laryngoscope*. 2015; 125(5):1137-1143. EBM level 4.....161-167

Summary: This study looked at 340 patients with tracheal or laryngeal stenosis at two different sites. The etiology categories were idiopathic, iatrogenic, autoimmune, and trauma. The trauma group had significantly younger patients, whereas the idiopathic group had significantly more females. Comorbidities such as cardiovascular disease, peripheral vascular disease, and diabetes were more prevalent in the iatrogenic group. The idiopathic group also had the least-severe degree of laryngotracheal stenosis, with significantly fewer patients (none in this study) having had tracheostomy. As expected, patients with higher-grade stenosis (Cotton-Myer grades III or IV) had higher odds of being tracheostomy-dependent. The presence of tracheomalacia increased the odds of requiring a tracheostomy in the iatrogenic group.

Kettunen WW, Helmer SD, Haan JM. Incidence of overall complications and symptomatic tracheal stenosis is equivalent following open and percutaneous tracheostomy in the trauma patient. *Am J Surg*. 2014; 208(5):770-774. EBM level 3.....168-172

Summary: This is a large (N = 616) retrospective comparative study of the rate of tracheal stenosis in trauma patients who underwent either percutaneous (N = 351) versus open (N = 265) tracheostomy. The authors found no significant difference in the rate of tracheal stenosis in the open (1.9%) versus percutaneous (1.1%) groups. They did find that patients who developed tracheal stenosis were younger ($p = 0.02$) and had longer mechanical ventilation periods ($p = 0.055$). In addition, mortality was significantly higher in open tracheostomy patients, but this may be secondary to selection bias since patients with higher acuity of illness may be more likely to undergo open procedures.

Kraft SM, Sykes K, Palmer A, Schindler J. Using pulmonary function data to assess outcomes in the endoscopic management of subglottic stenosis. *Ann Otol Rhinol Laryngol*. 2015; 124(2):137-142. EBM level 4.....173-178

Summary: This retrospective case series described the utility of using pulmonary function tests to evaluate the efficacy of interventions for idiopathic subglottic stenosis. The pulmonary function parameters of PEF, PIF, FEV1/PEF, and FIF50% appeared to be the most valuable in judging response to endoscopic management and were significantly improved after airway dilation. PIF was the only parameter that was significantly associated with balloon size used for dilation. This study suggests that changes in PFTs are individualistic and need to be compared pre- and postprocedure for each patient (ie, there was no proposed "cut off" for intervention).

Taylor SC, Clayburgh DR, Rosenbaum JT, Schindler JS. Clinical manifestations and treatment of idiopathic and Wegener granulomatosis-associated subglottic stenosis. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(1):76-81. EBM level 4.....179-184

Summary: This article compares the clinical presentation, treatments, and outcomes in patients with subglottic stenosis (SGS) due to granulomatosis with polyangiitis (GPA, previously Wegener granulomatosis) versus idiopathic SGS. Although retrospective, this article has some interesting comparisons between the groups after similar treatment. For example, following open airway reconstruction, no idiopathic SGS patients required subsequent endoscopic dilations, while all GPA-SGS patients required subsequent interventions and had a higher rate of tracheotomy.

2016 SECTION 6 ADDITIONAL REFERENCES

Badia PI, Hillel AT, Shah MD, et al. Computed tomography has low yield in the evaluation of idiopathic unilateral true vocal fold paresis. *Laryngoscope*. 2013; 123(1):204-207.

Belafsky PC, Rees CJ, Allen J, Leonard RJ. Pharyngeal dilation in cricopharyngeus muscle dysfunction and Zenker diverticulum. *Laryngoscope*. 2010; 120(5):889-894.

Calloway HE, Kimbell JS, Davis SD, et al. Comparison of endoscopic versus 3D CT derived airway measurements. *Laryngoscope*. 2013; 123:2136-2141.

Carroll TL, Fedore LW, Aldahlawi MM. pH impedance and high-resolution manometry in laryngopharyngeal reflux disease high-dose proton pump inhibitor failures. *Laryngoscope*. 2012; 122(11):2473-2481.

Cheng J, Elden L. Outcomes in children under 12 months of age undergoing adenotonsillectomy for sleep-disordered breathing. *Laryngoscope*. 2013; 123(9):2281-2284.

Chin CJ, George C, Lannigan R, Rotenberg BW. Association of CPAP bacterial colonization with chronic rhinosinusitis. *J Clin Sleep Med*. 2013; 9(8):747-750.

Cohen SM, Misono S. Use of specific neuromodulators in the treatment of chronic, idiopathic cough: a systematic review. *Otolaryngol Head Neck Surg*. 2013; 148(3)374-382.

Deckard N, Yeh J, Soares DJ, et al. Utility of two-stage laryngotracheal reconstruction in the management of subglottic stenosis in adults. *Ann Otol Rhinol Laryngol*. 2013; 122(5):322-329.

Del Pero MM, Chaudhry A, Rasmussen N, et al. A disease activity score for ENT involvement in granulomatosis with polyangiitis (Wegener's). *Laryngoscope*. 2013; 123(3):622-628.

Eun YG, Shin SY, Kim SW. Effects of uvulopalatopharyngoplasty with or without radiofrequency tongue base reduction on voice in patients with obstructive sleep apnea. *Laryngoscope*. 2013; 123(7):1806-1810.

Friedman AD, Kobler JB, Landau-Zemer T, et al. High-force simulated intubation fails to dislocate cricoarytenoid joint in ex vivo human larynges. *Ann Otol Rhinol Laryngol*. 2012; 121(11):746-753.

Hartzell LD, Richter GT, Glade RS, Bower CM. Accuracy and safety of tracheoscopy for infants in a tertiary care clinic. *Arch Otolaryngol Head Neck Surg*. 2010; 136(1):66-69.

Huang WJ, Shu CH, Chou KT, et al. Evaluating the autonomic nervous system in patients with laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2013; 148(6):997-1002.

Johnston N, Yan JC, Hoekzema CR, et al. Pepsin promotes proliferation of laryngeal and pharyngeal epithelial cells. *Laryngoscope*. 2012; 122(6):1317-1325.

Kelchner LN, Brehm SB, de Alarcon A, Weinrich B. Update on pediatric voice and airway disorders: assessment and care. *Curr Opin Otolaryngol Head Neck Surg*. 2012; 20(3):160-164.

Kezirian EJ, Hohenhorst W, de Vries N. Drug-induced sleep endoscopy: the VOTE classification. *Eur Arch Otorhinolaryngol*. 2011; 268(8):1233-1236.

Krausert CR, Ying D, Choi SH, et al. Effect of vocal fold injury location on vibratory parameters in excised canine larynges. *Otolaryngol Head Neck Surg*. 2013; 148(1):89-95.

Marklund M, Verbraecken J, Randerath W. Non-CPAP therapies in obstructive sleep apnoea: mandibular advancement device therapy. *Eur Respir J*. 2012; 39(5):1241-1247.

Maturo SC, Braun N, Brown DJ, et al. Intraoperative laryngeal electromyography in children with vocal fold immobility: results of a multicenter longitudinal study. *Arch Otolaryngol Head Neck Surg*. 2011; 137(12):1251-1257.

Murry T, Branski RC, Yu K, et al. Laryngeal sensory deficits in patients with chronic cough and paradoxical vocal fold movement disorder. *Laryngoscope*. 2010; 120(8):1576-1581.

Nouraei SAR, Sandhu GS. Outcome of a multimodality approach to the management of idiopathic subglottic stenosis. *Laryngoscope*. 2013; 123(10):2474-2484.

Rao DA, Dellaripa PF. Extrapulmonary manifestations of sarcoidosis. *Rheum Dis Clin N Am*. 2013; 39(2):277-297.

Shah RN, Deal AM, Buckmire RA. Multidimensional voice outcomes after type I Gore-Tex thyroplasty in patients with nonparalytic glottic incompetence: a subgroup analysis. *Laryngoscope*. 2013; 123(7):1742-1745.

Smith DF, Vikani AR, Benke JR, et al. Weight gain after adenotonsillectomy is more common in young children. *Otolaryngol Head Neck Surg*. 2013; 148(3):488-493.

Sufyan AS, Kincaid JC, Wannemuehler TJ, Halum SL. The interarytenoid spatial relationship: accuracy and interrater reliability for determining sidedness in cases of unilateral adductor paresis. *J Voice*. 2013; 27(1):90-94.

Ummarino D, Vandermeulen L, Roosens B, et al. Gastroesophageal reflux evaluation in patients affected by chronic cough: Restech versus multichannel intraluminal impedance/pH metry. *Laryngoscope*. 2012; 123(4):980-984.

Vaca M, Mora E, Cobeta I. The aging voice: influence of respiratory and laryngeal changes. *Otolaryngol Head Neck Surg*. 2015; 153(3):409-413.

Van de Heyning PH, Badr MS, Baskin JZ, et al. Implanted upper airway stimulation device for obstructive sleep apnea. *Laryngoscope*. 2012; 122(7):1626-1633.

Vertigan AE, Gibson PG. The role of speech pathology in the management of patients with chronic refractory cough. *Lung*. 2012; 190(1):35-40.

Wang X, Ouyang Y, Wang Z, et al. Obstructive sleep apnea and risk of cardiovascular disease and all-cause mortality: a meta-analysis of prospective cohort studies. *Int J Cardiol*. 2013; 169(3):207-214.

Zerbib F, Roman S, Bruley Des Varannes S, et al. Normal values of pharyngeal and esophageal 24-hour pH impedance in individuals on and off therapy and interobserver reproducibility. *Clin Gastroenterol Hepatol*. 2013; 11(4):366-372.

Reprinted by permission of Laryngoscope. 2014; 124(10):2371-2374.

The Laryngoscope
© 2014 The American Laryngological,
Rhinological and Otological Society, Inc.

Pharyngeal Weakness and Upper Esophageal Sphincter Opening in Patients With Unilateral Vocal Fold Immobility

Amanda S. Domer, MS, CCC-SLP; Rebecca Leonard, PhD, CCC-SLP; Peter C. Belafsky, MD, PhD, MPH

Objectives/Hypothesis: To evaluate pharyngeal strength and upper esophageal sphincter opening in patients with unilateral vocal fold immobility (UVFI).

Study Design: Case control study.

Methods: Charts of individuals with UVFI who underwent a videofluoroscopic swallow study were reviewed. To exclude confounding variables associated with pharyngeal weakness, inclusion was limited to patients with iatrogenic and idiopathic UVFI. Data abstracted included patient demographics, etiology of UVFI, pharyngeal constriction ratio (PCR), and upper esophageal sphincter (UES) opening (UESmax). Data were compared to age/gender-matched controls with no history of dysphagia or UVFI. Discrete variables were analyzed using a chi-square test of independence, and an independent samples *t* test was used to compare the UVFI and control groups ($P = 0.05$). A one-way analysis of variance (ANOVA) was used to compare iatrogenic and idiopathic UVFI groups.

Results: The mean age of the cohort (n = 25) was 61 (±14 SD) years and 52% was female. The etiologies of UVFI were iatrogenic (n = 17) and idiopathic (n = 8). Thirty-eight percent of UVFI patients (n = 25) aspirated compared to 0% of controls ($P < 0.05$). The mean PCR for the UVFI group was 0.14 (±0.02) compared to 0.06 (±.01) for controls ($P < 0.05$). The mean UESmax for the UVFI group was 0.82 cm (±0.04) compared to 1.0 cm (±0.05) for controls ($P > 0.05$).

Conclusion: Individuals with UVFI of iatrogenic and idiopathic etiologies with subjective dysphagia demonstrate objective evidence of pharyngeal weakness. The increased prevalence of aspiration in this population may not be solely the result of impaired airway protection.

Key Words: Dysphagia, aspiration, vocal fold immobility, swallowing disorder, unilateral vocal fold immobility, UVFI.

Level of Evidence: 3b.

***Laryngoscope*, 124:2371–2374, 2014**

INTRODUCTION

Glottal competence is essential in airway protection during deglutition. If glottal closure is ineffective as a result of unilateral vocal fold immobility (UVFI), airway protection during the swallow may be compromised. UVFI may result from damage to the 1) brainstem nuclei, 2) vagus nerve, or 3) recurrent laryngeal nerve. Etiologies of UVFI include surgical trauma/iatrogenic (40%), tumor/neoplasm (30%), unknown/idiopathic (11%), trauma (8%), central nervous system dysfunction (4%), radiation (3%), inflammatory conditions (2%), and cardiovascular disease (2%).[1] Individuals with UVFI may present with aphonia (i.e., absence of voice), dysphonia (i.e., voice impairment), and/or dysphagia (i.e., swallowing impairment).

The precise etiology of dysphagia in patients with UVFI is uncertain. It is generally accepted that UVFI results in diminished airway protection. If airway protection is ineffective, an individual may aspirate material into the lungs, which may result in respiratory infection and/or death due to aspiration pneumonia. Approximately 33% to 42% of individuals with UVFI have been identified to aspirate.[2–5] Diminished airway protection is presumed to be the primary cause of swallowing dysfunction in patients with UVFI.[3,5] The integrity of other important biomechanics of the swallow, such as upper esophageal sphincter opening and pharyngeal contractility, however, has not been adequately evaluated in this patient population. Due to the highly intricate nature of the nerves and muscles in the pharynx and larynx, as well as the complex kinematics of the swallow, we hypothesize that features aside from impaired airway protection alone may contribute to increased occurrence of aspiration in this population.

This has been hypothesized in previous studies, which have demonstrated subjective findings in addition to impaired glottic closure that the authors stated contributed to a patient's increased risk of aspiration. One study that included patients with UVFI of both central and peripheral origins identified poor pharyngeal movement in patients with peripheral (i.e., recurrent laryngeal nerve injury, vagus nerve injury, or idiopathic etiologies) UVFI.[4] Another study identified decreased

From the Department of Communication Sciences & Disorders, University of South Florida (A.S.D.), Tampa, Florida; and the Department of Otolaryngology–Head & Neck Surgery, University of California, Davis (R.L., P.C.B.), Sacramento, California, U.S.A.

Editor's Note: This Manuscript was accepted for publication May 20, 2014.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Peter C. Belafsky, MD, PhD, MPH, Department of Otolaryngology–Head & Neck Surgery, University of California, Davis, 2521 Stockton Blvd, Suite 7200, Sacramento, CA 95817.
E-mail: peter.belafsky@ucdmc.ucdavis.edu

DOI: 10.1002/lary.24779

TABLE I.
Objective Measures and Definitions for a Dynamic Videofluoroscopic Swallow Study.

Objective Measures on Dynamic Videofluoroscopic Swallow Study (DSS)	Definition
Total pharyngeal transit time (TPT)	The time between the head of the bolus passing the posterior nasal spine to the time the tail of the bolus passes through the UES.
Upper esophageal sphincter opening (UESmax)	UES opening. The narrowest point of opening between C3 and C6 during maximal distention for bolus passage.
Pharyngeal constriction ratio (PCR)	A surrogate measure of pharyngeal strength. Specifically, a ratio of pharyngeal area measured in lateral fluoroscopic view at the point of maximal pharyngeal constriction during the swallow to the pharyngeal area measured with the bolus held in the oral cavity.
Hyoid to larynx approximation (HLx)	The difference in distance between the anterior margin of the hyoid bone with a 1cc bolus held in the oral cavity to maximal approximation of the hyoid and larynx during swallow and maximal hyoid to larynx approximation. A clear and consistently visible landmark on the anterior thyroid cartilage, such as calcification, was used as an alternative to the subglottic air column if it could not easily be visualized. Maximal approximation usually occurred just after maximal hyoid excursion.
Hyoid excursion (Hmax)	The distance traveled by the hyoid to the point of maximal elevation during a swallow from its position during hold.

sensation as a contributing factor to aspiration.[5] Other studies have provided evidence to support the notion that aspiration may not be solely related to vocal fold immobility, as demonstrated by continued dysphagia and/or aspiration after patients underwent surgical intervention to improve glottic closure.[6,7] There have yet to be objective kinematic and temporal measurements obtained from patients with subjective dysphagia as a result of UVFI related solely to vagus nerve injury. The purpose of this investigation was to evaluate pharyngeal strength and upper esophageal sphincter (UES) opening in individuals with UVFI caused by idiopathic or iatrogenic injuries to the vagus nerve.

MATERIALS AND METHODS

A clinical swallowing database consisting of individuals with dysphagia who underwent a dynamic videofluoroscopic swallow study (DSS) was reviewed to identify individuals with UVFI between January 1, 1999, and June 1, 2012. The Institutional Review Board of the University of California, Davis, approved use of this database for clinical research. All patients with UVFI were confirmed by videolaryngoscopy and/or strobovideolaryngoscopy. Individuals were excluded if they were under 18 years of age, had suspected vagus nerve injury of central origin (e.g., cerebrovascular accident, neuromuscular disease, brain tumor, etc.), head or neck cancer (i.e., except that isolated to the thyroid, which resulted in only surgical removal of all or part of the thyroid), and/or anterior approach cervical spine surgery. The purpose for excluding these populations was the possibility of a more complex swallowing disorder not necessarily limited to injury of the vagus nerve. This resulted in patients with iatrogenic and idiopathic UVFI. None of the patients had undergone a vocal fold medialization procedure prior to the videofluoroscopic swallow study. The timing between onset of UVFI and time of evaluation was not recorded.

All swallow studies were preformed using a properly collimated OEC Medical Systems 9800 Radiographic/Fluoroscopic unit that provided a 63 kV, 1.2 mA-type output for the full field of view mode (12-inch input phosphor diameter). In accordance with our standard protocol, a metal ring of known diameter was taped to the chin or neck of the patient for measurement calibration purposes. Lateral views were obtained while the patient, seated in an examination chair, was administered liquid barium (EZpaque barium sulfate suspension, 60% w/v; EZ-EM, Inc., Westbury, NY) boluses of 1 cc, 3 cc, and 20 cc and a 3-cc paste bolus (EZ-paste, EZ-EM, Inc.) measured with a syringe or graduated medicine cup. The patient was then turned to obtain anteroposterior views and administered liquid barium boluses of 3 cc and 20 cc. Studies were recorded on a Sony Md-1000 DVD recorder (Sony Corp. America, New York, NY) and were played back using Quick Time (7.7.1; Apple, Cupertino, CA). Measures were obtained from digitized images using ImageJ software (National Institutes of Health, Bethesda, MD) and software tools from Iconico, Inc (New York, NY). Specific measurement techniques have been previously described in detail.[8,9] An experienced unblinded clinician (i.e., the same clinician who conducted the videofluoroscopic swallow studies) analyzed all studies; however, because this study was retrospectively completed, there was no information available related to this study at the time of evaluation. All measures for the current study were obtained from the lateral view.

The primary outcome measures were upper esophageal sphincter opening (UESmax) and the pharyngeal constriction ratio (PCR). The PCR is a validated surrogate measure of pharyngeal strength on fluoroscopy; and an elevated PCR suggests pharyngeal weakness.[9] The secondary outcome measures were larynx to hyoid approximation (HLx), hyoid displacement (Hmax), and total pharyngeal transit time (TPT) (see Table I for definitions). The data from each variable were compared to age and gender-matched controls with no history of dysphagia. Discrete variables were analyzed with a chi-square test of independence and an independent samples t test was used to compare the control and combined UVFI groups with alpha set at 0.05. A one-way analysis of variance (ANOVA) was used to compare the iatrogenic and idiopathic UVFI groups with the control group. A Bonferroni correction was applied to adjust for multiple comparisons with alpha set at 0.01. There is a probability of 0.05 that a type I error has been made in the set of tests.

RESULTS

A total of 137 individuals with UVFI were identified from the clinical database. There were 25 subjects who met strict inclusion and exclusion criteria (i.e., did not

TABLE II.
Mean, Standard Deviation, and *P* Value for Iatrogenic and Idiopathic Groups.

Iatrogenic UVFI vs. Idiopathic UVFI	Iatrogenic (n = 17) Mean (SD)	Idiopathic (n = 8) mean (SD)	*P* Value
UESmax (cm)	0.85 (0.05)	0.76 (0.07)	1.00
PCR	0.12 (0.02)	0.18 (0.04)	0.34
HLx (cm)	1.25 (0.14)	1.58 (0.15)	0.55
Hmax (cm)	1.91 (0.16)	1.84 (0.16)	1.00
TPT (seconds)	1.36 (0.09)	2.66 (0.93)	0.04

$\alpha = 0.01$.
*Statistical significance.
HLx = hyoid to larynx approximation; Hmax = hyoid excursion; PCR = pharyngeal constriction ratio; SD = standard error; TPT = total pharyngeal transit time; UESmax = upper esophageal sphincter opening; UVFI = unilateral vocal fold immobility.

have dysphagia complaints that could be explained by any other etiology in their medical history). The etiology of UVFI was identified as iatrogenic in 17 individuals and idiopathic in eight individuals. The cohort was 52% female with a mean age of 61 (±14) years.

Comparison of the Idiopathic Group, Iatrogenic Group, and Control Group

The idiopathic group (n = 8) was 50% female with a mean age of 64 (±14) years (Table II). The iatrogenic group (n = 17) was 53% female with a mean age of 60 years (±14) years. There was no difference in age or gender between groups ($P > 0.05$).

The mean UESmax was 0.76 (±0.07) cm for the idiopathic group, 0.85 (±0.05) cm for the iatrogenic group, and 1.0 (±0.05) cm for the control group. There were no significant differences between any groups for UESmax ($P > 0.01$).

The PCR was 0.18 (±0.04) for the idiopathic group, 0.12 (±0.02) for the iatrogenic group, and 0.06 (±0.01) for the control group. PCR was significantly greater for the idiopathic group compared with the control group ($P < 0.01$). PCR for the iatrogenic group was not significant, but less than the control group ($P > 0.01$). There was not a significant difference between the idiopathic or iatrogenic groups ($P > 0.01$).

The HLx was 1.58 (±0.15) cm for the idiopathic group, 1.25 (±0.14) cm for the iatrogenic group, and 1.42 (±0.12) cm for the control group ($P > 0.01$). There were no significant differences between any groups for HLx ($P > 0.01$).

The Hmax was 1.84 (±0.16) cm for the idiopathic group, 1.91 (±0.16) cm for the iatrogenic group, and 2.22 (±0.18) cm for the control group ($P > 0.01$). There were no significant differences between any groups for Hmax ($P > 0.01$).

The mean TPT was 2.66 (±0.93) seconds for the idiopathic group, 1.36 (±0.09) seconds for the iatrogenic group, and 1.01 (±0.06) seconds for the control group ($P > 0.01$). There was a significant difference between the idiopathic UVFI and iatrogenic UVFI groups ($P < 0.01$). There was a significant difference between the idiopathic group and the control group. There was not a significant difference between the iatrogenic and the control group.

Thirty-eight percent of individuals with idiopathic UVFI and 35% of individuals with iatrogenic UVFI aspirated at least once during the videofluoroscopic swallow study. There was not a significant difference between either of the UVFI groups ($P > 0.05$); however, there was a significant difference between the idiopathic group and control group ($P < 0.05$), as well as the iatrogenic group and control group ($P < 0.05$).

UVFI Group Compared With the Control Group

The mean UESmax opening was 0.82 (±0.04) cm for the UVFI group compared to 1.00 (±0.05) cm for controls ($P > 0.05$ see Table III). The pharyngeal constriction ratio was 0.14 (±0.02) for the UVFI group compared to 0.06 (±0.01) for controls ($P < 0.05$). Larynx to hyoid approximation was 1.35 (±0.11) cm for the UVFI group and 1.42 (±0.12) cm for the control group ($P > 0.05$). The mean for hyoid displacement was 1.89 (±0.12) cm for the UVFI group and 2.22 (±0.17) cm for the control group ($P > 0.05$). The mean TPT was 1.78 (±0.32) seconds for the UVFI group and 1.01 (±0.06) seconds for the control group ($P < 0.05$). Thirty-six percent of individuals with UVFI aspirated at least once during the videofluoroscopic swallow study compared to 0% of controls ($P < 0.05$).

DISCUSSION

The data in the current investigation provided evidence to suggest that individuals with UVFI of iatrogenic and idiopathic etiologies may present with additional biomechanical findings that may increase the prevalence of aspiration. The group of individuals with UVFI of idiopathic and iatrogenic etiologies demonstrated significantly prolonged TPT and elevated PCRs, suggesting delayed bolus transit and pharyngeal weakness. Additionally, individuals with UVFI of idiopathic etiology demonstrated significantly prolonged TPT, increased PCR (i.e., pharyngeal weakness), and decreased UESmax compared to controls. These findings support the notion that factors other than glottal

TABLE III.
Mean, Standard Deviation, and *P* Value for UVFI Group and Controls.

UVFI vs. Controls	UVFI (n = 25) Mean (SD)	Controls (n = 25) Mean (SD)	*P* Value
UESmax (cm)	0.82 (0.05)	1.00 (0.05)	0.94
PCR	0.14 (0.02)	0.06 (0.01)	0.03*
HLx (cm)	1.35 (0.11)	1.42 (0.12)	0.94
Hmax (cm)	1.89 (0.12)	2.21 (0.17)	0.16
TPT (seconds)	1.78 (0.32)	1.01 (0.06)	0.02*

$\alpha = 0.05$
*Statistical significance.
HLx = hyoid to larynx approximation; Hmax = hyoid excursion; PCR = pharyngeal constriction ratio; SD = standard error; TPT = total pharyngeal transit time; UESmax = upper esophageal sphincter opening; UVFI = unilateral vocal fold immobility.

competence influence the increased prevalence of aspiration observed in patients with UVFI.

The results of the present study confirm the qualitative findings of Jang et al. with objective data. In that study, individuals with UVFI of peripheral origin (i.e., recurrent laryngeal nerve injury, vagus nerve injury, or idiopathic etiologies) presented with subjective suggestion of abnormal laryngeal elevation and epiglottic inversion, residue in the valleculae, residue in the pyriform sinuses, and aspiration.[4]

The percentage of aspiration in this study was also comparable to previous studies, which identified aspiration in approximately 33% to 42% of individuals with UVFI.[2–5] It is important to note that the present study was performed in an outpatient tertiary care center and most of the previous work evaluating aspiration in patients with UVFI was performed in acute care settings. Although information about the length of time from the onset of vocal fold immobility was not available in these studies, we suspect that the individuals included in the present investigation may have had a more prolonged duration of UVFI in comparison to previous work. Nonetheless, the percentage of aspiration in the current investigation was similar to previous findings.

When evaluating UVFI between groups, the only significant finding was total pharyngeal transit time. This may be the result of a higher vagal injury in the idiopathic group, compared to the iatrogenic group. However, this confirmed that, although the iatrogenic group was more likely isolated to recurrent laryngeal nerve and/or superior laryngeal nerve injury than the idiopathic group, the finding of increased PCR did not vary significantly between groups. Additionally, the iatrogenic group was approaching significance compared to the control group, and significance may be achieved with a larger sample size. Therefore, pharyngeal weakness may exist in individuals with UVFI of both idiopathic and iatrogenic etiologies.

This study was not without limitations. Electromyography was not utilized to determine the site of lesion causing UVFI. Also, in an effort to keep the groups as homogenous as possible, the sample size was small. However, the groups that were chosen were intended to represent individuals with UVFI limited to vagus nerve injury. None of the individuals underwent a vocal fold medialization procedure prior to study, and all individuals presented with a dysphagia complaint. In addition, this study was retrospective, so a future prospective investigation with a larger sample size is required to confirm these results.

CONCLUSION

Individuals with UVFI of iatrogenic and idiopathic etiologies with subjective dysphagia demonstrate objective evidence of pharyngeal weakness. The increased prevalence of aspiration in this population may not be solely the result of impaired airway protection.

BIBLIOGRAPHY

1. Chen HC, Jen YM, Wang CH, Lee JC, Lin YS. Etiology of vocal cord paralysis. *ORL J Otorhinolaryngol Relat Spec* 2007;69:167–171.
2. Leder SB, Suiter DM, Duffey D, Judson BL. Vocal fold immobility and aspiration status: a direct replication study. *Dysphagia* 2012;27:265–270.
3. Heitmiller RF, Tseng E, Jones B. Prevalence of aspiration and laryngeal penetration in patients with unilateral vocal fold motion impairment. *Dysphagia* 2000;15:184–187.
4. Jang YY, Lee SJ, Jeon JY, Lee SJ. Analysis of video fluoroscopic swallowing study in patients with vocal cord paralysis. *Dysphagia* 2012;27:185–190.
5. Tabaee A, Murry T, Zschommler A, Desloge RB. Flexible endoscopic evaluation of swallowing with sensory testing in patients with unilateral vocal fold immobility: incidence and pathophysiology of aspiration. *Laryngoscope* 2005;115:565–569.
6. Bhattacharyya N, Kotz T, Shapiro J. Dysphagia and aspiration with unilateral vocal cord immobility: incidence, characterization, and response to surgical treatment. *Ann Otol Rhinol Laryngol* 2002;111:672–679.
7. Yip HT, Leonard R, Kendall, KA. Impact of vocal fold medialization on pharyngeal transit time and incidence of aspiration. *Otolaryngol Head Neck Surg* 2004; 131
8. Kendall K, McKenzie S, Leonard R. Dynamic swallow study: objective measures and normative data. In: Leonard R, Kendall K, eds. *Dysphagia Assessment and Treatment Planning: A Team Approach*. San Diego, CA: Singular Publishing; 1998.
9. Leonard R, Rees CJ, Belafsky P, Allen J. Fluoroscopic surrogate for pharyngeal strength: the pharyngeal constriction ratio (PCR). *Dysphagia* 2011;26:13–17.

Reprinted by permission of Laryngoscope. 2014; 124(8):1900-1905.

The Laryngoscope
© 2014 The American Laryngological,
Rhinological and Otological Society, Inc.

Effect of Duration of Denervation on Outcomes of Ansa-Recurrent Laryngeal Nerve Reinnervation

Meng Li, MD; Shicai Chen, MD; Wei Wang, MD; Donghui Chen, MD; Minhui Zhu, MD; Fei Liu, MD; Caiyun Zhang, MD; Yan Li, MD; Hongliang Zheng, MD

Objectives/Hypothesis: To investigate the efficacy of laryngeal reinnervation with ansa cervicalis among unilateral vocal fold paralysis (UVFP) patients with different denervation durations.

Study Design: We retrospectively reviewed 349 consecutive UVFP cases of delayed ansa cervicalis to the recurrent laryngeal nerve (RLN) anastomosis. Potential influencing factors were analyzed in multivariable logistic regression analysis. Stratification analysis performed was aimed at one of the identified significant variables: denervation duration.

Methods: Videostroboscopy, perceptual evaluation, acoustic analysis, maximum phonation time (MPT), and laryngeal electromyography (EMG) were performed preoperatively and postoperatively. Gender, age, preoperative EMG status and denervation duration were analyzed in multivariable logistic regression analysis. Stratification analysis was performed on denervation duration, which was divided into three groups according to the interval between RLN injury and reinnervation: group A, 6 to 12 months; group B, 12 to 24 months; and group C, $>$ 24 months.

Results: Age, preoperative EMG, and denervation duration were identified as significant variables in multivariable logistic regression analysis. Stratification analysis on denervation duration showed significant differences between group A and C and between group B and C ($P < 0.05$)—but showed no significant difference between group A and B ($P > 0.05$) with regard to parameters overall grade, jitter, shimmer, noise-to-harmonics ratio, MPT, and postoperative EMG. In addition, videostroboscopic and laryngeal EMG data, perceptual and acoustic parameters, and MPT values were significantly improved postoperatively in each denervation duration group ($P < 0.01$).

Conclusions: Although delayed laryngeal reinnervation is proved valid for UVFP, surgical outcome is better if the procedure is performed within 2 years after nerve injury than that over 2 years.

Key Words: Vocal cord paralysis, laryngeal reinnervation, ansa cervicalis, recurrent laryngeal nerve, denervation duration.

Level of Evidence: 4.

***Laryngoscope*, 124:1900–1905, 2014**

INTRODUCTION

Unilateral vocal fold paralysis (UVFP) is a condition commonly seen in otolaryngology clinics. The most frequent cause of UVFP is injury to the recurrent laryngeal nerve (RLN).[1] UVFP can present as various degrees of dysphonia and dysphagia, and has a significant impact on a patient's quality of life. There are various surgical methods for treating UVFP. These include vocal fold injection, thyroplasty, arytenoid adduction, and laryngeal reinnervation, which is an effective surgical procedure with better long-term outcomes because it restores neural connections to laryngeal muscles and thus maintains the bulk, tension, and position of the paralyzed vocal fold.[2,3] Our previous large-scale study, as well as reports from other researchers, demonstrated satisfactory or good voice outcomes in patients who underwent ansa cervicalis–RLN anastomosis.[4–6]

Experimental studies have shown that reinnervation surgery can restore laryngeal function with excellent results when performed immediately following nerve resection.[7,8] However, in clinical practice, there is usually a considerable delay between RLN injury and presentation for reinnervation surgery. In addition, 6 to 12 months are usually allowed for possible spontaneous recovery of the paralyzed vocal fold or compensation from the contralateral vocal fold, even when UVFP is diagnosed early. Surgical intervention is considered only in cases with unsatisfactory spontaneous recovery. At present, there are no definitive clinical data regarding the longest allowable period between the onset of RLN injury and nerve reconstruction to achieve functional recovery of the adductor muscle.

In the present study, we analyzed several potential influencing factors of laryngeal reinnervation, including gender, age of patients, preoperative maximal voluntary motor-unit recruitment (VMUR) of laryngeal muscles,

From the Department of Otolaryngology–Head and Neck Surgery, Changhai Hospital, Second Military Medical University, Shanghai, 200433, People's Republic of China

Editor's Note: This Manuscript was accepted for publication January 27, 2014.

This work was supported by Grants 81100724, 81170899, 81070775 for science research from National Natural Science Foundation of China, and Grant 10XD1405500 from the Science and Technology Commission of Shanghai Municipality. The authors have no other financial relationships, or conflicts of interest to disclose.

Send correspondence to Hongliang Zheng, MD, Department of Otolaryngology–Head and Neck Surgery, Changhai Hospital, Second Military Medical University, Shanghai 200433, People's Republic of China. E-mail: zheng_hl2004@163.com

DOI: 10.1002/lary.24623

and denervation duration using multivariate logistic regression analysis method in a large series of 349 UVFP patients who underwent delayed laryngeal reinnervation. We also performed further stratification analysis aimed at one of the identified significant variables—denervation duration—in order to investigate the effect of denervation duration on the surgical outcome of laryngeal reinnervation.

MATERIALS AND METHODS

Patient Characteristics

Our study was approved by the institutional review board of Second Military Medical University, Shanghai, People's Republic of China. The medical records of 349 UVFP patients (94 males and 255 females; mean age 44.0 years, ranging 17–69 years) who underwent anastomosis of the main branch of the ansa cervicalis to the RLN between January 1996 and January 2011, and who were followed for at least 2 years, were reviewed. The etiology of UVFP in this series of patients was RLN injury during thyroid surgery. Informed consent was obtained from all patients involved in this study. Patients who were lost to follow-up were excluded. There was a minimum waiting period of 6 months following onset of RLN injury to allow for possible spontaneous recovery or compensation. The median denervation course was 16.1 months (range, 6–45 months). When stratified by denervation duration, the number of patients in each denervation duration group was: 172 (49.3%) patients with a denervation duration 6 to 12 months (group A); 108 (30.9%) patients with a denervation duration 12 to 24 months (group B), and 69 (19.8%) patients with a denervation duration >24 months (group C). The median follow-up period after laryngeal reinnervation was 70.8 months (range, 24–156 months).

Surgical Procedure

The surgical procedure has been elaborated in our previous report.[4] Briefly, under general anesthesia, the ipsilateral ansa cervicalis was explored, and the main branch was transected at the bifurcation and freely mobilized for preparation of anastomosis. The RLN was dissected at a point sufficiently far from the injury site to provide a tension-free anastomosis and then transected. Under an operating microscope, the distal RLN stump was anastomosed to the main branch of the ansa cervicalis using nylon 11-0 thread in three to five epineural sutures.

Videostroboscopy

All patients were observed via a videostroboscope (RICHARD WOLF GmbH, model 5570, Knittlingen, Germany) during "eee" phonation at a comfortable loudness and pitch for as long as possible, and dynamic videos were recorded preoperatively and postoperatively. Three experienced laryngologists who had not performed any of the surgeries reviewed all of the videos. The videos were randomized, and the reviewers were blinded to whether the videos were preoperative or postoperative. Visual laryngeal analysis included glottal closure (0, complete; 1, slightly incomplete; 2, moderately incomplete; 3, severely incomplete), vocal fold position, vocal fold edge of paralyzed side, phrase symmetry, and regularity. Consensus of the reviewers was reached on the visual appearance of the larynx. Our previous studies demonstrated that the above parameters were consistent in presenting reinnervation outcome of vocal fold paralysis, among which the parameter glottal closure was the most representative one[4]; therefore, only the parameter "glottal closure" was included when performing statistical analysis using a univariable analysis and multivariable logistic regression analysis.

Vocal Function Assessment

Vocal function assessment included perceptual evaluation, acoustic analysis, and maximum phonation time (MPT) measurement. Preoperative and postoperative voice samples containing sustained vowels /a/ and connected speech samples were used for perceptual evaluation and acoustic analysis. The recording equipment consisted of a digital audiotape recorder and a dynamic microphone (Tiger Electronics Inc., North Reading, MA). Five laryngologists who had been trained in grade, roughness, breathiness, asthenia, and strain (GRBAS) rating performed voice perceptual evaluation using a perceptual rating scale (GRBAS) for voice quality and characteristics. The ratings were accomplished in a blinded fashion, with patient voice samples arranged in a random manner. Each listener was asked to score connected speech samples for overall grade, roughness, breathiness, asthenia, and strain using a voice-quality scale for each parameter (0, normal; 1, mild; 2, moderate; 3, severe). The values were averaged among the five listeners. Our previous studies demonstrated that the interrater and intrarater reliability was acceptable (interrater reliability >0.76; intrarater reliability >0.81).[4,9] In addition, the above five parameters of perceptual evaluation were consistent in presenting vocal outcome of vocal fold paralysis, among which the parameter overall grade was the most representative one.[4] Therefore, only the parameter "overall grade" was included when performing statistical analysis using a univariable analysis and multivariable logistic regression analysis.

The acoustic parameters of sustained vowel /a/ were evaluated using Praat software (Boersma, Paul & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.1.12, retrieved from http://www.praat.org/). The acoustic parameters were mean noise-to-harmonics ratio (NHR) and measures of phonatory stability—jitter (local) and shimmer (local). MPT was defined as the duration of sustained phonation of the vowel /a/ after maximum inspiration and was measured preoperatively and postoperatively.[4]

Laryngeal Electromyography

A four-channel electromyograph and concentric needle electrodes (Dantec Counterpoint, Copenhagen, Denmark) were used for the laryngeal electromyography (EMG) recordings. To test for proper needle position, the unaffected vocal fold was examined first. The electromyographic activity of the bilateral thyroarytenoid (TA) muscles was recorded during the following two stages: while breathing quietly when relaxed, and while pronouncing the vowel /eee/ with the greatest exertion, then sniff. One board-certified otolaryngologist performed the EMG, and a neurologist operated the EMG machine and interpreted the EMG results. The neurologist rated the VMUR using the following scale: 0, full interference; 1, mixed interference; 2, simple interference; and 3, without motor unit potential.[4]

Statistical Analysis

The perceptual evaluation, acoustic analysis, and MPT data did not follow normal distribution and were presented as median (low quartile, upper quartile). We sought to evaluate influencing factors for the surgical outcome of laryngeal reinnervation using multivariable logistic regression methods. Potential influencing factors were examined in univariable

TABLE I.
Univariable Analysis of Influencing Factors on the Surgical Outcome of Laryngeal Reinnervation.

	Overall Grade	Jitter	Shimmer	NHR	MPT	Post-VMUR	Glottal Closure
Sex	0.7131	0.1844	0.1394	0.3456	< ***0.001***	0.1804	0.5141
Age	< ***0.001***	0.3441	< ***0.001***	***0.0010***	< ***0.001***	< ***0.001***	< ***0.001***
Pre-VMUR	< ***0.001***	0.3807	0.0630	0.1554	< ***0.001***	< ***0.001***	< ***0.001***
Denervation duration	***0.0022***	< ***0.001***	< ***0.001***	< ***0.001***	***0.0007***	***0.0010***	0.5410

Data in Table I represent *P* value calculated by univariable analysis.
$P < 0.05$ is deemed as statistically significant.
MPT = maximum phonation time; NHR = noise-to-harmonics ratio; VMUR = voluntary motor-unit recruitment.

analyses before building the multivariable logistic regression model. Univariable logistic regression models containing each covariate were fit. Covariates of known clinical significance or with P value < 0.05 in the univariable logistic models were selected as candidates for the multivariable model. Statistical significance was considered α-level 0.05.

RESULTS

Univariable Results

Table I displays the results of demographic or preoperative medical variables in univariable analysis. Briefly, gender was only significantly associated with MPT. Age was significantly associated with overall grade, shimmer, NHR, MPT, postoperative VMUR, and glottal closure. Preoperative VMUR was significantly associated with overall grade, MPT, postoperative VMUR, and glottal closure, whereas denervation duration was significantly associated with overall grade, jitter, shimmer, NHR, MPT, and postoperative VMUR.

Multivariable Logistic Regression Model Results

Based on univariable logistic regression results, significant candidate variables were identified ($P < 0.05$) for assessment in the multivariable logistic regression model. Table II shows the results of multivariable logistic regression. These indicate that age of patients; severity of nerve injury, which was presented as preoperative VMUR; and denervation duration had impact on the surgical effect of laryngeal reinnervation. Stratification analysis on other significant variables such as age of patients and preoperative VMUR will be presented in another report (unpublished). In the present study, we performed further stratification analysis aimed at one of the identified significant variables—denervation duration. The results are as follows (see Table III).

Videostroboscopic Findings

On preoperative videostroboscopy, the majority of cases in each group had severely incomplete glottal closure, which did not differ significantly among the three groups. On videostroboscopy performed 2 years after the reinnervation operation, most patients showed complete glottal closure. The postoperative stroboscopic findings were significantly improved in total sample and within each group compared with the corresponding preoperative findings ($P < 0.01$). However, denervation duration was not a significant variable with regard to multivariable logistic regression analysis in glottal closure.

Vocal Function Assessment

The preoperative values of perceptual evaluation parameter (overall grade) and acoustic parameters—jitter (local), shimmer (local), and the NHR—showed no significant differences among the three groups. Two years after the reinnervation operation, the postoperative value of overall grade was significantly improved within each of the three groups compared with the corresponding preoperative values ($P < 0.001$). Postoperative values of jitter (local), shimmer (local), and the NHR were significantly lower within each group compared with the corresponding preoperative values ($P < 0.05$). The postoperative MPT was significantly longer than the preoperative MPT in each group ($P < 0.05$).

Denervation duration was identified as a significant variable with regard to parameters (overall grade), jitter, shimmer, NHR, and MPT in multivariable logistic

TABLE II.
Multivariable Analysis of Influencing Factors on the Surgical Outcome of Laryngeal Reinnervation.

	Overall Grade	Jitter	Shimmer	NHR	MPT	Post-VMUR	Glottal Closure
Sex	0.6692	0.1124	**0.0133**	0.2320	**< 0.001**	0.4301	0.7663
Age	**< 0.001**	0.2588	**< 0.001**	**< 0.001**	**< 0.001**	**0.0080**	**0.0004**
Pre-VMUR	**< 0.001**	0.3374	0.4935	**0.0091**	**< 0.001**	**< 0.001**	**< 0.001**
Duration of denervation	< **0.001**	< **0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**0.0004**	0.1809

Data represent *P* value calculated by multivariable analysis.
$P < 0.05$ is deemed as statistically significant.
MPT = maximum phonation time; NHR = noise-to-harmonics ratio; VMUR = voluntary motor-unit recruitment.

TABLE III.
Stratification Analysis on Denervation Duration.

Parameters	Groups	Changes Between Preoperative and Postoperative Values: Preoperative Median (QL, QU)	Changes Between Preoperative and Postoperative Values: Postoperative Median (QL, QU)	*P* Values of Preoperative and Postoperative Comparison	*P* Values of Intergroups Comparison
Overall Grade	A	2.2 (2.0–2.4)	0.0 (0.0–0.4)	< 0.001	*P* total < 0.01
	B	2.2 (2.0–2.4)	0.0 (0.0–0.6)	< 0.001	*P* ab > 0.05; *P* bc < 0.01
	C	2.2 (2.0–2.4)	0.4 (0.2–0.7)	< 0.001	*P* ac < 0.01
	Overall sample	2.2 (2.0–2.4)	0.2 (0.0–0.6)	< 0.001	
Jitter	A	1.87 (1.44–2.66)	0.26 (0.21–0.35)	< 0.001	*P* total < 0.001
	B	1.75 (1.33–2.34)	0.24 (0.15–0.30)	< 0.001	*P* ab > 0.05
	C	1.85 (1.48–2.05)	0.49 (0.39–0.83)	< 0.001	*P* ac < 0.05; *P* bc < 0.05
	Overall sample	1.82 (1.40–2.47)	0.27 (0.22–0.42)	< 0.001	
Shimmer	A	9.47 (8.49–10.90)	2.86 (2.44–3.61)	< 0.001	*P* total < 0.001
	B	9.55 (8.55–11.12)	3.12 (2.34–3.72)	< 0.001	*P* ab > 0.05; *P* bc < 0.05
	C	9.37 (8.37–10.87)	4.82 (4.04–5.54)	< 0.001	*P* ac < 0.05
	Overall sample	9.49 (8.49–10.90)	3.22 (2.52–4.29)	< 0.001	
NHR	A	0.18 (0.14–0.29)	0.02 (0.01–0.02)	< 0.001	*P* total < 0.001
	B	0.17 (0.14–0.25)	0.02 (0.01–0.02)	< 0.001	*P* ab > 0.05; *P* ac < 0.05
	C	0.20 (0.13–0.23)	0.04 (0.03–0.07)	< 0.001	*P* bc < 0.05
	Overall sample	0.18 (0.14–0.25)	0.02 (0.01–0.03)	< 0.001	
MPT	A	6.02 (4.68–6.79)	17.17 (14.97–21.46)	< 0.001	*P* total < 0.001
	B	5.59 (4.46–6.89)	17.37 (14.93–20.75)	< 0.001	*P* ab > 0.05; *P* ac < 0.05
	C	5.45 (4.54–6.69)	14.75 (11.10–17.22)	< 0.001	*P* bc < 0.05
	Overall sample	5.73 (4.60–6.79)	16.56 (14.42–20.33)	< 0.001	
Post-VMUR	A	2.0 (1.0–2.0)	0.0 (0.0–0.0)	< 0.001	*P* total < 0.01
	B	2.0 (1.0–2.0)	0.0 (0.0–0.0)	< 0.001	*P* ab > 0.05; *P* bc < 0.01
	C	2.0 (1.0–2.0)	0.0 (0.0–2.0)	< 0.001	*P* ac < 0.001
	Overall sample	2.0 (1.0–2.0)	0.0 (0.0–1.0)	< 0.001	
Glottal closure	A	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	*P* total > 0.05
	B	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	*P* ab > 0.05
	C	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	*P* bc > 0.05; *P* ac > 0.05
	Overall sample	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	

MPT = maximum phonation time; NHR = noise-to-harmonics ratio; QL= low quartile; QU= upper quartile; VMUR = voluntary motor-unit recruitment.
$P < 0.05$ is deemed as statistically significant.

regression analysis. In further stratification analysis, there was a significant difference between group A and C and group B and C ($P < 0.05$), but not between group A and B ($P > 0.05$).

Electromyographic Findings

All patients received laryngeal electromyographic examinations preoperatively. However, postoperative EMG results were available in only 148, 97, and 41 patients in group A, B, and C, respectively. The electrical activity in the TA muscles of the affected vocal folds during EMG in the present study was divided into two types: spontaneous activity and VMUR. Spontaneous activity, which included positive waves, fibrillations, and complex repetitive discharges, could not be recorded; and VMUR were improved in each group postoperatively ($P < 0.05$). Denervation duration was identified as a significant variable with regard to the parameter postoperative VMUR in multivariable logistic regression analysis. In further stratification analysis, there was a significant difference between group A and C and group B and C ($P < 0.05$), but not between group A and B ($P > 0.05$).

DISCUSSION

Laryngeal reinnervation with ansa cervicalis to RLN anastomosis was reintroduced by Crumley[2] and followed by many other investigators,[4–6] and it has been presented as a good procedure to achieve a normal voice quality. Previous studies in animal models[7,10,11] and in humans[12] have shown that reinnervation can potentially restore laryngeal function when it is performed immediately after RLN injury. In the clinical setting, most UVFP patients suffer from hoarseness and aspiration for several months or even years before they seek medical help. Even in UVFP patients who are diagnosed early,

surgery is usually postponed for at least 6 months to allow for possible spontaneous recovery or compensation from the contralateral vocal fold. However, the studies on regeneration of other peripheral nerves showed that the degree of functional reinnervation lessens as the period of denervation increases; and there appears to be a time range beyond which effective reinnervation declines dramatically.[13] But a significant body of evidence indicates that this does not necessarily apply to the larynx.[8,14,15] Clinical and experimental evidences have demonstrated that spontaneous regeneration commonly takes place after RLN injuries.[16–18] Although this type of reinnervation is usually nonfunctional and seldom occurs with laryngeal mobility (termed subclinical reinnervation), it can help to halt or even reverse muscle atrophy and/or fibrosis caused by denervation.[16] Therefore, the researchers think that denervation duration does not affect the surgical outcome of laryngeal reinnervation in a linear fashion. However, so far we have not seen any report regarding the stratification analysis of denervation duration on the laryngeal reinnervation effect. Thus, it is of great clinical importance to explore whether the same situation in the regeneration of other peripheral nerves also happens to the recurrent laryngeal nerve—that regeneration capacity declines progressively as the denervation duration increases.

Delayed reinnervation procedures have proven effective after peripheral nerve injury in animal experiments. For example, selective reinnervation of the posterior cricoarytenoid muscle with a phrenic nerve transfer has been feasible after a 9-month delay in cat models; however, functional recovery was less successful than with immediate reinnervation.[14] We previously reported that laryngeal reinnervation is still possible to some degree, even after an 18-month denervation period in dogs; however, the degree of RLN regeneration is less than those with an 8-month denervation period.[8] In the clinical studies, Maronian et al. reported on nine patients, eight of whom had an interval between RLN injury and surgery that exceeded 12 months. These patients had a normal or improved voice after laryngeal reinnervation. The longest denervation interval in that series was 9 years, and the postoperative voice in that case was improved.[15] Olson et al. reported excellent acoustic and perceptual results in patients with the maximal interval of 6 years between injury and surgery.[19] Our study of a large sample of UVFP patients, in which the longest denervation course was more than 3 years, confirmed that delayed reinnervation can be effective.[4] Nevertheless, the relationship between denervation duration and degree of functional recovery of the laryngeal muscle in UVFP patients remains to be elucidated.

Results of the present study showed that there was no significant difference with regard to glottal closure among the three groups. This was probably due to a lack of standardization of inspiratory effort while the patients were undergoing videostroboscopy examination. In addition, a difference in the vertical plane of the vocal folds can result in a significant glottic gap, even when the apparent closure as viewed from above seems adequate.[17] However, all of the parameters of vocal function assessment, including perceptual evaluation, objective acoustic analysis, and aerodynamics parameter MPT, showed that denervation duration was an influential factor to the surgical outcome of laryngeal reinnervation. Data of postoperative motor-unit recruitment also support the vocal function results. The perceptual and acoustic parameters showed no significant difference postoperatively among patients with denervation intervals of less than 24 months, and the parameter values in these patients were better than those in patients with longer denervation intervals. These results indicate that delayed reinnervation is still effective. There are several reasons that may support delayed laryngeal reinnervation. There may be an inherent cellular mechanism for preserving the structure of denervated laryngeal muscles.[20] Johns et al. found that 6 months after RLN resection there was no significant difference in maximal isometric force of the TA muscle between the experimental and control cats,[21] possibly due to spontaneous regeneration of the RLN. Our previous study indicated a strong tendency for regeneration in the RLN following injury, which may at least partially reinnervate the laryngeal muscle, helping to maintain its structural integrity and function and to alleviate excessive muscle atrophy and fibrosis.[16] In addition, laryngeal muscle stem cells provide persistent regenerative potential for delayed laryngeal reinnervation for up to 2 years after denervation, as revealed by our previous study.[22] The population of activated muscle stem cells in the laryngeal muscles may be more resistant to apoptosis than those in limb muscles, which may contribute to regenerative myogenesis in denervated laryngeal muscles through compensatory mechanisms. [22]

Nevertheless, after 2 years of denervation, the surgical outcomes were less favorable in the present study, although most postoperative parameters in these patients showed improvement compared with the corresponding preoperative values. As fixation of the cricoarytenoid joint was precluded preoperatively in these cases, the compromise of the recovery of voice quality might have been due to insufficient laryngeal reinnervation, which was confirmed by postoperative EMG.

One cause of poor functional recovery after excessive long-term muscle denervation is the failure of many regenerating axons to elongate and/or make synaptic connections with denervated muscle fibers. The ability of nerve sheaths to support axon regeneration to long-term denervated muscle fibers may progressively deteriorate because of: a decrease in the number of Schwann cells to a level that cannot provide adequate support for regenerating axons[23]; degeneration and collagenization of endoneurial tubes, which may obstruct axonal regeneration[24]; and an inability of the basal lamina to be renewed without Schwann cell–axon contact.[25] These factors contribute to a profound reduction in the number of axons that eventually reach denervated muscles.[26] Another possible explanation is occupation of the denervated muscle end plates by axons coming from adjacent nerves or by fibers of autonomous origin, precluding delayed reinnervation.[27] In addition, muscle fiber

atrophy and irreversible muscle fibrosis and degeneration of muscle endplates during long-term denervation may hinder successful reinnervation.[28] Even after reinnervation, muscle fibers may fail to resume their normal size, possibly because of the progressive exhaustion of satellite cells whose activity and status may be the key determinant of skeletal muscle regeneration potential. Our previous work showed that the levels of myoD and myogenin, which are markers for activated satellite cells, were upregulated 6 to 12 months after denervation and then downregulated over time, ultimately becoming undetectable by 2 years after denervation. This indicates a decreased myogenic ability after a 2-year denervation duration, [22] which might explain the better surgical outcome of laryngeal reinnervation on patients with a denervation duration less than 2 years than on patients with a denervation duration longer than 2 years. Therefore, for UVFP patients with a denervation course of more than 2 years, it may be better to combine reinnervation surgery with arytenoid adduction.[29,30]

In addition, there are some other factors that may also affect the surgical outcome of laryngeal reinnervation, such as the age of patients and the severity of nerve injury.[31–33] Crumley recommended that patients' age should be less than 70 in order to ensure the effectiveness of laryngeal reinnervation,[2] while Paniello et al. revealed that patients under age 52 had significantly better voice recovery than those over age 52.[34] Stratification analysis on age in the present series of patients revealed that laryngeal reinnervation is less effective when patients are older than 60 years old. Details of further stratification analysis on these two identified influential factors will be presented in other reports (unpublished).

CONCLUSION

The data from this study indicate that surgical outcome of laryngeal reinnervation is affected by denervation duration, the age of patients, and the severity of nerve injury. Although delayed reinnervation is effective, surgical outcome is better when the procedure is performed within 2 years after nerve injury than when the procedure is performed over 2 years.

BIBLIOGRAPHY

1. Crumley RL. Unilateral recurrent laryngeal nerve paralysis. *J Voice* 1994; 8:79–83.
2. Crumley RL. Update: ansa cervicalis to recurrent laryngeal nerve anastomosis for unilateral laryngeal paralysis. *Laryngoscope* 1991;101:384–387; discussion 388.
3. Zheng H, Li Z, Zhou S, Cuan Y, Wen W. Update: laryngeal reinnervation for unilateral vocal cord paralysis with the ansa cervicalis. *Laryngoscope* 1996;106:1522–1527.
4. Wang W, Chen D, Chen S, et al. Laryngeal reinnervation using ansa cervicalis for thyroid surgery-related unilateral vocal fold paralysis: a long-term outcome analysis of 237 cases. *PLoS One* 2011;6:e19128.
5. Lee WT, Milstein C, Hicks D, Akst LM, Esclamado RM. Results of ansa to recurrent laryngeal nerve reinnervation. *Otolaryngol Head Neck Surg* 2007;136:450–454.
6. Lorenz RR, Esclamado RM, Teker AM, et al. Ansa cervicalis-to-recurrent laryngeal nerve anastomosis for unilateral vocal fold paralysis: experience of a single institution. *Ann Otol Rhinol Laryngol* 2008;117:40–45.
7. Zheng H, Li Z, Zhou S, Cuan Y, Wen W, Lan J. Experimental study on reinnervation of vocal cord adductors with the ansa cervicalis. *Laryngoscope* 1996;106:1516–1521.
8. Zheng H, Zhou S, Chen S, Li Z, Cuan Y. An experimental comparison of different kinds of laryngeal muscle reinnervation. *Otolaryngol Head Neck Surg* 1998;119:540–547.
9. Wang W, Chen S, Chen D, et al. Contralateral ansa cervicalis-to-recurrent laryngeal nerve anastomosis for unilateral vocal fold paralysis: a long-term outcome analysis of 56 cases. *Laryngoscope* 2011;121:1027–1034.
10. Mahieu HF, van Lith-Bijl JT, Groenhout C, Tonnaer JA, de Wilde P. Selective laryngeal abductor reinnervation in cats using a phrenic nerve transfer and ORG 2766. *Arch Otolaryngol Head Neck Surg* 1993;119: 772–776.
11. van Lith-Bijl JT, Stolk RJ, Tonnaer JA, Groenhout C, Konings PN, Mahieu HF. Selective laryngeal reinnervation with separate phrenic and ansa cervicalis nerve transfers. *Arch Otolaryngol Head Neck Surg* 1997; 123:406–411.
12. Yumoto E, Sanuki T, Kumai Y. Immediate recurrent laryngeal nerve reconstruction and vocal outcome. *Laryngoscope* 2006;116:1657–1661.
13. Keilhoff G, Fansa H. Successful intramuscular neurotization is dependent on the denervation period. A histomorphological study of the gracilis muscle in rats. *Muscle Nerve* 2005;31:221–228.
14. van Lith-Bijl JT, Stolk RJ, Tonnaer JA, Groenhout C, Konings PN, Mahieu HF. Laryngeal abductor reinnervation with a phrenic nerve transfer after a 9-month delay. *Arch Otolaryngol Head Neck Surg* 1998; 124:393–398.
15. Maronian N, Waugh P, Robinson L, Hillel A. Electromyographic findings in recurrent laryngeal nerve reinnervation. *Ann Otol Rhinol Laryngol* 2003;112:314–323.
16. Chen D, Chen S, Wang W, Zhang C, Zheng H. Spontaneous regeneration of recurrent laryngeal nerve following long-term vocal fold paralysis in humans: histologic evidence. *Laryngoscope* 2011;121:1035–1039.
17. Woodson GE. Spontaneous laryngeal reinnervation after recurrent laryngeal or vagus nerve injury. *Ann Otol Rhinol Laryngol* 2007;116:57–65.
18. Kupfer RA, Old MO, Oh SS, Feldman EL, Hogikyan ND. Spontaneous laryngeal reinnervation following chronic recurrent laryngeal nerve injury. *Laryngoscope* 2013;123:2216–2227.
19. Olson DE, Goding GS, Michael DD. Acoustic and perceptual evaluation of laryngeal reinnervation by ansa cervicalis transfer. *Laryngoscope* 1998; 108:1767–1772.
20. Miyamaru S, Kumai Y, Ito T, Yumoto E. Effects of long-term denervation on the rat thyroarytenoid muscle. *Laryngoscope* 2008;118:1318–1323.
21. Johns MM, Urbanchek M, Chepeha DB, Kuzon WM Jr, Hogikyan ND. Thyroarytenoid muscle maintains normal contractile force in chronic vocal fold immobility. *Laryngoscope* 2001;111:2152–2156.
22. Donghui C, Shicai C, Wei W, et al. Functional modulation of satellite cells in long-term denervated human laryngeal muscle. *Laryngoscope* 2010; 120:353–358.
23. Salonen V, Lehto M, Kalimo M, Penttinen R, Aro H. Changes in intramuscular collagen and fibronectin in denervation atrophy. *Muscle Nerve* 1985;8:125–131.
24. Sunderland S, Bradley KC. Endoneurial tube shrinkage in the distal segment of a severed nerve. *J Comp Neurol* 1950;93:411–420.
25. Bunge MB, Williams AK, Wood PM. Neuron-Schwann cell interaction in basal lamina formation. *Dev Biol* 1982;92:449–460.
26. Fu SY, Gordon T. Contributing factors to poor functional recovery after delayed nerve repair: prolonged denervation. *J Neurosci* 1995;15:3886–3895.
27. Hydman J, Remahl S, Bjorck G, Svensson M, Mattsson P. Nimodipine improves reinnervation and neuromuscular function after injury to the recurrent laryngeal nerve in the rat. *Ann Otol Rhinol Laryngol* 2007; 116:623–630.
28. Anzil AP, Wernig A. Muscle fibre loss and reinnervation after long-term denervation. *J Neurocytol* 1989;18:833–845.
29. Chhetri DK, Gerratt BR, Kreiman J, Berke GS. Combined arytenoid adduction and laryngeal reinnervation in the treatment of vocal fold paralysis. *Laryngoscope* 1999;109:1928–1936.
30. Hassan MM, Yumoto E, Kumai Y, Sanuki T, Kodama N. Vocal outcome after arytenoid adduction and ansa cervicalis transfer. *Arch Otolaryngol Head Neck Surg* 2012;138:60–65.
31. Korte N, Schenk HC, Grothe C, Tipold A, Haastert-Talini K. Evaluation of periodic electrodiagnostic measurements to monitor motor recovery after different peripheral nerve lesions in the rat. *Muscle Nerve* 2011;44:63–73.
32. Kovacic U, Sketelj J, Bajrovic FF. Chapter 26: Age-related differences in the reinnervation after peripheral nerve injury. *Int Rev Neurobiol* 2009; 87:465–482.
33. van Lith-Bijl JT, Mahieu HF, Stolk RJ, Tonnaer JA, Groenhout C, Konings PN. Laryngeal abductor function after recurrent laryngeal nerve injury in cats. *Arch Otolaryngol Head Neck Surg* 1996;122:393–396.
34. Paniello RC, Edgar JD, Kallogjeri D, Piccirillo JF. Medialization versus reinnervation for unilateral vocal fold paralysis: a multicenter randomized clinical trial. *Laryngoscope* 2011;121:2172–2179.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 153(3):414-419.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–
Head and Neck Surgery
2015, Vol. 153(3) 414–419
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599815593268
http://otojournal.org
SAGE

Diagnostic Yield of Computed Tomography in the Evaluation of Idiopathic Vocal Fold Paresis

Paul M. Paddle, MD, MPH[1], Masaany B. Mansor, MD[1], Phillip C. Song, MD[1], and Ramon A. Franco Jr, MD[1]

No sponsorships or competing interests have been disclosed for this article.

Received October 26, 2014; revised May 6, 2015; accepted June 5, 2015.

Abstract

Objective. To determine the diagnostic yield of computed tomography (CT) in establishing an etiology in patients with idiopathic unilateral vocal fold paresis (IUVFP). To determine the proportion of CT scans yielding incidental findings requiring further patient management.

Study Design. Case series with chart review.

Setting. Tertiary laryngology practice.

Subjects. Laryngology clinic patients under the care of the 2 senior authors.

Methods. All clinic patients were identified who had a diagnosis of IUVFP and underwent CT of the skull base to the upper mediastinum from 2004 to 2014. Demographic, historical, examination, and investigation data were extracted. CT reports and endoscopic recordings were reviewed. Patients were excluded if there were insufficient clinical findings recorded or if there was a known neurologic disorder, complete vocal fold immobility, or bilateral involvement.

Results. A total of 174 patients with IUVFP who had also undergone contrast-enhanced CT were identified. Of the 174 patients, 5 had a cause for their paresis identified on CT. This equated to a diagnostic yield of 2.9% (95% confidence interval, 0.94% to 6.6%). Of the 174 patients, 48 had other incidental lesions identified that required further follow-up, investigation, or treatment. This equated to an incidental yield of 27.6% (95% confidence interval, 21.1% to 34.9%).

Conclusion. This is the second and largest study to evaluate the diagnostic yield of CT in the evaluation of IUVFP. It demonstrates a low diagnostic yield and a high incidental yield. These findings suggest that the routine use of CT in the evaluation of idiopathic vocal fold paresis should be given careful consideration and that a tailored approach to investigation with good otolaryngologic follow-up is warranted.

Keywords

idiopathic unilateral vocal fold paresis, computed tomography, diagnostic yield, incidental yield

Vocal fold paresis implies vocal fold hypomobility due to neurologic injury, with a peripheral etiology in 90% of cases. It may result from weakness of the vagus nerve or its superior or recurrent laryngeal branches. This may occur anywhere in its course—from the lower motor neurons in the nucleus ambiguus of the medulla through the jugular foramen, neck, and mediastinum. Vocal fold paresis is unilateral in 90% of cases.[1]

Paresis is the most common cause of vocal fold hypomobility, being present in 90% of cases.[1] Other causes of vocal fold hypomobility include myopathies (4%) and cricoarytenoid joint dysfunction (6%). Paresis is an increasingly recognized phenomenon in patients with laryngologic complaints. Previous studies reported mild vocal fold hypomobility in 46% of patients with vocal complaints, 71% of singing teachers with complaints of technical difficulties, and 23% of singing teachers with no vocal complaints.[1,2]

Diagnosis of vocal fold paresis requires a high index of suspicion. The symptoms of vocal fold paresis are more varied and subtle than paralysis. Classic symptoms of glottic insufficiency—such as breathy dysphonia, diplophonia, aspiration, and dysphagia—may be absent or muted. Instead, the patient may complain of a loss of quality volume and range, vocal instability, and increased phonatory effort.[3] Atypical symptoms, such as globus, chronic cough, and laryngospasm, are also described.

Unlike the findings of an established vocal fold paralysis, such as vocal fold atrophy, bowing, and arytenoid prolapse, the examination findings of unilateral vocal fold paresis are subtle and difficult to discern from nonpathologic asymmetries. Findings may include asymmetric vocal fold range and velocity of movement, decreased ipsilateral false vocal

[1]Harvard Medical School, Division of Laryngology, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, USA

Corresponding Author:
Paul M. Paddle, MD, Harvard Medical School, Division of Laryngology, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA.
Email: paulpaddle@me.com

fold and supraglottic contraction, pushing rotation of the petiole toward the weakened side, and asymmetric fatigability on repetitive movement. Findings associated with compensatory hyperfunction may also be present, including contralateral supraglottic contraction, and benign vocal fold lesions such as nodules, polyps, or pseudocysts.[4] On stroboscopy, asymmetry of mucosal wave motion may be the only manifestation. The diagnostic accuracy of these signs is controversial, but when they are identified on nasoendoscopy, along with a suggestive history, a diagnosis of paresis is made.

Idiopathic paresis is diagnosed when no cause is found on thorough history and examination. In our clinic, paresis with a history of preceding upper respiratory infection is defined as idiopathic, as there is no definitive test to confirm causality.

Laryngeal electromyography (LEMG) is used as a diagnostic and prognostic tool in cases of vocal fold paralysis. There is no consensus on the use of LEMG in the context of vocal fold hypomobility. When performed correctly, LEMG can confirm the presence and laterality of a neuropathy and identify neuromuscular junction abnormalities and myopathies, as well as ongoing degeneration or regeneration. Some advocates propose that it be employed systematically in paresis,[3] while others use it in situations where the results would alter patient management.[5] Certainly, it does not obviate the role of imaging studies in the evaluation of vocal fold paresis.

The rationale of imaging in paresis is twofold: First, paresis may be an early sentinel of an underlying pathology that, where identified, would require further investigation and management in its own right, particularly neoplasia. Second, finding an underlying pathologic process may guide management of the paresis itself. The role of computed tomography (CT) in the evaluation of vocal fold paralysis is well established, given a high overall diagnostic yield (35% to 62%)[6,7] and a high proportion of neoplastic causes (13% to 33%).[8-11] However, its role in the evaluation of paresis is not clearly established, and current practice seems to be extrapolated from the paralysis literature. A single previous study assessed the diagnostic yield of CT in the investigation of paresis.[12] In our institution, CT is performed when there is a clinical diagnosis of paresis but the cause remains ''idiopathic'' after thorough history and examination—that is, no clear history of preceding nerve injury or other compressive or infiltrative lesion and no evidence of a cause on otolaryngologic, neurologic, and chest examination and video endoscopy. Patients may also refuse or strongly desire a CT study.

Our study objectives are twofold: first, to establish a diagnostic yield in performing CT in patients with idiopathic vocal fold paresis; second, to establish a percentage yield of incidental lesions requiring further management in this cohort of patients. This has important clinical, cost, and medicolegal implications.

Method

This study was approved by the Massachusetts Eye and Ear Infirmary Institutional Review Board. With a precision-based sample size calculation based on an expected diagnostic yield of approximately 2.0%,[12] an acceptable precision of 1.99%, and a confidence level of 95%, an estimated 191 patients were required. The practice records from January 2004 to January 2014 of 2 senior laryngologists from a single tertiary practice were reviewed. All adult patients were identified who had a clinical diagnosis of idiopathic unilateral vocal fold paresis (IUVFP) and underwent contrast-enhanced CT from skull base to mediastinum. Patients were excluded if they had bilateral vocal fold hypomobility due to the decreased reliability of clinical assessment and the higher likelihood of a central etiology.[13] Patients were also excluded if there was a history of a neurologic diagnosis, such as laryngeal dystonia or tremor, myoclonus, parkinsonism, stroke, or other central neurologic process.

In each case, a diagnosis of paresis was made by a senior laryngologist, using the above-described symptoms and signs. CT images and reports were reviewed. In each case, any etiology for paresis and any incidental finding were recorded. An etiology for paresis was defined as any lesion along the expected extracranial course of the ipsilateral superior or recurrent laryngeal nerve or vagus, which could be causing pathologic compression, invasion, stretch, or inflammation. An incidental lesion was defined as any clinically silent lesion, not associated with the diagnosis of paresis, but that could lead to further diagnostic or therapeutic intervention.

Longitudinal review of files was also undertaken to identify evolution of findings or interval evidence of an etiology for the diagnosed paresis. Due to the evolution in endoscopic diagnostic criteria for paresis over the 10 years of the study, a sensitivity analysis was performed comparing the mean diagnostic yield of the first 5 years with that of the second 5 years. The null hypothesis of no difference between the means was tested with an unpaired 2-sample *t* test. Excel 2010 and Stata 10.0 were used for data storage and statistical analysis.

Results

Patients (n = 237) with unilateral paresis were identified over the period January 2004 to January 2014. Of these, 174 (73%) underwent contrast-enhanced CT scans of skull base to mediastinum and were included in the study. The other 63 (27%) did not undergo CT due to either a clear etiology of their paresis or patient refusal. There was no systematic difference in the demographic characteristics of the CT and non-CT workup patient populations. In the CT workup group of patients, the mean age at diagnosis was 54.5 years (range, 21 to 82). There were a greater proportion of women (56%), while laterality of paresis was evenly distributed, with 51% of lesions being left sided (**Table 1**).

The most common patient symptom was hoarseness. Symptoms of glottic inefficiency were also common, such as vocal fatigue, increased phonatory effort, and decreased projection (**Table 2**). Patients less commonly complained of loss of range, cough, laryngospasm, globus/dysphagia, and pain.

Table 1. Demographic Data of IUVFP by Workup Groups: CT and Non-CT.[a]

	CT	Non-CT
Total unilateral paresis	174	63
Age, y[b]	54.5 (21-82)	53.9 (18-75)
Women	98 (56)	33 (52)
Left laterality	88 (51)	33 (52)

Abbreviations: CT, computed tomography; IUVFP, idiopathic unilateral vocal fold paresis.
[a]Results presented as n (%), except where noted otherwise.
[b]Mean (range).

Table 2. Symptoms of Paresis Subjects Included in this Study.

Symptom	n (%)
Hoarseness	144 (83)
Vocal fatigue	113 (65)
Increased phonatory effort	67 (39)
Decreased vocal projection	63 (36)
Loss of range	45 (26)
Cough	44 (25)
Dysphagia	37 (21)
Pain: odynophagia / odynophonia / laryngeal strain	20 (11)
Breathlessness during voicing	13 (7)
Laryngospasm	11 (6)

Table 3. Videostroboscopic Findings of Paresis Subjects Included in this Study.

Examination Feature	n (%)
Asymmetry of velocity of movement	138 (79)
Increased glottic show	129 (74)
Asymmetry of range of movement	117 (67)
Phase asymmetry	91 (66)
Supraglottic hyperfunction	72 (41)
Incomplete closure	65 (37)
Bowing/atrophy of vocal fold	45 (26)
Deviation	26 (15)
Increased vibratory amplitude	17 (12)

On video endoscopy, common findings were subtle vocal fold range-of-motion asymmetries, asymmetrically increased glottic show, and asymmetric velocity of motion. These were seen in 79%, 74%, and 67% of patients, respectively. Asymmetries at rest were less prevalent as were signs of bowing and incomplete closure (**Table 3**).

Of the 174 patients, 5 had CT that revealed an etiologic lesion for their vocal fold paresis, a diagnostic yield of 2.9% (95% confidence interval, 0.94 to 6.6; **Table 4**). Of these 5 CT-positive cases, 1 was positive for malignancy. The patient had an exophytic thyroid nodule with possible compression of the recurrent laryngeal nerve in the ipsilateral tracheoesophageal groove. This nodule was positive for papillary thyroid carcinoma on fine-needle aspiration. The patient underwent total thyroidectomy and adjuvant radioactive iodine. His paresis did not improve on serial follow-up.

Four CT-positive cases were benign: 1 was due to previous thoracic aortic aneurysm repair with dense scarring on CT in the aortopulmonary window. Two cases were due to tracheoesophageal groove masses. One mass was an exophytic thyroid nodule and associated tracheoesophageal groove lymph node. The patient underwent a right hemithyroidectomy and prelaryngeal lymph node dissection. The final histopathology was a benign follicular adenoma. The other tracheoesophageal groove case was due to a large parathyroid adenoma that, on removal, was seen to be stretching the recurrent laryngeal nerve. A final case was due to an undiagnosed Arnold Chiari II malformation with tentorium crowding and tonsillar herniation. This patient was referred to neurosurgery and underwent urgent posterior fossa decompression. It is interesting to note that there were no other neurologic symptoms or signs nor evidence of bilateral paresis. In all of the above 3 benign cases undergoing surgery, there was no recovery of function of the nerve after surgical intervention. The diagnostic yield equates to a number needed to treat of 34. In other words, to find 1 patient with a vocal fold paresis-associated lesion, 34 patients had to undergo CT.

In contrast, 48 of 174 patients had a new incidental finding on CT that required further management. Further management was defined as serial clinical examination, repeat imaging, a diagnostic procedure, or operation. This equates to an incidental yield of 27.6% (95% confidence interval, 23.7% to 37.8%).

Of these 48 patients, 40 underwent clinical and or serial imaging follow-up alone; 5 underwent fine-needle aspiration alone; and 3 underwent surgery. The range of incidental lesions included pulmonary nodules, thyroid nodules, and other mediastinal and cervical lesions, predominantly lymphadenopathy (**Table 5**). Over the mean 2.95 years of follow-up (SD, 1.52), none of these patients developed a symptomatic or clinically significant pathology. Of the 3 patients who underwent surgery, 1 underwent hemithyroidectomy for a follicular adenoma that had no extracapsular extension and was not compressing on the tracheoesophageal groove. A second patient underwent total thyroidectomy for a dominant intrathyroid nodule that was positive for papillary carcinoma on fine-needle aspiration, and the third patient underwent resection of a benign, submucosal false fold lipoma. The number needed to ''harm'' was 4.

A sensitivity analysis of diagnostic yield revealed a yield of 2.2% for the first 5 years of the study, compared with a yield of 5.1% for the second 5 years of the study. An unpaired 2-sample *t* test of the difference between these 2 means (2.9%) resulted in a *P* value of .34.

Table 4. Yield of CT in IUVFP.

Yield	Positive CT, n	Yield, % (95% CI)	Needed to Treat/Harm, n
Diagnostic	5[a]	2.9 (0.94-6.6)	34
Incidental	48[b]	27.6 (21.1-34.9)	4

Abbreviations: CI, confidence interval; CT, computed tomography; IUVFP, idiopathic unilateral vocal fold paresis.
[a]Diagnostic of benign lesion, n = 4; diagnostic of a malignancy, n = 1.
[b]Underwent clinical or imaging follow-up alone, n = 40; underwent fine-needle aspiration alone, n = 5; required operation, n = 3.

Table 5. Proportion of Incidental Lesions by Anatomic Group.

Incidental Lesion Type	Proportion of All Incidental Lesions
Thyroid abnormalities: nodules, cysts, enlargement	29
Pulmonary lesions: nodules, granulomas, pleural plaques, hilar lymphadenopathy	40
Mediastinal lesions: thoracic aortic aneurysms, mediastinal lymphadenopathy	15
Cervical abnormalities: laryngocele, thyroglossal duct cyst, cervical lymphadenopathy	13
Miscellaneous: vertebral lesions	4

Table 6. Etiology of Vocal Fold Paralysis vs Paresis (in Percentages).

	Paralysis	Paresis			
Etiology	MacGregor[10] (n = 1308)	Koufman[4] (n = 50)	Heman-Ackah[1] (n = 46)	Badia[12] (n = 176)	Present Study (n = 237)
Total iatrogenic	22	20	4.3	39.2	3.4
Total neoplastic lesions	21.7	6	13	1.1	2.9
Total nonneoplastic benign disease	39.9	6	54.3	13.6	5.7
Idiopathic[a]	16.4	68	28.3	46	88

[a]Includes viral neuritis.

Discussion

The role of CT in the evaluation of vocal fold paralysis is well established. Its near routine use is justified by a high diagnostic yield (35% to 62%)[6,7] and a high proportion of cases due to underlying neoplasia (13% to 33%).[8,10] Its role in paresis, however, is unclear, partly because the prevalence of a neoplastic etiology in published studies is a comparatively low (1.1% to 6%; **Table 6**)[1,4,12] and partly because there is a paucity of studies examining the diagnostic yield of CT paresis. A previous study found the diagnostic yield of CT in IUVFP to be 1.7%.[12]

The overall diagnostic yield of CT in the evaluation of paresis in the current study was 2.9% (95% confidence interval, 0.94% to 6.6%). Such values may justify the use of routine CT evaluation. A similar percentage yield (1% to 4%) is seen by many authors as justification for the routine use of magnetic resonance imaging in the evaluation of asymmetric sensorineural hearing loss.[14] This argument for routine CT in paresis might be strengthened from a qualitative perspective, when one considers that 4 of 5 (80%) of our CT-positive cases had a pathology that required operation, which untreated may have led to serious morbidity. Conversely, the study by Badia et al[12] described a final yield of 0%, as the single CT-positive case was benign and required no intervention.

The argument against the routine use of CT in the evaluation of IUVFP is strengthened when one considers the discovery of incidental lesions and the potential morbidities due to the performance of additional diagnostic tests and interventions. This study identified an incidental yield of 27.6%. The majority of these patients with incidental imaging findings (83.3%) underwent serial examination and imaging alone (including serial thyroid ultrasound, CT of the neck and chest, and magnetic resonance imaging of the brain), thus exposing them to additional perhaps unnecessary radiation and expense. In all of these imaging-alone cases, the incidental lesions did not evolve into clinically significant pathologies over the mean laryngologic follow-up period of 2.95 years (SD, 1.52). Five patients underwent

fine-needle aspiration for thyroid nodules. In all but 1 case, these lesions were benign, and no direct complications from the aspiration procedure were recorded. Two patients underwent surgery for benign, otherwise asymptomatic lesions. In both cases, there were no complications from surgery. Only 1 of the 48 incidental cases was an incidental finding clearly beneficial to the patient. In this case, the patient had a serendipitous incidental finding of a thyroid malignancy, which was successfully treated with total thyroidectomy. Although it is not possible to definitively state how many of these lesions may have become clinically significant at a later date, none of the lesions became symptomatic or required further action over the period of follow-up. And while it can be argued that the incidental discovery of an early thyroid malignancy or a small descending thoracic aortic aneurysm is beneficial for the patient and may lead to improved survival, it cannot be argued that this justifies routine CT in the evaluation of IUVFP. One must be clear that, in this context, CT investigation is a diagnostic test, trying to answer the question of etiology of paresis, and is not a screening investigation.

While not a focus of this study, the underlying etiologies of vocal fold paresis further inform the finding of a low diagnostic yield. Compared with paresis, a larger proportion of paralysis cases is due to neoplastic causes, and a smaller proportion is idiopathic.[8,15] In the 4 published studies (including this one), there is an increasing proportion of idiopathic cases and a much smaller proportion of cases due to neoplasia. In our study, the proportion of idiopathic cases that remained idiopathic was 88%, and the proportion of cases due to neoplasia was 2.9%. Thus, there is already a low likelihood that CT will uncover an underlying malignancy.

This is the largest published study to evaluate the role of CT imaging in the investigation of unilateral vocal fold paresis. It has an adequate sample size for the desired precision and has a rigorous study method. Only 1 other study addressed the diagnostic yield of CT in the evaluation of IUVFP. Other strengths of this study include the long-term follow-up of patients with incidental findings and the use of 2 laryngologists from a single institution. Given the controversy in the literature in making a clinical diagnosis of paresis, having only 2 assessors may limit generalizability, but this singular clinical definition of paresis maximizes internal validity.

A possible weakness of this study is the lack of routine LEMG in the confirmation of a diagnosis of paresis, possibly leading to misclassification bias. Certain authors suggested that LEMG is an essential diagnostic tool in the evaluation of paresis and espoused its use in every case,[3,4] emphasizing that there is a marked discrepancy between clinical observations of paresis and LEMG findings, with discordance in 25% to 40% of cases.[2,4] This discrepancy, however, is not in the absence or presence of neuropathy in the larynx but on the paretic side and the nerve involved. In fact, there is excellent concordance between a clinical diagnosis of paresis and LEMG findings of the presence of a neuropathy, with 1 study demonstrating LEMG confirmation of a clinically diagnosed mild paresis in 86.4% of cases[1,5] and another demonstrating an LEMG-confirmed neuropathy in 100% of clinically diagnosed pareses. Dursun et al demonstrated that a thorough neurolaryngeal examination in the hands of an experienced laryngologist can diagnose superior laryngeal nerve paresis in 98% of cases with characteristic examination findings.[16] Merati et al demonstrated that 92% of patients with clinical vocal fold motion impairment had a neuropathy.[17] Furthermore, the ability to obtain reliable and accurate LEMG results in the larynx requires a significant degree of experience and expertise, specifically in the interpretation of laryngeal data, and it is heavily dependent on accurate and consistent needle placement, which limits its usefulness and availability in many centers. Last, it is an invasive test, not without attendant morbidity, and it does not often address the etiology of the paresis. Thus, it would not obviate the need for further imaging and laboratory testing in confirmed cases.

A possible source of selection bias lies in the fact that 27% of paresis cases did not undergo CT. These cases likely had a clear etiology on history and examination, such as preceding surgical injury or known cervicothoracic malignancy. While this may raise the diagnostic yield of CT by excluding cases with a higher likelihood of a negative finding, they are not true idiopathic cases, and CT in this context would not likely provide additional diagnostic information, which is the key clinical question.

Last, a possible source of misclassification bias lies in the fact that the clinical diagnosis of vocal fold paresis has evolved over the 10-year sample period. That is, we have had a higher index of suspicion and perhaps a lower threshold for the diagnosis of paresis in recent years, and patients with more subtle findings of paresis may have been excluded in earlier years of the study. However, a sensitivity analysis was performed comparing the diagnostic yields of the first 5 years of the study (2.2%) and the second 5 years (5.1%), and it revealed no statistically significant difference between these 2 periods (2.9%, $P = .34$).

In the light of the above findings, our current investigation of IUVFP involves a tailored approach. Patients who have additional localizing symptoms and examination findings are investigated in a targeted manner with imaging and appropriate blood tests. In true clinically idiopathic cases, we now offer the patient an informed choice between repeat videostroboscopic evaluation and initial CT, having discussed the diagnostic and incidental yields of imaging. If there is a convincing history of a significant preceding upper respiratory infection and definitive coincident onset of laryngologic symptoms, we are more likely to advise repeat observation over up-front CT, as the underlying etiology is more likely a viral neuritis. Reliability of follow-up is also an important consideration.

On serial examination, if a paresis evolves into paralysis or bilateral findings, our suspicion of an underlying nefarious lesion is heightened, and we then perform CT and other investigations as appropriate. Future prospective multicenter studies validating clinical diagnosis of paresis against laryngeal EMG and CT are warranted.

Conclusion

This study demonstrates that contrast-enhanced CT of the base of the skull to the upper mediastinum has a low yield of 2.9% in the initial evaluation of idiopathic vocal fold paresis. CT has a high rate (27.6%) of incidental asymptomatic cervicothoracic findings. These findings suggest that routine use of CT in the evaluation of idiopathic vocal fold paresis should be given careful consideration and that a tailored approach to investigation with good otolaryngologic follow-up is warranted.

Author Contributions

Paul M. Paddle, study design, data acquisition, analysis and interpretation, drafting and revision, final approval and accountability for all aspects of the work; **Masaany B. Mansor**, study design, data acquisition, manuscript drafting, final approval and accountability for all aspects of the work; **Phillip C. Song**, study design, manuscript revision, final approval and accountability for all aspects of the work; **Ramon A. Franco Jr**, study design, data interpretation, manuscript revision, final approval and accountability for all aspects of the work.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Heman-Ackah YD, Batory M. Determining the etiology of mild vocal fold hypomobility. *J Voice*. 2003;17:579-588.
2. Heman-Ackah YD, Barr A. Mild vocal fold paresis: understanding clinical presentation and electromyographic findings. *J Voice*. 2006;20:269-281.
3. Sulica L, Blitzer A. Vocal fold paresis: evidence and controversies. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:159-162.
4. Koufman JA, Postma GN, Cummins MM, Blalock PD. Vocal fold paresis. *Otolaryngol Head Neck Surg*. 2000;122:537-541.
5. Heman-Ackah YD, Barr A. The value of laryngeal electromyography in the evaluation of laryngeal motion abnormalities. *J Voice*. 2006;20:452-460.
6. Terris DJ, Arnstein DP, Nguyen HH. Contemporary evaluation of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1992;107:84-90.
7. Ramadan HH, Wax MK, Avery S. Outcome and changing cause of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1998;118:199-202.
8. Yumoto E, Minoda R, Hyodo M, Yamagata T. Causes of recurrent laryngeal nerve paralysis. *Auris Nasus Larynx*. 2002; 29:41-45.
9. Rosenthal LH, Benninger MS, Deeb RH. Vocal fold immobility: a longitudinal analysis of etiology over 20 years. *Laryngoscope*. 2007;117:1864-1870.
10. MacGregor FB, Roberts DN, Howard DJ, Phelps PD. Vocal fold palsy: a re-evaluation of investigations. *Journal Laryngol Otol*. 1994;108:193-196.
11. Richardson BE, Bastian RW. Clinical evaluation of vocal fold paralysis. *Otolaryngol Clin North Am*. 2004;37:45-58.
12. Badia PI, Hillel AT, Shah MD, Johns MM 3rd, Klein AM. Computed tomography has low yield in the evaluation of idiopathic unilateral true vocal fold paresis. *Laryngoscope*. 2013; 123:204-207.
13. Hillel AD, Benninger M, Blitzer A, et al. Evaluation and management of bilateral vocal cord immobility. *Otolaryngol Head Neck Surg*. 1999;121:760-765.
14. Jiang ZY, Mhoon E, Saadia-Redleaf M. Medicolegal concerns among neurotologists in ordering MRIs for idiopathic sensorineural hearing loss and asymmetric sensorineural hearing loss. *Otol Neurotol*. 2011;32:403-405.
15. Altman JS, Benninger MS. The evaluation of unilateral vocal fold immobility: is chest X-ray enough? *J Voice*. 1997;11:364-367.
16. Dursun G, Sataloff RT, Spiegel JR, Mandel S, Heuer RJ, Rosen DC. Superior laryngeal nerve paresis and paralysis. *J Voice*. 1996;10:206-211.
17. Merati AL, Shemirani N, Smith TL, Toohill RJ. Changing trends in the nature of vocal fold motion impairment. *Am Journal Otolaryngol*. 2006;27:106-108.

Reprinted by permission of Otolaryngol Head Neck Surg. 2014; 151(2):286-293.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–Head and Neck Surgery
2014, Vol. 151(2) 286–293
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2014
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599814531733
http://otojournal.org
SAGE

Etiology and Time to Presentation of Unilateral Vocal Fold Paralysis

Emily A. Spataro, MD[1], David J. Grindler, MD[1], and Randal C. Paniello, MD[1]

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract

Objective. To determine the etiology, laterality, and time to presentation of unilateral vocal fold paralysis (UVFP) at a tertiary care institution over 10 years.

Study Design. Case series with chart review.

Setting. Academic medical center.

Subjects and Methods. All patients seen between 2002 and 2012 by the Department of Otolaryngology at the Washington University School of Medicine (WUSM), with a diagnosis of unilateral vocal fold paralysis, were included. Medical records were reviewed for symptom onset date, presentation date(s), and etiology of UVFP.

Results. Of the patients, 938 met inclusion criteria and were included. In total, 522 patients (55.6%) had UVFP due to surgery; 158 (16.8%) were associated with thyroid/parathyroid surgery, while 364 (38.8%) were due to nonthyroid surgery. Of the patients, 416 (44.4%) had nonsurgical etiologies, 124 (13.2%) had idiopathic UVFP, and 621 (66.2%) had left-sided UVFP. The diagnosis was more common on the left side in cases of intrathoracic surgeries and malignancies, as expected, but also in idiopathic, carotid endarterectomy, intubation, and skull base tumors. In total, 9.8% of patients presented first to an outside otolaryngologist at a median time of 2.1 months after onset, but these patients presented to WUSM at a median time of 9.5 months. Overall, 70.6% of patients presented to a WUSM otolaryngologist within 3 months of onset.

Conclusion. Iatrogenic injury remains the most common cause of UVFP. Thyroidectomy remains the leading cause of surgery-related UVFP. Patients are typically seen within 3-4 months of onset; however, a significant delay exists for those referred to WUSM.

Keywords

vocal cord, paralysis, etiology

Received October 31, 2013; revised February 3, 2014; accepted March 25, 2014.

The etiology of unilateral vocal fold paralysis (UVFP) is of great interest to the otolaryngologist and has been reported in many studies over the past 40 years.[1-19] Etiologies include thyroid surgery, nonthyroid surgery, trauma, neurologic disease, malignancy, intubation, infection, inflammatory diseases, and idiopathic causes. Among past studies, there is great discrepancy between the most common causes of vocal fold paralysis. The most common have included lung malignancies,[1-5] idiopathic causes,[2,6-10] thyroid surgery,[11-16] and nonthyroid surgeries.[15-18] In 2 recent large retrospective chart review studies, thyroid surgery was the single most common cause of UVFP, but nonthyroid surgeries as a group more commonly cause UVFP.[15,16]

The etiology of UVFP is important because it affects the natural course, treatment, and outcome of the condition. Both the mechanism and degree of injury are important, ranging from neuropraxia, where complete recovery is expected, to complete transection, which may require surgical intervention.[20] Outcomes are affected by contralateral vocal fold compensation, as well as the degree of reinnervation and synkinesis established.[20] In a recent review of idiopathic UVFP, most improvement of vocal fold function and voice occurred within the first year of injury.[21]

Treatment of UVFP includes voice therapy, permanent and nonpermanent medialization procedures, and reinnervation. If the etiology suggests the nerve was not transected, then some degree of recovery of laryngeal nerve function is expected, and nonpermanent treatments are generally recommended until 6 to 12 months after onset of paralysis, whereas if complete transection has occurred, permanent medialization or reinnervation procedures may be undertaken sooner.[20] In addition, voice and airway are affected by the degree of synkinesis present. Synkinesis is caused by

[1]Department of Otolaryngology–Head and Neck Surgery, Washington, University School of Medicine, St Louis, Missouri, USA

This article was presented at the 2013 AAO-HNSF Annual Meeting & OTO EXPO; September 29–October 3, 2013; Vancouver, British Columbia, Canada.

Corresponding Author:
Randal C. Paniello, MD, Department of Otolaryngology–Head and Neck Surgery, Washington University School of Medicine, 660 S. Euclid Avenue, Campus Box 8115, St Louis, MO 63110, USA.
Email: paniellor@ent.wustl.edu

reinnervation of opposing muscle groups by the same nerve, leading the muscles to contract simultaneously. In canine models, reinnervating axons begin reaching the vocal fold muscles within 3 months of injury. Therefore, treatments to prevent unfavorable synkinesis would need to be administered within this time frame.[22,23]

The objective of this study was to determine how often patients with UVFP present to the Washington University School of Medicine (WUSM) within 3 to 4 months of symptom onset. These patients might be eligible for a clinical trial of early intervention for the prevention of synkinesis. A review of literature revealed no previous studies focusing on time of presentation of patients with UVFP. Etiology and laterality data were also collected and reported.

Methods

Approval for the study was obtained from the WUSM Institutional Review Board. Adult patients (>18 years of age) seen between January 1, 2002, and January 1, 2012, with a diagnosis of unilateral vocal fold paresis or paralysis (based on *International Classification of Diseases, Ninth Revision* and *Current Procedural Terminology* codes) were identified via a query of electronic medical records. In most cases, the diagnosis of UVFP was made by the attending physician based only on examination of the larynx; electromyography or cricoarytenoid joint palpation was performed in only a small percentage of patients. In all cases, the diagnosis was confirmed by flexible fiberoptic examination or videostroboscopy.

Each patient's chart was reviewed and data were collected and stored in an electronic database. Data included age at initial visit, sex, date of visit to a WUSM otolaryngologist, date of presentation to a non-WUSM otolaryngologist (in any), primary symptom, date of symptom onset, side of paralysis, etiology of paralysis, initial and subsequent treatment received, date of initial treatment, voice improvement, and fiberoptic examination vocal fold movement outcomes. From this information, the time intervals from symptom onset to initial presentation to an otolaryngologist and the time interval from initial presentation to referral to WUSM (if made), were calculated. The specific reason for referral to the university was not usually recorded in the chart, but very few had been previously treated. Diagnosis of idiopathic UVFP was confirmed by negative imaging along the course of the vagus and recurrent laryngeal nerves. Patients with incomplete records were excluded.

Data were analyzed by first determining percentages of men and women in the study population, laterality of UVFP, and etiology. To evaluate laterality, a calculation for left-right skew was devised as follows:

$$\text{Skew} = (\#\ \text{cases on right}/\text{total}\ \#\ \text{cases}) - 0.5.$$

With this formula, it can be seen that if there is a perfect 50-50 split, the skew is zero; as the proportion of left-sided cases increases, the value becomes more negative (moves to the left), and so on, with a maximum/minimum value of ± 0.5. This value was determined and plotted for each etiology. Statistical significance was determined based on a null hypothesis that the frequency on each side was the same, using a χ^2 test.

Median presentation time to an outside otolaryngologist was compared with median referral presentation time to a WUSM otolaryngologist. For this analysis, patients who initially presented to the WUSM were excluded. Due to wide variance in presentation times and nonnormally distributed data, the mean presentation times were not presented. To compare median presentation times between the 2 groups, a paired Wilcoxon rank sum test was used. Histograms were also generated incorporating all data, displaying number of patients and cumulative percentage of patients by time of presentation. Percentages of patients per etiology who presented at 2, 3, and 4 months after symptom onset were also calculated.

Results

Of the charts reviewed, 938 patients met inclusion criteria for this study; 497 (53%) were women and 441 (47%) were men. The average age of patients was 56.9 years (range, 18-93 years). Overall, 621 (66.2%) patients had left-sided UVFP, while 317 (33.8%) patients had right-sided UVFP.

Table 1 displays the etiologies of unilateral vocal fold paralysis. In total, 522 (55.6%) patients had UVFP due to iatrogenic effects related to surgery. The most frequently observed surgery related to UVFP was thyroid/parathyroid surgery, noted in 158 (16.8%) patients. Lung surgery (n = 73 [7.8%]), cardiac surgery (n = 58 [6.2%]), and cervical spine surgery (n = 48 [5.1%]) were the next most common surgical causes of UVFP.

In total, 358 patients (38.2%) had UVFP due to causes not directly related to surgical intervention. Malignancy was the cause of UVFP in 167 (17.8%) of patients. Lung malignancy (n = 73 [7.8%]), metastatic malignancy (n = 24 [2.6%]), skull base malignancy (n = 18 [1.9%]), and direct invasion by thyroid malignancy (n = 14 [1.5%]) were most common. Idiopathic UVFP was noted in 124 patients (13.2%). Other less common causes of UVFP included intubation (n = 58 [6.2%]), trauma (n = 30 [3.2%]), cerebral vascular accident (CVA; n = 18 [1.9%]), and neck radiation (n = 8 [0.9%]).

Table 2 shows the laterality of UVFP based on etiology. In total, 622 (66.2%) patients had left-sided UVFP. This table shows the difference between right- and left-sided UVFP. Left-skewed etiologies of UVFP, represented by negative values, and right-skewed etiologies of UVFP, represented by positive values, are plotted in **Figure 1**. In addition to the expected left-sided predominance of intrathoracic etiologies (lung surgery, cardiac surgery, esophageal surgery, and lung malignancy), other significantly left-sided causes included idiopathic, intubation, carotid surgery, and skull base malignancy. There were no etiologies that were significantly skewed to the right.

Table 3 shows the median time of presentation for the 92 patients (9.8% of study population) who initially

Table 1. Etiology of Unilateral Vocal Fold Paralysis.

Etiology	No. (% of Total)
Surgery	
Cardiac surgery	58 (6.2)
Carotid surgery	22 (2.3)
Cervical spine surgery	48 (5.1)
Tracheostomy	2 (0.2)
Esophageal surgery	37 (3.9)
Lung surgery	73 (7.8)
Mediastinal surgery	17 (1.8)
Laryngeal surgery	4 (0.4)
Lateral neck surgery	61 (6.5)
Parathyroid surgery	18 (1.9)
Thyroid surgery	140 (14.9)
Skull base surgery	18 (1.9)
Intracranial surgery	24 (2.6)
Total	522 (55.6)
Malignancy	
Laryngeal cancer	20 (2.1)
Esophageal cancer	11 (1.2)
Lung cancer	73 (7.8)
Skull base tumor	18 (1.9)
Lymphoma	1 (0.1)
Mediastinal mass	5 (0.5)
Metastatic cancer	24 (2.6)
Parotid cancer	1 (0.1)
Thyroid cancer—direct invasion	14 (1.5)
Total	167 (17.8)
Idiopathic	124 (13.2)
Intubation	58 (6.2)
Trauma	30 (3.2)
CVA	18 (1.9)
Transesophageal echocardiogram	1 (0.1)
IJ catheter placement	1 (0.1)
Infected vagal nerve stimulator	1 (0.1)
Neck infection	1 (0.1)
Right skull base osteomyelitis	1 (0.1)
Neck radiation	8 (0.9)
Lung radiation	2 (0.2)
Thoracic deformity	1 (0.1)
Ankylosing spondylitis	1 (0.1)
Sarcoidosis	2 (0.2)
Total	938 (100.0)

Abbreviations: CVA, cerebral vascular accident; IJ, internal jugular.

presented to an outside otolaryngologist and were later referred. The median time of presentation to an outside otolaryngologist was 2.1 months, while the median time of presentation to a WUSM otolaryngologist was 9.5 months ($P < .001$). Given the very low rate of patients previously treated, we conclude that this 7.4-month difference in median is the period during which patients were observed by the outside otolaryngologist for possible recovery prior to referral to WUSM. The etiology with the largest delay between symptom onset and treatment was trauma at 563 months, while several etiologies had delays of only 0.5 months. Etiologies with the greatest percentage of patients presenting to an outside otolaryngologist before a WUSM otolaryngologist included carotid surgery (18.2%), laryngeal surgery (25.0%), parathyroid surgery (16.7%), thyroid surgery (12.1%), laryngeal cancer (20.0%), thyroid cancer (28.6%), idiopathic causes (16.9%), CVA (22.2%), and neck radiation (25.0%).

Figure 2 displays histograms of time of presentation to any otolaryngologist and to a WUSM otolaryngologist within 3 years of symptom onset, as well as cumulative percentages of patients who presented within this time window. It can be seen that 81% of patients present within 6 months, 89% within 1 year, and 93% within 2 years. In **Figure 3**, the same data focus on the first 4 months after onset, during which 44% present within the first month (many during the same hospital stay during which the paralysis began), 63% within 2 months, 71% within 3 months, and 75% within 4 months. These are the patients for whom an early intervention strategy might be an option. The cumulative plots for all patients and for WUSM-only have similar contours because the WUSM referral group comprises 90.2% of the patients.

Table 4 shows the first 4-month presentation data by etiology, excluding those groups with less than 10 patients. Etiologies with the greatest percentage of patients presenting to the WUSM within a 4-month period included esophageal cancer (90.9% present within 4 months), skull base surgery (88.9%), esophageal surgery (86.5%), intubation (86.2%), lung surgery (84.9%), and lung cancer (82.2%). Etiologies with the lowest percentage of patients presenting to the WUSM within a 4-month period included idiopathic causes (54.8%), CVA (55.6%), thyroid cancer (57.1%), and carotid surgery (59.1%). In patients who had UVFP caused by thyroid surgery, 66.4% presented to the WUSM within 4 months, and in patients with parathyroid surgery, 72.2% presented to the WUSM within 4 months.

Discussion

Etiology of UVFP

In this large retrospective study of UVFP, most of the etiologic findings were similar to 2 other large series, by Rosenthal et al[15] and Takano et al,[16] as shown in **Table 5**. Surgical/iatrogenic causes of UVFP are more common than nonsurgical causes, and thyroid/parathyroid surgeries are implicated more often than other types of surgery but do not comprise most surgical etiologies overall. Intubation injuries and idiopathic UVFP frequencies are similar in all 3 series, and the condition occurs on the left side in nearly two-thirds of cases. Among nonsurgical cases, malignancy was the most common category, most often lung cancer. Malignancy of the lung was the most common cause in 3 previous studies.[1-3]

The risk of iatrogenic injury to the recurrent laryngeal nerve in different surgical procedures has been widely

Table 2. Unilateral Vocal Fold Paralysis Laterality.[a]

Etiology	n	Right	Left	Skew	*P*-value[b]
Thyroid surgery	140	64	76	–0.04	.31
Lung surgery	73	15	58	–0.29	**<.01**
Lateral neck surgery	61	25	36	–0.09	.16
Cardiac surgery	58	9	49	–0.34	**<.01**
Cervical spine surgery	48	27	21	0.06	.39
Esophageal surgery	37	6	31	–0.34	**<.01**
Carotid surgery	22	6	16	–0.23	**.03**
Skull base surgery	18	12	6	0.17	.16
Parathyroid surgery	18	9	9	0	1.00
Mediastinal surgery	17	10	7	0.09	.47
Idiopathic	124	39	85	–0.19	**<.01**
Lung cancer	73	6	67	–0.42	**<.01**
Intubation	58	17	41	–0.21	**<.01**
Trauma	30	14	16	–0.03	.72
Metastatic cancer	24	10	14	–0.08	.41
Laryngeal cancer	20	7	13	–0.15	.18
Skull base tumor	18	4	14	–0.28	**.02**
CVA	18	7	11	–0.11	.35
Thyroid cancer—direct invasion	14	6	8	–0.07	.59
Esophageal cancer	11	4	7	–0.14	.37

Abbreviation: CVA, cerebral vascular accident.
[a]Etiologies with less than 10 patients not shown.
[b]*P* values based on χ^2 comparison with 50-50 L-R split with same N. Bold, $P < .05$.

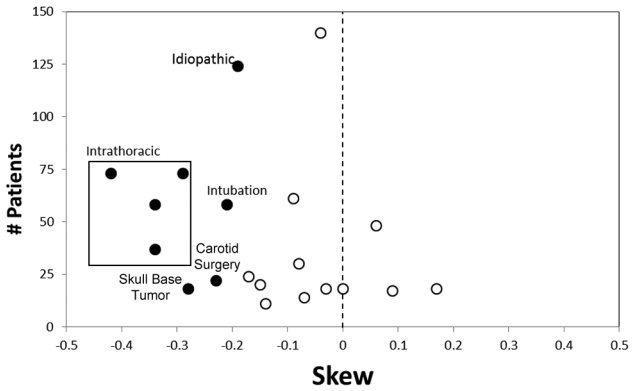

Figure 1. Left-right skew by etiology. Solid circles, statistically significantly skewed; open circles, not significant. Skew = (# cases on right / total cases for etiology) – 0.5. If left = right, skew = 0 (as seen for 18 parathyroid cases); if all cases occurred on left, skew = –0.5. Grouped are the intrathoracic causes of unilateral vocal fold paralysis (lung malignancy, thoracic, cardiac, and esophageal surgery).

reported. In a recent review by Misono and Merati,[24] the risk of vocal fold paralysis in thyroidectomy was between 0.8% and 2.3%; anterior cervical spine injury, less than 1%; cardiac/aortic surgery, 2%; mediastinoscopy, 0.2% to 6%; esophagectomy, 11%; and carotid endarterectomy, 4%.

Idiopathic UVFP was the etiology in 124 (13.2%) patients. Some older studies reported idiopathic etiologies to be the most common cause of UVFP,[2,6-10] but a recent review of the literature reported a rate of idiopathic UVFP of 24% ± 10%.[21] The decrease in idiopathic UVFP is likely due to better imaging capabilities to find small lesions along the nerve, as well as the shift toward surgery-related UVFP, which increases the proportion of cases with a clear etiology.

The data in this study by comparison to earlier studies of UVFP etiologies show the trend toward increasing surgical rather than malignant or idiopathic causes. This trend may be due to several factors. Since this study was conducted at a large tertiary referral center, as were the other recent studies by Rosenthal et al[15] and Takano et al,[16] a greater amount of surgery, as well as more complicated surgery, was likely being performed. An increased number of complicated surgeries both increase the risk of injury to the recurrent laryngeal nerve, as well as the relative numbers of UVFP due to surgical causes. In addition, as diagnostic imaging capabilities continue to improve, UVFP due to malignant or idiopathic causes will continue to decrease, as tumors are identified before causing UVFP and the course of the vagus and recurrent nerve can be imaged to determine the etiology of cases formerly diagnosed as idiopathic.

In analysis of laterality, 621 (66.2%) patients had left-sided UVFP, consistent with previous studies showing left-sided UVFP ranging from 59% to 81%.[1,2,7,8,11,15,16] Intrathoracic etiologies had the expected left-sided predominance, due to the anatomic course of the left recurrent laryngeal nerve. The statistically significant left-sidedness of some other etiologies may be a little harder to explain:

Table 3. Presentation Time for 92 Patients Referred to the WUSM.[a]

Etiology	n	Fraction of Original Total (%)	Outside ENT Presentation, Median, mo	WUSM ENT Presentation, Median, mo	Difference	Range in Difference
Carotid surgery	4	4/22 (18.2)	1.8	8.5	6.7	2-178
Cervical spine surgery	5	5/48 (10.4)	3.0	7.0	4.0	0.5-43
Lateral neck surgery	5	5/61 (8.2)	4.0	15.0	11.0	2-116
Parathyroid surgery	3	3/18 (16.7)	3.0	8.0	5.0	5-49
Thyroid surgery	17	17/140 (12.1)	1.5	12.0	10.5	0.5-111
Lung surgery	5	5/73 (6.8)	2.0	92.0	90.0	1-104
Cardiac surgery	4	4/58 (6.9)	2.5	6.5	4.0	2-5
Laryngeal cancer	4	4/20 (20.0)	1.5	4.5	3.0	3-7
Thyroid cancer	4	4/14 (28.6)	2.1	6.0	3.9	1.75-19
Idiopathic	21	21/124 (16.9)	3.0	12.0	9.0	0.5-51
Intubation	4	4/58 (6.9)	2.0	3.8	1.8	1-3
CVA	4	4/18 (22.2)	2.5	14.0	11.5	3-150
Total[b]	92	92/938 (9.8)	2.1	9.5	7.4	0.5-563

Abbreviations: ENT, ear, nose, and throat; WUSM, Washington University School of Medicine.
[a]Excludes etiologies with less than 3 patients.
[b]Includes 12 outside referral patients not listed above.

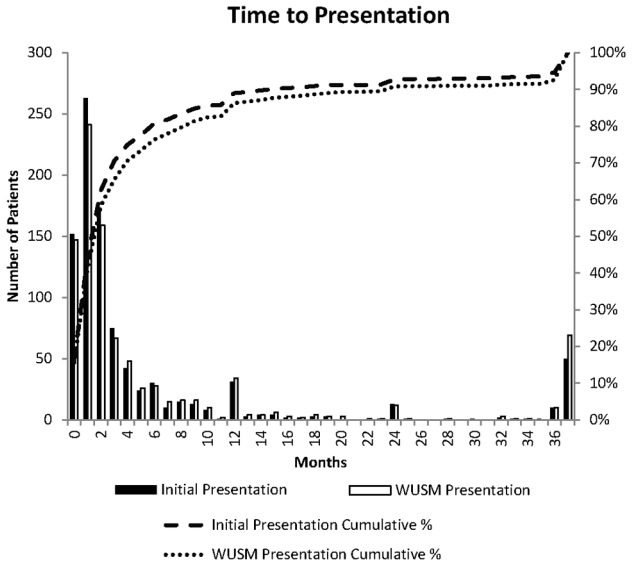

Figure 2. Histogram and cumulative plot of time from unilateral vocal fold paralysis symptom onset to presentation to the Washington University School of Medicine (WUSM; open bars) or to outside otolaryngologist prior to referral to the WUSM (solid bars; 9.8% of cases). Data shown for first 3 years following onset.

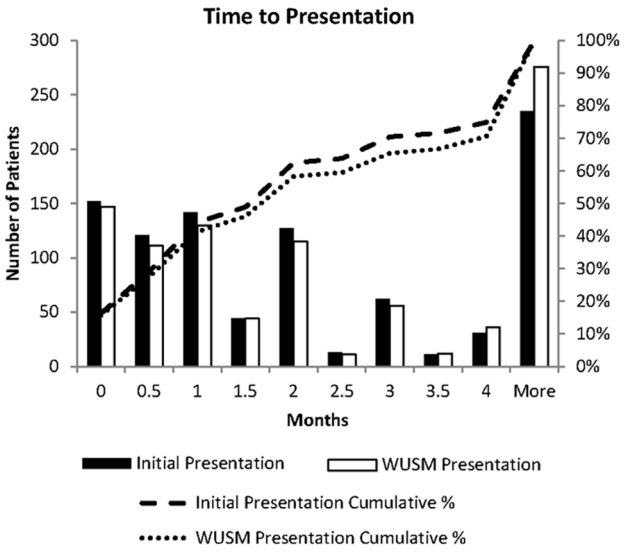

Figure 3. Histogram and cumulative plot from **Figure 2**, focusing on the first 4 months after symptom onset. WUSM, Washington University School of Medicine.

Carotid endarterectomy. Stroke and death rates in carotid endarterectomy have been shown to be higher for left-sided vs right-sided surgeries. Explanations for this observation include that due to the dominance of the left cerebral hemisphere, left-sided events are more symptomatic than right-sided events, and therefore more surgeries occur on the left side. There also may be a role in surgeon handedness, making left-sided carotid endarterectomies more technically difficult for right-handed surgeons.[25]

Intubation. More people in general are right-handed; there may be some greater tendency for right-handed anesthetists to traumatize the left hemilarynx more than the right during intubation. The mechanism of UVFP with intubation is not known but may relate to an acute event (at the time of intubation) or a longer-term event (while the tube is in place) that may be related to the cuff pressure or tube positioning. Some postintubation vocal fold immobility may

Table 4. Percentage of Patients Presenting to Washington University School of Medicine within 2, 3, and 4 Months from Symptom Onset.[a]

Etiology	n	No. (%)		
		Within 2 Months	Within 3 Months	Within 4 Months
Intracranial surgery	24	17 (70.8)	17 (70.8)	17 (70.8)
Skull base surgery	18	15 (83.3)	15 (83.3)	16 (88.9)
Carotid surgery	22	10 (45.5)	13 (59.1)	13 (59.1)
Cervical spine surgery	48	23 (47.9)	26 (54.2)	31 (64.6)
Lateral neck surgery	61	33 (54.1)	36 (59.0)	40 (65.6)
Parathyroid surgery	18	13 (72.2)	13 (72.2)	13 (72.2)
Thyroid surgery	140	81 (57.9)	86 (61.4)	93 (66.4)
Lung surgery	73	51 (69.9)	57 (78.1)	62 (84.9)
Mediastinal surgery	17	12 (70.6)	12 (70.6)	13 (76.5)
Cardiac surgery	58	19 (32.8)	41 (70.7)	44 (75.9)
Esophageal surgery	37	29 (78.4)	31 (83.8)	32 (86.5)
Skull base tumor	18	10 (55.6)	10 (55.6)	12 (66.7)
Laryngeal cancer	20	11 (55.0)	13 (65.0)	13 (65.0)
Thyroid cancer—direct invasion	14	6 (42.9)	6 (42.9)	8 (57.1)
Lung cancer	73	46 (63.0)	54 (74.0)	60 (82.2)
Esophageal cancer	11	9 (81.8)	10 (90.9)	10 (90.9)
Metastatic cancer	24	17 (70.8)	19 (79.2)	19 (79.2)
Idiopathic	124	48 (38.7)	64 (51.6)	68 (54.8)
Intubation	58	40 (69.0)	46 (79.3)	50 (86.2)
Trauma	30	14 (46.7)	18 (60.0)	19 (63.3)
CVA	18	7 (38.9)	9 (50.0)	10 (55.6)

Abbreviation: CVA, cerebral vascular accident.
[a]Etiologies with less than 10 patients not shown.

Table 5. Some Comparisons between the 3 Largest Series of Unilateral Vocal Fold Paralysis Etiologies.[a]

Etiology	Rosenthal et al[15] (n = 643)	Takano et al[16] (n = 797)	Present Study (n = 938)
Surgery, total	36.5	51.1	55.6
Thyroid/parathyroid	12.4	14.1	16.8
Nonthyroid	24.1	37.1	38.8
Intubation	5.8	7.3	6.2
Malignancy	18.4	9.9	17.8
Idiopathic	18.5	16.8	13.2
Left side	60.9[b]	64.0	66.2

[a]All values are percentages.
[b]Data reported for only 56.4% of patients in study.

have actually been cricoarytenoid joint ankylosis or dislocation that was misdiagnosed as UVFP.

Idiopathic. This is commonly presumed to be a post-viral neuropathy, but this would not explain a left-sided preference unless the distal left recurrent laryngeal nerve (RLN) is somehow more susceptible to viral attack than the right. Alternately, there may be some mechanism in which sudden neck extension or other postural change causes the left RLN to be stretched around the aorta. There may also be differences in intrathoracic connective tissue near the aorta that predispose to this sort of injury.

Skull base tumor. This finding may simply be related to the relatively low number of patients in this series.[18] We can find no references indicating a biological explanation for a left-sided predilection for skull base tumors.

Time to Presentation

To our knowledge, this study is the first large series to investigate time to presentation of these patients. It was found that three-fourths of patients with UVFP seen at the WUSM present within 3 or 4 months of the onset of symptoms. While this is encouraging, it also indicates that there

is potential for earlier presentation in the remaining 25% of patients. In patients who initially presented to an outside otolaryngologist, there was a significant referral delay of 7.4 months in median presentation time to the WUSM. This also highlights an opportunity for education of the referring physicians of the value of early intervention. For example, a recent study by Young et al[14] found that patients with UVFP had better vocal function after undergoing temporary vocal fold injection (VFI) even after direct benefit of VFI had dissipated, regardless of whether vocal fold mobility had recovered. Yung et al,[26] Arviso et al,[27] and Friedman et al[28] all reported lower rates of thyroplasty in patients who underwent temporary VFI. Bhattacharyya et al[29] found that early medialization within 1 to 4 days after onset of UVFP after thoracic surgery decreased the rate of pneumonia and led to a shorter length of stay compared with late medialization. Early intervention for UVFP clearly improves patient outcomes.

Early intervention could also involve treatment to prevent synkinesis. In animal studies, it has been found that following UVFP, some degree of reinnervation is evident within 3 months of injury.[22] One strategy that has been proposed is to perform a chemical blockade of reinnervation of the posterior cricoarytenoid muscle using a neurotoxic drug such as vincristine.[30,31] In an animal model, this was found to improve adductor recovery if given at 3 months postinjury but not at 5 months,[23] indicating there is a window of opportunity for treatment, after which it becomes too late for effective early intervention strategies. The present study shows that 71% of patients would be eligible for such intervention with current referral patterns. A clinical trial is the next step to determine whether this approach can help these patients.

A limitation of this study is that the surgical care at a tertiary care referral center skews data due to the greater number of difficult cases with a greater likelihood of nerve injury during surgery. In addition to missing data from those with incomplete charts, there are also an unknown number of patients who may have had a vocal fold paralysis but, due to quick recovery of voice, never sought treatment at a tertiary care facility.

Future directions of this study include analysis of the initial treatment and outcomes for each etiology of UVFP. Outcomes include voice improvement and return of vocal fold motion by fiberoptic examination. This analysis would allow further correlation with specific UVFP etiologies with the natural history of the disease, effectiveness of treatment, and type of treatment received. Outcomes specific to the length of time from symptom onset to treatment can also be assessed. This assessment would determine if delay in treatment adversely affects outcomes.

Conclusion

This retrospective medical record review of 938 patients with UVFP over the past 10 years is the largest series to date. It expands on the previous reports of UVFP etiology, with surgery and specifically thyroid surgery being the most common causes of UVFP. This study also reflects the growing contribution of nonthyroid surgeries accounting for a significant amount of injury to the recurrent laryngeal nerve, especially on the left side. Presently, 71% of patients with UVFP are seen within 3 months of RLN injury and would be eligible for early intervention procedures. Patients referred from outside otolaryngologists present, on average, after a significant delay.

Author Contributions

Emily A. Spataro, data analysis, manuscript preparation; **David J. Grindler**, data collection and analysis, manuscript preparation; **Randal C. Paniello**, original idea, final manuscript approval and editing.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: NIH (R01 DC010884)—salary support for senior author (R.C.P.).

References

1. Titche L. Causes of recurrent laryngeal nerve paralysis. *Arch Otolaryngol*. 1976;102:249-261.
2. Benninger MS, Gillen JB, Altman JS. Changing etiology of vocal fold immobility. *Laryngoscope*. 1998;108:1346-1350.
3. Ramadan HH, Wax MK, Avery S. Outcome and changing cause of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1998;118:199-202.
4. Kearsley JH. Vocal Cod paralysis—an aetiologic review of 100 cases over 20 years. *Aust NZ J Med*. 1981;11:663-666.
5. Barondess JA, Pompel P, Schley W. A study of vocal cord paralysis. *Trans Am Clin Climatol Assoc*. 1985;97:141-148.
6. Maisel RH, Ogura JH. Evaluation and treatment of vocal cord paralysis. *Laryngoscope*. 1974;84:302-316.
7. Hirose H, Sawashima M, Yoshioka H. Clinical observations on 750 cases of laryngeal palsy. *Tokyo J Med Sci*. 1981;88:33-37.
8. Hirose H. Clinical observations on 600 cases of recurrent laryngeal nerve paralysis. *Auris Nasus Larynx*. 1978;5:39-48.
9. Leon X, Venegas MP, Orus C, Quer M, Maranillo E, Sanudo JR. Glottic immobility: retrospective study of 229 cases. *Acta Otolaryngol Esp*. 2001;52:486-492.
10. Yamada M, Hirano M, Ohkubo H. Recurrent laryngeal nerve paralysis: a 10-year review of 564 patients. *Auris Nasus Larynx*. 1983;10(suppl):S1-S15.
11. Parnell FW, Bandenbug JH. Vocal cord paralysis: a review of 100 cases. *Laryngoscope*. 1970;80:1036-1045.
12. Chen HC, Jen YM, Wang CH, et al. Etiology of vocal cord paralysis. *ORL J Otorhinolaryngol Relat Spec*. 2007;69:167-171.
13. Sielska-Badurek E, Domeracka-Kolodziej A, Zawadzka R, Debowska-Jarzebska E. Vocal fold paralysis in the medical university of Warsaw's ambulatory of phoniatry in years 2000-2011. *Otolaryngol Pol*. 2012;66:313-317.
14. Young VN, Smith LJ, Rosen C. Voice outcome following acute unilateral vocal cord paralysis. *Ann Otol Rhinol Laryngol*. 2013;2122:197-204.

15. Rosenthal LHS, Benninger MS, Deeb RH. Vocal fold immobility: a longitudinal analysis of etiology over 20 years. *Laryngoscope*. 2007;117:1864-1870.
16. Takano S, Nito T, Tamaruya N, Kimura M, Tayama N. Single institutional analysis of trends over 45 years in etiology of vocal fold paralysis. *Auris Nasus Larynx*. 2012;39:597-600.
17. Terris D, Anstein D. Nguyen H. Contemporary evaluation of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1992;107:84-90.
18. Tucker H. Vocal cord paralysis-1979: etiology and management. *Laryngoscope*. 1980;90:585-590.
19. Myssiorek D. Recurrent laryngeal nerve paralysis: anatomy and etiology. *Otolaryngol Clin North Am*. 2004;37:25-44.
20. Rubin AD, Sataloff RT. Vocal fold paresis and paralysis. *Otolaryngol Clin North Am*. 2007;40:1109-1131.
21. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118: 1303-1307.
22. Shindo M, Herzon G, Hanson D, et al. Effects of denervation on laryngeal muscles: a canine model. *Laryngoscope*. 1992; 102:663-669.
23. Paniello RC. Effect on laryngeal adductor function of vincristine block of posterior cricoarytenoid muscle 3-5 months after recurrent laryngeal nerve injury. *Annals Otol Rhinol Laryngol*. In press.
24. Misono S, Merati A. Evaluation and management of unilateral vocal cord paralysis. *Otolaryngol Clin North Am*. 2012;45: 1083-1108.
25. Girard LP, Feasby TE, Eliasziw M, et al. Complication rates after left- versus right-sided carotid endarterectomy. *Circ Cardiovasc Qual Outcomes*. 2009;2:642-647.
26. Yung KC, Likhterov I, Courey MS. Effect of temporary vocal fold injection medialization on the rate of permanent medialization laryngoplasty in unilateral vocal fold paralysis patients. *Laryngoscope*. 2011;121:2191-2194.
27. Arviso LC, Johns MM III, Mathison CC, Klein AM. Longer-term outcomes of injection laryngoplasty in patients with potentially recoverable vocal fold paralysis. *Laryngoscope*. 2010;120:2237-2240.
28. Friedman AD, Burns JA, Henton JT, Zeitels SM. Early versus late injection medialization for unilateral vocal cord paralysis. *Laryngoscope*. 2010;120:2042-2046.
29. Bhattacharyya N, Batirel H, Swanson SJ. Improved outcomes with early vocal fold medialization for vocal fold paralysis after thoracic surgery. *Auris Nasus Larynx*. 2003;30:71-75.
30. Paydarfar JA, Paniello RC. Functional study of four neurotoxins as inhibitors of post-traumatic nerve regeneration. *Laryngoscope*. 2001;111:844-850.
31. McRae BR, Kincaid JC, Illing EA, et al. Local neurotoxins for prevention of laryngeal synkinesis after recurrent laryngeal nerve injury. *Ann Otol Rhinol Laryngol*. 2009;118:887-893.

Reprinted by permission of Laryngoscope. 2015; 125(4):904-908.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Diagnosis of Vocal Fold Paresis: Current Opinion and Practice

Amy P. Wu, MD; Lucian Sulica, MD

Objectives/Hypothesis: No accepted standard exists for the diagnosis of vocal fold paresis (VFP). Laryngeal specialists are surveyed to establish expert opinion on diagnostic methodology and criteria.
Study Design: Cross-sectional survey.
Methods: Questionnaires were distributed at laryngology conferences in fall 2013. Responses were collated anonymously and subjected to cross-tabulated data analysis.
Results: Fifty-eight responses completed by posttraining physicians whose practice focused in laryngology $\geq 75\%$ were analyzed. One (1.7%) relied principally on laryngeal electromyography, one (1.7%) on history, 10 (17%) on laryngoscopy, and 42 (72%) on strobovideolaryngoscopy for diagnosis. Only 12 (21%) performed laryngeal electromyography on $> 50\%$ of vocal fold paresis patients. Laryngeal electromyography sensitivity was considered moderate ($61 \pm 3.7\%$, $\sigma = 28$). Laryngoscopic/stroboscopic findings considered to have the strongest positive predictive value for VFP were slow/sluggish vocal fold motion ($75 \pm 3.0\%$, $\sigma = 23$), decreased adduction ($67 \pm 3.5\%$, $\sigma = 27$), decreased abduction ($65 \pm 3.4\%$, $\sigma = 26$), and decreased vocal fold tone ($61 \pm 3.5\%$, $\sigma = 26$). Asymmetric mucosal wave amplitude ($52 \pm 4.2\%$, $\sigma = 32$), asymmetric mucosal wave phase ($60 \pm 4.1\%$, $\sigma = 31$), hemilaryngeal atrophy ($60 \pm 4.0\%$, $\sigma = 31$), and asymmetric mucosal wave frequency ($49 \pm 4.0\%$, $\sigma = 30$) generated greatest disagreement.
Conclusions: Surveyed expert laryngologists diagnose vocal fold paresis predominantly on stroboscopic examination. Gross motion abnormalities had the highest positive predictive value. Laryngeal electromyography was infrequently used to assess for vocal fold paresis.
Key Words: Vocal fold paresis, laryngeal electromyography, laryngoscopy, stroboscopy.
Level of Evidence: 5

***Laryngoscope*, 125:904–908, 2015**

INTRODUCTION

Vocal fold paresis (VFP) is a partial motor denervation of the vocal fold causing variable degrees of compromised glottal function.[1] Appreciation of its clinical significance has grown hand in hand with an increasingly sophisticated understanding of laryngeal neuropathy, and it continues to evolve. Although not different in kind but only in degree from vocal fold paralysis, VFP is often considered separately; the spectrum of difficulties it causes is different, and perhaps most important, its diagnosis is more challenging and controversial. Paresis is usually diagnosed based on qualitative findings on laryngoscopy, stroboscopy, and/or laryngeal electromyography (LEMG). Endoscopic diagnosis typically rests on the observation of asymmetries of laryngeal motion. However, some asymmetry in laryngeal motion may be without clinical significance. Electromyographic findings may be indistinct as well.[1–4] In the absence of a standard for diagnosis, investigation of important aspects of paresis including causes, incidence, natural history, and effectiveness of treatment is challenging.

In the context of these limitations, expert clinical consensus may provide a useful basis to initiate discourse regarding VFP. The purpose of this investigation is to describe expert opinion regarding the diagnosis of VFP by means of a survey of practicing laryngeal specialists.

MATERIALS AND METHODS

A 29-item, 4-part questionnaire (see Appendix 1) was designed to characterize responders' experience, training and practice setting (part 1), assess diagnostic strategy (part 2), evaluate opinion regarding the positive predictive value of various laryngoscopic signs (part 3), and evaluate option regarding the sensitivity of LEMG (part 4). The roster of laryngoscopic signs was compiled from clinical experience, consultation with colleagues, and a review of the literature regarding VFP diagnosis and VFP-associated lesions. It included vocal fold hypomobility,[1,2,5] glottic insufficiency,[1,5] unilateral atrophy,[1] supraglottic hyperfunction,[2,6] mucosal wave asymmetries,[5,7] glottic axis deviation,[2,5] vocal fold height mismatch, arytenoid rotation,[8] and presence of a contact lesion[2,9,10] or pseudocyst.[2,11] Examples are illustrated in Figures 1 to 3. Opinion regarding each sign and LEMG sensitivity was assessed using a visual analog scale.

Additional Supporting Information may be found in the online version of this article.

From the New York Center for Voice and Swallowing Disorders, Department of Otolaryngology–Head and Neck Surgery, St Luke's–Roosevelt Hospital (A.P.W.); and the Sean Parker Institute for the Voice, Department of Otolaryngology–Head and Neck Surgery, Weill Cornell Medical College (L.S.), New York, New York, U.S.A.

Editor's Note: This Manuscript was accepted for publication October 7, 2014.

Scheduled to be presented at the 136th Annual Meeting of the American Laryngological Association, Las Vegas, NV, U.S.A., May 14–15, 2014.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Lucian Sulica, MD, 1305 York Avenue, 5th Floor, New York, NY 10021. E-mail: lus2005@med.cornell.edu

DOI: 10.1002/lary.25004

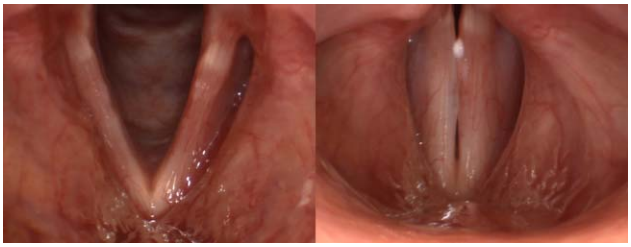

Fig. 1. A 44-year-old man with atrophy of the left hemilarynx, manifested as an enlarged laryngeal ventricle, and the beginning of a right vocal fold pseudocyst.

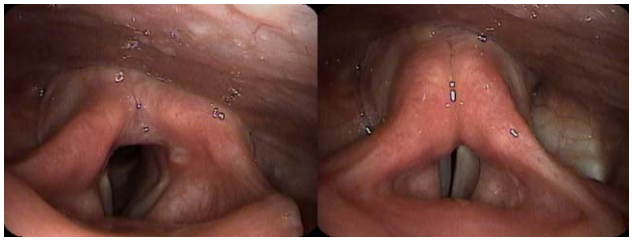

Fig. 3. A 30-year-old woman with a left-sided contact lesion and decreased left vocal fold adduction.

This questionnaire was distributed to attendees at three national laryngology conferences that took place in 2013: 1) the Neurolaryngology Study Group at the American Academy of Otolaryngology Annual Meeting, Vancouver, BC; 2) the Fall Voice Conference, Atlanta, GA; and 3) Advances and Controversies in Laryngology, Elk Grove, IL. Duplicate responses were avoided, and confidentiality was preserved by separately tracking responder identity. Responses were collated anonymously and subjected to cross-tabulated data analysis. Only otolaryngologists were surveyed.

Physicians whose practices did not predominantly focus on laryngology (< 75% laryngology), who did not see patients with dysphonia as a chief complaint, or who never diagnosed patients with VFP were excluded. Responders still in training, whether in residency or fellowship, were also excluded. Incomplete surveys were removed from analysis. One response with internally conflicting responses (a greater number of monthly paresis diagnoses than new dysphonic patients) was excluded. Statistical analysis using descriptive frequencies and cross-tabulations were performed with 2011 Microsoft Excel.

RESULTS

A total of 88 responses were received. After applying the exclusion criteria, a total of 58 entries were included for data analysis. Respondent demographics are summarized in Table I. The vast majority of respondents are laryngology fellowship-trained otolaryngologists with predominantly adult practices in the United States. Practice specifics related to VFP are summarized in Table II. Most respondents (72%) base diagnosis principally on stroboscopy. Only 12 (21%) respondents performed LEMG on greater than 50% of their VFP patients. Figure 4 and Table III summarize the respondents' opinion regarding positive predictive value of laryngoscopic findings in VFP patients. In general, abnormalities of gross vocal fold motion were considered to be most highly predictive, followed by stroboscopic signs. Associated lesions, namely contact lesions and pseudocysts, were not considered highly diagnostic of paresis. Respondents rated the sensitivity of LEMG at $61 \pm 3.7\%$, $\sigma = 28$.

Fig. 2. A 38-year-old woman with atrophy of the left vocal fold, incomplete glottic closure, unilateral (right) supraglottic hyperfunction, and impairment of arytenoid rotation on the left.

DISCUSSION

The prevalence and clinical importance of VFP is not yet established, which should not be surprising when diagnostic criteria remain under discussion, but it has the potential to be substantial. Among our respondents, VFP was diagnosed 8 times in a typical month, or roughly 100 times per year, and was found in about one-sixth of new patients presenting with voice complaints. By comparison, the largest reported series of VFP only consist of under 50 patients per year reviewed.[2,3,5] The prevalence of paresis has been proposed to be similar to paralysis on the basis of similar pathophysiology.[1] The prevalence of paralysis varies widely from report to report and is dependent on practice environment, geographic location, patient selection, and a host of other factors. Studies from one center have reported 46% prevalence of mild vocal hypomobility among patients with vocal complaints, 15% to 23% among singing teachers without vocal complaints, and 71% among singing teachers with technical difficulty complaints.[12–14] Simpson et al.[5] drew cases from a series of 739 patients presenting to their tertiary laryngology practice over a 4-year period with a chief complaint of dysphonia. Of 195 (26.4%) patients initially diagnosed with VFP or paralysis by videostroboscopy, only 13 or 1.8% of the overall dysphonic patients had LEMG-confirmed unilateral or bilateral VFP. Koufman et al.[15] reviewed 415 patients who underwent LEMG over a 5-year period. This group found "abnormal diagnostic LEMG" results (presumed to

TABLE I.
Respondent Demographics (N = 58).

Years in practice posttraining	11 ± 1.1 ($\sigma = 8.6$, range 0.1–31)
Laryngology fellowship trained	54 (93%)
Practicing in the United States	56 (97%)
Percent of practice which is laryngology	$93 \pm 1.0\%$
Percent adult patients	$91 \pm 1.1\%$

TABLE II.
Practice Related to VFP

New patients with voice-related complaint/month	49 ± 3.2, $\sigma = 25$
VFP diagnosis/month	8.5 ± 1.6, $\sigma = 12$
Diagnosis of VFP Rests Principally On:	
History	1 (1.7%)
Laryngoscopy (continuous light)	10 (17%)
Strobovideolaryngoscopy	42 (72%)
LEMG	1 (1.7%)
% Patients diagnosed with VFP who had videostroboscopy	$96 \pm 1.6\%$, $\sigma = 12$
% patients diagnosed with VFP who had LEMG	$26 \pm 4.0\%$, $\sigma = 31$

LEMG = laryngeal electromyography; VFP = vocal fold paresis.

represent paresis or paralysis in large part) in 346 (83%) patients. Sataloff et al.[14] reviewed 751 patients who underwent LEMG for all causes over a 4-year period. This series contained 689 suspected cases of paresis/paralysis by videostrobscopy, with LEMG confirming the diagnosis in 661 patients (95.9%). The variation among these three series reveals substantial differences among practitioners regarding diagnosis and testing.

Respondents indicated that they principally relied on laryngoscopy, usually under stroboscopic light, to make the diagnosis of VFP. Although LEMG is the only way to definitively diagnose laryngeal neuropathy objectively in vivo, the vast majority of respondents evidently felt that laryngoscopic criteria were sufficiently reliable to support the diagnosis alone. Only one respondent routinely relied on LEMG for diagnosis, and only the minority of patients ever had LEMG at all. Many reasons may prevent the use of LEMG, lack of availability and expertise prominent among them, but respondents felt that the sensitivity of LEMG was not high. There is little doubt that LEMG is highly specific for neuropathy. Findings of fibrillations, positive sharp waves, or polyphasic motor unit action potentials are unambiguous signs of neurologic impairment. Unfortunately in paresis, such clearly abnormal findings may be absent or obscured. Decreased recruitment of otherwise normal-appearing motor unit action potentials may be the only abnormality present. Because this relative change may be small and mimicked by incomplete muscle activation or suboptimal needle placement, there remains a role for physician judgment and inevitably error. Moreover, the maximal interference pattern in striated muscle is typically present at only 30% of maximum isometric contraction, which creates the possibility that even substantial paresis may be obscured during testing. Thus, although LEMG can provide important information that laryngoscopy cannot, it is not clear that it is a more accurate diagnostic tool than laryngoscopy in the diagnosis of VFP.

Reliance on laryngoscopy begs the question of which findings are considered important. To say that one may find signs of paresis in virtually every larynx is only a mild exaggeration. Unlike systems such as the extraocular muscles, mild discoordination in the larynx probably carries little functional and evolutionary disadvantage as long as glottic closure for airway protection is brisk and effective. Thus, much asymmetry in vocal fold

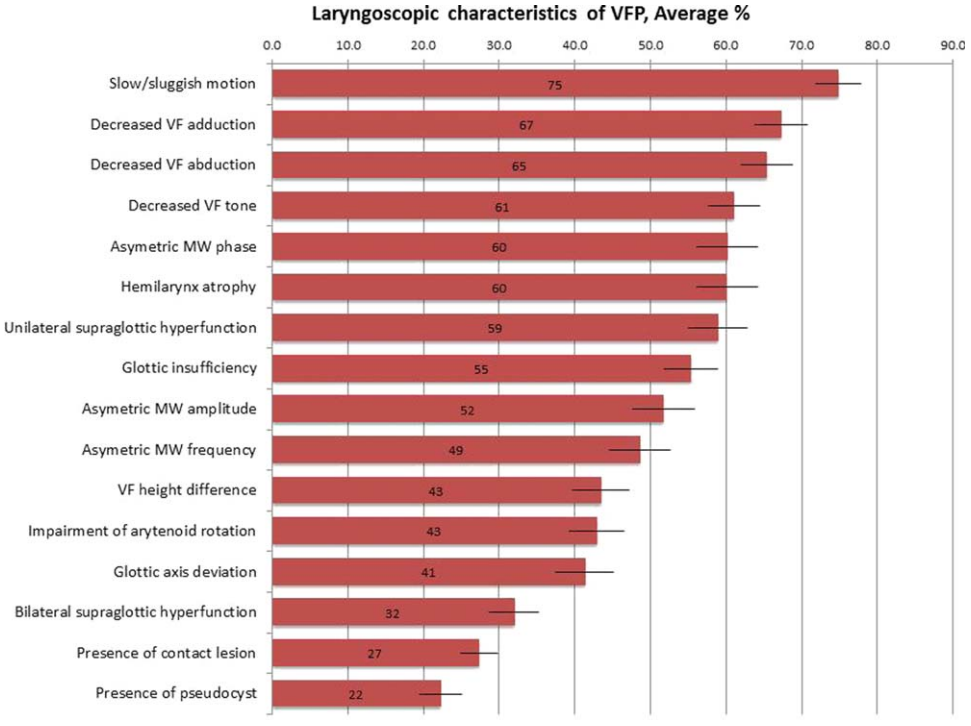

Fig. 4. Positive predictive value of laryngoscopic signs for vocal fold paresis. MW = mucosal wave; VF = vocal fold.

TABLE III.
Respondent Opinion Regarding Positive Predictive Value of Laryngoscopic Findings in VFP.

Examination finding	Average %	Error	σ
Slow/sluggish motion	74.9	3.0	22.8
Decreased VF adduction	67.3	3.5	26.7
Decreased VF abduction	65.4	3.4	26.1
Decreased VF tone	61.1	3.5	26.3
Asymetric MW phase	60.2	4.1	31.3
Hemilarynx atrophy	60.1	4.0	30.8
Unilateral supraglottic hyperfunction	58.9	3.9	29.4
Glottic insufficiency	55.4	3.5	26.8
Asymetric MW amplitude	51.7	4.2	31.7
Asymetric MW frequency	48.6	4.0	30.6
VF height difference	43.5	3.8	28.6
Impairment of arytenoid rotation	42.9	3.6	27.8
Glottic axis deviation	41.3	3.8	29.3
Bilateral supraglottic hyperfunction	32.0	3.3	25
Presence of contact lesion	27.3	2.4	18.1
Presence of pseudocyst	22.3	2.8	21.4

MW = mucosal wave; VF = vocal fold; VFP = vocal fold paresis.

motion may be without clinical significance. Further, Roy et al. showed that laryngoscopic findings are not consistent from case to case, even in experimentally induced isolated unilateral superior laryngeal nerve palsy, a condition probably more homogeneous than that which presents clinically.[17] Respondents identified deficits of gross motion as having the highest positive predictive value for VFP, despite reporting heavy reliance on stroboscopic examination in practice. This may reflect merely that stroboscopy is the standard clinical examination for patients with a voice complaint in the specialized practices of these physicians rather than the use of examination under stroboscopic light to identify VFP. Mucosal wave phase asymmetry was deemed the most useful stroboscopic sign, ranking only fifth in order of preference despite a report that identified it as correlating very well with LEMG abnormalities.[7] Stroboscopic signs (phase, amplitude, and frequency) were also marked by the greatest divergence of opinion regarding significance, as reflected by the standard error. The few signs that have been the subject of systematic analysis in the literature, namely arytenoid rotation[8] and unilateral ventricular fold hyperfunction,[6] were not regarded as among the most useful. Reports have proposed a relationship between contact lesions[9,10] and vocal fold pseudocysts[11] and VFP. Despite this, respondents thought that the potential for VFP to be present when such lesions were identified was very low.

Overall, this investigation reveals that paresis is frequently diagnosed and appears to be a significant clinical entity in laryngology practices. Diagnosis appears to be made on the basis of qualitative findings on laryngoscopy, principally deficits of gross vocal fold motion. Although stroboscopy is widely used, stroboscopic signs are not considered the most reliable signs to identify VFP. Electrophysiologic testing is not used often. Plainly, there exists no clear consensus on how the diagnosis of VFP should be made in a given patient, and establishing one will be a challenge in the absence of tests or findings that are both reasonably specific and sensitive. Under these circumstances, and given the frequency of asymmetric motion in the larynx, VFP is at risk of being diagnosed uncritically when no other obvious reason for a patient's complaint is evident to the examiner.

The survey format is subject to substantial recall bias and may give a false impression—likely falsely elevated—of the prevalence of paresis. This survey explicitly did not distinguish between superior laryngeal nerve paresis and recurrent laryngeal nerve paresis, frequently separated in the literature, which may have caused surveyed physicians to assign less positive predictive value to the signs under consideration than a more specific diagnosis. Reasons for the relatively rare use of LEMG were not investigated; these may have little to do with reservations regarding LEMG utility. Most importantly, the format of the survey necessarily does not well reflect the method of diagnosis of VFP in clinical practice. Such a diagnosis is rarely made on the basis of a single element of the evaluation or a single sign considered by itself, but depends on an educated critical synthesis of the clinical evidence. Physicians may form an impression of the likelihood of a given diagnosis based on the history, which then informs the physical examination. In fact, the perceived likelihood of VFP based on symptoms and clinical evolution of the complaint may significantly affect the perceived positive predictive value of a given laryngoscopic sign. Despite these limitations, this data may form a useful basis for further consideration of this challenging topic.

CONCLUSION

Surveyed laryngologists diagnose VFP frequently, relying principally on laryngeal strobovideolaryngoscopy to make the diagnosis. Among laryngoscopic signs, gross motion abnormalities were judged to have the highest positive predictive value for VFP, followed by abnormalities in the mucosal wave. Opinion varied most about the importance of these. LEMG was infrequently used to assess for VFP and was considered to have only moderate sensitivity for the diagnosis. Given the perceived clinical importance of VFP, directed investigation is necessary to refine diagnostic accuracy.

BIBLIOGRAPHY

1. Sulica L., Vocal Fold Paresis: An Evolving Clinical Concept. *Curr Otorhinolaryngol Rep* 2013;1:158–162.
2. Koufman JA, Postma GN, Cummins MM, Blalock PD. Vocal fold paresis. *Otolaryngol Head Neck Surg* 2000;122:537–541.
3. Heman-Ackah YD, Barr A. Mild vocal fold paresis: understanding clinical presentation and electromyographic findings. *J Voice* 2006;20:269–281.
4. Rubin AD, Praneetvatakul V, Heman-Ackah Y, Moyer CA, Mandel S, Sataloff RT. Repetitive phonatory tasks for identifying vocal fold paresis. *J Voice* 2005;19:679–686.
5. Simpson CB, Cheung EJ, Jackson CJ. Vocal fold paresis: clinical and electrophysiologic features in a tertiary laryngology practice. *J Voice* 2009; 23:396–398. doi: 10.1016/j.jvoice.2007.10.011.
6. Belafsky PC, Postma GN, Reulbach TR, Holland BW, Koufman JA. Muscle tension dysphonia as a sign of underlying glottal insufficiency. *Otolaryngol Head Neck Surg* 2002;127:448–451.

7. Simpson CB, May LS, Green JK, Eller RL, Jackson CE. Vibratory asymmetry in mobile vocal folds: is it predictive of vocal fold paresis? *Ann Otol Rhinol Laryngol* 2011;120:239–242.
8. Sufyan AS, Kincaid JC, Wannemuehler TJ, Halum SL. The interarytenoid spatial relationship: accuracy and interrater reliability for determining sidedness in cases of unilateral adductor paresis. *J Voice* 2013;27:90–94. doi: 10.1016/j.jvoice.2012.08.003.
9. Carroll TL, Gartner-Schmidt J, Statham MM, Rosen CA. Vocal process granuloma and glottal insufficiency: an overlooked etiology? *Laryngoscope* 2010;120:114–120. doi: 10.1002/lary.20656.
10. Halum SL, Miller P, Early K. Laryngeal granulomas associated with superior laryngeal nerve paresis. *J Voice* 2010;24:490–493. doi: 10.1016/j.jvoice.2008.12.010.
11. Koufman JA, Belafsky PC. Unilateral or localized Reinke's edema (pseudocyst) as a manifestation of vocal fold paresis: the paresis podule. *Laryngoscope* 2001;111:576–580.
12. Heman-Ackah YD, Batory M. Determining the etiology of mild vocal fold hypomobility. *J Voice* 2003;17:579–588.
13. Heman-Ackah YD, Dean CM, Sataloff RT. Strobovideolaryngoscopic findings in singing teachers. *J Voice* 2002;16:81–86.
14. Sataloff RT, Hawkshaw MJ, Johnson JL, Ruel B, Wilhelm A, Lurie D. Prevalence of abnormal laryngeal findings in healthy singing teachers. *J Voice* 2012;26:577–583. doi: 10.1016/j.jvoice.2012.01.002.
15. Koufman JA, Postma GN, Whang CS, et al., Diagnostic laryngeal electromyography: The Wake Forest experience 1995–1999. *Otolaryngol Head Neck Surg* 2001;124:603–606.
16. Brewer DW, et al., Unilateral recurrent laryngeal nerve paralysis: A reexamination. *J Voice* 1991;5:178–185.
17. Roy N, Barton ME, Smith ME, Dromey C, Merrill RM, Sauder C. An in vivo model of external superior laryngeal nerve paralysis: laryngoscopic findings. *Laryngoscope* 2009;119:1017–1032. doi: 10.1002/lary.20193.

Reprinted by permission of J Voice. 2014; 28(2):262.e13-262.e21.

Stroboscopy in Detection of Laryngeal Dysplasia Effectiveness and Limitations

***,†Vojko Djukic, *,†Jovica Milovanovic, *,†Ana D. Jotic, and †Milan Vukasinovic,** **†Belgrade, Serbia*

Summary: Vocal fold pathology changes the appearance and vibratory patterns observed during stroboscopic examination, but a strict correlation between the vibratory pattern and the dysplasia type does not exist. The aims of this study were to determine the role of stroboscopy in vocal fold dysplasia assessment and to determine whether stroboscopy is the deciding factor when performing laryngomicroscopy with biopsy in suspicious lesions. This prospective controlled study involved 112 patients with laryngeal dysplasia treated over a 2-year period at a tertiary medical center. Patient data and clinical, stroboscopy, laryngomicroscopy, and histopathologic reports were reviewed. During the stroboscopy, glottic occlusion, phase symmetry, periodicity, amplitude, mucosal wave, and nonvibratory segments were followed. Laryngomicroscopy with different types of endoscopic cordectomies (types I–III) was performed as a therapeutic measure, with a 12-month follow-up period. Nonvibrating segments were present in 15.1% of the patients with mild dysplasia and in 38.5% of the patients with moderate dysplasia. In 45.5% of the patients with severe dysplasia (carcinoma *in situ*), nonvibrating segments were absent. The amplitude of vocal fold vibrations in patients with mild dysplasia ($P = 0.03$) was a significant factor indicative of recurrent disease, but none of the stroboscopic signs was significant for the disease progression. Severe dysplasia can be related to both nonvibrating and vibrating vocal fold segments. Stroboscopy cannot be used reliably for classifying laryngeal dysplasia and may indicate the need to perform laryngomicroscopy with biopsy in suspicious vocal fold lesions. The warning factors for recurrence and progression of dysplasia are treatment modality, abnormal amplitude of vibration, and nonvibrating segment.

Key Words: Laryngeal dysplasia–Stroboscopy–Nonvibrating segment.

INTRODUCTION

Despite all the efforts made in discovering and classifying vocal fold lesions, uncertainty exists when determining which lesions are malignant or premalignant. These lesions are usually described as chronic laryngitis, parakeratosis, leukoplakia, erythroplakia, or dyskeratosis. A number of histologic results can be found under the same clinical appearance; therefore, the histologic nature of these lesions is completely unpredictable until a biopsy is performed. Malignant transformation rates range from 6% to 22%, and the rates increase with the severity of the precancerous lesion.[1–3] Therefore, the early detection of these lesions is of paramount importance.

Another difficulty in diagnosing these lesions is that there is no universally accepted histopathologic classification system. In the current literature and clinical practice, there are several widely accepted classification systems: the 2005 World Health Organization (WHO), Squamous Intraepithelial Neoplasia, Laryngeal Intraepithelial Neoplasia, and the Ljubljana Classification of Squamous Intraepithelial Lesions systems. This disparity makes it difficult to compare the diagnostic and follow-up studies. The WHO system uses three tiers of dysplasia: mild, moderate, and severe. Severe dysplasia includes what has been previously reported as noninvasive carcinoma (carcinoma *in situ* [CIS]) and severe dysplasia.[4] The progression and transformation to invasive carcinoma is one of the important outcome measures for intraepithelial lesions. Correlating molecular parameters with clinical outcome was recently suggested as a gold standard for classifying dysplasia. Some authors have stated that any histopathologic classification of this millennium should also depend on additional evidence, such as the genetic and molecular structural changes of the cells that contribute to the malignant transformation.[5]

Stroboscopy is considered to be an important part of diagnosing patients with laryngeal dysplasia. Nevertheless, we must note that a strict correlation between a vocal fold vibratory pattern and a certain type of lesion does not exist. Vocal fold pathology may produce changes in the appearance and vibratory patterns observed during stroboscopic examination. Interpreting the stroboscopic examination involves systematic judgment and describing the different vibratory pattern signs. These signs, which were first identified by Hirano and Bless,[6] included the fundamental frequency and periodicity, amplitude of horizontal excursion, glottal closure, symmetry of bilateral movement, mucosal wave, and nonvibrating portions of the vocal fold. Recently, Kelley et al have attempted to improve or refine the basic stroboscopic rating form and develop criteria to improve the reliability of selected stroboscopic signs.[7] Few studies have indicated which stroboscopic signs are more significant than others in evaluating the vibratory pattern of vocal folds with premalignant lesions. The aim of this study was to determine the importance of stroboscopy in diagnosing vocal fold dysplasia and ascertain if it can reliably estimate a level of dysplasia and be the deciding factor when performing laryngomicroscopy with biopsy. We also wanted to determine whether other factors, such as treatment modality and stroboscopic

Accepted for publication July 16, 2013.

From the *Medical Faculty Belgrade, University of Belgrade, Belgrade, Serbia; and the †Clinic for Otorhinolaryngology and Maxillofacial Surgery, Clinical Centre of Serbia, Belgrade, Serbia

Address correspondence and reprint requests to Ana D. Jotic, Clinic for Otorhinolaryngology and Maxillofacial Surgery, Clinical Centre of Serbia, Pasterova 2, 11000 Belgrade, Serbia. E-mail: anajotic@yahoo.com

Journal of Voice, Vol. 28, No. 2, pp. 262.e13-262.e21
0892-1997/$36.00
© 2014 The Voice Foundation
http://dx.doi.org/10.1016/j.jvoice.2013.07.006

signs, could be used to anticipate if disease recurrence or progression will occur.

MATERIALS AND METHODS

This prospective study included 112 patients who were treated over a 2-year period (between January 1, 2010 and December 31, 2011, with a 12-month follow-up period) in the Clinic for Otorhinolaryngology and Maxillofacial Surgery at the Clinical Centre of Serbia in Belgrade. This study was approved by the Institutional Ethical Committee, and all patients provided written informed consent before their inclusion in the study.

The following inclusion criteria were applied: the presence of a vocal fold lesion of any grade of dysplasia according to the WHO classification (mild, moderate, and severe dysplasia), a vocal fold lesion on the superior surface and free edge of the membranous part of the vocal fold, lesions ranging in size from 2 to 10 mm and up to 2 mm in thickness, normal motility of the vocal folds and arytenoid, no previous or simultaneous vocal fold lesions (inflammatory, dysplastic, carcinoma, or otherwise), and no previous laryngeal surgery, radiotherapy, or endotracheal intubation. All patient data, including clinical, stroboscopy, and laryngomicroscopy examinations and histopathologic reports were evaluated.

Stroboscopy was performed with the ATMOS Strobo 21 LED, ATMOS Cam 31 DV Data, and Laryngoscope 70° resp. 90° (ATMOS MedizinTechnik GmbH & Co., Lenzkirch, Germany) during modal pitch at comfortable intensity on sustained vowel /i/. The following parameters were rated:

1. glottic occlusion (1, sufficient or 2, insufficient),
2. phase symmetry (1, symmetrical or 2, asymmetrical opening and closing of the other vocal fold mirrors),
3. periodicity (1, regular or 2, irregular successive vibrations),
4. amplitude (1, normal; 2, decreased; or 3, increased),
5. mucosal wave (1, normal with 30–50% lateral travel; 2, increased with lateral travel greater than 50%; or 3, decreased with lateral travel less than 30%),
6. nonvibratory segment (1, presence or 2, absence of nonvibratory segment in the vocal fold or a portion thereof).

Laryngomicroscopy and different types of endoscopic cordectomy with cold instruments (types I–III according to recommended European Laryngological Society (ELS) classification for endoscopic cordectomies)[8] were performed using a Carl Zeiss Surgical OPMI Sensera optical microscope (Carl Zeiss Meditec Inc, Dublin, CA) under general endotracheal anesthesia.

The follow-up period for every patient was 12 months. During this period, a control examination with stroboscopy was performed monthly, and all patients with established recurrent vocal fold lesions on their control examinations underwent a laryngomicroscopy with complete lesion removal and histopathologic analysis. Any histologic progression of the lesions was noted.

PASW Statistics 18 program (IBM Corporation, New York, NY) was used for the data analysis. To determine the statistical significance of change in dynamics between the stroboscopic signs before the treatment and after the follow-up period, the McNemar and the Wilcoxon signed-rank tests were used. To determine a correlation between the chosen predicting factors and dysplasia, a multivariate regression analysis was performed. To assess which of the stroboscopic signs was most useful in predicting the histopathologic outcome and the degree of dysplasia, logistical regression was used. P values <0.05 were considered statistically significant.

RESULTS

The study included 98 males (87.5%) and 14 females (12.5%), with an average age of 55.65 years. There were 105 (93.7%) smokers, 95 (90.5%) of whom were males and 10 (9.5%) were females. Considering histopathologic results according to the WHO classification, 53 (47.3%) patients were classified as mild, 26 (23.2%) as moderate, and 33 (29.5%) as severe dysplasia.

Stroboscopic signs for patients with mild dysplasia before any treatment and after 12 months of follow-up because of recurrent disease are shown in Table 1. Considering phase symmetry, periodicity, amplitude of the vocal fold vibrations, and mucosal wave appearance, there were significant changes in the number of patients before the treatment and after the follow-up (McNemar or Wilcoxon signed-rank test, $P < 0.00$). Nonvibrating segments were present in eight (15.1%) patients before the treatment and in nine (17.0%) patients after the treatment ($P = 1.000$, McNemar test).

Considering the number of patients in the group with moderate dysplasia (Table 2), the changes in glottic occlusion and the presence of nonvibrating segment were not statistically significant, but the changes in the number of patients considering phase symmetry, periodicity, amplitude of vocal fold vibrations, and the mucosal wave appearance were statistically significant (McNemar or Wilcoxon signed-rank test, $P < 0.00$). In the group with moderate dysplasia, nonvibrating segments were present in 38.5% of the patients before the treatment and in 23.1% of the patients after the 12-month follow-up.

The results were similar in a group with severe dysplasia (Table 3). There were significant changes in the number of patients considering periodicity, amplitude of vocal fold vibrations, mucosal wave appearance, and the existence of nonvibrating segments (McNemar or Wilcoxon signed-rank test, $P < 0.00$). In this group, McNemar test could not be performed for the phase symmetry because all patients had asymmetric vibrations of the vocal fold vibrations before the treatment. Nonvibrating segments were present in 54.5% patients before the treatment and in 24.2% of patients after the 12-month follow-up. Most stroboscopic parameters were statistically significantly improved in all three patient groups.

Considering the treatment options, our patients underwent cordectomy types I–III, according to ELS classification for endoscopic cordectomies, the microscopic appearance of the change, and the assessment of the vertical expansion of the lesion (Table 4). Type I cordectomy was performed in 64.1% of the patients with mild dysplasia, 25.4% of the patients with

TABLE 1.
Stroboscopic Signs for Patients With Mild Dysplasia Before Treatment and After 12 Months of Follow-Up or Before Retreatment

	Mild Dysplasia		
Stroboscopic Signs	Before Treatment (%)	After 12 Months (%)	Sig.
Glottic occlusion			0.791
Sufficient	38/53 (71.7)	40/53 (75.5)	
Insufficient	15/53 (28.3)	13/53 (24.5)	
Phase symmetry			0.000*
Symmetrical	6/53 (11.3)	36/53 (67.9)	
Asymmetrical	47/53 (88.7)	17/53 (32.1)	
Periodicity			0.000*
Regular	6/53 (11.3)	36/53 (67.9)	
Irregular	47/53 (88.7)	17/53 (32.1)	
Amplitude			0.000*
Normal	10/53 (18.9)	35/53 (66)	
Decreased	25/53 (47.1)	14/53 (26.4)	
Increased	18/53 (34)	4/53 (7.6)	
Mucosal wave			0.000*
Normal with 30–50% lateral travel	13/53 (24.5)	35/53 (66)	
Increased with lateral travel greater than 50%	32/53 (60.4)	17/53 (32.1)	
Decreased with lateral travel less than 30%	8/53 (15.1)	1/53 (1.9)	
Nonvibratory segment			1.000
Presence	8/53 (15.1)	9/53 (17.0)	
Absence	45/53 (84.9)	44/53 (83.0)	

Abbreviation: Sig., statistical significance.
*$P < 0.05$.

TABLE 2.
Stroboscopic Signs for Patients With Moderate Dysplasia Before Treatment and After 12 Months of Follow-Up or Before Retreatment

	Moderate Dysplasia		
Stroboscopic Signs	Before Treatment (%)	After 12 Months (%)	Sig.
Glottic occlusion			1.000
Sufficient	16/26 (61.5)	17/26 (65.4)	
Insufficient	10/26 (38.5)	9/26 (34.6)	
Phase symmetry			0.007*
Symmetrical	3/26 (11.6)	14/26 (53.8)	
Asymmetrical	23/26 (88.4)	12/26 (46.2)	
Periodicity			0.021*
Regular	4/26 (15.4)	14/26 (53.8)	
Irregular	22/26 (84.6)	12/26 (46.2)	
Amplitude			0.001*
Normal	3/26 (11.6)	14/33 (42.4)	
Decreased	18/26 (69.2)	12/33 (57.6)	
Increased	5/26 (19.2)	0/33 (0)	
Mucosal wave			0.029*
Normal with 30–50% lateral travel	6/26 (23.1)	14/33 (42.4)	
Increased with lateral travel greater than 50%	18/26 (69.2)	11/33 (54.6)	
Decreased with lateral travel less than 30%	2/26 (7.7)	1/33 (3)	
Nonvibratory segment			0.344
Presence	10/26 (38.5)	6/26 (23.1)	
Absence	16/26 (61.5)	20/26 (76.9)	

Abbreviation: Sig., statistical significance.
*$P < 0.05$.

TABLE 3.
Stroboscopic Signs for Patients With Severe Dysplasia Before Treatment and After 12 Months of Follow-Up or Before Retreatment

	Severe Dysplasia		
Stroboscopic Signs	Before Treatment (%)	After 12 Months (%)	Sig.
Glottic occlusion			1.000
Sufficient	25/33 (75.8)	24/33 (72.7)	
Insufficient	8/33 (24.2)	9/33 (27.3)	
Phase symmetry			—
Symmetrical	0/33 (0)	19/33 (57.6)	
Asymmetrical	33/33 (100.0)	14/33 (42.4)	
Periodicity			0.000*
Regular	1/33 (3.0)	19/33 (57.6)	
Irregular	32/33 (97.0)	14/33 (42.4)	
Amplitude			0.000*
Normal	1/33 (3)	19/33 (57.6)	
Decreased	29/33 (87.9)	14/33 (42.4)	
Increased	3/33 (9.1)	0/33 (0)	
Mucosal wave			0.000*
Normal with 30–50% lateral travel	1/33 (3)	19/33 (57.6)	
Increased with lateral travel greater than 50%	0/33 (0)	0/33 (0)	
Decreased with lateral travel less than 30%	32/33 (97)	14/33 (42.4)	
Nonvibratory segment			0.013*
Presence	18/33 (54.5)	8/33 (24.2)	
Absence	15/33 (45.5)	25/33 (75.8)	

Abbreviation: Sig., statistical significance.
*$P < 0.05$.

moderate dysplasia, and 36.4% of the patients with severe dysplasia. Type II cordectomy was performed in 35.9% of the patients with mild dysplasia, 34.6% of the patients with moderate dysplasia, and 57.6% of the patients with severe dysplasia. Type III cordectomy was performed in only two (6%) patients with severe dysplasia.

Recurrence of the disease occurred after 8 months in six patients, after 9 months in five patients, after 10 months in four patients, after 11 months in eight patients, and after 12 months in five patients. Most patients with recurrence were in a group with moderate dysplasia (Table 5). Disease progression was noted in 10 patients with recurrence. Invasive carcinoma developed in four patients: one from group with moderate dysplasia and three from group with severe dysplasia.

Multivariate regression analysis was performed to determine how factors such as cordectomy type and stroboscopic signs (glottic occlusion, phase symmetry, periodicity, amplitude of vibrations, mucosal wave, and nonvibrating segment) correlated with the histopathologic verification of different types of dysplasia (Table 6). Some factors, such as the type of cordectomy and the existence of nonvibrating segment, were set apart from others and were proven to be significantly different in various levels of dysplasia ($P < 0.05$).

Logistic regression was used to determine whether the recurrence and progression of the disease could be anticipated by cordectomy type and stroboscopic signs (glottic occlusion, phase symmetry, periodicity, amplitude of vibrations, mucosal wave, and nonvibrating segment) (Table 7). Of all the analyzed factors, only the amplitude of vocal fold vibrations in group with mild dysplasia ($P = 0.03$) was statistically significant for recurrent disease. In this group, some factors indicated a higher risk of recurrence but not statistically significant enough. Patients with asymmetry in vocal fold vibrations and irregular vocal fold vibrations, with abnormal amplitude of vocal fold vibrations, and the existing nonvibrating segment were at higher risk of recurrence. None of the factors was statistically

TABLE 4.
Treatment Modalities for the Different Grades of Dysplasia

Cordectomy Types	Mild Dysplasia (%)	Moderate Dysplasia (%)	Severe Dysplasia (%)	All Patients (%)
I	34/53 (64.1)	17/26 (65.4)	12/33 (36.4)	63/112 (56.2)
II	19/53 (35.9)	9/26 (34.6)	19/33 (57.6)	47/112 (42.0)
III	0/53 (0)	0/26 (0)	2/33 (6.0)	2/112 (1.8)

TABLE 5.
Patients With Recurrent Disease, Progression of the Disease, and Developed Invasive Carcinoma After 12 Months of Follow-Up

	Recurrence (%)	Progression of the Disease (%)	Malignant Transformation (%)
Mild dysplasia	12/53 (22.6)	4/53 (7.5)	0/53 (0)
Moderate dysplasia	9/26 (34.6)	3/26 (11.5)	1/26 (3.8)
Severe dysplasia	7/33 (21.2)	3/33 (9.1)	3/33 (9.1)
All patients	28/112 (25)	10/112 (8.9)	4/112 (3.6)

significant in patients with moderate dysplasia, but factors, such as cordectomy type, abnormal amplitude of the vocal fold vibrations, and the existence of nonvibrating segment, placed the patients at greater risk of recurrent disease. In patients with severe dysplasia type of cordectomy, insufficient glottic occlusion and abnormal amplitude of the vocal fold vibrations carried higher risk of recurrent disease than others, but this result was not statistically significant. In this group, logistic regression could not be performed for phase symmetry because all patients had asymmetric vibrations of the vocal fold vibrations. For all patients, regardless of the degree of dysplasia, abnormal amplitude of vocal fold vibrations ($P = 0.01$) was a significant factor connected with recurrence. Considering the other factors, the type of cordectomy and the existence of vibratory segment placed the patients at greater risk of recurrence, but this result was not statistically significant.

Regarding the disease progression, none of the considered factors proved to be statistically significant. Some factors placed the patients at greater risk of progression (Table 8). The existence of nonvibrating segment carried higher risk of progression in group with mild and moderate dysplasia and in all patients regardless of the degree of dysplasia. In a group of patients with severe dysplasia, logistic regression could not be performed for the phase symmetry because all patients had asymmetric vibrations of the vocal fold vibrations.

DISCUSSION

In our study, there is male predominance in laryngeal dysplasia (87.5% males). The average age of our patients was 55.63 years, and most patients were in their sixth and seventh decades of life. There were 105 (93.7%) smokers. These facts aligned with other studies, with slight variations; therefore, we can generally expect this demography.[2,9–11] Malignant transformation occurred in 3.6% of the patients after 12 months of follow-up. Most patients were from the group with severe dysplasia. Ricci et al[2] observed a recurrence rate of 14.1% for all dysplasia patients. Dispenza et al observed a recurrence rate of 13.2% for patients with LIN1 and 29% for patients with LIN2 after a 1-year follow-up.[9] Malignant transformation in those studies ranged from 6.48%[2] to 16%.[9,10,12] Weller et al observed a progression rate (according to severity) of 21% with severe dysplasia compared with 14% for mild and moderate dysplasia, which differed from our results.

Gamboa et al[13] conducted a study on stroboscopic assessment of chronic laryngitis in 27 patients (eight of whom had different degrees of laryngeal dysplasia). Among the 15 cases with absence of mucosal wave in the stroboscopic exploration, 60% of the cases had severe dysplasia with squamous cell carcinoma. The authors concluded that the stroboscopic results were related to the pathologic results.

Atypical mucosal waves, as viewed through stroboscopy, should travel one-half of the width of the superior surface of the vocal fold during modal phonation. A reduced mucosal wave and decreased amplitude during modal phonation signifies stiffness, which may result from a lesion, edema, or scar. The vocal fold epithelium normally shows five to 10 cell layers and a thickness of 100–200 μm. Arens et al[14] determined that vocal fold mucosa shows progressive thickening from normal epithelium (147 μm) over the different epithelial dysplasia grades (grade I epithelial dysplasia, 258 μm; grade II epithelial dysplasia, 301 μm; and CIS, 445 μm) up to early invasive carcinoma (974 μm). This result can explain the increasing number of patients with decreased mucosal wave in the three dysplasia groups. Colden et al[15] also conducted a study to determine whether stroboscopy is a reliable method for differentiating invasive glottic carcinoma from intraepithelial atypia and determining the depth of cancer invasion. The authors examined 62 keratotic lesions (45 intraepithelial and 17 carcinomas). The reduced amplitude of vocal fold vibration and/or mucosal wave propagation associated with keratosis did not reliably predict the presence of cancer or the depth of cancer invasion. Reductions in the amplitude of vocal fold vibration and in mucosal wave magnitude were noticed in intraepithelial atypia, despite the fact that there was no invasion into the superficial lamina propria. The authors concluded that the reduced epithelial flexibility could be caused by voluminous keratosis without dysplasia and that abnormalities of the superficial lamina propria could be provoked by inflammation or fibrovascular scarring; for this reason, the absence of mucosal wave was not synonymous with malignancy.

The existence of atypical vocal fold vibration patterns was also reported in normophonic speakers.[16] Nonvibrating segments were associated with the existence of malignant infiltration of the vocal fold epithelium and basal membrane. In our patients, nonvibrating segments were present in 15.1% of the patients with mild hyperplasia, 38.5% of the patients with moderate hyperplasia, and 54.5% of the patients with severe hyperplasia. Shaw and Deliyski[17] determined the presence of atypical magnitude and symmetry of the mucosal waves in the vocal fold vibration of normophonic speakers. In their study, mucosal wave absence was noted in at least 21% of vocal fold vibration

TABLE 6.
Multivariate Regression Analysis of Correlation of Dysplasia With Stroboscopic Signs and Type of Treatment

	Cordectomy Type	Glottic Occlusion	Phase Symmetry	Periodicity	Amplitude	Mucosal Wave	Nonvibratory Segment
t	5.31	7.65	1.39	2.36	6.69	5.288	10.39
OR (95% CI)	1.22 (0.76 to 1.68)	1.88 (1.39 to 2.37)	0.80 (−0.34 to 1.94)	1.24 (0.19 to 2.28)	2.03 (1.43 to 2.63)	1.68 (1.05 to 2.30)	2.950 (2.39 to 3.51)
Sig.	0.006*	0.791	0.076	0.265	0.474	0.636	0.000*

Abbreviations: CI, confidence interval; OR, odds ratio.
* $P < 0.05$.

samples from normophonic speakers. The authors strongly advised that caution should be used when determining the abnormality of mucosal wave variations during clinical visualization procedures. One concern in our study is that in 45.5% of the patients with histopathologically determined CIS, nonvibrating segments were absent, which is a significant number.

Treatment involves removing the lesion with epithelium, basal membrane, and lamina propria and, depending on the type of cordectomy, deeper underlying structures. Vocal fold scaring was examined on animal models. Rousseau et al[18,19] described the development of a vocal fold scar 6 months after surgical injury in canine and rabbit models. As early as 2 months after the surgical removal of the epithelium and lamina propria, no significant difference in collagen density was noted, but at 6 months after injury, collagen density was significantly increased in the surgically injured animals compared with those with normal vocal folds. By 6 months, the procollagen and elastin levels had achieved the densities observed in normal vocal folds, although the elastin fibers remain fragmented and disorganized. The basal layer of the mucosal epithelium continues to experience remodeling in the later stages of wound healing, whereas the intercellular epithelial space undergoes remodeling earlier during the acute stage of wound healing.[20]

Kishimoto et al[21] investigated the maturation process of vocal fold scarring after cordectomy in 10 patients (eight with early laryngeal carcinoma and two with laryngeal dysplasia) using videostroboscopy. The patients were treated with cordectomy types I–III. Improvements in amplitude of mucosal wave were visible 6 months after the procedure and continued to improve up to 14 months after the procedure. Twelve months after the initial treatment was a reasonable time to assess the treatment results in our study. Indeed, there were improvements in phase symmetry, periodicity, amplitude of vocal fold vibrations, and the regularity of mucosal wave. The number of patients with nonvibratory segment decreased. At the end of the follow-up period, there were 23 (20.53%) patients with detected nonvibrating segment. Four patients who developed invasive carcinoma were among these patients. In other patients, this result could be explained by the vocal fold scarring process, particularly because in these patients, type II and type III cordectomies were performed as a treatment of choice. This is yet another limiting factor for stroboscopy use because it cannot reliably distinguish the vocal fold process resulting from the existence of a nonvibrating segment.

Many voice disorders are marked by either aperiodicity or fluctuating frequency and, therefore, cannot be visualized with stroboscopy.[22] There are a growing number of articles that emphasize the importance of different and more effective methods in evaluating irregular vocal fold vibrations and the propagation and existence of the mucosal wave, such as electroglottography, high-speed digital imaging, videokymography, or digital kymography. Mucosal wave propagates in both vertical and horizontal directions, and quantifying the vertical displacement is crucial for understanding the effect of pathologies on the mucosal wave. Stroboscopy, videokymography, and high-speed digital imaging only provide a two-dimensional image

TABLE 7.
Logistic Regression Analysis of Different Stroboscopic Signs and Type of Treatment and Recurrent Disease

	Pretreatment			
	Mild Dysplasia	Moderate Dysplasia	Severe Dysplasia	All Patients
Cordectomy type				
Sig.	0.84	0.34	0.17	0.13
OR (95% CI)	1.15 (0.29 to 4.49)	2.45 (0.39 to 15.49)	3.03 (0.61 to 15.02)	1.96 (0.82 to 4.67)
Glottic occlusion				
Sig.	0.66	0.2	0.49	0.40
OR (95% CI)	0.73 (0.18 to 2.92)	0.33 (0.06 to 1.78)	2.21 (0.22 to 21.78)	0.68 (0.27 to 1.68)
Phase symmetry				
Sig.	0.71	0.96	—	0.84
OR (95% CI)	0.65 (0.07 to 6.21)	0.93 (0.07 to 11.99)		0.85 (0.16 to 4.33)
Periodicity				
Sig.	0.71	0.66	1	0.58
OR (95% CI)	0.65 (0.07 to 6.21)	0.58 (0.05 to 6.58)	0	0.64 (0.13 to 3.16)
Amplitude				
Sig.	0.03*	0.22	0.6	0.01*
OR (95% CI)	2.93 (1.08 to 7.95)	2.80 (0.53 to 14.7)	2.03 (0.14 to 28.86)	2.79 (1.25 to 6.22)
Mucosal wave				
Sig.	0.94	0.77	1	0.83
OR (95% CI)	0.96 (0.34 to 2.71)	0.79 (0.17 to 3.66)	0	0.91 (0.39 to 2.11)
Nonvibratory segment				
Sig.	0.06	0.65	0.49	0.35
OR (95% CI)	4.62 (0.95 to 22.51)	1.47 (0.28 to 7.63)	0.55 (0.10 to 2.97)	1.53 (0.63 to 3.72)

Abbreviations: CI, confidence interval; OR, odds ratio.
*$P < 0.05$.

TABLE 8.
Logistic Regression Analysis of the Different Stroboscopic Signs and Type of Treatment and Disease Progression

	Pretreatment			
	Mild Dysplasia	Moderate Dysplasia	Severe Dysplasia	All patients
Cordectomy type				
Sig.	0.13	0.99	0.92	0.78
OR (95% CI)	0.16 (0.02 to 1.68)	0	1.10 (0.14 to 8.73)	0.84 (0.25 to 2.79)
Glottic occlusion				
Sig.	0.88	0.85	0.7	0.97
OR (95% CI)	1.2 (0.11 to 12.54)	1.28 (0.10 to 16.34)	0.61 (0.048 to 7.76)	0.97 (0.23 to 4.01)
Phase symmetry				
Sig.	0.99	0.99	—	0.99
OR (95% CI)	0	0		0
Periodicity				
Sig.	0.99	0.99	0.99	0.99
OR (95% CI)	0	0	0	0
Amplitude				
Sig.	0.66	0.79	0.75	0.55
OR (95% CI)	1.37 (0.33 to 5.7)	1.33 (0.15 to 12.09)	1.84 (0.04 to 73.47)	1.41 (0.46 to 4.26)
Mucosal wave				
Sig.	0.75	0.59	1	0.56
OR (95% CI)	0.77 (0.15 to 3.95)	0.53 (0.05 to 5.48)	0	0.68 (0.19 to 2.48)
Nonvibratory segment				
Sig.	0.07	0.31	0.45	0.21
OR (95% CI)	7.17 (0.84 to 60.79)	3.75 (0.29 to 47.99)	0.38 (0.03 to 4.69)	2.29 (0.62 to 8.48)

Abbreviations: CI, confidence interval; OR, odds ratio.

of the mucosal wave, whereas digital kymography can be used to provide a complete three-dimensional profile of vocal fold vibration dynamics.[23,24] Although these new techniques are superior to stroboscopy, there are some limitations to their use. High-speed imaging systems are still too expensive to be widely used in clinical practice, gathering larger data sets is problematic because of that fact, and there are no general accepted clinical protocols in laryngology for these techniques.[25]

After performing the multivariate regression analysis, some factors, such as the type of cordectomy and the existence of a nonvibrating segment, were set apart from others and were proven to significantly correlate with various levels of dysplasia ($P < 0.05$). However, after logistic regression of all chosen factors (ie, cordectomy type and stroboscopic signs—glottic occlusion, phase symmetry, periodicity, amplitude of vibrations, mucosal wave, and nonvibrating segment), only the abnormal amplitude of vocal fold vibrations was observed to occur significantly more frequently for recurrent disease in the group with mild dysplasia and in all patients. None of the factors was significant for disease progression. The presence of some factors placed the patients at higher risk of recurrence and progression of the disease. The patient group with mild dysplasia and abnormal vocal fold vibration amplitudes had a 2.93 times greater risk of recurrence, and the group with nonvibrating segments was at 4.62 times greater risk compared with patients without those stroboscopic signs. Nonvibrating segment placed those patients at a 7.17 times greater risk of disease progression than those patients without nonvibrating segment during stroboscopic examination. In the group with moderate dysplasia, patients with insufficient glottic occlusion and abnormal amplitude of vocal fold vibrations were at a greater risk of recurrence. In that group, the patients with insufficient glottic occlusion, with abnormal amplitude of vocal fold vibrations, and the existence of nonvibrating segment were at a greater risk of disease progression. In the patient group with severe dysplasia, the greater risk of recurrence and disease progression aligned with the type of cordectomy and abnormal amplitude of vocal fold vibrations. These findings could also be the result of a relatively small number of patients in the different dysplasia groups, which is one of the limitations of this study. With a larger number of patients, some of the stroboscopic signs could be more prominent. Chang et al[26] conducted a study on a small (18 patients) and nonhomogenous group of patients with laryngeal dysplasia and carcinoma to determine whether the clinical features and clinical appearance of the lesions at presentation correlated with the outcomes of treatment in terms of cure rate and voice outcome. They noted that the clinical appearance of the lesion at presentation, as judged by either still light endoscopy or stroboscopy, did not correlate with disease recurrence. The lesion appearance on still light endoscopy and vibratory characteristics on stroboscopy also did not correlate with the disease-free interval or voice outcome after endoscopic resection.

Stroboscopy is a subjective method in terms of a stroboscopic parameter rating system, and the person conducting the procedure should be well trained to reduce variation and bias. Because of the increasing popularity of stroboscopy equipment in the general otolaryngology office, it is useful to point out some limitations of stroboscopy that can benefit less experienced examiner. In this article, we showed that a large and clinically significant number of cases with CIS with absence of nonvibrating segments can be overlooked when relying solely on stroboscopy. Caution must be exercised when assessing stroboscopic findings, particularly during the posttreatment follow-up period, or if other more sophisticated means of diagnostics are unavailable.

CONCLUSION

Stroboscopy cannot be used reliably for classifying laryngeal dysplasia. Some stroboscopic signs cannot be used as an indication for performing or not performing laryngomicroscopy with biopsy in cases of any suspicious vocal fold lesions. In the absence of more expensive and advanced diagnostic methods, vocal fold dysplasia could be precisely classified only by histopathology analysis. The patient age, treatment modality, and stroboscopic signs, such as abnormal amplitude of vocal fold vibration and the existence of nonvibrating segment, can be considered as warning factors for recurrence and disease progression.

REFERENCES

1. Gallo A, De Vincentiis M, Della Rocca C, Moi R, Simonelli M, Minni A, Shaha AR. Evolution of precancerous laryngeal lesions: a clinicopathological study with long-term follow-up on 259 patients. *Head Neck*. 2001;23:42–47.
2. Ricci G, Molini E, Faralli M, Simoncelli C. Retrospective study on precancerous laryngeal lesions: long-term follow-up. *Acta Otorhinolaryngol Ital*. 2003;23:362–367.
3. Isenberg JS, Crozier DL, Dailey SH. Institutional and comprehensive review of laryngeal leukoplakia. *Ann Otol Rhinol Laryngol*. 2008;117:74–79.
4. Shanmugaratnam K, Sobin LH. The World Health Organization histological classification of tumours of the upper respiratory tract and ear. A commentary on the second edition. *Cancer*. 1993;71:2689–2697.
5. Bosman FT. Dysplasia classification: pathology in disgrace? *J Pathol*. 2001;194:143–144.
6. Hirano M, Bless DM. *Videostroboscopic Examination of the Larynx*. San Diego, CA: Singular Publishing Group Inc; 1993.
7. Kelley RT, Colton RH, Casper J, Paseman A, Brewer D. Evaluation of stroboscopic signs. *J Voice*. 2011;25:490–495.
8. Remacle M, Van Haverbeke C, Eckel H, Bradley P, Chevalier D, Djukic V, de Vicentiis M, Friedrich G, Olofsson J, Peretti G, Quer M, Werner J. Proposal for revision of the European Laryngological Society classification of endoscopic cordectomies. *Eur Arch Otorhinolaryngol*. 2007;264:499–504.
9. Dispenza F, De Stefano A, Marchese D, Martines F, Dispenza C. Management of laryngeal precancerous lesions. *Auris Nasus Larynx*. 2012;39:280–283.
10. Montgomery J, White A. A decade of laryngeal dysplasia in Paisley, Scotland. *Eur Arch Otorhinolaryngol*. 2012;269:947–951.
11. Spielmann PM, Palmer T, McClymont L. 15-Year review of laryngeal and oral dysplasias and progression to invasive carcinoma. *Eur Arch Otorhinolaryngol*. 2010;267:423–427.
12. Weller M, Nankivell P, McConckey C, Paleri V, Mehanna M. The risk and interval to malignancy of patients with laryngeal dysplasia: a systematic review of case series and meta-analysis. *Clin Otolaryngol*. 2010;35:364–372.
13. Gamboa J, Echeverría L, Molina B, Cobeta I. Stroboscopic assessment of chronic laryngitis. *Acta Otorrinolaringol Esp*. 2006;57:266–269.
14. Arens C, Glanz H, Wönckhaus J, Hersemeyer K, Kraft M. Histologic assessment of epithelial thickness in early laryngeal cancer or precursor lesions and its impact on endoscopic imaging. *Eur Arch Otorhinolaryngol*. 2007;264:645–649.
15. Colden D, Jarboe J, Zeitels SM, Bunting G, Hillman RE, Spanou K. Stroboscopic assessment of vocal fold keratosis and glottic cancer. *Ann Otol Rhinol Laryngol*. 2001;110:293–298.

16. Haben CM, Kost K, Papagiannis G. Mucosal wave asymmetries in the clinical voice laboratory. *J Otolaryngol*. 2002;31:275–280.
17. Shaw HS, Deliyski DD. Mucosal wave: a normophonic study across visualization techniques. *J Voice*. 2006;22:23–33.
18. Rousseau B, Hirano S, Scheidt TD, Welham NV, Thibeault SL, Chan RW, Bless DM. Characterization of vocal fold scarring in a canine model. *Laryngoscope*. 2003;113:620–627.
19. Rousseau B, Hirano S, Chan RW, Welham NV, Thibeault SL, Ford C, Bless DM. Characterization of chronic vocal fold scarring in a rabbit model. *J Voice*. 2004;18:116–124.
20. Hansen JK, Thibeault SL. Current understanding and review of the literature: vocal fold scarring. *J Voice*. 2006;20:110–120.
21. Kishimoto Y, Hirano S, Tateya I, Kanemaru S, Ito J. Temporal changes in vocal functions of human scarred vocal folds after cordectomy. *Laryngoscope*. 2010;120:1597–1601.
22. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008;117:413–424.
23. George NA, de Mul FF, Qiu Q, Rakhorst G, Schutte HK. Depth kymography: high-speed calibrated 3D imaging of human vocal fold vibration dynamics. *Phys Med Biol*. 2008;53:2667–2675.
24. Krausert CR, Olszewski AE, Taylor LN, McMurray JS, Dailey SH, Jiang JJ. Mucosal wave measurement and visualization techniques. *J Voice*. 2011; 25:395–405.
25. Verikas A, Uloza V, Bacauskiene M, Gelzinis A, Kelertas E. Advances in laryngeal imaging. *Eur Arch Otorhinolaryngol*. 2009;266: 1509–1520.
26. Chang J, Fang TJ, Yung K, et al. Clinical and histologic predictors of voice and disease outcome in patients with early glottic cancer. *Laryngoscope*. 2012;122:2240–2247.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2016; 125(4):331-335.

Article

Annals of Otology, Rhinology & Laryngology
2016, Vol. 125(4) 331–335
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489415613801
aor.sagepub.com
SAGE

Prognostic Relevance of Mucosal Waves in Patients With Unilateral Vocal Fold Paralysis

Rudolf Reiter, MD, PhD[1], Anja Pickhard, MD, PhD[2], Sylvia Sander[3], and Sibylle Brosch, MD, PhD[1]

Abstract

Objectives: To analyze the prognostic relevance of mucosal waves (MWs) for recovery of unilateral vocal fold paralysis (UVP).

Methods: The charts and stroboscopic examinations of 100 consecutive patients with a complete UVP were reviewed retrospectively. All had a minimal (estimated <3 mm) mucosal gap on stroboscopy. A positive or negative MW on the paralyzed vocal fold was associated with complete recovery to full adduction and abduction. All patients were followed for at least 12 months.

Results: Causes of the paralysis were iatrogenic/traumatic (n = 82), malignancy associated (n = 10), and idiopathic (n = 8). In patients with positive MW at diagnosis (n = 80), the chance of recovery of unilateral vocal fold paralysis was 91.25%, whereas the chance of recovery with a negative mucosal wave (n = 20) was only 10%.

Conclusion: Positive MWs in stroboscopy are a predictor for recovery of (iatrogenic/traumatic) unilateral vocal fold paralysis and should be used in routine diagnostic assessment.

Keywords

unilateral vocal fold paralysis, mucosal wave, microstroboscopy, recovery

Introduction

Unilateral vocal fold paralysis (UVP) is a challenge for otolaryngologists and phoniatricians. Insufficient glottic closure during phonation can lead to severe vocal impairment with dysphonia and reduced vocal intensity. The etiology of UVP is often traumatic (especially in thyroid surgery) followed by malignancy-associated and idiopathic paralysis.[1] Recovery of vocal fold paralysis is mainly observed within 12 months after onset, as shown in a literature review of 717 cases with an idiopathic UVP.[2,3] Recovery of postoperative vocal fold paralysis in patients with thyroidectomy usually occurs within the first 6 months,[4] but according to a literature review, up to 11% do not recover.[5] Knowing the prognosis of UVP is helpful in planning therapy such as voice therapy and augmentation, either early temporary or permanent.[6-8] Stroboscopy is a tool for imaging the vibration of the vocal folds during phonation, especially the mucosal waves.[9,10] By analyzing the vibrations of the vocal folds, an assessment can be made as to the state of the mucosa (cover) and the underlying laryngeal muscle tone (body).[9,11] In patients with UVP, reduced vocal fold movement (adduction or abduction), vocal fold bowing, incomplete glottis closure, or a vibratory asymmetry, and absent mucosal waves (MW) can be observed in videostroboscopy. Recurrence of MW suggests reinnervation of the paretic vocal fold.[12-15] The aim of this study was to analyze the prognostic relevance of MW for recovery of mobility to full adduction and abduction in unilateral vocal fold paralysis in 100 patients.

Material and Methods

A retrospective review was made of a select group of 100 consecutive patients who had a UVP with complete vocal

[1]Section of Phoniatrics and Pedaudiology, Department of Otolaryngology, Head and Neck Surgery, University of Ulm, Ulm, Germany
[2]Department of Otolaryngology Head and Neck Surgery, Technical University, Munich, Germany
[3]Institute of Epidemiology and Medical Biometry, University of Ulm, Germany

Corresponding Author:
Rudolf Reiter, Section of Phoniatrics and Pedaudiology, Department of Otolaryngology, Head and Neck Surgery, University of Ulm, Frauensteige 12, 89070 Ulm, Germany.
Email: rudolf.reiter@uniklinik-ulm.de

fold immobility and a minimal mucosal gap (estimated <3 mm by 2 experienced phoniatricians) in videolaryngostroboscopy between January 2008 and January 2013. They were identified in our "Electronic Patient Record of the University Hospital Ulm"—a specially designed electronic health record. As we have a close connection to a large over-regional center of (thyroid) surgery, all their patients with voice problems are routinely seen in our department, usually the day after surgery, and could be included in this study. Taken together, all stroboscopic examinations were done 1 day to a maximum of 3 days after surgery (iatrogenic paralysis) and about up to 3 weeks after onset of the symptoms (other paralyses).

The videolaryngostroboscopic vocal fold examination was done (90° endostroboscope 5052, Wolf, Hamburg, Germany) and documented (rpSzene, Rehder, Hamburg, Germany). For measurement of the glottal gap and MW presence on the paralyzed vocal fold by a phoniatrician, phonation trials were performed with a sound pressure level of 65 dB and 100 Hz for men and 200 Hz for women. Only patients diagnosed with a UVP who had a vocal fold positioned with a minimal mucosal gap (estimated <3 mm) by videolaryngostroboscopy were included in this study group. It was only in these patients that evaluation by videolaryngostroboscopy was technically feasible because the vocal fold could get into contact (couple). Analysis of the presence of a MW in the paralyzed vocal fold was done by 2 experienced phoniatricians.

Patients with chronic laryngitis/leucoplakia, carcinoma, and scar formation of the vocal folds were not included as these pathologies influence mucosal waves. Similarly, patients with a medialization or augmentation of the paralyzed vocal fold were excluded. Positive MW (pMW) and negative MW (nMW) on the paralyzed vocal fold in stroboscopy were associated with complete recovery of the paralysis to full adduction and abduction. All patients were followed for at least 12 months at an interval of 3 to 6 months.

Statistical data analysis was performed with Microsoft Excel 2003 and SAS 9.3 (SAS Institute, Cary, North Carolina, USA). In the descriptive statistical analysis median, minimum and maximum were calculated for quantitative variables. For qualitative variables, absolute and relative frequencies with corresponding exact 95% confidence interval were calculated. The study was approved by the local Ethics Committees at the University of Ulm.

Results

Patients, Etiology of UVP, and Therapy

A hundred patients (median age 49.6 years; range, 16-81 years; 28 males/72 females) with a UVP were included. The group was divided iatrogenic/traumatic group (n = 82; median age 46.5 years; range, 16-76 years; 18 males/64 females), a malignancy-associated UVP (n = 10; median age 51.3 years; range, 16-70 years; 4 males/6 females), and an idiopathic UVP (n = 8; median age 43.5 years; range, 20.5-65.7 years; 6 males/2 females). The etiology of the 82 patients with iatrogenic/traumatic was thyroidectomy in 61 patients (56 thyroid hypertrophy and 5 patients with thyroid carcinoma), cardiac/carotid surgery in 14 patients, and spine surgery in 7 patients. In 10 patients, a UVP was observed as a primary symptom of a malignancy (8 carcinoma of the thyroid gland, 2 bronchial carcinoma), and in 8 cases, idiopathic paralysis was diagnosed. These patients received voice therapy (n = 88) or observation (n = 12). Patients with a medialization or augmentation of the paralyzed vocal fold were not included.

Time of Recovery

Out of the 100 patients, a complete recovery was seen in 75 patients with none in the remaining 25 patients. This recovery was observed at a median of 6.6 months (range, 1-14 months) after diagnosis. In the group of the 67 patients with iatrogenic/traumatic paresis, complete recovery occurred at a median of 6.5 months (range, 2-14 months). In thyroid hypertrophy surgery associated paralysis (n = 43), recovery was observed at a median of 4.6 months (range, 2-12 months) and in surgery because of a thyroid carcinoma (n = 5) in 10.4 months (range, 4-14 months). The 1 malignancy-associated UVP recovered after 10 months and the idiopathic paralysis (n = 7) after a median of 5.7 months (range, 3-9).

Relationship of Recovery With MW

As mentioned before, recovery was observed in 75% (75/100; 95% CI, 65.3-83.1). In total, 80 patients had pMW and 20 nMW.

In all patients with pMW at diagnosis, the chance of recovery of UVP was 91.25% (73/80; 95% CI, 82.8%-96.4%) (Figure 1), whereas the chance of recovery of the UVP was only 10% (2/20; 95% CI, 1.2%-31.7%) in patients with nMW (Figure 2, Table 1).

In the subgroup of the 82 patients with iatrogenic/traumatic UVP, a recovery was observed in 81.7% (67/82; 95% CI, 71.6%-89.4%), especially in 90.4% (66/73; 95% CI, 81.2%-96.1%) in patients with pMW. In 3 cases with nMW that did not resolve, the nerve was cut.

The best chance of recovery was in the group of the patients with idiopathic UVP in 87.5% (7/8; 95% CI, 47.4%-99.7%). All patients who recovered had pMW. Only 1 patient (10%, 1/10; 95% CI, 0.3%-44.5%) recovered in malignancy-associated UVP. This patient had nMW. Unilateral vocal fold paralyses with pMW were not observed in this group (Tables 2-4).

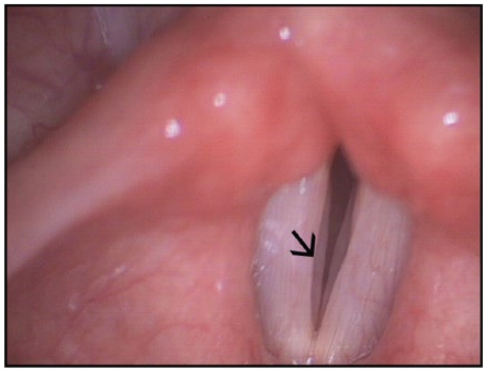

Figure 1. A patient with a unilateral vocal fold paralysis (UVP) on the right side. Positive mucosal waves (MWs) are seen in microstroboscopy on the paralyzed right (→) and the non-paralyzed vocal left fold.

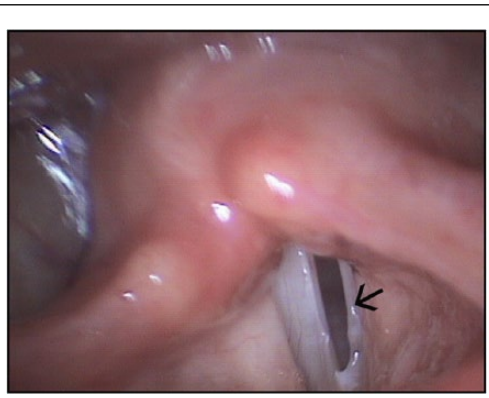

Figure 2. A patient with a unilateral vocal fold paralysis (UVP) on the left side. Positive mucosal waves (MWs) are seen in microstroboscopy on the non-paralyzed right side but are absent in the paralyzed left vocal fold (→).

Table 1. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in All Patients.

UVP	Positive MW	Negative MW	Total
Recovery	73	2	75
No recovery	7	18	25
Total	80	20	100

Discussion

The main reasons for UVP are iatrogenic/traumatic paralysis (about 4/5, especially in thyroid surgery) followed by malignancy-associated and idiopathic paralysis, as shown recently in a cohort of 400 patients[1] and confirmed in our study. Iatrogenic paralysis after (thyroid) surgery represents the majority of patients because we are associated with a large supra-regional center of thyroid surgery where difficult cases including revision operations are performed. All patients with voice problems after surgery were sent to our department and could be included in this study.

Unilateral vocal fold paralysis often causes severe impairment of the voice with dysphonia and reduced intensity because of insufficient glottic closure during phonation. However, UVPs have a potential of resolution that usually occurs within 12 months and in most cases within the first 6 months[2-4] (also confirmed in our study).

The overall rate of recovery was 75 of 100 (75%) in all patients and 67 of 82 (82%) in the iatrogenic/traumatic group, respectively. Studies have shown that permanent paralysis remains in 15% after thyroid surgery and in up to 20% in 717 patients with an idiopathic paralysis—similar to our study.[2,4]

In order to plan therapy and keep the patient informed, it is advantageous to know if any resolution of paralysis is likely. Laryngeal electromyography (LEMG) is an established method of assessing the neuromuscular status of the paralyzed vocal fold. Early evidence of reinnervation or innervation in paresis can be shown by this method. Laryngeal electromyography criteria for poor prognosis were the presence of spontaneous activity and absence or reduced recruitment of motor unit potentials.[16] In a meta-analysis by Rickert et al[16] and an analysis by Sittel et al,[17] LEMG predicted defective recovery defined as absence of completely free vocal fold mobility in up to 94%. However, in a study with a small cohort, it was shown that LEMG findings predict only 44.4% of the resolved cases,[18,19] which is less satisfactory. A drawback of the LEMG is the fact that it takes time and is an invasive procedure not well tolerated by patients. Furthermore, it requires expensive equipment and an experienced investigator.[9]

By analyzing the vibrations of the vocal folds, the state of the mucosa (cover), and the underlying laryngeal, muscle tone (body) can be assessed.[9,11] In microlarygostroboscopy, the mucosal wave represents the clinical correlate for muscular tonicity of the vocal cord. When present, recurrent nerve function is at least in a sense present, that the epithelium of the vocal cord can exactly follow the musculus vocalis movements in pitch as well as in intensity. However, mucosal waves can only be judged when there is no phonation gap ≤3 mm in UVP, enabling the mucosa to get in contact with each other. Furthermore, MWs are reduced or absent if mucosa (cover) adheres/sticks to the musculus vocalis and/or ligamentum vocale (body). This phenomenon is observed in patients with chronic laryngitis/leucoplakia, carcinoma, and scar formation of the vocal folds. The warning factors for progression of dysplasia to invasive carcinomas is a nonvibrating segment/absence or reduction of the MW in videostroboscopy.[20,21] A scarred vocal fold has an absent or limited MW[22] as well, and for this reason, such patients were excluded from our study.

To our knowledge, this is the first clinical study to report the prognostic relevance of MW in laryngostroboscopy for

Table 2. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in Iatrogenic/Traumatic UVP.

UVP	Positive MW	Negative MW	Total
Recovery	66 (42 thyroid hypertrophy, 5 thyroid carcinoma, 12 cardiac/carotid surgery, 7 spine surgery)	1 (1 thyroid hypertrophy)	67
No recovery	7 (6 struma, 1 cardiac/carotds surgery)	8 (7 thyroid hypertrophy, 1 cardiac/carotid surgery)	15
Total	73	9	82

Table 3. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in Malignancy-Associated UVP.

UVP	Positive MW	Negative MW	Total
Recovery	0	1 (1 thyroid carcinoma)	1
No recovery	0	9 (7 thyroid carcinoma, 2 bronchial carcinoma)	9
Total	0	10	10

Table 4. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in Idiopathic UVP.

UVP	Positive MW	Negative MW	Total
Recovery	7	0	7
No recovery	0	1	1
Total	7	1	8

recovery of a (iatrogenic/traumatic) UVP. It was shown that the chance of recovery of UVP was over 90% when pMW were present at diagnosis in patients with a UVP, whereas only in 10% if not. Laryngostroboscopy is always done routinely in patients with UVP by an experienced phoniatrician or laryngologist. It is a noninvasive procedure that takes only a few minutes.

A disadvantage of laryngostrobsocopy is the fact that in our experience, reliable stroboscopic signals are only obtained in patients with the paralyzed vocal fold close to midline during phonation (glottis gap ≤3 mm during phonation). This is the reason why only 61 out of 100 patients could be analysed in a study by Harries and Morrison.[9] In a recent study with 400 patients with a vocal fold paralysis, it was shown that microstroboscopy was technically feasible in 76% of the patients because the vocal folds could get into contact (couple).[1] These data show that a routinely performed microlarygostroboscopy is a very good, noninvasive alternative to EMG in patents with UVP and—in addition—much better tolerated and less expensive.

Conclusion

The authors conclude that positive mucosal waves in microstroboscopy are a simple predictor for recovery of (iatrogenic/traumatic) UVP and should be used in routinely for diagnosis and prognosis

Acknowledgment

We thank John Bennett, FRCS, for critically reading this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Reiter R, Pickhard A, Smith E, et al. Vocal cord paralysis—analysis of a cohort of 400 patients. *Laryngorhinootologie*. 2015;94(2):91-96.
2. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118(7):13030-13037.
3. Misono S, Merati AL. Evidence-based practice: evaluation and management of unilateral vocal fold paralysis. *Otolaryngol Clin North Am*. 2012;45(5):1083-1108.
4. Steurer M, Passler C, Denk DM, Schneider B, Niederle B, Bigenzahn W. Advantages of recurrent laryngeal nerve identification in thyroidectomy and parathyroidectomy and the importance of preoperative and postoperative e laryngoscopic examination in more than 1000 nerves at risk. *Laryngoscope*. 2002;112(1):124-133.
5. Dralle H, Sekulla C, Lorenz K, Brauckhoff M, Machens A; German IONM Study Group. Intraoperative monitoring of the recurrent laryngeal nerve in thyroid surgery. *World J Surg*. 2008;32(7):1358-1366.
6. Reiter R, Brosch S. Laryngoplasty with hyaluronic acid in patients with unilateral vocal fold paralysis. *J Voice*. 2013;26:785-791.
7. Friedman AD, Burns JA, Heaton JT, Zeitels SM. Early versus late injection medialization for unilateral vocal cord paralysis. *Laryngoscope*. 2010;120(10):2042-2046.
8. Yung KC, Likhterov I, Courey MS. Effect of temporary vocal fold injection medialization on the rate of permanent medialization laryngoplasty in unilateral vocal fold paralysis patients. *Laryngoscope*. 2011;121(10):2191-2194.

9. Harries ML, Morrison M. The role of stroboscopy in the management of a patient with a unilateral vocal fold paralysis. *J Laryngol Otol*. 1996;110(2):141-143.
10. Sercarz JA, Berke GS, Ming Y, Gerratt BR, Natividad M. Videostroboscopy of human vocal fold paralysis. *Ann Otol Rhinol Laryngol*. 1992;101(7):567-577.
11. Hirano M, Yoshida Y, Yoshida T, Tateishi O. Strobofiberscopic video recording of vocal fold vibration. *Ann Otol Rhinol Laryngol*. 1985;94:588-590.
12. Simpson CB, Cheung EJ, Jackson CJ. Vocal fold paresis: clinical and electrophysiologic features in a tertiary laryngology practice. *J Voice*. 2009;23(3):396-398.
13. Simpson CB, May LS, Green JK, Eller RL, Jackson CE. Vibratory asymmetry in mobile vocal folds: is it predictive of vocal fold paresis? *Ann Otol Rhinol Laryngol*. 2011;120(4):239-242.
14. Altman KW. Laryngeal asymmetry on indirect laryngoscopy in a symptomatic patient should be evaluated with electromyography. *Arch Otolaryngol Head Neck Surg*. 2005;131(4):356-360.
15. Kosztyła-Hojna B, Rogowski M. Usefulness of videolaryngo-stroboscopy in the diagnosis of laryngeal pathology. *Pol Merkur Lekarski*. 2003;14(83):413-416.
16. Rickert SM, Childs LF, Carey BT, Murry T, Sulica L. Laryngeal electromyography for prognosis of vocal fold palsy: a meta-analysis. *Laryngoscope*. 2012;122(1):158-161.
17. Sittel C, Stennert E, Thumfart WF, Dapunt U, Eckel HE. Prognostic value of laryngeal electromyography in vocal fold paralysis. *Arch Otolaryngol Head Neck Surg*. 2001;127(2):155-160.
18. Blitzer A, Crumley RL, Dailey SH, et al. Recommendations of the Neurolaryngology Study Group on laryngeal electromyography. *Otolaryngol Head Neck Surg*. 2009;140(6):782-793.
19. Munin MC, Rosen CA, Zullo T. Utility of laryngeal electromyography in predicting recovery after vocal fold paralysis. *Arch Phys Med Rehabil*. 2003;84(8):1150-1153.
20. Djukic V, Milovanovic J, Jotic AD, Vukasinovic M. Stroboscopy in detection of laryngeal dysplasia effectiveness and limitations. *J Voice*. 2014;28(2):262.e13-262.
21. El-Demerdash A, Fawaz SA, Sabri SM, Sweed A, Rabie H. Sensitivity and specificity of stroboscopy in preoperative differentiation of dysplasia from early invasive glottic carcinoma. *Eur Arch Otorhinolaryngol*. 2015;272(5):1189-1193.
22. Chhetri DK, Berke GS. Injection of cultured autologous fibroblasts for human vocal fold scars. *Laryngoscope*. 2011;121(4):785-792.

Reprinted by permission of Laryngoscope. 2010; 120(7):1376-1382.

The Laryngoscope
© 2010 The American Laryngological, Rhinological and Otological Society, Inc.

Laryngoscopy of Vocal Fold Paralysis: Evaluation of Consistency of Clinical Findings

David E. Rosow, MD; Lucian Sulica, MD

Objectives/Hypothesis: Laryngoscopy is the principal tool for the clinical assessment of vocal fold paralysis (VFP). Yet no consistent, unified vocabulary to describe laryngoscopic findings exists, compromising the evaluation and comparison of cases, outcomes, and treatment. The goal of this investigation was to evaluate laryngoscopic findings in VFP for inter- and intra-rater consistency.

Study Design: Prospective survey-based study.

Methods: Half-minute excerpts from stroboscopic exams of 22 patients with VFP were mailed to 22 fellowship-trained laryngologists. Each reviewer was sent exams in randomized order, with three random repeats included to determine intra-rater reliability. Twelve laryngoscopic criteria were assessed and recorded on preprinted sheets. Eleven criteria were binary in nature (yes/no); glottic insufficiency was rated on a four-point scale (none/mild/moderate/severe). Raters were blinded to clinical history, each other's ratings, and to their own previous ratings. Inter-rater agreement was calculated by Fleiss' kappa.

Results: Twenty reviewers (91%) replied. Intra-rater reliability by reviewer ranged from 66% to 100% and by laryngoscopic criterion from 77% to 100%. Of the laryngoscopic criteria used, glottic insufficiency (κ = 0.55), vocal fold bowing (κ = 0.49), and salivary pooling (κ = 0.45) showed moderate agreement between reviewers. Arytenoid stability (κ = 0.1), arytenoid position (κ = 0.12), and vocal fold height mismatch (κ = 0.12) showed poor agreement. The remainder showed slight to fair agreement.

Conclusions: Inter-rater agreement on commonly used laryngoscopic criteria is generally fair to poor. Glottic insufficiency, vocal fold bowing, and salivary pooling demonstrated the most agreement among responding laryngologists. These findings suggest a need for a standardized descriptive scheme for laryngoscopic findings in VFP.

Key Words: Vocal fold paralysis, vocal cord paralysis, laryngoscopy, reliability, agreement.

Level of Evidence: 2b

***Laryngoscope*, 120:1376–1382, 2010**

INTRODUCTION

Laryngoscopy is the mainstay investigation in the diagnosis of vocal fold paralysis (VFP), and not infrequently the sole diagnostic evaluation on which direct treatment (as opposed to treatment of underlying cause) is based. For much of the history of laryngology, complex nosological schemes have been constructed around the laryngoscopic appearance of VFP. Semon's law, for instance, held that differences in vocal fold position were the product of differential vulnerability of adductor and abductor fibers of the recurrent laryngeal nerve.[1] Wagner and Grossman maintained that the position of the paralyzed vocal fold was indicative of the integrity of the superior laryngeal nerve.[2,3] Such constructs were abandoned as increasing anatomical knowledge and careful physiological investigations invalidated their assumptions. In the course of this progress, systematic analysis of the laryngoscopic appearance of VFP has apparently been abandoned too, as unrewarding in the face of the evident complexity of the neuropathology underlying VFP.

Yet, it is clear to any clinician that VFP manifests itself laryngoscopically in many different ways. Terms like height and length mismatch, arytenoid prolapse, flaccidity, posterior gap, and others that plainly refer to physical characteristics of the appearance of the paralyzed vocal fold make their appearance in the professional discourse with some frequency. Woodson, in a seminal study of the paralyzed vocal fold, described several such features: foreshortening, arytenoid displacement, decreased vocal process contact, bowing, and ventricular hyperfunction.[4] Recent literature has addressed vocal process height asymmetry.[5,6] Both the configuration and degree of glottic insufficiency related to VFP have been presented as important in the selection of treatment techniques.[7,8]

From the Department of Otorhinolaryngology, Weill Cornell Medical College, New York, New York, U.S.A.

Editor's Note: This Manuscript was accepted for publication March 17, 2010.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Lucian Sulica, MD, Director, Laryngology/Voice Disorders, Department of Otorhinolaryngology, Weill Cornell Medical College, 1305 York Avenue, New York, NY 10021. E-mail: lus2005@med.cornell.edu

DOI: 10.1002/lary.20945

Rating Sheet for Evaluators

Examiner #_____

Patient #_____

Volitional Adduction	Present	Not Present
Vocal Process Contact	Normal	Impaired/Decreased
Arytenoid Stability ("Jostle Sign")	Normal (no jostle)	Impaired/Decreased
Arytenoid Position	Normal	Displaced
Vocal Fold Tone	Normal	Decreased
Vocal Fold Atrophy	Not Present	Present
Vocal Fold Bowing	Not Present	Present
Vocal Fold Shortening	Not Present	Present
Salivary Pooling	Not Present	Present
Height Mismatch	Not Present	Present
Ventricular (False vocal fold) contraction	Not Present	Present
Glottic Insufficiency	None Mild	Moderate Severe

Fig. 1. Rating sheet for evaluators.

Despite the acknowledgment of variability in the laryngoscopic appearance of VFP implicit in these terms, no widely accepted rating system, or even a consistent, unified vocabulary to describe such variability exists. Consistency and reproducibility is fundamental in the evaluation and comparison of cases, their outcomes, and their treatment; even a brief reflection on the House-Brackmann scale for grading facial paralysis reveals the broad potential utility of such a standardized approach.

The goal of this investigation was to evaluate characteristics of the laryngoscopic appearance of VFP with respect to inter- and intra-rater consistency, and to identify features for which clinical consensus exists, which might lend themselves to a useful, standardized description system for VFP.

MATERIALS AND METHODS

Selection of Examinations

Strobovideolaryngoscopies of patients with a clinical diagnosis of VFP based on history, physical examination, and laryngoscopy were selected from a corpus of such examinations recorded during the course of routine evaluation. All recording was made under stroboscopic light using either a rigid glass rod peroral laryngoscope (Model 9106; KayPentax, Lincoln Park, NJ) or a distal chip flexible transnasal laryngoscope (VNL-1170K; Pentax Medical, Montvale, NJ). Only patients with VFP of known cause were included. Nineteen had pathology or injury limited to the recurrent laryngeal nerve, and three had paralysis from vagal neuropathy. For inclusion, the examination had to feature a sustained, unobstructed view of the glottis, arytenoids, aryepiglottic folds, and pyriform sinuses. A 20-second sample of each exam, containing at least one example each of phonatory adduction and postphonatory abduction as well as several cycles of phonatory vibration, was selected and saved. The pitch and intensity capabilities of patients were variable from exam to exam, as one would expect in cases of VFP. However, as evaluators' ratings of identical examinations were assessed in this study, no effort was made to standardize these parameters among examinations.

Patients who had been treated for their VFP in any way, including injection augmentation, framework surgery, and reinnervation were excluded. Cases of vocal fold paresis in which significant gross vocal fold mobility remained, even if it was clearly less than normal, were excluded. The authors recognize that the distinction between paralysis and paresis is not always sharply defined and does not necessarily reflect the underlying neurologic status.

Each exam was numbered, and randomly ordered lists of these exams, one for each potential reviewer, were generated. In each list, three exams were selected to be repeated by a random number generator, which created a new set of three numbers for each of the reviewers and brought the total number of exams to 25. Therefore, every reviewer had a randomly selected series of repeat examinations to test their intra-rater reliability. The purpose of de novo, random selection of the repeat exams for every reviewer was to eliminate any possibility that one exam might be more easily identified on repeat viewing than the others. Such a scenario would bias the entire sample and yield an artificially

TABLE I.
Examiners and Their Intra-Rater Reliabilities Determined Via the Three Repeat Examinations Given to All Participants.

Examiner	% Agreement	Pearson	Spearman
A	0.944	0.932	0.898
B	0.806	0.794	0.775
C	0.889	0.836	0.839
D	0.833	0.849	0.768
E	0.917	0.906	0.906
F	0.861	0.846	0.764
G	1.000	1.000	1.000
H	0.944	0.949	0.957
I	0.667	0.616	0.446
J	0.778	0.700	0.678
K			
L	0.694	0.605	0.526
M	0.778	0.711	0.661
N	0.778	0.752	0.652
O			
P	0.833	0.773	0.681
Q	0.861	0.893	0.775
R	0.750	0.700	0.623
S	0.889	0.890	0.803
T	0.750	0.675	0.650
U	1.000	1.000	1.000
V	0.833	0.631	0.622
Average	0.840	0.803	0.751

Examiners K and O did not participate in the study.

high intra-rater reliability. By randomizing the repeat exams across all reviewers, this potential source of bias was eliminated. Each list of exams was then burned to a DVD with all identifying information removed; the file names on the disc simply appeared in order as 01.avi, 02.avi, and so on.

Selection of Reviewers

Twenty-two fellowship-trained laryngologists were asked to participate in the study. A package containing the disc of videos, 25 prelabeled rating sheets, an institutional review board waiver, and a return mailer was sent to each. Results from returned rating sheets were entered into Excel spreadsheets (Microsoft Corp., Redmond, WA) for analysis.

Examination Rating

Each stroboscopic exam was rated with respect to 12 features (Fig. 1). A standard vocabulary to describe the paralyzed vocal fold does not exist. Therefore, criteria were selected from a literature review of the diagnosis and treatment of VFP, including those articles cited above[4–8] and others,[9,10] and informal consultation with colleagues. No formal definition of each term was presented; the study relied on common clinical usage to inform raters' perception of the meaning of each term rather than any formal training. We also acknowledge that some terms might be in part redundant or overlapping—for example, tone, atrophy and bowing—but sought to evaluate the utility of each of these concurrently. Responses to 11 of 12 categories were binary (normal/not normal or present/not present); the exception was glottic insufficiency, which was rated on a four-point scale of none/mild/moderate/severe. However, for the purposes of statistical analysis, answers in this category were grouped in binary fashion into none/mild or moderate/severe. Reviewers were instructed to view and rate examinations sequentially; individual exams could be reviewed an unlimited number of times, but once a new examination was started, examiners were instructed not to return to any prior examinations or alter ratings. Raters were therefore blinded to clinical history, each other's ratings, and their own previous ratings. Reviewers were also asked not to include written justification for their answers on the rating forms.

Statistical Analysis

Inter-rater reliability was determined using the kappa statistic as described by Fleiss[11,12]; values closer to 0 represent poor agreement, whereas those close to 1 represent near-perfect agreement. Although no uniformly agreed-upon scale exists for Fleiss' kappa, Fleiss described a scale where values >0.75 represent excellent agreement, 0.40 to 0.75 represent fair to good agreement, and values <0.40 represent poor agreement.[12] When an examiner omitted a rating for one of the 12 categories for an examination (accidental or otherwise), all of that examination's rankings in that category were excluded from kappa

TABLE II.
Laryngoscopic Criteria and Their Associated Inter- and Intra-Rater Reliability Scores.

	Inter-Rater	Intra-Rater		
Laryngoscopic Criterion	Fleiss	% Agreement	Pearson	Spearman
Volitional adduction	0.335	0.900	−0.053	−0.053
Vocal process contact	0.303	0.817	0.445	0.445
Arytenoid stability (jostle)	0.097	0.833	0.615	0.615
Arytenoid position	0.119	0.817	0.629	0.629
Vocal fold tone	0.310	0.900	0.744	0.744
Vocal fold atrophy	0.326	0.867	0.726	0.726
Vocal fold bowing	0.488	0.883	0.714	0.714
Vocal fold shortening	0.225	0.817	0.610	0.610
Salivary pooling	0.454	0.900	0.762	0.762
Height mismatch	0.123	0.733	0.457	0.457
Ventricular contraction	0.217	0.883	0.756	0.756
Glottic insufficiency	0.550	0.733	0.818	0.798

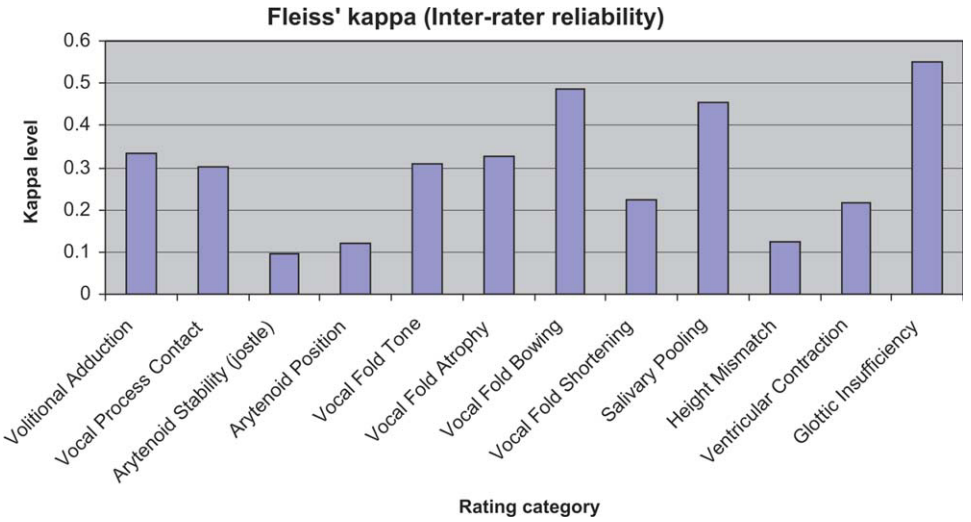

Fig. 2. Inter-rater reliability as determined by Fleiss' kappa. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

analysis. This exclusion assured a constant denominator in all the statistical calculations.

Intra-rater reliability for each examiner was determined by comparing the 12 laryngoscopic criteria in each of the three repeated patients, for a total of 36 comparison points. Intra-rater reliability for each criterion was determined in a similar fashion, with the denominator determined by adding up the 20 examiners' three repeated tests, by criterion, for a total of 60 comparison points. Three complementary methods were used to assess intra-rater reliability, both for each examiner and for each laryngoscopic criterion. The overall percent agreement was calculated, simply as the number of points of agreement divided by the total. This was compared to two known measures of correlation, Pearson product moment coefficient and Spearman corrected rank correlation coefficient.

This investigation was approved by the institutional review board of Weill Cornell Medical College.

RESULTS

Twenty of 22 examiners returned the survey, for a 91% response rate. The overall intra-rater reliability for each individual examiner varied between 44% and 100%, with an average internal consistency of 75% to 84%, depending on the statistical method used (Table I). As measured by all three statistics, 18 of 20 examiners (90%) showed >60% internal consistency (Table I).

The intra-rater reliability for each stroboscopic criterion had, for the most part, a very similar range of 44% to 100% (Table II), The single exception was volitional adduction, a category in which not a single examiner rated an exam as normal on both viewings. As a result, despite a 90% rate of intra-rater agreement, this category was found to have near-0 intra-rater correlation by both Spearman and Pearson correlation coefficients. Overall, height mismatch, vocal fold shortening, and vocal process contact had the lowest intra-rater reliability, whereas the ratings of salivary pooling, glottic insufficiency, ventricular contraction, and vocal fold tone were generally consistent.

Inter-rater reliability for each stroboscopic criterion was determined by kappa analysis. As represented in Figure 2, these kappa values ranged from 0.10 (poor

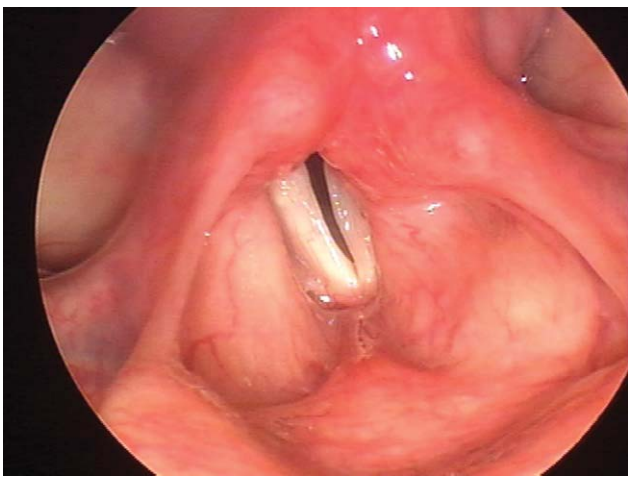

Fig. 3. Vocal process contact impaired. This case generated the most consistent rating for impaired vocal process contact. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

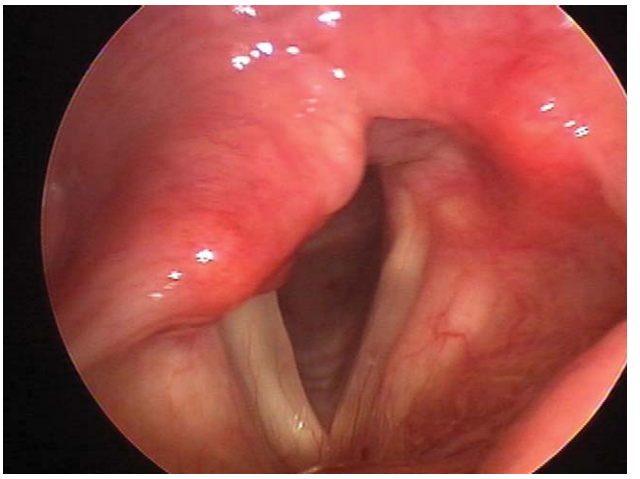

Fig. 4. Arytenoid position displaced. This case generated the most consistent rating for displaced arytenoid position. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

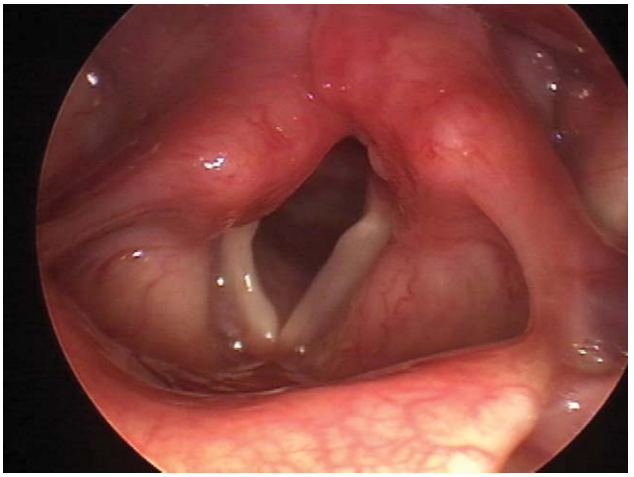

Fig. 5. Vocal fold tone decreased. This case generated the most consistent rating for decreased vocal fold tone. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

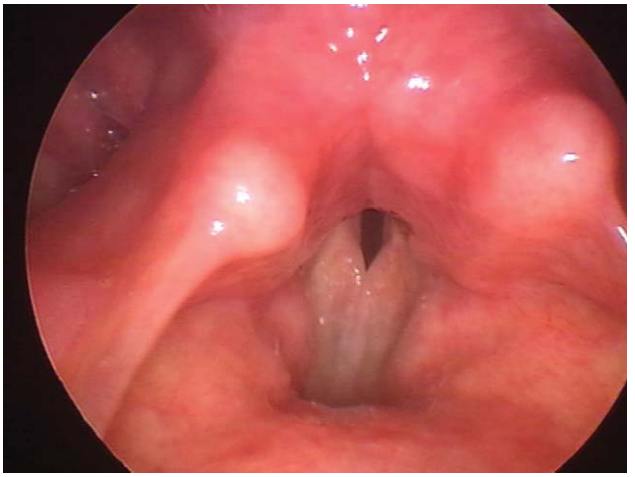

Fig. 7. Glottic insufficiency: none to mild. This case generated the most consistent rating for no or mild glottic insufficiency. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

agreement) to 0.55 (moderate agreement). Ratings for arytenoid stability, arytenoid position, and height mismatch generated the poorest values, whereas glottic insufficiency, vocal fold bowing, and salivary pooling resulted in the best (Fig. 3–Fig. 8].

DISCUSSION

Traditionally conceptualized as an all-or-none phenomenon, VFP has been shown by ample clinical and laboratory investigation to represent a continuum of neurogenic dysfunction encompassing partial denervation, complete denervation, and variable degrees and patterns of reinnervation. It should come as no surprise, then, that its clinical appearance too is highly variable. This is not synonymous with random, however; this variability no doubt reflects the considerable heterogeneity in the neurologic dysfunction that underlies the immobile vocal fold. Historical efforts to decipher the laryngoscopic appearance have fallen short, compromised by oversimplifications and an incomplete understanding of the relevant pathophysiology. Given these limitations, we find these failures neither particularly surprising nor discouraging.

Lest a re-examination of laryngoscopy in VFP be considered unnecessary or irrelevant, it is important to appreciate that existing neurodiagnostic techniques have also been defeated by the complex neurologic picture underlying the paralyzed vocal fold. Electromyography has proved to be as qualitative as laryngoscopy, and it is similarly susceptible to individual variation in interpretation. Although it has yielded crucial insight in the

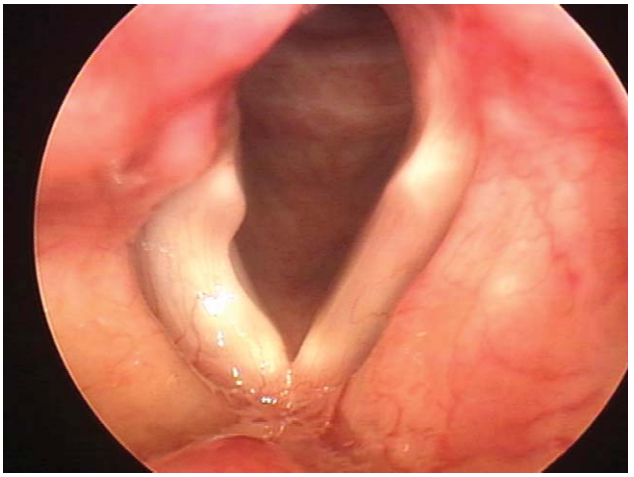

Fig. 6. Vocal fold shortened. This case generated the most consistent rating for shortened vocal fold. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

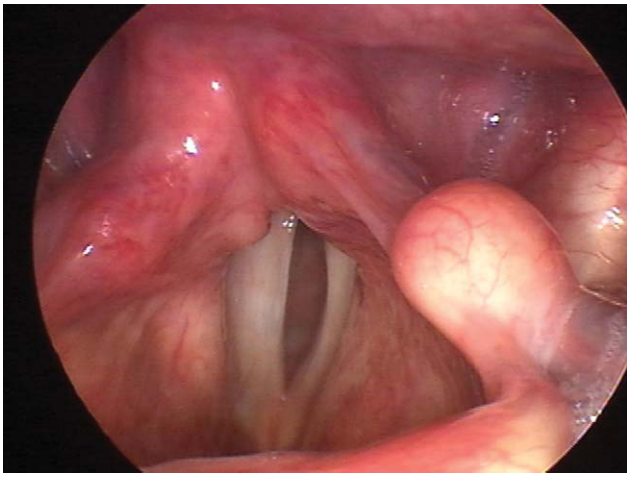

Fig. 8. Glottic insufficiency: moderate to severe. This case generated the most consistent rating for moderate to severe glottic insufficiency. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

pathophysiology of VFP, it has proved particularly disappointing as a prognostic tool, at best only partially helpful in informing treatment of individual patients.

This study was undertaken with the understanding that any re-examination of the laryngoscopy of VFP must be carried out using a commonly agreed-upon array of findings. The evolving discussion about vocal fold paresis, or partial paralysis of the vocal fold, reveals clearly that lack of consensus regarding clinical findings prevents conclusions regarding diagnosis, much less prognosis, treatment, and outcomes.[13]

No definitions of the terms used were provided to the reviewers for the simple reason that no such definitions exist. The terms used in this study have entered the literature informally, and the medical discourse has generally assumed that a broad and consistent understanding of these descriptors exists. This study is a formal testing of that assumption.

This investigation reveals that the evaluation of laryngoscopic appearance of VFP remains a personal and individual activity. As demonstrated by multiple correlation calculations, most evaluators were relatively consistent in their own evaluations across the entire range of features presented. These results were roughly equivalent to those of Rosen, who found that two thirds of voice professionals reviewing stroboscopic exams had intra-rater reliability scores less than 0.80.[14] Thus, it is possible that individual practitioners might use laryngoscopic features to analyze cases for diagnosis, as for example to identify degree of denervation, distribution of involvement across laryngeal muscles, and for selection and timing of treatment in a reasonably reliable manner. It remains to be proven, of course, that pathophysiologic aspects such as degree and distribution of neural compromise indeed have consistent laryngoscopic correlates. Incidentally, this study does not demonstrate whether individual observations across multiple examinations are reliable, or if changes over time in the same case can be consistently identified.

On the other hand, inter-rater variability revealed considerable lack of consensus regarding all aspects but salivary pooling, bowing, and a simplified rating of the degree of glottic insufficiency. Our study might even have been biased in favor of greater inter-rater agreement by the inclusion of the audio track in the video samples sent to reviewers. Such additional information might provide clues to blinded reviewers who are ultimately being studied for their video perceptual analysis alone. Future work in this area will require removal of all audio from samples sent to reviewers. Not only is this lack of agreement discouraging for the development of a unified rating system for this disorder, it also calls into question existing assumptions in the literature about consensus in the rating of features such as posterior gap (an important factor in the selection of patients for arytenoid adduction surgery), vocal fold height (hypothesized to be relevant to rehabilitation technique), and other features referred to in the discussion of the evaluation of unsatisfactory results of medialization.[7,15–17] Generalizations from study to study might be compromised by patient populations that are not comparable or equivalent. Also, treatment recommendations or descriptions of outcome based on laryngoscopic features are likely to be of limited utility. The prospects for agreement on vocal fold paresis, where clinical variability would be expected to be greater than in VFP at the same time that the degree of abnormality would be less, appear to be extremely poor.

Based on our results, degree of glottic insufficiency, vocal fold bowing, salivary pooling, and perhaps to a lesser extent volitional adduction, vocal fold tone, and vocal fold atrophy appear to be the best candidates for development into a standardized system of rating VFP. A rating or classification system for VFP based on only the three most consistently appreciated criteria might not be discriminating enough to be useful in diagnosis or treatment. We hypothesize that more formal development of rating categories, including explicit definitions and examples, would generate greater inter-rater agreement, for the terms and concepts evaluated in this investigation have received relatively little formal attention despite commonplace clinical use. We intend to explore this further before trialing an integrated rating system. At the same time, we recognize the possibility that individual variation in laryngeal anatomy and possibly in innervation, and the heterogeneity of neuropathic dysfunction might yet defeat such an effort.

CONCLUSION

However, although individuals are often consistent in their own evaluation of laryngoscopic features of VFP, little consensus appears to exist among physicians regarding these same findings. This raises the possibility that many assumptions about the significance of laryngoscopic features might not be reliable. This is an obvious challenge to arriving at a unified understanding of the laryngoscopic appearance of the disorder and will need to be addressed.

Results suggest that degree of glottic insufficiency, vocal fold bowing, and salivary pooling appear to be laryngoscopic features in cases of VFP with the highest inter-rater reliability. With further investigation and standardization, these might form a basis for the development of a clinically useful rating scheme.

Acknowledgments

We gratefully acknowledge the participation of the following as raters or consultants in this study: Drs. Linda Gerber, Gregory Postma, Michael Johns, Edward Damrose, Marshall Smith, Tanya Meyer, Gaelyn Garrett, Richard Kelly, Mark Courey, Seth Dailey, Ajay Chitkara, Sid Khosla, Pieter Noordzij, Libby Smith, Phillip Song, James Thomas, Albert Merati, Andrew McWhorter, Joel Blumin, Michiel Bove, and Milan Amin.

BIBLIOGRAPHY

1. Semon F. Clinical remarks on the proclivity of the abductor fibres of the the recurrent laryngeal nerve to become

affected sooner than the adductor fibres, or even exclusively, in cases of undoubted central or peripheral injury or disease of the roots or trunks of the pneumogastric, spinal accessory, or recurrent nerves. *Arch Laryngol New York* 1881;2:197–222.
2. Wagner R. Die medianstellung des stimmbandes bei Recurrenslahmung. *Archiv fur Pathologische Anatomie und Physiologie* 1890;120:437–459.
3. Grossman M. Experimentelle beitrage zur lehre von der posticuslahmung. *Archiv Laryngologie und Rhinologie* 1897;6:282–360.
4. Woodson GE. Configuration of the glottis in laryngeal paralysis. I: clinical study. *Laryngoscope* 1993;103(11 pt 1): 1227–1234.
5. Tsai V, Celmer A, Berke GS, Chhetri DK. Videostroboscopic findings in unilateral superior laryngeal nerve paralysis and paresis. *Otolaryngol Head Neck Surg* 2007;136: 660–662.
6. Rubin AD, Praneetvatakul V, Heman-Ackah Y, Moyer CA, Mandel S, Sataloff RT. Repetitive phonatory tasks for identifying vocal fold paresis. *J Voice* 2005;19: 679–686.
7. Netterville JL, Stone RE, Luken ES, Civantos FJ, Ossoff RH. Silastic medialization and arytenoid adduction: the Vanderbilt experience. A review of 116 phonosurgical procedures. *Ann Otol Rhinol Laryngol* 1993;102: 413–424.
8. Woodson GE, Picerno R, Yeung D, Hengesteg A. Arytenoid adduction: controlling vertical position. *Ann Otol Rhinol Laryngol* 2000;109:360–364.
9. Benninger MS, Crumley Rl, Ford CN, et al. Evaluation and treatment of the unilateral paralyzed vocal fold. *Otolaryngol Head Neck Surg* 1994;111:497–508.
10. Woo P. Arytenoid adduction and medialization laryngoplasty. *Otolaryngol Clin North Am* 2000;33:817–839
11. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–382.
12. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd ed. New York, NY: Wiley; 1981.
13. Sulica L, Blitzer A. Vocal fold paresis: evidence and controversies. *Curr Opin Otolaryngol Head Neck Surg* 2007;15: 159–162.
14. Rosen CA. Stroboscopy as a research instrument: development of a perceptual evaluation tool. *Laryngoscope* 2005; 115:423–428.
15. Woo P, Pearl AW, Hsiung MW, Som P. Failed medialization laryngoplasty: management by revision surgery. *Otolaryngol Head Neck Surg* 2001;124:615–621.
16. Anderson TD, Spiegel JR, Sataloff RT. Thyroplasty revision: frequency and predictive factors. *J Voice* 2003;17: 442–448.
17. Cohen JT, Bates DD, Postma GN. Revision Gore-Tex medialization laryngoplasty. *Otolaryngol Head Neck Surg* 2004;131:236–240.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2011; 120(4):239-242.

Annals of Otology, Rhinology & Laryngology 120(4):239-242.
© 2011 Annals Publishing Company. All rights reserved.

Vibratory Asymmetry in Mobile Vocal Folds: Is It Predictive of Vocal Fold Paresis?

C. Blake Simpson, MD; Linda Seitan May, MD; Jill K. Green, MS; Robert L. Eller, MD; Carlayne E. Jackson, MD

Objectives: The purpose of this study was to determine whether the videostroboscopic finding of vibratory asymmetry in mobile vocal folds is a reliable predictor of vocal fold paresis. In addition, the ability of experienced reviewers to predict the distribution (left/right/bilateral) of the paresis was investigated.

Methods: This is a retrospective chart review of all patients who presented to our clinic during a 3-year period with symptoms suggestive of glottal insufficiency (vocal fatigue or reduced vocal projection) accompanied by the videostroboscopic findings of bilateral normal vocal fold mobility and vibratory asymmetry. Twenty-three of these patients underwent diagnostic laryngeal electromyography of the thyroarytenoid and cricothyroid muscles to determine the presence of vocal fold paresis.

Results: Nineteen of the 23 patients (82.6%) were found to have electrophysiological evidence of vocal fold paresis, either unilaterally or bilaterally, when videostroboscopic asymmetry was present in mobile vocal folds. However, the three expert reviewers' ability to predict the distribution (left/right/bilateral) of the paresis was poor (26.3%, 36.8%, and 36.8%, respectively).

Conclusions: The videostroboscopic finding of vibratory asymmetry in mobile vocal folds is a reliable predictor of vocal fold paresis in most cases. However, the ability of expert reviewers to determine the distribution (left/right/bilateral) of the paresis using videostroboscopic findings is poor. This study highlights the value of laryngeal electromyography in arriving at a correct diagnosis in this clinical situation.

Key Words: electromyography, videostroboscopy, vocal fold paralysis, vocal fold paresis.

INTRODUCTION

Vocal fold paresis (VFP) is a well-established, albeit controversial, entity. Its incidence is not well established, but it is likely rare. The few reports that are available in the literature have shown a range of as many as 29 cases in a year to as few as 13 cases over 4 years in tertiary laryngology practices.[1-4] Although all of these studies used laryngeal electromyography (LEMG) to confirm the diagnosis, clinicians often use subtle asymmetries on videostroboscopy as indicators that paresis is likely present. During videostroboscopic examination, reduced vocal fold movement (adduction or abduction), vocal fold bowing, incomplete glottal closure, and vibratory asymmetry can all be associated with VFP.[4,5] Rubin et al[6] have also described the use of repetitive phonatory tasks to induce fatigue as a means of bringing out hypomobility in paretic vocal folds. As pointed out by Sulica and Blitzer, however, "Separating innocent asymmetries [on laryngoscopy] from significant findings may present the greatest challenge in defining vocal fold paresis."[7(p159)] The clinical setting of glottal insufficiency symptoms and grossly intact vocal fold mobility has previously been described. In these cases, vibratory asymmetry may be the only laryngoscopic clue to suggest VFP.[7] Identification of the asymmetry may help guide the clinician toward performing LEMG and eventually confirming a diagnosis of VFP.

The purpose of this study was to determine whether the videostroboscopic finding of vibratory asymmetry in mobile vocal folds was a reliable predictor of VFP. In addition, the ability of experienced reviewers to predict the distribution (left/right/bilateral) of the paresis was investigated.

METHODS

Institutional Review Board approval was obtained

From the Departments of Otolaryngology–Head and Neck Surgery (Simpson, May, Green) and Neurology (Jackson), University of Texas Health Science Center–San Antonio, and the Department of Otolaryngology–Head and Neck Surgery, Wilford Hall Medical Center, Lackland Air Force Base (Eller), San Antonio, Texas.

Presented at the meeting of the American Broncho-Esophagological Association, Las Vegas, Nevada, April 28-29, 2010.

Correspondence: C. Blake Simpson, MD, Dept of Otolaryngology–Head and Neck Surgery, University of Texas Health Science Center, Medical Arts and Research Center, 8300 Floyd Curl Dr, San Antonio, TX 78229.

from our institution before the study period. A retrospective chart review was carried out for all patients who presented to our clinic during a 3-year period and underwent LEMG for suspected vocal fold paresis.

Over the study period, 48 patients with suspected VFP underwent diagnostic LEMG. Of those, 23 patients met the study criteria with symptoms of VFP (vocal fatigue or reduced vocal projection) accompanied by the videostroboscopic findings of bilateral normal vocal fold mobility and vibratory asymmetry. The diagnostic LEMG examinations included an evaluation of the motor unit morphology and recruitment of motor unit potentials (MUPs) for the thyroarytenoid and cricothyroid muscles. Interpretation of the LEMG findings was done by a neurologist (C.E.J.) who was blinded to the findings of the laryngoscopic examination. In all cases, abnormal LEMG findings were considered to be present when there were large-amplitude polyphasic MUPs and incomplete recruitment of MUPs. All abnormal LEMG findings were then classified as left, right, or bilateral, depending on the side of involvement. We did not distinguish between recurrent laryngeal nerve (RLN) and superior laryngeal nerve (SLN) neuropathy for the purposes of this portion of the study. In other words, if the RLN, SLN, or both showed electrophysiological evidence of denervation, the findings were considered "abnormal" for that side.

Our endoscopic clinical examination protocol was as follows. All of the patients underwent videostroboscopy by means of a flexible laryngoscope with a distal chip (Olympus ENF-VQ, Olympus Surgical, Orangeburg, New York) rhinolaryngoscope, and most also had rigid laryngoscopy with a 70° rigid endoscope (KayPENTAX, Lincoln Park, New Jersey). The patients were instructed to phonate /i/ at low, modal, and high frequencies. When indicated, the technique of "unloading" as described by Koufman[8] was also used to help reveal more subtle vibratory asymmetry that may have been hidden under compensatory muscle tension patterns.

When retrospective evaluation of the endoscopic segments was carried out, the following protocol was used. The best-quality videostroboscopic examination (either flexible or rigid) was used for each case. Of the 48 cases in which LEMG was performed for suspected paresis, 23 examinations that were considered to show isolated vibratory asymmetry were selected for the study. The other 25 cases, which showed vocal fold immobility, partial immobility, videostroboscopic evidence of incomplete closure, or vocal fold lesions, were excluded.

TABLE 1. VOCAL FOLD PARESIS DEMOGRAPHICS AND LEMG FINDINGS

Age (y)	*Gender*	*Duration*	*LEMG Findings*	*Cause of Paresis*
62	F	1 y	B RLN + SLN	Idiopathic
67	F	1 y	B RLN	Idiopathic
30	M	9 y	L RLN + SLN	Idiopathic
36	M	36 y	B RLN	Congenital
28	M	4 mo	B RLN	Idiopathic
65	M	6 y	B RLN + SLN	Idiopathic
36	F	10 y	B RLN	Idiopathic
69	F	2 mo	B RLN	Idiopathic
35	F	1 y	B RLN	Idiopathic
36	M	7 y	B RLN	Idiopathic
44	F	9 y	R RLN	Idiopathic
29	F	1.5 y	L RLN	Idiopathic
58	F	9 mo	L RLN	Idiopathic
37	F	1 y	B RLN	Idiopathic
51	F	5 y	L RLN	Idiopathic
43	F	16 mo	R RLN	Idiopathic
76	M	6 mo	B RLN	Idiopathic
58	M	14 mo	B RLN	Idiopathic
54	F	4 mo	L SLN	Traumatic

LEMG — laryngeal electromyography; B — bilateral; RLN — recurrent laryngeal nerve paresis; SLN — superior laryngeal nerve paresis; L — left; R — right.

The videos were edited to include only segments in which the vocal folds were in a fully adducted position and were engaged in vibratory activity. We decided not to show footage of vocal fold mobility, in order to help exclude any possible bias that could occur from interpreting vocal fold movement. The video segments were then randomized and were interpreted by three reviewers with extensive experience in videostroboscopic interpretation. Each video segment was reviewed, and the following questions were addressed: 1) Is asymmetry of vibration (amplitude or mucosal wave) present? 2) If vibration is asymmetric, which side has the increased amplitude and/or mucosal wave? and 3) On which side would you predict the paresis to be present?

The LEMG results were used as the gold standard for the diagnosis of VFP. Interpretation of the videostroboscopic findings by our reviewers was then compared to this gold standard to determine the predictive value of subjective vibratory asymmetry on videostroboscopic examination.

RESULTS

Of the 19 patients with a diagnosis of LEMG-confirmed VFP (Table 1), the mean patient age was 48.5 years (range, 28 to 76 years). Twelve of the patients were female (63.2%) and had a mean age of 48.8 years, and 7 patients were male (36.8%) and had a mean age of 47 years. The mean time interval from

TABLE 2. LEMG RESULTS AND REVIEWERS' INTERPRETATION

Patient	*Reviewer 1*	*Reviewer 2*	*Reviewer 3*	*LEMG*
1	L	R	R	B
2	R	R	R	B
3	R	R	R	L
4	R	B	R	Normal
5	R	B	B	B
6	R	R	R	Normal
7	R	R	R	B
8	L	L	L	B
9	L	R	B	R
10	B	B	B	B
11	R	B	B	B
12	L	L	L	B
13	R	B	L	R
14	L	B	B	Normal
15	R	B	R	L
16	L	L	R	L
17	R	B	B	B
18	R	L	L	Normal
19	L	B	B	L
20	R	B	R	R
21	R	L	R	B
22	R	B	B	B
23	R	B	L	L

L — left-sided paresis; R — right-sided paresis; B — bilateral paresis.

the onset of symptoms to presentation to our clinic was 4.8 years (range, 2 months to 36 years). The cause of the paresis was idiopathic in the vast majority of cases (17 of 19 or 89.5%), and the remaining cases were congenital (1 of 19 or 5.2%) or traumatic (1 of 19 or 5.2%). In terms of neural involvement, the majority of cases involved the RLN only. Ten cases were bilateral RLN paresis, and 5 cases were unilateral RLN paresis. The remaining cases were 2 cases of bilateral combined RLN and SLN paresis, 1 case of unilateral combined RLN and SLN paresis, and 1 case of unilateral SLN paresis.

Of the 23 patients with symptoms of glottal insufficiency and isolated vibratory asymmetry on videostroboscopy, 19 (82.6%) were found to have electrophysiological evidence of denervation of one or both vocal folds (Table 2). However, the individual reviewers' ability to correctly predict the distribution of the paresis was quite poor. Given three options (bilateral, left, or right), each reviewer was unable to correctly predict the side in most cases (reviewer 1, 5 of 19 correct; reviewer 2, 7 of 19 correct; and reviewer 3, 7 of 19 correct). With all examination evaluations combined, the side of paresis was correctly predicted in only 33.3% of cases (19 of 57).

DISCUSSION

The idea behind this study was to answer a common question that is posed in our multidisciplinary clinics. As a general rule, the voice team (which includes the senior author, speech pathologist, and resident physician) reviews the videostroboscopic examination of the patient and discusses the subjective interpretation of the vibratory parameters. In most cases of suspected VFP, the clinicians can agree that vibratory asymmetry is present, and LEMG will later confirm the diagnosis. However, the reliability of using vibratory asymmetry to correctly predict the presence of VFP has not been examined. Although we can usually agree on the presence of vibratory asymmetry, there is often a debate about the sidedness of the suspected paresis. Conventional thinking suggests that the denervated side will have an increased amplitude and/or mucosal wave due to the laxity of the paretic vocal fold. Despite this consensus, we have noted that many times the clinicians do not agree as to which side(s) is involved.

Obviously, the clinical diagnosis of some cases of VFP is fairly straightforward when based on videostroboscopic findings and clinical history. In the setting of gross hypomobility and glottal insufficiency, the diagnosis is not often in question. However, when there are no readily apparent differences in vocal fold mobility, the diagnosis can be more difficult to make, or may not be suspected by the clinician at all. In these cases, vibratory asymmetry may be the only clue that VFP is present.[7] This finding may help guide the clinician toward performing LEMG and establishing a correct diagnosis.

Our clinical protocol for patients with symptoms suggestive of glottal insufficiency and an increased amplitude and/or mucosal wave or "chasing wave" (asymmetry of vibration) is to recommend LEMG. Obviously, not all patients with this combination of symptoms and findings agree to undergo or follow up for diagnostic LEMG, so we are not able to comment on the positive predictive value of vibratory asymmetry in these cases. Nonetheless, when vibratory asymmetry prompted LEMG testing in our series, the clinical "hunch" ended up being correct in 83% of cases. However, the ability of experienced clinicians to correctly predict which side was involved was quite poor (33.3%). This is exactly the percentage one would expect if the clinician's determination were randomly generated; ie, there is a 1-in-3 chance of predicting the outcome correctly. The difficulty partially arises from using the subjective observation that one side demonstrates increased vibratory amplitude (often thought to be a manifestation of reduced muscular tone in a denervated vocal fold). By necessity, that determination involves using the contralateral side as a control, ie, the side with the "normal tone." In many cas-

es, however, this side may also be affected, making the assumption of unilaterality erroneous. Despite this problem, there were many cases in which the reviewer correctly predicted that the paresis was unilateral, but the predicted side (ie, distribution of involvement) was incorrect.

Relying solely on laryngoscopic findings to predict VFP continues to be problematic. Other studies have shown that 25% to 40% of patients had LEMG findings that were not predicted by their laryngoscopic examination.[2,3] Although vibratory asymmetry is fairly predictive of VFP (83% of cases in our study), determining the distribution (left/right/bilateral) of the paresis is very poorly predictive.

Interpretation of videostroboscopic examinations is by nature subjective. We have observed that vibratory asymmetry can sometimes be difficult to detect on routine stroboscopy. The best method of accentuating asymmetry is to have the patient phonate at a modal or low fundamental frequency at a high intensity. In addition, extinguishing any secondary supraglottic muscular tension seems to be beneficial, as this allows for the differential tension of the true vocal folds to be observed. Last, recording the examination and playing it back in slow motion, or performing frame-by-frame analysis, is yet another method to aid in the detection of vibratory asymmetry.

CONCLUSIONS

The videostroboscopic finding of vibratory asymmetry in mobile vocal folds is a reliable predictor of VFP in most cases. However, the ability of expert reviewers to determine the distribution (left/right/bilateral) of the paresis using videostroboscopic findings is poor. This finding highlights the value of LEMG in arriving at a correct diagnosis in this clinical situation.

REFERENCES

1. Merati AL, Shemirami N, Smith TL, Toohill RJ. Changing trends in the nature of vocal fold motion impairment. Am J Otolaryngol 2006;27:106-8.

2. Heman-Ackah YD, Barr A. Mild vocal fold paresis: understanding clinical presentation and electromyographic findings. J Voice 2006;20:269-81.

3. Koufman JA, Postma GN, Cummins MM, Blalock DP. Vocal fold paresis. Otolaryngol Head Neck Surg 2000;122:537-41.

4. Simpson CB, Cheung EJ, Jackson CJ. Vocal fold paresis: clinical and electrophysiologic features in a tertiary laryngology practice. J Voice 2009;23:396-8.

5. Altman KW. Laryngeal asymmetry on indirect laryngoscopy in a symptomatic patient should be evaluated with electromyography. Arch Otolaryngol Head Neck Surg 2005;131:356-60.

6. Rubin AD, Praneetvatakul V, Heman-Ackah Y, Moyer CA, Mandel S, Sataloff RT. Repetitive phonatory tasks for identifying vocal fold paresis. J Voice 2005;19:679-86.

7. Sulica L, Blitzer A. Vocal fold paresis: evidence and controversies. Curr Opin Otolaryngol Head Neck Surgery 2007;15:159-62.

8. Koufman JA. Evaluation of laryngeal biomechanics by fiberoptic laryngoscopy. In: Rubin JA, Sataloff RT, Korovin GS, Gould WJ, eds. Diagnosis and treatment of voice disorders. New York, NY: Igaku-Shoin, 1995:122-34.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2014; 123(4):271-278.

Article

Annals of Otology, Rhinology & Laryngology
2014, Vol. 123(4) 271–278
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489414525022
aor.sagepub.com
SAGE

Reliability of Clinical Office-Based Laryngeal Electromyography in Vocally Healthy Adults

Daniel J. Croake, MM, MS[1], Joseph C. Stemple, PhD[1], Timothy Uhl, PhD[1], Sanford Archer, MD, FACS, FAAOA[2], and Richard D. Andreatta, PhD[1]

Abstract

Objective: This study aimed to conduct a 3-session reliability assessment of the laryngeal electromyography (LEMG) signal in healthy participants during intensity controlled vocalization tasks. We hypothesized that vocal intensity level and testing session would affect LEMG measures.

Methods: This prospective study used a 2-factor repeated measures design. Seven participants underwent bipolar needle LEMG of the right thyroarytenoid muscle. Data were collected over 3 testing sessions using vocalization tasks performed with visually guided intensity feedback targets (65 and 75 dB SPL). Root mean square amplitudes in microvolts were analyzed for within-session and between-session reliability.

Results: The main effect for intensity was found to approach significance (F = 5.71, P = .054). However, intraclass correlation coefficients (ICCs) using a 2-factor mixed random effect model indicated poor to fair signal reliability between testing sessions (ICC = 0.56 at 65 dB, 0.40 at 70 dB). Intraclass correlation coefficients for within-session data indicated excellent reliability for all testing conditions (0.84–0.98).

Conclusion: Using a quantitative analysis protocol to inform an essentially qualitative technique, our results indicated that there was generally poor to fair reliability in the LEMG signal over testing sessions. Vocal intensity was an important variable that affected LEMG signal reliability. Standardization of LEMG protocols using vocal control parameters and quantitative analyses may help improve LEMG reliability in clinical settings.

Keywords

LEMG, motor units, neuromuscular, thyroarytenoid

Introduction

Laryngeal electromyography (LEMG) is commonly used for the assessment of neuromuscular disorders of the larynx.[1] Laryngeal electromyography is the only direct measure of laryngeal muscle activity, and although it provides general information about the function of the laryngeal musculature, its in-office clinical utility beyond general appreciation of gross neuromuscular function is debatable.[2,3] Studies regarding clinical usefulness of LEMG have not addressed the question of LEMG reliability both within a patient and across clinical testing sessions. Nonclinical and experimental investigations using LEMG have the advantage of signal processing software, control of environmental conditions to reduce electromagnetic field noise, tasks that provide graded control of laryngeal muscle recruitment, and less constrained time frames under which to perform LEMG. Because many variables may affect LEMG reliability, it remains unknown whether clinical office-based LEMG, often performed qualitatively, without vocalization control parameters and under less than optimal recording conditions, can be considered a reliable and clinically meaningful diagnostic tool.

Laryngeal electromyography has been used for research and assessment of the function of the intrinsic laryngeal muscles for more than 60 years and has been shown to be useful in revealing the function of the laryngeal musculature and demonstrating the dynamic control of this musculature during voicing.[4-7] Laryngeal electromyography is frequently used in conjunction with stroboscopic/

[1]College of Health Sciences, University of Kentucky, Lexington, Kentucky, USA
[2]College of Medicine, University of Kentucky, Lexington, Kentucky, USA

Corresponding Author:
Daniel J. Croake, MM, MS, Department of Rehabilitation Sciences, College of Health Sciences, University of Kentucky, Room 120H Wethington Building, 900 South Limestone Street, Lexington, KY 40536-0200, USA.
Email: djcroa2@uky.edu

laryngoscopic evaluation of vocal fold mobility to assess a variety of laryngeal disorders such as dystonia and vocal fold paresis and paralysis, to differentiate among various neurological disorders,[5] and for guiding the placement of botulinum toxin for the treatment of spasmodic dysphonia.[1,8] Laryngeal electromyography may also be potentially useful as a tool for the prognosis of laryngeal nerve disorders.[9,10]

In general, electromyography (EMG) recordings are affected by multiple confounding variables including electrode type and placement, level of muscle activation, left and right side dominance, artifact from electrode movement, and so on, which may all, or in part, compromise the accuracy of the data necessary for diagnostic evaluation.[11,12] Electromyographic investigations of between-session and intra-session reliability for some limb muscles have revealed high reliability for both between- and within-session measurements.[13,14] However, similar data for laryngeal-based EMG are absent and cannot be directly interpolated from limb studies due to significant differences in anatomical structure and the ability to control for muscle length and loading.

Unlike most limb muscles that have skeletal support and firm attachment points, the larynx is suspended in the neck, surrounded by soft muscle tissue, a series of membranes, and a somewhat yielding cartilaginous framework. Distinct muscle force and leverage points are difficult to determine in the laryngeal complex because of these flexible attachment points. Placing a consistent and measureable isotonic load on laryngeal muscles for accurate and reliable activation is difficult, complicating replication of motor unit activation in these muscles.

Another factor to be considered regarding the reliability of clinical in-office LEMG is that phonation is an emergent behavior, arising through the complex interaction of respiratory, phonatory, and resonance subsystems of the vocal tract. These vocal subsystems function synergistically, integrating properties of tissue elasticity, muscle activation, and aerodynamics toward normal vocal function. A change in any subsystem's dimension will potentially alter vocal output. These additional confounding variables have the potential to further complicate in-office LEMG interpretation.[15]

In general, reliable LEMG measurements are dependent on consistent muscle activation tasks. These tasks must be carefully controlled and performed for measurement reliability. For example, to describe relative recruitment of motor unit potentials for the thyroarytenoid muscle (TA), maximum voluntary contraction (MVC) strategies have been used for comparison. Typically, a maximal voluntary contraction is assigned a 100% possible recruitment value whereby subsequent muscle contractions during voicing tasks are given a percentage of decreased recruitment. Maximum voluntary contraction in the laryngeal system is typically accomplished through performance of a Valsalva maneuver (hard breath hold).[8] However, it has been shown that vocal fold closure is not consistently accomplished during Valsalva maneuvers up to 14% of the time, potentially leading to significant diagnostic error.[16] Other qualitative ratings such as decreased recruitment scales are not comparable across offices due to their highly subjective nature and lack of standardized between-office collection protocols.[17]

Another commonly used clinical LEMG technique is comparison of recruitment against the contralateral muscle. Unfortunately, this technique does not take into account the notion that the contralateral TA muscle is dependent on the co-contraction of neighboring intrinsic muscles. Thus, TA contraction may be altered in the presence of a contralateral paresis or paralysis. In this scenario, compensatory muscle activation is a likely confounder.[18] In addition, a large-scale retrospective study reported unexpected contralateral neuropathy in 26% of patients with laryngeal movement disorders.[9] Electromyography studies of limb muscle also indicate significant contralateral differences in motor unit recruitment even during simultaneously controlled muscle contractions.[12]

Because raw EMG signals are quasi-random in nature, they cannot be directly compared. Thus, a principle goal of this study was to use quantitative methodology with the addition of control parameters and measures, to characterize the reliability of a primarily qualitative clinical evaluation. One such measure was quantification of the LEMG signal via calculation of the root mean square (RMS). Root mean square is considered to be the current "gold standard" for quantitative electromyographic analysis[11,12] and allows for rapid quantitative comparisons among groups of signals. Root mean square was chosen as a measurement metric because it provides an indication of mean muscle activity and signal power and is the analog to voltage output. Because RMS is also considered a data smoothing technique, it is not well suited for visualization of waveform transients and morphology characteristics such as polyphasic or nascent potentials; however, it is useful to quantify and compare LEMG across samples in terms of signal voltage and power. Because EMG is a time-varying signal containing positive and negative values, RMS is an ideal quantitative measure that can be easily calculated post hoc or in real time with many commercially available data acquisition software programs.

Determining LEMG data reliability within the context of an in-office clinical environment is important to make careful and useful clinical interpretations and to potentially improve clinical protocols. To our knowledge, in-office clinical LEMG reliability has not been systematically investigated in a cohort of vocally healthy adults. As such, the purpose of this study was to prospectively investigate LEMG signal reliability recorded from the thyroarytenoid muscle over multiple testing sessions using a common in-office clinical routine. We modeled our basic methodology

after the University of Iowa head and neck protocol for LEMG diagnostics.[21] Laryngeal electromyography was performed on the right thyroarytenoid muscle of 7 participants with normal vocal function on 3 separate occasions with the application of an additional behavioral control for vocal intensity. We hypothesized that time of testing and vocal intensity would significantly alter quantitative measures of thyroarytenoid EMG signals.

Materials and Methods

Seven participants, ages 18 to 40 years (3 male and 4 female), were recruited and volunteered for this study. Ten participants were initially recruited based on power analysis for the repeated measures design. Two participants did not complete the entire study protocol, and 1 participant had unusable data. All were nonsmokers, English speaking, and free of laryngeal pathology and movement disorders as judged by stroboscopic examination of the larynx. Exclusion criteria were professional voice training, diagnosed bleeding disorder, prior head/neck/spinal surgery, and/or intubation within the past 5 years. All procedures were approved by the University of Kentucky Institutional Review Board, and informed consent was obtained from all participants. All assessments were performed at the University of Kentucky Laryngeal & Speech Dynamics Laboratory.

Prior to data collection, all participants received a laryngeal exam via videostroboscopy to ensure normal vocal function and structure (Kay Elemetrics Rhino-Laryngeal Stroboscope [Model RLS 9100 B], Kay Elemetrics 70 degree rigid endoscope [Model SN 1541]; PENTAX Medical, Montvale, New Jersey, USA). This examination was performed by a certified/licensed speech-language pathologist and reviewed by a board certified otolaryngologist.

After imaging was completed, electromyographic biopotentials (μV) were acquired from the right TA muscle via a 25-mm, 30-gauge concentric bipolar needle electrode (XLTEK 101468; Natus Medical, San Carlos, California, USA). The raw EMG signal was routed to a biopotential amplifier (Grass Model 15A54; Natus Neurology, Warwick, Rhode Island, USA) and serially coupled to a 16-bit analog-to-digital converter sampling at 10 kHz (PowerLab 1630; ADInstruments, Inc, Colorado Springs, Colorado, USA). Prior to digitization, analog signals were band-pass filtered (30 Hz–3 kHz @ –3 dB), with an in-line notch filter applied to reduce 60 Hz line contamination. The LEMG analog output was paralleled to a stereo amplifier and played over loud speakers to provide the otolaryngologist with auditory feedback of muscle activity during needle insertion. Audio (volts) and sound intensity levels (dB SPL) from the participant's vocalizations were recorded by a lapel-style microphone (Sony ECM44B; Sony Corporation, New York, New York, USA) and a commercially available sound level meter (REED ST-8850; REED Instruments, Sainte-Anne-De-Bellvue, Quebec, Canada). Both signals were also digitized by the A/D system (audio sampling rate = 2 kHz; sound pressure level = 1 kHz). All signals were recorded in calibrated units using a proprietary 2-point interpolation method found in our digitization software package (LabChart 7; ADInstruments, Inc). Post-acquisition LEMG signal processing was completed using custom-coded LabChart routines.

Laryngeal electromyography testing was performed in a custom-built Faraday booth to reduce electromagnetic field effects. Participants were seated in an examination chair, reclined to approximately 60 degrees, with their heads comfortably supported by a neck pillow. No sedation or anesthetic was used during the needle insertion and recording procedure. A ground electrode was placed on the participant's neck, below the mastoid process. Thyroarytenoid needle insertion by an otolaryngologist with 10 years of experience performing clinical LEMG procedures was accomplished using a para-medial approach percutaneously with the electrode directed in a superolateral direction through the cricothyroid ligament with the muscle entered submucosally. Needle electrode placement was confirmed using the following behavioral tasks: normal rest breathing, phonation on sustained /i/, sniff, and sustained phonation of falsetto /i/. Upon completion of the study, the needle electrode was removed by the physician and participants were monitored for 15 minutes after the study in case of complications.

Digitized raw LEMG signals were full-wave rectified and RMS signal amplitude values were calculated online with the LabChart software package. Root mean square was calculated as the square root of the mean of a series of squared LEMG amplitude values. Root mean square amplitudes were used in all analyses to determine if significant variance in the LEMG signal existed as a function of vocal intensity and time of data acquisition.

Experimental Protocol

Electromyography of the right TA was performed on each participant on 3 different occasions with a minimum of 1 month between adjacent procedures (mean duration = 2.5 months) to allow for tissue healing. Each session took place at approximately the same time of day and lasted no more than 30 minutes (including videostroboscopy). After needle electrode placement was confirmed within the TA, LEMG signals were recorded under 2 task conditions. First, a confirmation condition was completed by recording baseline LEMG signals while the participant was instructed to relax and breathe normally, sustain the vowel /i/ at a comfortable modal pitch, gently sniff through the nose, and sustain a falsetto /i/. Second, a feedback condition was completed with the participant producing a sustained /i/ at 2 different

intensity targets (65 and 75 dB [± 3 dB]). These intensity targets were chosen to represent a typical healthy vocalization intensity range present in everyday speech. In this condition, participants received visual feedback of their intensity level by monitoring a dB sound level meter placed 16 inches from the individual's mouth. The participant was asked to hold the intensity constant for a minimum of 3 seconds. Trials less than 3 seconds were not accepted for analysis. A total of 10 trials of each task condition were recorded. The initial and final 2 trials were discarded, leaving 6 trials for post-hoc data analysis.

Habitual fundamental frequency (F_o) for the vocalization tasks was initiated by the participant without prompting from the investigators. The F_o chosen by the participant was recorded and played back via an auto-tuner before data collection so that the participant could remain in an acceptable modal pitch range. Before all data collection procedures and to ensure performance consistency, each participant was trained and given time to practice all tasks before data collection.

Data Analysis

A 2-factor repeated measures analysis of variance (RM-ANOVA) was used to compare the effect of the 2 independent variables among participants. The independent variables included (1) LEMG data recording sessions (3 sessions) and (2) the task-related vocal intensity levels (2 dB levels). To evaluate the reliability of LEMG signals across testing times, intraclass correlation coefficients (ICCs) using a 2-way mixed random effects model were calculated. For between-session measurements, standard error of the measurement (SEM) was calculated to determine the minimum detectable change (MDC) in microvolts with 70% and 95% confidence boundaries. Intraclass correlations were also calculated to measure intra-session reliability. All statistical calculations were performed using SPSS version 18 (IBM, Armonk, New York, USA).

Data Selection and Calculations

A 1-second window from the mid-portion of each 3-second task recording window was selected for detailed analysis. Root mean square amplitude values were automatically calculated for each of the 1-second windows using a custom sub-routine in LabChart (ADInstruments, Inc). To be included for further processing and analysis, digitized samples had to fall within +/– 1 dB of the desired intensity levels (65 or 75 dB). In certain cases, the RMS amplitude was not stable for 1 second at +/– 1 dB. In these cases, a shorter window was averaged to eliminate portions of the signal with poor signal quality. The inclusion of smaller analysis windows for these cases was preferable to averaging poor signal quality. Smaller sampling windows do not significantly affect analyses, as comparisons of EMG sample windows using RMS have been shown to produce moderate to excellent ICC reliability data from 1000 ms down to 100 ms window sizes.[13]

Table 1. Means for Intensity Level at 65 dB.

Variable	Session	Mean	SD	Min	Max
Root mean square (μV)	1	80.2	34.8	44.7	151.5
	2	91.8	28.4	45.4	117.6
	3	73.08	19.64	46.31	98.48

Table 2. Means for Intensity Level at 75 dB.

Variable	Session	Mean	SD	Min	Max
Root mean square (μV)	1	84.0	37.4	46.0	155.5
	2	106.5	29.4	69.2	136.9
	3	85.04	16.82	55.42	101.56

Table 3. Repeated Measures Analysis of Variance for Baseline Laryngeal Electromyography Across Time.

Source	*df*	*F* Value	*P* Value
Participant	6	0.85	.555
Time	2	0.02	.978

Table 4. Repeated Measures Analysis of Variance for Root Mean Square Values.

Source	*df*	*F* Value	*P* Value
Participant	6	1.91	.160
Time	2	1.38	.289
Intensity	1	5.71	.054
Time*Intensity	2	2.23	.150

Results

Statistical Analysis

Group data, including means and standard deviations for the dependent variable (RMS) for intensities at 65 dB and 75 dB, are presented in Tables 1 and 2, respectively. Results of the 2-factor RM-ANOVA during baseline and the task conditions are presented in Tables 3 and 4. Repeated measures ANOVA for the baseline data did not reveal a significant main or interaction effect for testing time (session) versus RMS, indicating stable LEMG baseline activity across recording sessions. Results of the RM-ANOVA for the feedback condition were not significant for the main effect of testing time ($df = 2$, $F = 1.38$, $P = .289$). However, the main effect for intensity level closely approached our a

Table 5. Intraclass Correlation Coefficients (ICCs) Between Sessions, Standard Error of Measurement (SEM), and Minimum Detectable Change (MDC) in μV Necessary to Detect True Change Between Measurements.

ICC Sessions 1–3		SEM Sessions 1–3	MDC in μV @ 70% Confidence	MDC in μV @ 95% Confidence
65 dB	0.56	18.57	26.26	51.47
75 dB	0.40	23.0	32.5	63.7

Table 6. Intraclass Correlation Coefficients (ICCs) for Within-Session Data.

Within-Session ICC	Session 1	Session 2	Session 3
65 dB	0.95	0.93	0.84
75 dB	0.88	0.84	0.98

priori significance level of α = .05 (df = 1, F = 5.71, P = .054). This indicated that vocal intensity may play a role in LEMG signal reliability. No interaction effects between intensity and time factors were found.

To further evaluate the reliability of LEMG signals across the 3 testing days, ICCs using a 2-factor mixed random effect model were calculated. Average measures from 6 sample trials per participant were compared across the 3 testing sessions to evaluate reliability. The ICCs across Sessions 1 through 3 were 0.56 for the 65 dB condition and 0.40 for the 75 dB condition. These between-session ICCs were low, suggesting an interpretation of poor to, at best, fair reliability across testing sessions. Although not standardized, typical calculated values for ICC interpretation are as follows: less than 0.40 = poor reliability, 0.40 to 0.75 = fair to good reliability, and greater than 0.75 = excellent reliability.[19] In addition, the SEM was calculated. The SEM was then used to determine the MDC in microvolts necessary to demonstrate a true difference if the muscle was tested on multiple days with 95% confidence boundaries. Results indicated that a change of 51μV would be necessary to determine a true difference in LEMG activity between testing sessions. Numerical results of ICC, SEM, and MDC are presented in Table 5. Intraclass correlations for within-session data revealed strong reliability among participants ranging from 0.84 to 0.95 and from 0.88 to 0.98 for the 65 dB and 75 dB conditions, respectively. Comparisons of within-session data are located in Table 6.

Discussion

The use of clinical in-office LEMG has been incorporated into the diagnostic routine for the evaluation and treatment of voice disorders in many practices across the country.[8] Although some evidence supports LEMG use in the diagnosis and prognosis of certain neuromuscular disorders, the general reliability of the LEMG signal in normal participants has not been carefully evaluated. This study measured the reliability of the LEMG signal in normal, vocally healthy participants over time with the central aim of determining if significant LEMG signal variance occurred as a function of multiple testing sessions. Our second aim was to determine if vocalization intensity affected the LEMG signal. In limb studies, control over the degree of muscle contraction is necessary to achieve results that are comparable within and across participants. Both maximal and submaximal contractions have been shown to demonstrate strong reliability in limb muscle.[20] We used vocal intensity as a method to control laryngeal muscle contraction levels among participants. Our results indicated that between-session LEMG reliability was poor to fair and that control of vocal intensity may be an important performance variable to help improve the reliability of these measurements.

This study mirrored the University of Iowa head and neck protocol for LEMG diagnostics.[21] In addition to this basic protocol, we used vocal intensity control and a Faraday booth to reduce ambient electrical noise to improve the fidelity of the data and provide the optimal set of circumstances under which to perform our LEMG clinical evaluation. The intent of this study was not to quantify LEMG precisely but rather to use quantitative means to measure LEMG in an ideal environment to test the hypothesis that clinical LEMG data are variable across testing sessions even with added control parameters in place.

Repeated measures analysis of variance indicated a nonsignificant effect for time of testing, suggesting that LEMG signals for pooled data did not vary significantly across testing sessions. Intraclass correlation coefficient analysis for within-session reliability was considered excellent for both intensity conditions ranging from 0.84 to 0.95 and from 0.88 to 0.89 for the 65 dB and 75 dB conditions, respectively. However, the between-session ICC revealed poor to fair reliability for both intensity conditions. It should be noted that the most qualitatively consistent data from our study were collected when the participants vocalized at 65 dB, indicating a less reliable measure at greater loudness levels. Data from Sessions 1 and 2 at 65 dB represented the strongest reliability association with an R^2 value of 0.048. This indicated poor reliability even across the most consistent recording sessions (see Figures 1 and 2).

Reliability debates concerning the clinical usefulness of LEMG for diagnostic and prognostic applications have been raised.[1] According to a recent evidence-based review and clinical recommendations, LEMG data have been considered questionable for clinical uses such as diagnosing paresis/paralysis from joint fixation, for accuracy diagnosing diseases of the neuromuscular junction, and for providing accurate diagnostic information of neuropathic and myopathic disorders.[1,8] To address these questions, it has

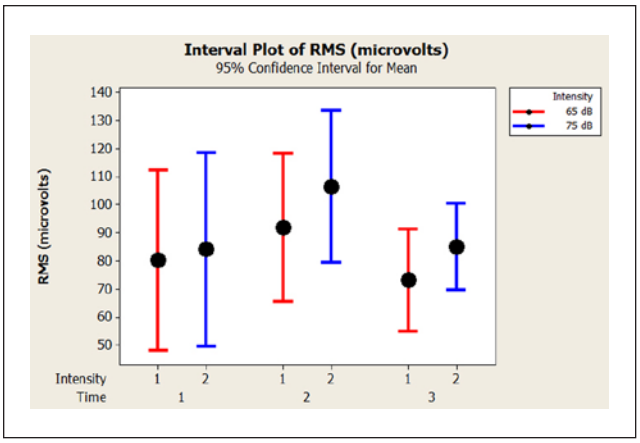

Figure 1. Interval plot with the mean of each data set presented with 95% confidence bars. The x-axis is scaled with regard to intensity and testing session. The y-axis for root mean square (RMS) is scaled in microvolts.

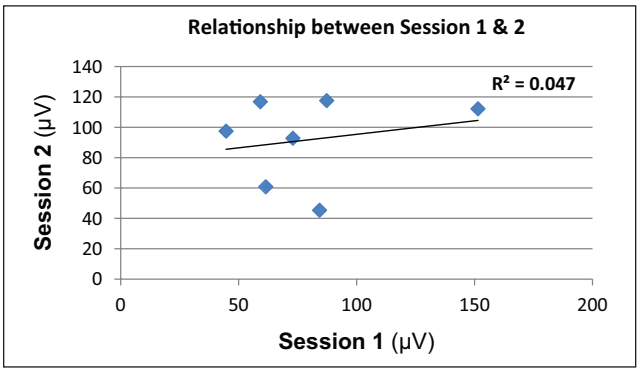

Figure 2. Scatter plot for voice task at the 65 dB target between sessions 1 and 2. These sessions represent the most consistent data from this study. Data points are mean root mean square (RMS) values in μV for each participant. The R^2 value does not indicate a strong association between the data for the 2 sessions. Approximately 5% of the data from Session 1 can be explained by Session 2.

been suggested that additional evidence-based research concerning LEMG methodology and validity be conducted.[8] Data from this study suggest that variables such as data collection time (multiple sessions) and possibly vocal intensity may play a role in the outcomes of LEMG assessment, suggesting methodological limitations of LEMG in terms of its clinical accuracy.

Our results indicated that to be 95% confident that a true detectable change could be observed between testing sessions, a change of 51 μV RMS was necessary with intensity level held constant. It is likely, then, that uncontrolled vocal intensity during LEMG procedures may operate as a confounding variable. Careful regulation of vocal intensity during LEMG may be necessary if the clinical utility of LEMG is to be determined with any degree of accuracy. In fact, intensity would be expected to contribute to changes in the RMS value of the LEMG signal. A near linear relationship between muscle force and EMG activity has been found in classic EMG studies.[22-24] Although EMG does not measure muscle force directly, vocal intensity can be viewed as a global indicator of performance effort and muscle loading on the vocal apparatus. It was not surprising, then, that controlling for intensity revealed changes in our calculated RMS values.

It has been previously demonstrated that both intensity and vocal frequency contribute to variability in quantitative LEMG output with frequency being the greater factor in TA recruitment variability, suggesting the need for control of both parameters for improved clinical assessment.[25,26] In this study, participants were generally able to maintain and regulate the intensity of their vocalization constant at 65 dB across all trials. Although participants were trained to reach the 75 dB target, many could not produce this intensity level consistently for 1 second with the LEMG needle in place. In the 75 dB condition, intensity levels actually ranged from 66.70 dB to 75.22 dB with a mean value of 70.42 dB. Thus, it can be inferred that not only does intensity play a role in LEMG signal stability but relatively small changes in intensity level (approximately 5 dB) can strongly affect RMS values, further arguing for the need and importance of regulating vocal intensity during LEMG diagnostics.

Limitations

The small sample size of 7 participants in this repeated measures study limits the ability to generalize our results to a larger clinical population. Changes in vocal intensity were limited to a 10 dB interval. Larger intensity intervals and additional participant data may better demonstrate differences in mean RMS values across testing conditions. The standard deviations of the RMS values in this study were large. This is an inherent problem with attempting to quantify LEMG because it is difficult to determine which variable(s), such as ambient noise, movement artifact, interpersonal differences in phonation, and so on, may be causing deviations in the signal.[11,27] Needle electrodes, as used in this study, have been shown to demonstrate greater artifact at greater intensities.[28] Movement/vibration artifact cannot be alleviated but is a concern because of the unsteadiness of the needle electrode and the vibration of the vocal fold mucosa. A solution to this problem may be to consider the use of hooked wire electrodes in clinical LEMG studies to ameliorate these concerns.

Clinical Relevance and Future Directions

The results of this study demonstrate that even during controlled laboratory conditions, the LEMG signal appears

significantly variable across testing sessions. It is reasonable to expect greater variability in an office setting without these control parameters. Laryngeal electromyography holds much potential to be a useful clinical tool available for diagnosing movement disorders of the larynx. To obtain the maximum benefit from clinical LEMG, a universal standardized protocol that is feasible within a typical in-office setting should be developed. Further prospective research studies should consider the evaluation of (1) hooked wire electrode use for contralateral comparisons and (2) vocal frequency and intensity controls to regulate muscle recruitment to maximize the interpretability of LEMG measures. Quantitative LEMG techniques that may be used for clinical application are now feasible and relatively inexpensive and should be explored. Laryngeal electromyography follow-up diagnostics would also be useful for data comparisons. A recent article by Sataloff et al stated that follow-up LEMG is not performed in up to 90% of cases if visual inspection of the larynx demonstrates improved vocal fold mobility.[29] Follow-up testing could provide useful reliability data for LEMG as performed in the clinic.

Data from this study offer insight into the importance of using loading controls (control of intensity and frequency) in order to obtain the most accurate data from clinical LEMG. Methodologically, the use of hooked wire LEMG may be a good alternative to needle electrodes for several reasons, including freeing the clinician to direct the patient to control pitch and loudness levels, allowing for simultaneous measures, and reducing the possibility of artifact from needle electrode movement. With today's technology, control of vocal frequency and intensity can be easily accomplished through visual feedback to the patient using an inexpensive headset microphone connected to a laptop computer or other mobile device running commercially available sound intensity applications. In addition, quantitative signal processing tools are becoming more accessible and less expensive, allowing for the real-time use of quantitative techniques such as RMS to improve the quality of in-office assessments and the use of rise-time functions to confirm optimal electrode placement.[30]

In closing, LEMG is a clinical assessment tool that has not yet reached its full potential. The means to make LEMG a more quantitative and reliable assessment method are available and ready for usage to improve the clinical reliability and usefulness of this potentially important diagnostic method.

Acknowledgments

The authors would like to thank the participants for undergoing multiple LEMG procedures for this study, and the University of Kentucky College of Health Sciences Office of Research for supporting this project through pilot funding.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The University of Kentucky College of Health Sciences Office of Research supported this project through pilot funding.

References

1. Sataloff RT, Mandel S, Mann EA, Ludlow CL. Practice parameter: laryngeal electromyography (an evidence-based review). *J Voice*. 2004;18(2):261-274.
2. Ludlow CL, Yeh J, Cohen LG, Van Pelt F, Rhew K, Hallett M. Limitations of electromyography and magnetic stimulation for assessing laryngeal muscle control. *Ann Otol Rhinol Laryngol*. 1994;103(1):16-27.
3. Woo P. Laryngeal electromyography is a cost-effective clinically useful tool in the evaluation of vocal fold function. *Arch Otolaryngol Head Neck Surg*. 1998;124(4):472-475.
4. Heman-Ackah YD, Mandel S, Manon-Espaillat R, Abaza MM, Sataloff RT. Laryngeal electromyography. *Otolaryngol Clin North Am*. 2007;40(5):1003-1023.
5. Sataloff R, Mandel S, Heman-Ackah Y, Manon-Espaillat R, Abaza M. *Laryngeal Electromyography*. 2nd ed. San Diego, CA: Plural Publishing; 2006.
6. Basmajian J. *Muscles Alive: Their Functions Revealed by Electromyography*. 3rd ed. Baltimore, MD: Williams and Wilkins Company; 1974.
7. Hillel AD. The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *Laryngoscope*. 2001;111(4 Pt 2 Suppl 97):1-47.
8. Blitzer A, Crumley RL, Dailey SH, et al. Recommendations of the Neurolaryngology Study Group on laryngeal electromyography. *Otolaryngol Head Neck Surg*. 2009;140(6): 782-793.e786.
9. Koufman JA, Postma GN, Whang CS, et al. Diagnostic laryngeal electromyography: The Wake Forest experience 1995-1999. *Otolaryngol Head Neck Surg*. 2001;124(6): 603-606.
10. Sittel C, Stennert E, Thumfart WF, Dapunt U, Eckel HE. Prognostic value of laryngeal electromyography in vocal fold paralysis. *Arch Otolaryngol Head Neck Surg*. 2001;127(2):155-160.
11. Basmajian J, De Luca C. *Muscles Alive: Their Functions Revealed by Electromyography*. 5th ed. Baltimore, MD: Williams & Wilkins; 1985.
12. De Luca CJ. *Surface Electromyography: Detection and Recording*. Natick, MA: Delsys Inc.; 2002. http://www.delsys.com/Attachments_pdf/WP_SEMGintro.pdf. Accessed November 2008.
13. Bamman MM, Ingram SG, Caruso JF, Greenisen MC. Evaluation of surface electromyography during maximal

voluntary contraction. *Journal of Strength & Conditioning Research*. 1997;11(2):68-72.
14. Yang JF, Winter DA. Electromyography reliability in maximal and submaximal isometric contractions. *Arch Phys Med Rehabil*. 1983;64(9):417-420.
15. Huber JE, Stathopoulos ET. Respiratory and laryngeal responses to an oral air pressure bleed during speech. *J Speech Lang Hear Res*. 2003;46(5):1207-1220.
16. Donzelli J, Brady S. The effects of breath-holding on vocal fold adduction: implications for safe swallowing. *Arch Otolaryngol Head Neck Surg*. 2004;130(2):208-210.
17. Smith LJ, Rosen CA, Niyonkuru C, Munin MC. Quantitative electromyography improves prediction in vocal fold paralysis. *Laryngoscope*. 2012;122(4):854-859.
18. Rubin AD, Sataloff RT. Vocal fold paresis and paralysis. *Otolaryngol Clin North Am*. 2007;40(5):1109-1131.
19. Rosner B. *Fundamentals of Biostatistics*. Belmont, CA: Duxbury Press; 2005.
20. Burden A. How should we normalize electromyograms obtained from healthy participants? What we have learned from over 25 years of research. *J Electromyogr Kinesiol*. 2010;20(6):1023-1035.
21. University of Iowa Health Care. Iowa head and neck protocols: laryngeal electromyography. http://www.healthcare.uiowa.edu/otolaryngology/protocols//Parts/Part1/Part1C/P1C4.htm. Accessed July 14, 2008.
22. Abbott BC, Bigland B. The effects of force and speed changes on the rate of oxygen consumption during negative work. *J Physiol*. 1953;120(3):319-325.
23. Bigland B, Lippold OCJ. The relation between force, velocity and integrated electrical activity in human muscles. *J Physiol*. 1954;123(1):214-224.
24. Bigland B, Lippold OCJ. Motor unit activity in the voluntary contraction of human muscle. *J Physiol*. 1954;125(2):322-335.
25. Lindestad PÅ, Fritzell B, Persson A. Quantitative analysis of laryngeal EMG in normal subjects. *Acta Otolaryngol*. 1991;111(6):1146-1152.
26. Statham MM, Rosen CA, Nandedkar SD, Munin MC. Quantitative laryngeal electromyography: turns and amplitude analysis. *Laryngoscope*. 2010;120(10):2036-2041.
27. Lindestad PÅ, Persson A. Quantitative analysis of EMG interference pattern in patients with laryngeal paresis. *Acta Otolaryngol*. 1994;114(1):91-97.
28. Hirano M, Vennard W, Ohala J. Regulation of register, pitch and intensity of voice. An electromyographic investigation of intrinsic laryngeal muscles. *Folia Phoniatr*. 1970;22(1):1-20.
29. Sataloff RT, Praneetvatakul P, Heuer RJ, et al. Laryngeal electromyography: clinical application. *J Voice*. 2010;24(2):228-234.
30. Statham MM, Rosen CA, Nandedkar SD, Munin MC. Quantitative laryngeal electromyography: turns and amplitude analysis. *Laryngoscope*. 2010;120(10):2036-2041.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(6):1075-1081.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–Head and Neck Surgery
2015, Vol. 152(6) 1075–1081
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599815577104
http://otojournal.org
SAGE

Office-Based Photoangiolytic Laser Treatment of Reinke's Edema: Safety and Voice Outcomes

Ian J. Koszewski, MD[1], Matthew R. Hoffman, PhD[1], W. Greg Young, MD[1], Ying-Ta Lai, MD[1], and Seth H. Dailey, MD[1]

No sponsorships or competing interests have been disclosed for this article.

Abstract

Objective. To evaluate the safety, tolerability, and voice outcomes of office-based photoangiolytic laser treatment of Reinke's edema.

Study Design. Case series with chart review.

Setting. Academic medical center.

Subjects and Methods. We performed a retrospective analysis of patients undergoing office-based laser treatment of endoscopy-proven Reinke's edema. Safety and tolerability were evaluated by reviewing complications. Voice outcomes were analyzed by comparing pre- and postprocedural acoustic, aerodynamic, and Voice Handicap Index measurements. Complete data sets were not available for all subjects; sample size for each parameter is reported with the corresponding result.

Results. Nineteen patients met inclusion criteria. There were no minor or major complications. Five procedures were truncated due to patient intolerance. Phonatory frequency range increased (n = 12, $P = .003$), while percent jitter decreased (n = 12, $P = .004$). Phonation threshold pressure decreased after treatment (n = 4, $P = .049$). Voice Handicap Index also decreased (n = 14, $P < .001$).

Conclusion. This study represents the largest series of patients undergoing office-based photoangiolytic laser treatment specifically for Reinke's edema. Our data suggest that this is a safe and effective modality to treat dysphonia associated with Reinke's edema, although patient intolerance of the procedure may represent a barrier.

Keywords

Reinke's edema, photoangiolytic laser, office-based treatment, voice

Received October 24, 2014; revised January 14, 2015; accepted February 20, 2015.

Reinke's edema is a benign disease of the true vocal folds, characterized by edema and vascular congestion within Reinke's space, leading to diffuse polypoid degeneration. Patients are generally middle-aged women, and they have voice complaints of abnormally low speaking pitch and a rough voice quality.[1] Multiple risk factors have been identified, including tobacco use, vocal abuse, and laryngopharyngeal reflux.[2-5] Critical to preventing progression of the disease and managing the dysphonia are nonsurgical strategies, such as smoking cessation, voice therapy, and treatment of underlying laryngopharyngeal reflux.[1,6] When a nonsurgical approach is inadequate, surgery is employed. Traditional surgical options have focused on mucosal microflap elevation with removal of tissue,[7,8] microdebridement,[9] carbon dioxide laser ablation,[10,11] and cold knife cytoreduction.[12]

Recently, photoangiolytic laser treatment has been proposed as an alternative treatment.[13-16] Unlike traditional methods that involve the physical removal of tissue, the potassium titanyl phosphate (KTP) laser and pulsed dye laser (PDL) target oxyhemoglobin and are thought to address the vascular congestion characteristic of Reinke's edema. Importantly, both the KTP and the PDL have flexible fibers that can be passed through a flexible endoscope for office-based interventions, thus avoiding risks associated with general anesthesia.[17] Performing procedures in the office rather than the operating room has several notable advantages, including decreased cost[18] and avoidance of the potential complications of microlaryngoscopy, such as dental injury and dysgeusia.[19]

Office-based use of lasers for the treatment of Reinke's edema has gained popularity with the advent of improved instrumentation and evidence that the procedures are safe and cost-effective. Koufman et al and Sheu et al reported large series demonstrating support for the use of photoangiolytic lasers in the treatment of an array of laryngeal

[1]Department of Surgery, Division of Otolaryngology–Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA

This article was presented at the 2014 AAO-HNSF Annual Meeting & OTO EXPO; September 21-24, 2014; Orlando, Florida.

Corresponding Author:
Seth H. Dailey, MD, Division of Otolaryngology–Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, 600 Highland Ave, Clinical Science Center-H4, Madison, WI 53792-3284, USA.
Email: dailey@surgery.wisc.edu

lesions, with the latter demonstrating preserved or improved mucosal wave and glottic closure after treatment.[14,15] Pitman et al evaluated the effect of office-based KTP laser treatment in 7 patients with Reinke's edema and found significant improvement in subjective measures for voice quality as well as a trend toward improvement in objective measures.[13]

Although these studies provided important initial support for the use of office-based photoangiolytic laser therapy in the management of Reinke's edema, they are limited by a modest sample size and measurement of few objective voice parameters. As such, larger studies addressing the safety, tolerability, and voice outcomes for this approach are necessary. We report on 19 patients who underwent office-based photoangiolytic laser treatment of Reinke's edema. We hypothesized that no complications would occur, patients would tolerate the procedure, Voice Handicap Index would decrease, and objective voice parameters would move toward the normal ranges. Specifically, we hypothesized that posttreatment assessment would show increased fundamental frequency range, decreased jitter, decreased phonation threshold pressure, and increased maximum phonation time (MPT).

Materials and Methods

Subjects

Approval for this study was obtained from the University of Wisconsin Health Sciences Institutional Review Board. The study was designed as a retrospective case series of patients treated at the University of Wisconsin–Madison. Patient data were obtained from the University of Wisconsin–Madison Voice and Swallow Outcomes Database. Appropriate patients were identified by the University of Wisconsin–Madison Voice and Swallow Outcomes database manager by cross-matching diagnosis with the existence of a procedure, and data extraction was performed by a separate researcher. All patients presented to the University of Wisconsin Hospital and Clinics between January 2007 and November 2013 and underwent voice analysis by a speech-language pathologist, as well as evaluation and treatment by an otolaryngologist. All patients with documented Reinke's edema who underwent at least 1 office-based treatment were considered for inclusion. Patients were excluded if they were <18 years old, had a history of laryngeal malignancy, had a history of a neurolaryngologic disorder (eg, cerebrovascular accident, amyotrophic lateral sclerosis, vocal tremor, or recurrent laryngeal or superior laryngeal nerve injury), had a history of airway stenosis, or were unable to provide consent.

Treatment

All procedures were performed in a clinic setting without sedation. Patients are positioned sitting upright. Local nasal anesthesia is obtained by placing 4% lidocaine and 0.05% oxymetazoline-soaked sponges in the nasal cavities. A flexible endoscope is passed through the nasal cavity for indirect visualization of the endolarynx. Laryngeal anesthesia is obtained by instilling 3 aliquots of 0.5 mL of 4% lidocaine through the working channel of the endoscope during sustained phonation (the ''laryngeal gargle''). The laser fiber is passed through the working channel and advanced until the tip of the fiber is visualized. Laser energy is then applied to the involved tissues. Of note, tissue ablation is not desired during these procedures; rather, enough energy is applied to blanch the superficial tissues (**Figure 1**). Importantly, no immediate reduction is tissue size is desired. Twelve procedures were performed using the KTP laser, and 13 were performed using the PDL.

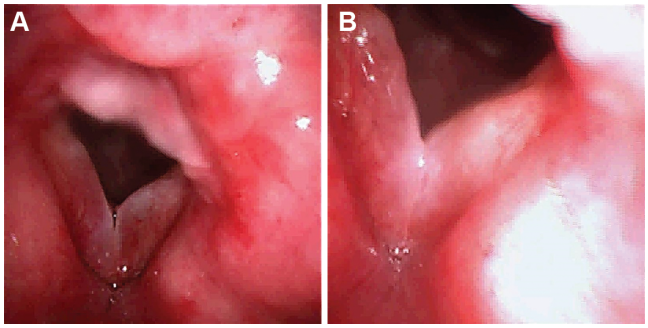

Figure 1. Endoscopic view of glottis immediately before (A) and immediately after (B) treatment. Note the superficial blanching without reduction in tissue mass immediately following application of laser energy.

Experimental Data

Patient demographics, chief complaint, and social history were collected. Endoscopic findings and physician impression were recorded to ensure diagnosis, as well as to document unilateral versus bilateral involvement. Treatment of laryngopharyngeal reflux disease was also documented. Reported complications and patient tolerance data were also collected by extracting physician documentation from the University of Wisconsin–Madison Voice and Swallow Outcomes database. This included whether a procedure was truncated, as well as the number of procedures performed for each patient and whether operative interventions were ultimately required. Minor complications included nasal or pharyngeal pain, minor nosebleed, and vasovagal events. Major complications included need for emergent airway intervention, hospitalization or presentation to the emergency department after treatment, airway bleeding, airway stenosis, reported myocardial infarction or cerebrovascular accident, and extralaryngeal tissue injury. Patients were requested to report complications at follow-up visits; however, our data set did not allow for standardized follow-up of all possible complications.

Objective voice measures were obtained as part of a standard comprehensive voice assessment. Aerodynamic parameters included MPT, mean airflow rate, laryngeal resistance, phonation threshold pressure, and mean peak air pressure. Aerodynamic parameters were measured using the Phonatory Aerodynamic System (model 6600, KayPENTAX, Montvale, New Jersey); the device was calibrated prior to each use according to manufacturer specifications. For MPT, patients produced a sustained /a/ at a modal pitch for as long as possible; this was repeated 3 times, and the longest trial taken as the

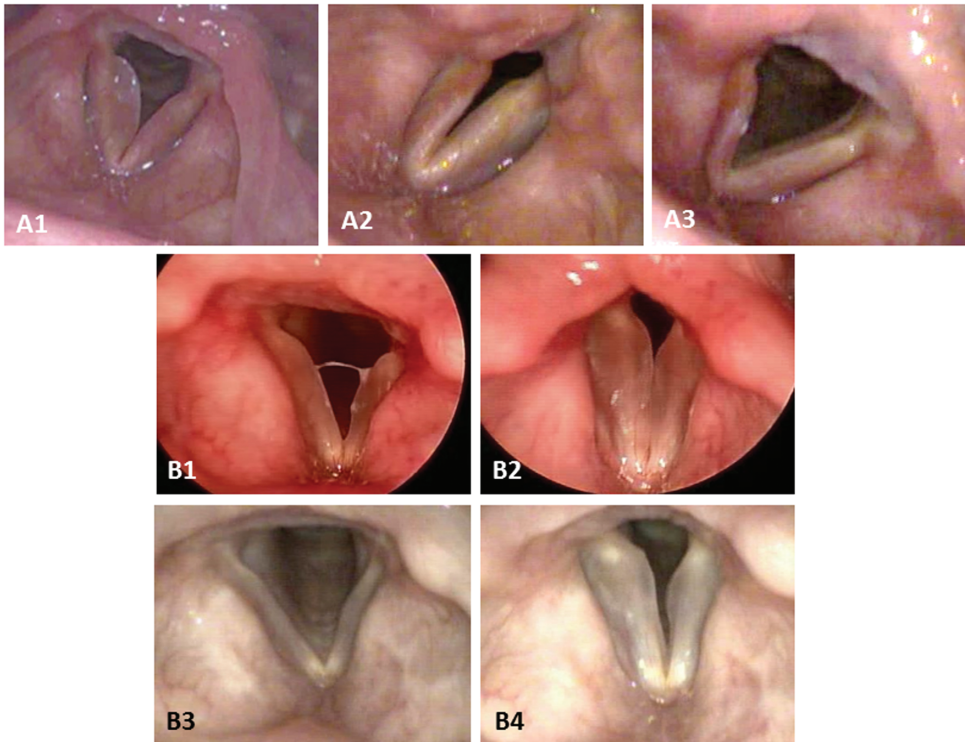

Figure 2. Sample pre- and posttreatment images from 2 patients. A, subject 1: A1, pretreatment, normal inspiration; A2, posttreatment, normal inspiration; A3, posttreatment, vocal fold abduction. B, subject 2: B1, pretreatment, vocal fold abduction; B2, pretreatment, normal inspiration; B3, posttreatment, vocal fold abduction; B4, posttreatment, normal inspiration.

MPT. For mean airflow rate, laryngeal resistance, and peak air pressure, patients produced 3 /pa/ syllable trains at comfortable pitch and loudness; the first and last /pa/ token within each train were removed and the average values computed. For phonation threshold pressure, patients produced 3 /pi/ syllable trains, each beginning with a soft whisper, followed by incremental increases in subglottal pressure until comfortable voicing was achieved; the lowest subglottal pressure at which phonation occurred was recorded as the phonation threshold pressure. Acoustic parameters included minimum and maximum fundamental frequency, phonatory frequency range, and percent jitter. Acoustic data were recorded using the Computerized Speech Lab (model 4150B, KayPENTAX) and Multi-Dimensional Voice Program (model 5105, KayPENTAX). For fundamental frequency values, patients performed ascending and descending glides on the vowel /a/ and were instructed to achieve the lowest and highest frequencies possible, inclusive of falsetto phonation; this was repeated 3 times and the extrema recorded. Phonatory frequency range was calculated as the difference between maximum and minimum fundamental frequency. For percent jitter, patients produced a stable /a/. Dysphonia severity index was calculated as described by Wuyts et al.[20] Subjective voice changes were quantified using pre- and postprocedural scores from the Voice Handicap Index.[21] This instrument measures the impact of one's voice in 3 separate subcategories: functional, physical, and emotional. Finally, total energy delivered was collected.

Although a standardized clinical protocol is followed for collection of voice measures at clinical visits, occasionally not all voice measures are available in the database. Accordingly, a complete data set including all voice parameters was not available for every subject. Analyses were performed using the data that were available, and the number of subjects included in each analysis has been specified. A subject was not included in a given analysis of efficacy if he or she did not have a measurement of that parameter before and after the procedure.

Statistical Analysis

Evaluation of treatment efficacy was performed using paired t tests. If data did not meet assumptions for parametric testing, a Wilcoxon-Mann-Whitney matched pairs signed-rank test was performed. All tests were 2-tailed with a significance level of $\alpha = 0.05$. As complete data sets were not available for all subjects, sample size for each parameter is reported with the corresponding result.

Results

Subject Characteristics

Nineteen patients underwent 25 in-office endoscopic laser treatments of Reinke's edema between January 2007 and November 2013. All but 1 patient was a woman, and all were smokers at the time of presentation. Average age at presentation was 53.9 $\pm$ 7.7 years (range, 43-67 years). All but 1 patient had bilateral involvement, with 1 demonstrating polypoid change affecting only 1 vocal fold. Sample pre- and posttreatment images are provided in **Figure 2**.

Table 1. Voice Outcome Data.[a]

Parameter	Pretreatment	Posttreatment	No.	*P* Value
Dysphonia severity index	−7.0 ± 3.3	−3.0 ± 2.6	12	.003
Acoustic				
Maximum F_0	290 ± 53	482 ± 272	12	<.001
Minimum F_0	110 ± 35	119 ± 95	12	.147
Frequency range	180 ± 67	363 ± 295	12	.003
Percent jitter	4.05 ± 2.83	1.66 ± 1.10	12	.004
Aerodynamic				
Maximum phonation time	8.77 ± 4.28	9.29 ± 3.71	13	.674
Phonation threshold pressure	8.21 ± 3.10	6.69 ± 2.59	4	.049
Mean airflow rate	0.30 ± 0.07	0.27 ± 0.13	4	.536
Laryngeal resistance	47.36 ± 16.97	46.46 ± 24.29	4	.918
Peak pressure	14.04 ± 4.58	10.92 ± 4.07	4	.069
Voice handicap index				
Functional	18 ± 10	12 ± 9	14	<.001
Physical	21 ± 8	15 ± 10	14	.001
Emotional	17 ± 10	11 ± 10	14	.005
Total	56 ± 26	37 ± 27	14	<.001

Abbreviation: F_0, fundamental frequency.
[a]Data are presented as mean ± standard deviation. Complete data sets with measurements of all parameters were not available for every subject; sample size is therefore variable.

Safety and Tolerability

There were no reported minor or major complications. Five procedures had to be truncated due to patient intolerance.

Voice Outcomes

Summary data are presented in **Table 1**. After treatment, dysphonia severity index changed significantly, with a move toward normal voice (P = .003). Phonatory frequency range increased (P = .003), and percent jitter decreased (P = .004). Phonation threshold pressure decreased (P = .049), but there were no significant changes in MPT, mean airflow rate, or laryngeal resistance. Total VHI ($P < .001$) as well as each component of the VHI decreased significantly after treatment (functional: $P < .001$; physical: P = .001; emotional: P = .005; **Figure 3**).

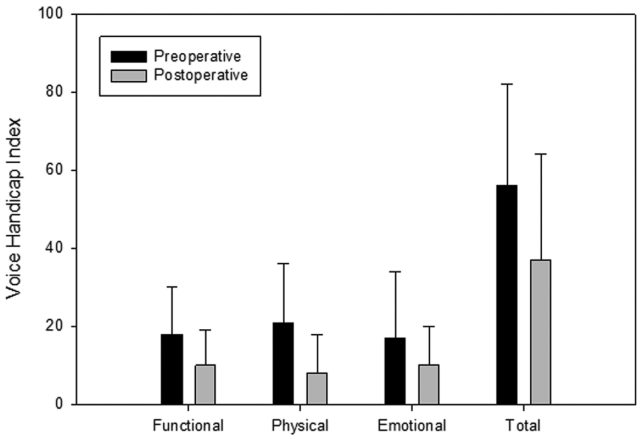

Figure 3. Each component, as well as the total Voice Handicap Index, decreased significantly after treatment. Bar height represents average reported voice handicap; error bars represent standard deviation.

Energy Delivered

Energy delivery data were available on 21 procedures performed to treat bilateral disease. Average energy delivered per procedure was 132 ± 68 J (range, 23-268 J). There was no meaningful difference between the amounts of energy delivered with each laser. For KTP procedures, 126 ± 63 J (range, 47-246 J) were applied; for PDL procedures, 128 ± 75 J (range, 23-268 J) were applied. In 2 procedures for unilateral disease, 108 and 45 J were delivered with the KTP and PDL, respectively.

Discussion

We present a retrospective case series of patients who underwent office-based laser treatment of Reinke's edema. To our knowledge, this study is the largest such series to date.

The increasingly common use of lasers in otolaryngology reflects a general trend toward rendering treatment in the office rather than the operating suite. Office-based treatments offer several advantages. In addition to avoiding the risks of general anesthesia, including myocardial infarction and stroke, unsedated office-based treatment of patients with airway limitations allows the patient to remain in control of his or her own airway throughout the procedure, reducing the risk of airway compromise during induction of general anesthesia. Office procedures cost less,[18] require less time, and avoid the potential complications of microlaryngoscopy, such as dental injury and dysgeusia.[19] Moreover, attempting

an office procedure does not preclude subsequent operative intervention. Still, patients with advanced airway compromise or concerning medical comorbidities are not appropriate for treatment in an office setting, and some patients will demonstrate recalcitrant anxiety to these procedures. In our study, all patients without an obvious airway concern were at least offered an office-based procedure; we do not, however, include patients in the present study who were not amenable to office treatment due to the above limitations.

Surgical lasers fall into 2 broad categories: cutting/ablating lasers and photoangiolytic lasers.[22] Photoangiolytic lasers, including KTP and PDL, selectively target hemoglobin and are therefore most often used to manage highly vascular lesions. Reinke's edema is characterized in part by vascular congestion and stasis within the superficial lamina propria.[2,23] While the exact mechanism of the laser-tissue interaction in benign lesions remains under investigation, it is theorized that photoangiolytic laser energy is effective in improving polypoid degeneration by ablating damaged microvasculature within the SLP, ultimately inducing regression of nonvascular pathologic tissue.[24] It has been proposed that localized energy delivery causes a nonspecific inflammatory response, leading to selective and time-dependent expression of inflammatory cytokines such as transforming growth factor beta 1 and cyclooxygenase 2,[25,26] as well as procollagen/collagenase genes such as matrix metalloproteinases.[26,27] These changes are thought to result in favorable alterations in tissue remodeling. As such, in contrast to classical surgical interventions designed to physically remove excessive tissue, laser therapy is thought to induce a favorable biochemical shift—a biological solution for a biological problem.

In our procedures, energy delivery is titrated to a point of superficial blanching of tissues. No immediate reduction of tissue mass is seen; instead, functional improvement is expected after a period of tissue remodeling. In our study, an average of 132 J was delivered per procedure; however, optimal laser settings and energy titration end points remain undefined. Efforts are underway to characterize these parameters. A recent study examined outcomes for Reinke's edema as a function of laser parameters and initial treatment effects; the average energy applied was 157 J delivered over a 0.369-second exposure time, and voice outcomes were favorable.[28] In an effort to standardize measurement, a validated classification schema was recently proposed to establish a consistent means for measuring response to the KTP laser.[24,29] The present study adds to this growing body of work beginning to evaluate the relationship between amount of energy delivered and treatment outcome.

The patients in our series underwent comprehensive voice analyses before and after completing an intervention, allowing for detailed evaluation of treatment effect. Acoustic measures improved significantly; patients demonstrated improved frequency range due to a higher posttreatment maximum fundamental frequency, representing an improvement in the classic ''low pitched voice'' reported by many patients. Percent jitter also improved after treatment, perhaps reflecting improved vocal fold symmetry after tissue remodeling. Changes in aerodynamic parameters were less pronounced. MPT, laryngeal resistance, mean airflow rate, and peak pressure did not significantly improve following treatment. This may in part be influenced by selection bias, as patients with significant airway compromise—and thus, likely, the most abnormal pretreatment aerodynamic profiles—were not offered office procedures. Phonation threshold pressure did, however, improve after treatment. Finally, our patients demonstrated improvement in all subcategories of the Voice Handicap Index—functional, physical, and emotional. This perhaps more than other measures suggests the utility of these procedures.

Some aspects of the present study may require clarification. First, some individuals showed worsening of certain voice measures after treatment. For example, 1 patient showed increased phonation threshold pressure and airway resistance after a second laser treatment, and 3 patients had decreased MPT after treatment. Also, note that 6 patients underwent multiple procedures. The decision for repeat treatment was based on clinical assessment of recurrent or persistent Reinke's edema with ongoing dysphonia rather than on objective voice data.

While unsedated endoscopic procedures are possible in the majority of patients, anatomic and physiologic limitations as well as anxiety-related factors will represent a barrier in some patients. Of the 25 procedures presented here, 5 were truncated due to patient intolerance. All patients were active smokers at the time of treatment; it is possible that reactive airway physiology contributed to this high rate of intolerance. Our database did not include which patients ultimately underwent operative interventions, but it is likely that some did. Given this limitation, we are unable to assess voice changes related to subsequent surgical intervention and therefore cannot comment on voice outcomes in these patients. Importantly, no patients required emergent airway intervention during or immediately after the procedure, had significant bleeding, or required hospitalization immediately following the procedure. As such, office-based laser treatments in our series were safe.

The present study has several important limitations. As a retrospective analysis without a control group, we cannot determine whether the changes in voice parameters observed after treatment were actually due to the intervention or simply reflect normal temporal variation of the disease. Second, although this is the largest series of patients undergoing photoangiolytic laser therapy for Reinke's edema, our sample size is still modest. Further, complete data sets were not available for all patients. This reduced our effective sample size for the pre- and posttreatment analyses and precluded detailed analysis of parameters over longer periods. Finally, our data set did not provide a standardized means for follow-up; as such, patients who developed complications following the conclusion of their procedure may not be included.

As office-based procedures become increasingly more common, there are many points for further study. Definition of laser settings to optimize tissue remodeling remains an important and active area of investigation. Prospective

studies with larger numbers will help define effect of treatment and clarify the appropriate frequency for follow-up evaluation. To reduce potential cofounders, further studies may benefit from tightly controlling for untreated voice abuse and LPR. Additionally, analyzing differences in outcomes between patients who continue to smoke and those who have quit may define the impact of continued tobacco use on Reinke's edema and healing following photoangiolytic laser intervention. Finally, attention toward procedural details to optimize comfort and minimize anxiety may improve procedure tolerability in a broader population.

Author Contributions

Ian J. Koszewski, study design; data interpretation; data analysis; drafting manuscript; final approval of manuscript; accountable for all aspects of work; **Matthew R. Hoffman**, study design; data interpretation; data analysis; drafting manuscript; final approval of manuscript; accountable for all aspects of work; **W. Greg Young**, study design; data acquisition; data analysis; critical revision of manuscript; final approval of manuscript; accountable for all aspects of work; **Ying-Ta Lai**, study design; data acquisition; data analysis; critical revision of manuscript; final approval of manuscript; accountable for all aspects of work; **Seth H. Dailey**, study design; performed surgical procedures; data interpretation; critical revision of manuscript; final approval of manuscript; accountable for all aspects of work.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Dursun G, Ozgursoy OB, Kemal O, Coruh I. One-year follow-up results of combined use of CO_2 laser and cold instrumentation for Reinke's edema surgery in professional voice users. *Eur Arch Otorhinolaryngol*. 2007;264:1027-1032.
2. Marcotullio D, Magliulo G, Pezone T. Reinke's edema and risk factors: clinical and histopathologic aspects. *Am J Otolaryngol*. 2002;23:81-84.
3. Chung JH, Tae K, Lee YS, et al. The significance of laryngopharyngeal reflux in benign vocal mucosal lesions. *Otolaryngol Head Neck Surg*. 2009;141:369-373.
4. Cipriani NA, Martin DE, Corey JP, et al. The clinicopathologic spectrum of benign mass lesions of the vocal fold due to vocal abuse. *Int J Surg Path*. 2011;19:583-587.
5. Kamargiannis N, Gouveris H, Katsinelos P, et al. Chronic pharyngitis is associated with severe acidic laryngopharyngeal reflux in patients with Reinke's edema. *Ann Otol Rhinol Laryngol*. 2011;120:722-726.
6. Hojslet PE, Moesgaard-Nielsen V, Karlsmore M. Smoking cessation in chronic Reinke's oedema. *J Laryngol Otol*. 1990; 104:626-628.
7. Hirano M, Shin T, Morio M, et al. An improvement in surgical treatment for polypoid vocal cord: sucking technique. *Otologia (Fukuoka)*. 1976;22:583-589.
8. Tan NCW, Pittore B, Puxeddu R. The ''M'' shaped microflap for treatment of complex Reinke's space oedema of the vocal cords. *Acta Otorhinolaryngologica Italica*. 2010;30:259-263.
9. Sant'Anna GD, Mauri M.Use of the microdebrider for Reinke's edema surgery. *Laryngoscope*. 2000;110:2114-2116.
10. Yates A, Dedo HH.Carbon dioxide laser enucleation of polypoid vocal cords. *Laryngoscope*. 1984;94:731-736.
11. Remacle M, Lawson G, Watelet JB. Carbon dioxide laser microsurgery of benign vocal fold lesions: indications, techniques, and results in 251 patients. *Ann Otol Rhinol Laryngol*. 1999;108:156-164.
12. Lumpkin SM, Bishop SG, Bennett S. Comparison of surgical techniques in the treatment of laryngeal polypoid corditis. *Ann Otol Rhinol Laryngol*. 1987;96:254-257.
13. Pitmann MJ, Lebowitz-Cooper A, Iacob C, Tan M. Effect of the 532nm pulsed KTP laser in the treatment of Reinke's edema. *Laryngoscope*. 2012;122:2786-2792.
14. Sheu M, Sridharan S, Kuhn M, et al. Multi-institutional experience with the in-office potassium titanyl phosphate laser for laryngeal lesions. *J Voice*. 2012;26:806-810.
15. Koufman JA, Rees CJ, Frazier WD, et al. Office-based laryngeal laser surgery: a review of 443 cases using three wavelengths. *Otolaryngol Head Neck Surg*. 2007;137:146-151.
16. Rees CJ, Halum S, Wijewickrama RC, Koufman JA, Postma GN. Patient tolerance of in-office pulsed dye laser treatments to the upper aerodigestive tract. *Otolaryngol Head Neck Surg*. 2006;134:1023-1027.
17. Zeitels SM, Burns JA. Office-based laryngeal laser surgery with the 532-nm pulsed-potassium-titanyl-phosphate laser. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:394-400.
18. Rees CJ, Postma GN, Koufman JA. Cost savings of unsedated office-based laser surgery for laryngeal papillomas. *Ann Otol Rhinol Laryngol*. 2007;116:45-48.
19. Klussmann JP, Knoedgen R, Wittekindt C, Damm M, Eckel HE. Complications of suspension laryngoscopy. *Ann Otol Rhinol Laryngol*. 2002;111:972-976.
20. Wuyts FL, de Bodt MS, Molenberghs G, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res*. 2000;43: 796-809.
21. Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI): development and validation. *Amer J Speech Lang Pathol*. 1997;6:66-70.
22. Yan Y, Olszewski AE, Hoffman MR, et al. Use of lasers in laryngeal surgery. *J Voice*. 2010;24:102-109.
23. Martins RH, Dominques MA, Fabro AT, Dias NH, Santana MF. Reinkes' edema: immunoexpression study of fibronectin, laminin and collagen IV in 60 cases by immunohistochemical techniques. *Braz J Otorhinolaryngol*. 2009;75:821-825.
24. Mallur PS, Tajudeen BA, Aaronson N, Branski RC, Amin MR. Quantification of benign lesion regression as a function of 532-nm pulsed potassium titanyl phosphate laser parameter selection. *Laryngoscope*. 2011;121:590-595.
25. Branski RC, Barbieri SS, Weksler BB, et al. Effects of transforming growth factor-beta1 on human vocal fold fibroblasts. *Ann Otol Rhinol Laryngol*. 2009;118:218-226.

26. Lin Y, Yamashita M, Zhang J, Ling C, Welham NV. Pulsed dye laser-induced inflammatory response and extracellular matrix turnover in rat vocal folds and vocal fold fibroblasts. *Lasers Surg Med*. 2009;41:585-594.
27. Mallur PS, Branski RC, Amin MR. 532-nanometer potassium titanyl phosphate (KTP) laser-induced expression of selective matrix metalloproteinases (MMP) in the rat larynx. *Laryngoscope*. 2011; 121:320-324.
28. Young VN, Mallur PS, Wong AW, et al. Analysis of potassium titanyl phosphate laser settings and voice outcomes in the treatment of Reinke's edema. *Ann Otol Rhinol Laryngol*. 2015; 124:216-220.
29. Mallur PS, Johns MM 3rd, Amin MR, Rosen CA. Proposed classification system for reporting 532-nm pulsed potassium titanyl phosphate laser treatment effects on vocal fold lesions. *Laryngoscope*. 2014;124:1170-1175.

Reprinted by permission of Laryngoscope. 2015; 125(4):909-912.

The Laryngoscope
© 2014 The American Laryngological,
Rhinological and Otological Society, Inc.

The Utility of Office-Based Biopsy for Laryngopharyngeal Lesions: Comparison with Surgical Evaluation

Amanda L. Richards, MBBS, FRACS; Manikandan Sugumaran, MD; Jonathan E. Aviv, MD; Peak Woo, MD; Kenneth W. Altman, MD, PhD

Objectives/Hypothesis: Advances in flexible endoscopy with working-channel biopsy forceps have led to excellent visualization of laryngopharyngeal lesions with capability for in-office awake biopsy. Potential benefits include prompt diagnosis without risk of general anesthesia, preoperative counseling, and avoiding an anesthetic should the lesion return benign. We evaluate the accuracy of these biopsies in order to determine their role and diagnostic value.

Study Design: Retrospective chart review.

Methods: Medical records were reviewed from January 1, 2010, through July 31, 2013, of patients who underwent office-based current procedural terminology code 31576 and were taken to the operating room for direct microlaryngoscopy with biopsy/excision. Clinical diagnoses and pathology reports were reviewed. For statistical analysis, we considered three groups: 1) malignant and premalignant, 2) lesions of uncertain significance, and 3) benign lesions.

Results: In the study period, 76 patients with an office biopsy had a clinical picture to warrant direct microlaryngoscopy and biopsy/excision. Kendall's coefficient for each group indicated moderate correlation only. When groups 1 and 2 were considered together, there was a substantial and statistically significant correlation. For malignant and premalignant lesions, the office biopsy analysis was as follows: sensitivity = 60%, specificity = 87%, positive predictive value = 78%, and negative predictive value = 74%.

Conclusion: Office biopsy may offer early direction and avoid operative intervention in some cases; however, for suspected dysplastic or malignant lesions, direct microlaryngoscopy should be the standard of care to ensure adequate full-thickness sampling and staging. For benign pathology, office biopsy is a safe and viable alternative to direct microlaryngoscopy and biopsy/excision.

Key Words: Office biopsy, lesion, leukoplakia, dysplasia, microlaryngoscopy, medical decision making.

Level of Evidence: 4.

***Laryngoscope*, 125:909–912, 2015**

INTRODUCTION

Advances in flexible laryngoscopy, imaging technology, instrument miniaturization, and changes to procedure reimbursement have led to an increase in office-based management in laryngology. Since the introduction of the fiber optic laryngoscope in 1976, there have been steady advances in the quality of lighting and imaging for office laryngeal examinations from fiber optic to distal chip endoscopes.[1] Also, adaptations in the design of the flexible scopes have allowed for the use of a side channel port or disposable sheath for passage of a cupped laryngeal biopsy forceps.[2] The combination of these forceps with optimal imaging has provided an option to obtain tissue for pathology during an outpatient office visit with topical anesthesia. Traditionally, these patients would require a visit to the operating room (OR) with general anesthesia for a direct microlaryngoscopy and biopsy or excision of the lesion. However, regardless of technique, all biopsies need to provide a representative sample of the lesion to demonstrate cell morphology. In addition, sample depth is also important, particularly in cases of dysplasia for which deeper levels may determine a different diagnosis and prognosis.[3]

There are a proposed number of conditions for which office biopsy alone has been proposed as sufficient: 1) confirmed diagnosis of carcinoma when clinically suspected; 2) complete excision of a lesion at the time of office biopsy; 3) benign pathology and resolution of the lesion with treatment; 4) evidence for keratosis, papilloma, or mild dysplasia with stable clinical examination; and 5) the risks of surgical evaluation with general anesthesia outweigh the potential diagnostic or therapeutic benefits of the procedure. Potential benefits include the following: 1) avoiding the risk of general anesthesia, 2) reduced duration from clinical suspicion to histologic confirmation, 3) negating patient anatomic limitations, and 4) avoiding the costs of general anesthesia and the OR.[4]

From the Department of Otolaryngology–Head and Neck Surgery, Icahn School of Medicine, Mount Sinai Hospital, New York, New York, U.S.A.

Editor's Note: This Manuscript was accepted for publication October 7, 2014.

Dr. Aviv is shareholder in Vision Sciences. The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Kenneth W. Altman, MD, PhD, FACS, Professor and Vice Chair for Clinical Affairs, Director, The Institute for Voice and Swallowing at BCM, Bobby R. Alford Department of Otolaryngology - HNS, Baylor College of Medicine, 6501 Fannin Street, Room NA 102, Houston, TX 77030. E-mail: Kenneth.Altman@bcm.edu

DOI: 10.1002/lary.25005

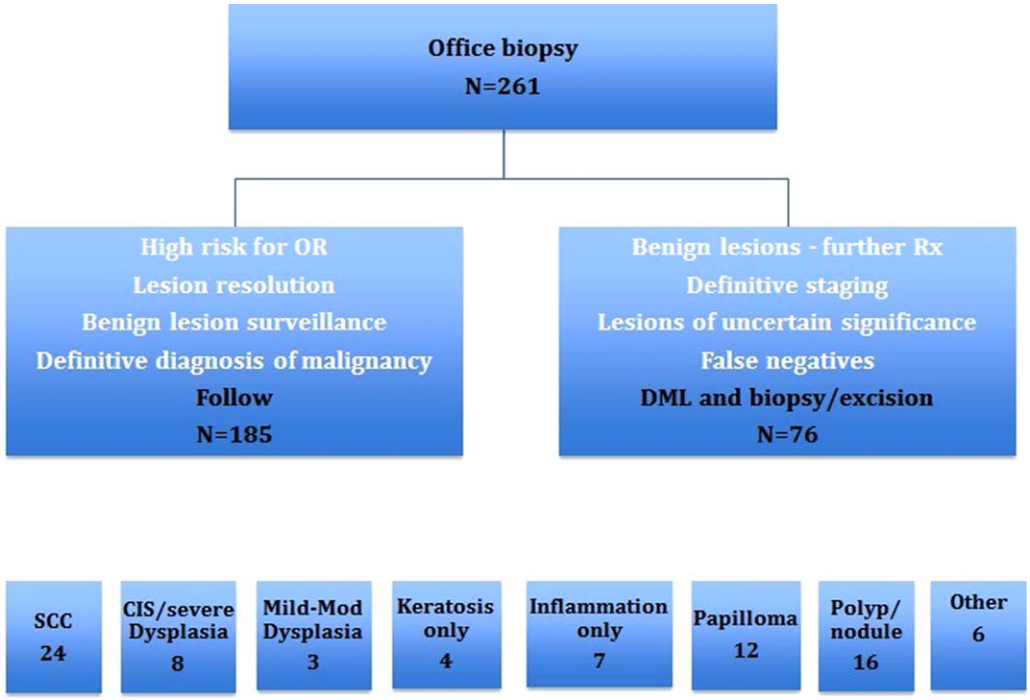

Fig. 1. Flow of office biopsy patients. (Operating room biopsy diagnoses are listed in the last row). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Traditionally, patients warranting direct microlaryngoscopy after office biopsy are those with limited tissue obtained during attempted office biopsy, a concern regarding false-negative office biopsy results, a requirement for disease volume reduction to avoid respiratory or swallowing impairment, and a need for excision of the lesion to improve the voice. Other advantages of direct microlaryngoscopy include a more detailed examination of the extent of a tumor, more accurate biopsy capabilities, and the option for definitive treatment by excision for many lesions.

Despite the popularity of office biopsy, there is a paucity of data in the literature evaluating the accuracy compared to histologic diagnosis at operation. The goals of this study are to determine the accuracy of office biopsies when compared to direct microlaryngoscopy and to evaluate its role and diagnostic value.

MATERIALS AND METHODS

A retrospective medical chart review was performed from January 1, 2010, to July 31, 2013, after receiving approval from the Institutional Review Board Human Subjects Committee. This review identified 261 patients in the clinical practices of the authors who underwent office biopsy (current procedural terminology code 31576) for laryngeal and pharyngeal lesions. Patients' records were then reviewed to determine those who underwent direct microlaryngoscopy with biopsy.

Patients who had resolution of the lesion following biopsy, surveillance of a previously histologically proven benign diagnosis, and a definitive diagnosis of cancer who proceeded to nonsurgical definitive treatment were excluded from the study. We also excluded current anticoagulation, anterior commissure lesions, submucosal lesions, and anatomically obstructive pathology. Patients with brush biopsy alone were also excluded. The pathology reports were reviewed for consistency between office and surgical specimens and compared to clinical diagnoses. The flow of the patients is summarized in Figure 1.

Office biopsies were performed using distal chip video endoscopes (ENT-5000, Vision Sciences, Inc. or VNL-1570STK, KayPENTAX Montvale, NJ) in conjunction with a 2-mm channel endosheath and 1.8-mm nonserrated cup biopsy forceps. The nasal cavity was anesthetized with aerosolized 4% lidocaine with epinephrine 1:100,000 or 4% lidocaine with phenylephrine hydrochloride. The channel-sheathed video endoscope was then passed transnasally into the laryngopharynx. Topical laryngopharyngeal anesthesia was achieved by delivering 0.5 cc of plain 4% lidocaine to the laryngeal surface of the epiglottis. Once supraglottic anesthesia was achieved, 1 to 2 cc of plain 4% lidocaine was then delivered topically to the glottis. The 1.8-mm biopsy forcep was then passed under videoendoscopic guidance and biopsies were performed.

Direct microlaryngoscopy with biopsy was performed under general anesthesia, and lesions were visualized with a zero-degree telescope and binocular microscope. Lesions were excised or sampled for pathologic evaluation using phonosurgical instruments. The procedures included a submucosal dissection in order to obtain epithelial basement membrane in the specimen.

Office biopsy results were divided into clinically relevant groups that would normally used to direct patient care algorithms. For example, mild to moderate dysplasia was separated from severe dysplasia and carcinoma in situ (CIS)/squamous cell carcinoma (SCC). For statistical analysis, we considered three groups: 1) malignant and premalignant (SCC, CIS, and severe dysplasia); 2) lesions of uncertain significance (mild–moderate dysplasia and hyperkeratosis); and 3) benign lesions. Patients who were noted to have a dual diagnosis on histology (e.g., inflammation with mild dysplasia) were analyzed within the group that would direct their final treatment.

To test interrater reliability, we utilized Kendall's coefficient of concordance for the numerically coded ordinal responses

TABLE I.
Summary of Results.

Office Biopsy Pathology	Office Biopsy N =	Accuracy Compared to Final Pathology	Pathology of Missed Diagnosis (False Negatives)	Rate of False Negatives%
SCC	4	100.0%	N/A	0.0%
Severe dysplasia/ CIS	23	17.4%	SCC	56.5%
			Mild–Moderate dysplasia	8.7%
			Inflammation only	8.7%
			Polyps or nodule	4.3%
			Keratosis	4.3%
Mild to moderate dysplasia	12	25.0%	SCC	25.0%
			Severe dysplasia/CIS	33.3%
			Polyps or nodule	8.3%
			Keratosis	8.3%
Keratosis without dysplasia	7	14.3%	SCC	28.6%
			Severe dysplasia/CIS	14.3%
			Polyps or nodule	28.6%
			Inflammation only	14.3%
Inflammation only	8	12.5%	SCC	12.5%
			Severe dysplasia/CIS	25.0%
			Keratosis	12.5%
			Polyps or nodule	25.0%
			Papilloma	12.5%
Polyp/nodule	10	100.0%	N/A	0.0%
Papilloma	11	81.8%	Inflammation only	9.1%
			Other	9.1%
Other benign	6	0.0%	SCC	16.7%
			Keratosis	33.3%
			Inflammation only	33.3%
			Papilloma	16.7%
Inadequate	4	N/A	SCC	50.0%
			Polyp or nodule	50.0%

CIS = carcinoma in situ; SCC = squamous cell carcinoma.

via the SAS macro MAGREE. Kendall's coefficients do not treat all misclassifications equally. For instance, Kendall's coefficients considers the consequences of misclassifying a perfect (rating = 5) object as bad (rating = 1) as more serious than misclassifying it as very good (rating = 4). For all statistical analysis, data was analyzed using the SAS System software (SAS Institute, Inc., Cary, NC).

RESULTS

Seventy-six subjects underwent evaluation with 81 office biopsies with subsequent direct microlaryngoscopy under general anesthesia. The age range of the subjects was 21 to 84 years, with a median age of 62 and a male to female ratio of 5:1. There were 76 laryngeal biopsies and five oropharyngeal biopsies. The oropharynx subsites included two for tonsil, two for tongue base, and one for soft palate. There were no complications from any of the office or operative procedures performed.

The results of office biopsy and their subsequent direct microlaryngoscopy (DML) and biopsy/excision are summarized in Table I.

When considered as separate groups, Kendall's coefficient for each group was all 0.5, indicating "moderate correlation" only. None of these approached statistical significance ($P = 0.5$). When groups 1 and 2 (i.e., lesions of uncertain significance and premalignant/malignancy) were considered together, the coefficient was 0.64 ($P = 0.029$), indicating "substantial correlation." For malignant/premalignant lesions, the office biopsy analysis was as follows: sensitivity = 60%, specificity = 87%, positive predictive value = 78%, and negative predictive value = 74%.

DISCUSSION

Medical decision making, patient counseling, and surgical planning benefit from understanding the nature of a lesion based on patient demographics, clinical history, the physical examination, and cytologic or pathologic diagnosis. The nature of a laryngeal lesion will affect prioritizing the different surgical goals of the following: 1) confidence in a pathologic diagnosis, 2) control of disease, and 3) voice preservation or improvement. See Table II for proposed office biopsy candidacy.

While office biopsy increases in popularity, there needs to be further clarification regarding its utility. The results of this study highlight the concern of office biopsy being used as a substitute for traditional DML. The small tissue sample, limited depth past the basement membrane, and ability to sample only portions of a suspicious lesion are known disadvantages to this technique, especially with leukoplakia and lesions that have some degree of dysplasia. Serrated or other forceps, however, may favourably influence the sensitivity and specificity of the findings, and appropriate selection of forceps is a major consideration in the management of early laryngeal malignancy by any method. Therefore, just as in the OR under direct laryngoscopic conditions for which the surgeon would choose the forceps for biopsy carefully, the same considerations must be applied when performing transnasal biopsy.

Although it may be acceptable for some screening tests to have a high specificity and lower sensitivity, this is not appropriate for this diagnostic test. Sensitivity for malignancy/premalignancy was only 60%, indicating that it is inadequate as a diagnostic test: clinical suspicion alone in this setting would seem at least equivalent. Only 15% of invasive SCC was identified at office biopsy, and it is evident that any clinically suspicious neoplasm must

TABLE II.
Candidates for Office Biopsy.

Inclusion	Exclusion
Anatomic limitations for DML	Anticoagulation
Voice less critical	Anterior commissure location
Following a known benign diagnosis	Submucosal lesion
High risk for general anesthesia	Lesion associated with obstruction

DML = direct microlaryngoscopy.

be evaluated with direct microlaryngoscopy as possible. Cohen et al. noted a 33% false negative rate of in-office flexible laryngeal biopsy; and in their series, CIS on office biopsy was most often SCC at DML.[5] It should be noted, however, that when an office biopsy showed a diagnosis of SCC, it was correlated with the final histologic diagnosis in 100% of patients. For patients with a suspected malignant lesion who are unable to undergo a general anesthetic, reassurance can be given that a diagnosis of cancer in the office correlates with a diagnosis of cancer in OR. The substantial correlation for malignancy/any degree of dysplasia may be adequate to counsel certain patients, and early diagnosis with office biopsy may offer extra confidence in the treatment paradigm, particularly when multiple treatment options are available. It may also reduce time to definitive treatment. When a histologic diagnosis of severe dysplasia or CIS in a lesion with malignant suspicion is added to the diagnostic/staging capabilities of imaging modalities such as CT scanning, radiation may be employed with increased confidence of the clinical diagnosis and stage.[6] Seventy-eight percent of our patients with severe dysplasia, CIS, or SCC on office biopsy were counseled appropriately in advance of surgery and underwent microlaryngoscopy with CO2 laser cordectomy at the time of initial surgical evaluation.

It is an accepted standard of care that mild to moderate dysplasia may be observed, and so it follows that this diagnosis at office biopsy may reassure the clinician. However, lesions of uncertain significance have only moderate correlation to their final pathologic diagnosis. Mild to moderate dysplasia frequently represents more sinister disease, and it may be false reassurance to rely on office biopsy alone. Similar to brush biopsy,[7] a dysplastic office biopsy result indicates a need for further investigation, without providing a definitive diagnosis.

The diagnosis of hyperkeratosis/parakeratosis (pathologic correlates of leukoplakia) and inflammation also frequently correlate with a final histopathologic diagnosis of malignancy and highlight the limitations on biopsy of the awake patient with the absence of tactile feedback. It is well known that potentially malignant and CIS epithelia are associated with pronounced stromal reaction,[8] which is reflected in the 37% of patients with only an office biopsy of inflammation, however, with severe dysplasia/CIS or SCC on final diagnosis. Clinical judgment is paramount in ensuring that patients with suspected false negative biopsies undergo DML.

When considering benign lesions such as papilloma and vocal fold nodules or polyps, office biopsy may play a more definitive role. A clinical suspicion of papilloma or polyps/nodules correlated well with final diagnosis. When clinically indicated, these lesions are appropriate for office management alone.

In the current environment of financially accountable medicine, consideration should be given to cost. Naidu et al.[9] and workers found an average cost saving of $6,970.56 per patient when office biopsy cost was compared to OR biopsy; however, it is instructive to consider that any patient who undergoes both office biopsy and OR biopsy does so at considerable financial burden.

There are limitations to our study, including selection bias determining the candidacy or need for office biopsy. In addition, the technique and types of cupped forceps used were not compared to other approaches. The utility of serrated forceps or rigid endoscopy with nonflexible forceps should also be considered and may enhance sampling. The ability to obtain high diagnostic yield biopsies in the office is also dependent on the experience of the surgeon; our group obtaining the biopsies has over 5 years experience with this method. Classification of the severity of dysplasia is also variable from different classification systems and pathologists,[2] which can skew the accuracy of office biopsies when different grades of dysplasia are considered. A multicenter study is warranted, especially in the realm of medical decisionmaking in complex laryngeal lesions, as well as outcomes of prioritizing disease eradication over voice outcomes with dysplastic lesions.

CONCLUSION

Our study shows that office biopsy has the highest utility in clinically benign lesions and those with SCC. Laryngopharyngeal biopsies have only a moderate correlation with final pathology, although the potential utility increases in certain clinical scenarios and with careful choice of forceps. Office biopsy has a tendency to underestimate the severity of dysplastic lesions, and any degree of dysplasia should be considered as potentially malignant until, as possible, proven otherwise with operative assessment. However, preoperative patient counseling, surgical planning, and therapy may be positively impacted by information from office biopsies, comparable to management of other head and neck neoplasms. For benign pathology that clinically harbors no suspicion for malignancy, office biopsy is a safe and viable alternative to direct microlaryngoscopy and biopsy/excision.

Acknowledgment

We would like to acknowledge Alan Weinberg, MS, for the statistical analysis, and Laurence T. Blanchard, BA, for data preparation.

BIBLIOGRAPHY

1. Eller R, Ginsburg M, Lurie D, Heman-Ackah Y, Lyons K, Sataloff R. Flexible laryngoscopy: a comparison of fiber optic and distal chip technologies. Part 1: vocal fold masses. *J Voice* 2008;22:746–750.
2. Rosen CA, Amin MR, Sulica L, et al. Advances in office-based diagnosis and treatment in laryngology. *Laryngoscope* 2009;119(suppl 2):S185–S212.
3. Zhang HK, Liu HG. Is severe dysplasia the same lesion as carcinoma in situ? 10-Year follow-up of laryngeal precancerous lesions. *Acta Otolaryngol* 2012;132:325–328.
4. Rees CJ, Postma GN, Koufman JA. Cost savings of unsedated office-based laser surgery for laryngeal papillomas. *Ann Otol Rhinol Laryngol* 2007; 116:45–48.
5. Cohen JT, Safadi A, Fliss DM, Gil Z, Horowitz G. Reliability of a transnasal flexible fiberoptic in-office laryngeal biopsy. *JAMA Otolaryngol Head Neck Surg* 2013;139:341–345.
6. Chu MM, Kositwattanarerk A, Lee DJ, et al. FDG PET with contrast-enhanced CT: a critical imaging tool for laryngeal carcinoma. *Radiographics* 2010;30:1353–1372.
7. Nichols ML, Quinn FB, Jr, Schnadig VJ, et al. Interobserver variability in the interpretation of brush cytologic studies from head and neck lesions. *Arch Otolaryngol Head Neck Surg* 1991;117:1350–1355.
8. Hellquist H, Cardesa A, Gale N, Kambic V, Michaels L. Criteria for grading in the Ljubljana classification of epithelial hyperplastic laryngeal lesions. A study by members of the Working Group on Epithelial Hyperplastic Lesions of the European Society of Pathology. *Histopathology* 1999;34:226–233.
9. Naidu N, Noordzij JP, Samin A, Jalisi S, Grillone GA. Comparison of efficacy, safety, and cost-effectiveness of in-office cup forcep biopsies versus operating room biopsies for laryngopharyngeal tumours. *J Voice* 2012; 26:604–606.

Reprinted by permission of J Voice. 2014; 28(3):382-386.

Office-Based Injection Laryngoplasty for the Management of Unilateral Vocal Fold Paralysis

***Sunil P. Verma and †Seth H. Dailey,** **Irvine, California and †Madison, Wisconsin*

Summary: Objective. Office-based injection laryngoplasty (OBIL) is a common method of addressing glottal insufficiency. This retrospective chart review identifies the demographics, laterality, technique, success rate, injectates, and complications of OBIL performed over a 3-year period at a single institution.

Study Design. Retrospective chart review.

Methods. All OBILs performed for the management of UVFP by the senior author over 3 years (2007–2009) were identified from billing records. The age, gender, laterality, underlying disease process, augmentation material, route of injection, and complications were recorded.

Results. Eighty-two OBILs were attempted on 57 patients. The most common route of access was transoral (85.6%). All OBILs were able to be completed. Injectates used were hyaluronic acid derivatives (57.3%), calcium hydroxyapatite (16%), and Cymmetra (16.5%). Three complications (3.7%) occurred. Thirty percent of patients ultimately elected for thyroplasty or ansa reinnervation, 22% found their condition to self-resolve, 14% died, and 25% were lost to follow-up.

Conclusions. Using a variety of approaches, OBIL is possible in almost all patients. The single surgeon transoral route using a rigid angled telescope and curved injection needle was the most commonly used approach. Multiple injectates can be used and have good safety records. The final disposition of patients may be variable and warrants further investigation.

Key Words: Laryngology–Laryngeal surgery–Office-based–Procedures–Surgery–Vocal fold paralysis–Hoarseness–Thyroplasty–Reinnervation.

INTRODUCTION

Injection laryngoplasty (IL) has been a cornerstone in the management of unilateral vocal fold paralysis (UVFP) since its first description.[1] During the majority of the last century, IL was commonly performed in the operating room (OR). However, with the advent of "chip-tip" endoscopes, refinements in the ability to deliver anesthesia to the larynx[2,3] and the development of numerous injectables,[4,5] there has been a move toward IL performed in the office.[6] Advantages of OBIL include markedly decreased cost, avoidance of the risks of general anesthesia, and the ability titrate injectate delivery for optimized voice outcomes, among others.[7]

As the population ages and grows and as some of the most common causes of UVFP increase,[8] including the number of thyroid cancers,[9] cervical spine surgeries,[10] lung cancer resections, and aortic valve replacements,[11] one may expect the incidence of UVFP to increase as well. As the paradigm of OBIL for UVFP continues to evolve, there are questions which remain to be answered.

The first involves the safety profile of both OBIL and the numerous injectables which are being used for the treatment. UVFP often occurs secondary to malignancy, complications from surgery, or both. As such, patients with UVFP often possess multiple morbidities including general health concerns, cardiopulomonary compromise, need for anticoagulation, among other medical and psychosocial concerns. With this in mind, it is critical to evaluate the safety of OBIL as has been done for other office-based laryngeal surgeries.[12,13] In an effort to avoid general anesthesia, another question to be answered is how often OBIL can actually be completed. Finally, there is an active discussion regarding the ultimate disposition of patients after injection.[14–16]

To answer these questions, a retrospective chart review was performed of all OBILs performed for UVFP over a 3-year period at an academic tertiary care institution.

MATERIALS AND METHODS

After obtaining approval by the institutional review board, all OBILs attempted for UVFP by the senior author over 3 years (2007–2009) were identified from billing records. The age, gender, laterality, underlying disease process, route of injection, procedural success rate, amount and type of augmentation material used, complications, and patient disposition were recorded.

All procedures were performed in the otolaryngology clinic examination suite containing a powered examination chair, video tower with photodocumentation capability. Informed consent was obtained and a procedural "time-out" was performed before each procedure. Patient vital signs were collected before the visit; however, no cardiopulomonary monitoring was performed during the procedure. All injectates were directed toward the paraglottic space musculature. Approaches used were transoral,[17] transcricothyroid membrane,[18] transthyrohyoid membrane,[19] and transthyroid ala.

Accepted for publication October 10, 2013.
Financial Disclosures: None.
Conflicts of interest: None.
Presented at the 2011 UCSF Fall Voice Conference/International Association of Phonosurgery; November 3, 2011; San Francisco, California.

From the *University Voice and Swallowing Center, Department of Otolaryngology-Head and Neck Surgery, University of California, Irvine School of Medicine, California; and the †Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin.

Address correspondence and reprint requests to Seth H. Dailey, Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, University of Wisconsin School of Medicine and Public Health, K4/760 Clinical Science Center, Madison, WI, 53792-7395. E-mail: dailey@surgery.wisc.edu

Journal of Voice, Vol. 28, No. 3, pp. 382-386
0892-1997/$36.00
© 2014 The Voice Foundation
http://dx.doi.org/10.1016/j.jvoice.2013.10.006

For the transoral approach, the oral cavity is first anesthetized with topical lidocaine spray applied with an atomizer. The tonsillar pillars, base of tongue, and posterior pharyngeal wall are sprayed with lidocaine. The patient is asked to assume the "sniffing" position and directed to hold his tongue with gauze. Visualization of the laryngopharynx is obtained with a transoral rigid 70° angled telescope held by the surgeon. The view from the scope is transmitted to a screen on the video tower (Figure 1).

An Abraham cannula attached to a syringe with 4% lidocaine is placed along the patient's lingual sulcus and directed over the larynx. A "laryngeal gargle" is performed with 4% lidocaine dripping lidocaine to the endolarynx during sustained phonation. The surgeon then advances a syringe with injectate attached to an orotracheal injector needle (model # 1650030 and 1650050; Medtronic, Minneapolis, MN) along the patient's lingual sulcus and directs it to the larynx. The needle may be used to lateralize the patient's false vocal fold. The needle is inserted through the superior surface of the vocal fold into its body. Injectate is applied within the paraglottic space with approximately 20% overinjection to account for reabsorption.

The percutaneous techniques are performed with a surgeon and an assistant. The skin is anesthetized with 1% lidocaine. After the nasal cavity is anesthetized, a channeled flexible laryngoscope is advanced into laryngopharynx. A laryngeal gargle is performed by dripping 4% lidocaine to the endolarynx *via* the channel of the laryngoscope during sustained phonation. A 25 gauage 1.25-in needle is passed through the skin into the larynx by the surgeon and is directed into the vocal fold.

RESULTS

Eighty-two OBILs were attempted on 57 patients. Patients injected were aged between 16 and 83 years, with a mean age of 60 years. Thirty-five males and 22 females were treated. UVFP occurred on the left side in 40 patients and on the right side in 17. Tables 1 and 2 list the etiology of paralysis and approach used for injection, respectively. No procedure had to be terminated early and all procedures were able to be performed to the intended completion point. On average, 0.64 mL of injectate was used in each setting. The augmentation material used is listed in Table 3.

Three complications (3.7%) were noted during or after OBIL. One patient had a hypersensitivity reaction to Restylane. One patient had calcium hydroxyapatite injected superficially requiring microdirect laryngoscopy and removal at a later date. One patient experienced vocal fold edema after injection and was observed in the office without incident.

Figure 2 details the disposition of patients after OBIL.

FIGURE 1. Surgeon and patient positioning for transoral vocal fold injection.

TABLE 1.
Etiology of UVFP

Etiology	Percentage of Patients
Thoracic	36
Idiopathic	30
Cervical	21
Cerebral	10
Intubation	3

DISCUSSION

UVFP is an entity often encountered by otolaryngologists-head and neck surgeons. Management options include voice therapy, OBIL, and injection laryngoplasty performed under general anesthesia in the OR, reinnervation, thyroplasty, and arytenoid repositioning maneuvers. Definitive treatment typically is deferred for the first 9 months after onset and during that time, patients' options are observation, voice therapy, or IL.

IL has an important role in the management of glottal insufficiency. It provides immediate treatment of symptoms related to voice and cough. OBIL offers some advantages over IL performed in the OR. OBIL permits an unobstructed view of the vocal folds, allowing the surgeon to clearly visualize the change in configuration during injection.[7] There is room for immediate analysis of results permitting simultaneous modification if necessary.[20] Performing the procedure under local anesthesia not only reduces the risks associated with general anesthesia but also allows patients to return to normal activities immediately, preventing lost time from work.

Another advantage of OBIL is cost savings. Grant et al estimated increased charges of $8250 for IL performed in the OR compared with the office.[21] Similarly, other authors have noted significant financial savings associated with performance of IL in office as opposed to the OR.[22,23]

Surgeon preference for performance of IL in the OR versus the office for management of UVFP varies tremendously. A recent multi-institution analysis reported equal numbers of IL performed in the OR and in the office.[6] Recent reports of UVFP management show IL performed entirely in the office[24,25] and entirely in the OR.[26] Rationale beyond surgeon preference drives the decision of where to perform IL, including

TABLE 2.
Approach Used for OBIL

Approach	Number of Times (Percent of Total)
Transoral	71 (86.6)
Transcricothyroid membrane	8 (9.8)
Transthyrohyoid membrane	2 (2.4)
Transthyroid ala	1 (1.2)

TABLE 3.
Injectate Used During OBIL

Injectate	Number of Times Used (Percent of Total)
Hyaluronic acid (Hylaform, Allergan-Inamed Crop, Irvine, CA)	33 (40.2)
Calicium hydroxyapetite (Radiesse Voice, BioForm Medical, San Mateo, CA)	20 (24.4)
Micronized dermis (Cymetra, LifeCell Corp, Branchburgh, NJ)	14 (17.1)
Hyaluronic acid gel (Juvederm Ultra Plus, Allergan, Santa Barbara, CA)	8 (9.8)
Hyaluronic acid (Restylane, Q Med, Uppsala, Sweden)	6 (7.3)
Teflon	1 (1.2)

access to resources. In this series, all patients were treated in office. One reason for this is the fact that University of Wisconsin Clinics is a hospital-based practice in which injectables may be billed to the insurance. In a stand-alone clinic, patients are responsible for cost of the injectate, which causes many to elect for procedures in the OR. Additionally, the office laryngeal surgery suite is located within the hospital building, allowing both inpatients and outpatients to be examined and treated using the same setup.

The average age of patients treated in this series was 60 years which is similar to other reports.[25,27] The left vocal fold was affected more often, which is also consistent with large studies.[8] The most common etiology of paralysis was thoracic which included injury to the recurrent laryngeal nerve (RLN) from mass effect of benign and malignant disease or complications after chest surgery. All patients in this series were able to be injected to the intended completion point using a transoral, transcricothyroid membrane, transthyrohyoid membrane, or transthyroid ala approach. The transoral approach was preferred by the authors as it can be performed by one surgeon, without the need for an assistant. It also allows for the entirety of the needle to be visualized during the injection.

The average amount of injectate applied in this population was 0.64 mL. Mau and Courey[28] demonstrated that on average 0.62 and 0.41 mL of calcium hydroxyapatite were necessary to medialize a cadaveric vocal fold *via* a lateral injection. The increase may be a result of the overinejction necessary to account for reabsorption of injectate. Numerous injectates were used in this study, which were tolerated well by most patients. However, two complications noted in this study were related to the injectate used. The first was a hypersensitivity reaction to Restylane. A study of rabbit vocal folds injected with Restylane revealed that at 1 week and 3 months after injection, the vocal folds experienced "low fibrinogenesis," "a slight inflammatory reaction and absence of necrosis," and "granuloma formation and low fibrinogenesis."[29] However, within the Dermatology literature, injection site inflammation resulting in transient redness and edema of the injected site immediately after injection has been noted in 0.02% of individuals who underwent injection of hyaluronic acid gel for soft tissue augmentation.[30] Additionally, hypersensitivity and inflammatory reactions to hyaluronic acid gel have been noted after cutaneous injections for management of facial rhytids.[30–32] It is very possible that the patient treated in this series experienced a similar reaction in the vocal fold after injection.

The other complication resulted from an injection of calcium hydroxyapatite into the superficial lamina propria, requiring removal under general anesthesia during microlaryngoscopy. This was removed in a manner similar to techniques described by others.[33] Ensuring placement of the injectate into the correct portion of the larynx is paramount in OBIL.

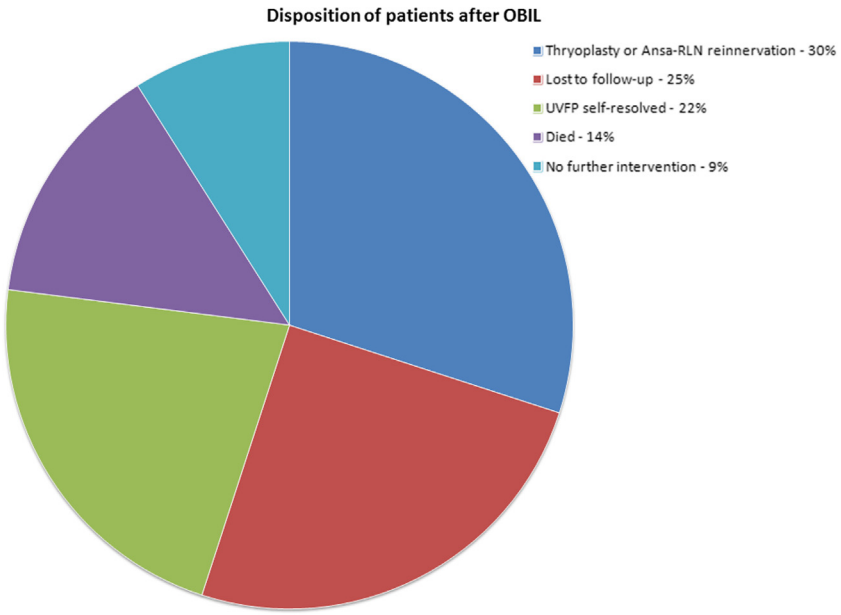

FIGURE 2. Disposition of patients after OBIL.

UVFP often results from malignancy, surgery, and sometimes both. As mentioned earlier, patients may also have general health concerns, cardiopulomonary compromise, anticoagulation needs, as well as psychosocial stressors. As such, the safety of any intervention for this patient population must critically be evaluated. These data, in combination with other data sets, confirm the notion that OBIL is a safe procedure for patients with UVFP.[24,25] One patient had a complication in which vocal fold edema was noted and that the procedure was terminated without incident. There were no complications requiring hospital admission. Patients who were on aspirin prophylactically to prevent cardiac events were asked to stop taking medication 1 week before injection. However, those patients who were taking anticoagulants for therapeutic treatments did not stop taking medications for IL. No complications with hematoma or airway compromise occurred with this approach. For most patients who had injection performed transorally, IL was performed using a 27 gauge needle, in which little, if any, bleeding was noted even if patients were anticoagulated. For this reason, it was deemed safe to continue blood thinners for patients in whom it was medically necessary and do not report any complications with this approach. Others have also shown that procedures performed while a patient is taking anticoagulants are safe.[3,34]

There are risks associated with general anesthesia, which is one of the major motivators to performing office-based laryngeal surgery. Graboyes et al[26] recently published their experience with IL performed under general anesthesia for patients with UVFP after thoracic surgery. Although the majority of their patients did quite well, one of the 20 patients did have intraoperative bile reflux on induction of anesthesia resulting in pneumonitis that may have been avoided with OBIL.

The disposition of patients after injection is shown in Figure 1. Thirty percent of patients sought a definitive intervention in the form of thyroplasty or ansa cervicalis-RLN reinnervation. These results are similar to a study performed by Arviso et al,[16] in which 29% of patients who underwent IL (in the OR or the office) for UVFP required further definitive intervention with medialization thryoplasty. Sixteen percent of the patients treated by Damrose[25] for UVFP required thryoplasty and/or arytenoid adduction after OBIL.

There are multiple reasons why this may have occurred. The concept of laryngeal synkinesis describes abnormal reinnervation of the laryngeal muscles after injury to the RLN.[35–37] After deinnervation of the vocal fold, regeneration of RLN motor axons place the vocal fold in either a favorable or unfavorable position.[37] It has been posited that early medialization of the vocal fold with IL places the vocal fold in a favorable position that is maintained by synkinetic reinnervation.[15] Another consideration is that fibrosis and scarring secondary to IL assist in placing the vocal fold in a permanent medial position.[14,38] Perhaps due to a combination of these reasons, only 30% of the patients in this study required definitive treatment.

In the present study, 22% of individuals had a documented return of function and normal voice noted during stroboscopic examination of the larynx. Fourteen percent of the individuals died, and 9% returned to the office, were noted not to have full recovery of vocal fold motion, and opted for no further intervention. One-quarter of patients did not follow-up. Although this is a sizable number, it is similar to the results of other retrospective studies.[13–16,39] One reason for this is likely due to the large draw of the University of Wisconsin where patients may choose to follow-up with a local otolaryngologist or primary care physician. Some of these patients may have had return of normal or near normal voicing and not found a reason to follow-up. Sulica[40] noted that in idiopathic vocal fold paralysis, which was the second most common reason for UVFP in this series, 52% ± 17% of individuals affected regained complete recovery of voice.

There are limitations to this study which should be recognized. All patients were treated by a single-physician and the data were analyzed in a retrospective fashion. Outcome measures were not obtained in this study, so it is not possible to examine how effective OBIL is. However, other studies have demonstrated improvements in voice quality, swallowing ability, and voice-related quality-of-life after OBIL.[13,25]

From these data, further questions remain to be answered. Multiple injectates were used and it would be interesting to determine which of these is the most durable. The reasons for patients not opting for a more definitive surgery would also be helpful to know.

CONCLUSIONS

OBIL is a safe procedure that is well tolerated in the management of UVFP. Multiple injectates may be used, and familiarity with multiple approaches is beneficial to be able to treat the most number of individuals in the office setting. As noted in this and other studies, a minority of patients who undergo IL require laryngeal framework surgery or a reinnervation procedure.

REFERENCES

1. Bruening W. Uber eine neue Behandlungsmethode der Rekurrenslahmung. *Verhandl Ver Dtsch Laryngol*. 1911;18:93–151. [In German].
2. Bastian RW, Collins SL, Kaniff T, Matz GJ. Indirect videolaryngoscopy versus direct endoscopy for larynx and pharynx cancer staging. Toward elimination of preliminary direct laryngoscopy. *Ann Otol Rhinol Laryngol*. 1989;98:693–698.
3. Bastian RW, Delsupehe KG. Indirect larynx and pharynx surgery: a replacement for direct laryngoscopy. *Laryngoscope*. 1996;106:1280–1286.
4. Rosen CA, Thekdi AA. Vocal fold augmentation with injectable calcium hydroxylapatite: short-term results. *J Voice*. 2004;18:387–391.
5. Ford CN, Bless DM, Loftus JM. Role of injectable collagen in the treatment of glottic insufficiency: a study of 119 patients. *Ann Otol Rhinol Laryngol*. 1992;101:237–247.
6. Sulica L, Rosen CA, Postma GN, et al. Current practice in injection augmentation of the vocal folds: indications, treatment principles, techniques, and complications. *Laryngoscope*. 2010;120:319–325.
7. Rubin HJ. Dysphonia due to unilateral nerve paralysis. Treatment by the intracordal injection of synthetics—a preliminary report. *Calif Med*. 1965; 102:105–109.
8. Rosenthal LH, Benninger MS, Deeb RH. Vocal fold immobility: a longitudinal analysis of etiology over 20 years. *Laryngoscope*. 2007;117: 1864–1870.
9. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973-2002. *JAMA*. 2006;295:2164–2167.

10. Patil PG, Turner DA, Pietrobon R. National trends in surgical procedures for degenerative cervical spine disease: 1990-2000. *Neurosurgery.* 2005; 57:753–758. discussion 753–758.
11. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. *N Engl J Med*. 2011;364: 2128–2137.
12. Yung KC, Courey MS. The Effect of office-based flexible endoscopic surgery on hemodynamic stability. *Laryngoscope*. 2010;120:2231–2236.
13. Mathison CC, Villari CR, Klein AM, Johns MM 3rd. Comparison of outcomes and complications between awake and asleep injection laryngoplasty: a case-control study. *Laryngoscope*. 2009;119:1417–1423.
14. Yung KC, Likhterov I, Courey MS. Effect of temporary vocal fold injection medialization on the rate of permanent medialization laryngoplasty in unilateral vocal fold paralysis patients. *Laryngoscope*. 2011;121: 2191–2194.
15. Friedman AD, Burns JA, Heaton JT, Zeitels SM. Early versus late injection medialization for unilateral vocal cord paralysis. *Laryngoscope*. 2010;120: 2042–2046.
16. Arviso LC, Johns MM 3rd, Mathison CC, Klein AM. Long-term outcomes of injection laryngoplasty in patients with potentially recoverable vocal fold paralysis. *Laryngoscope*. 2010;120:2237–2240.
17. Ford CN, Roy N, Sandage M, Bless DM. Rigid endoscopy for monitoring indirect vocal fold injection. *Laryngoscope*. 1998;108:1584–1586.
18. Ward PH, Hanson DG, Abemayor E. Transcutaneous Teflon injection of the paralyzed vocal cord: a new technique. *Laryngoscope*. 1985;95: 644–649.
19. Amin MR. Thyrohyoid approach for vocal fold augmentation. *Ann Otol Rhinol Laryngol*. 2006;115:699–702.
20. Dedo HH, Urrea RD, Lawson L. Intracordal injection of Teflon in the treatment of 135 patients with dysphonia. *Ann Otol Rhinol Laryngol*. 1973;82: 661–667.
21. Grant JR, Hartemink DA, Patel N, Merati AL. Acute and subacute awake injection laryngoplasty for thoracic surgery patients. *J Voice*. 2008;22: 245–250.
22. Bove MJ, Jabbour N, Krishna P, et al. Operating room versus office-based injection laryngoplasty: a comparative analysis of reimbursement. *Laryngoscope*. 2007;117:226–230.
23. Andrade Filho PA, Carrau RL, Buckmire RA. Safety and cost-effectiveness of intra-office flexible videolaryngoscopy with transoral vocal fold injection in dysphagic patients. *Am J Otolaryngol*. 2006;27: 319–322.
24. Lee SW, Kim JW, Chung CH, et al. Utility of injection laryngoplasty in the management of post-thyroidectomy vocal cord paralysis. *Thyroid*. 2010;20: 513–517.
25. Damrose EJ. Percutaneous injection laryngoplasty in the management of acute vocal fold paralysis. *Laryngoscope*. 2010;120:1582–1590.
26. Graboyes EM, Bradley JP, Meyers BF, Nussenbaum B. Efficacy and safety of acute injection laryngoplasty for vocal cord paralysis following thoracic surgery. *Laryngoscope*. 2011;121:2406–2410.
27. Bhattacharyya N, Kotz T, Shapiro J. Dysphagia and aspiration with unilateral vocal cord immobility: incidence, characterization, and response to surgical treatment. *Ann Otol Rhinol Laryngol*. 2002;111:672–679.
28. Mau T, Courey MS. Influence of gender and injection site on vocal fold augmentation. *Otolaryngol Head Neck Surg*. 2008;138:221–225.
29. Perazzo PS, Duprat Ade C, Lancellotti CL. Histological behavior of the vocal fold after hyaluronic acid injection. *J Voice*. 2009;23:95–98.
30. Friedman PM, Mafong EA, Kauvar AN, Geronemus RG. Safety data of injectable nonanimal stabilized hyaluronic acid gel for soft tissue augmentation. *Dermatol Surg*. 2002;28:491–494.
31. Lupton JR, Alster TS. Cutaneous hypersensitivity reaction to injectable hyaluronic acid gel. *Dermatol Surg*. 2000;26:135–137.
32. Soparkar CN, Patrinely JR. Managing inflammatory reaction to restylane. *Ophthal Plast Reconstr Surg*. 2005;21:151–153.
33. Chheda NN, Rosen CA, Belafsky PC, Simpson CB, Postma GN. Revision laryngeal surgery for the suboptimal injection of calcium hydroxylapatite. *Laryngoscope*. 2008;118:2260–2263.
34. Luu Q, Tsai V, Mangunta V, Berke GS, Chhetri DK. Safety of percutaneous injection of bovine dermal crosslinked collagen for glottic insufficiency. *Otolaryngol Head Neck Surg*. 2007;136:445–449.
35. Crumley RL. Experiments in laryngeal reinnervation. *Laryngoscope*. 1982; 92(9 Pt 2 suppl 30):1–27.
36. Crumley RL. Laryngeal synkinesis: its significance to the laryngologist. *Ann Otol Rhinol Laryngol*. 1989;98:87–92.
37. Crumley RL. Laryngeal synkinesis revisited. *Ann Otol Rhinol Laryngol*. 2000;109:365–371.
38. Carroll TL, Rosen CA. Long-term results of calcium hydroxylapatite for vocal fold augmentation. *Laryngoscope*. 2011;121:313–319.
39. Tan M, Woo P. Injection laryngoplasty with micronized dermis: a 10-year experience with 381 injections in 344 patients. *Laryngoscope*. 2010;120: 2460–2466.
40. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118:1303–1307.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(2):302-307.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–
Head and Neck Surgery
2015, Vol. 152(2) 302–307
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2014
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599814562727
http://otojournal.org
SAGE

Alterations in Extracellular Matrix Composition in the Aging Larynx

Anete Branco, SLP, PhD[1], Alexandre Todorovic Fabro, MD, PhD[2], Tatiana Maria Gonçalves, MD[1], and Regina Helena Garcia Martins, MD, PhD[1]

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract

Objective. To study by immunohistochemistry the alterations of collagens I, III, IV, and V and elastin in the aging process of the human larynx.

Study Design. Cadaver study.

Setting. Universidade Estadual Paulista, Botucatu Medical School, São Paulo State University (UNESP), Brazil.

Subjects and Methods. Thirty vocal folds were obtained at autopsy from 10 adult men (aged 30 to 50 years) and 20 geriatric men (10 aged 60 to 75 years and 10 aged >75 years). Mid membranous vocal fold slides were subjected to immunohistochemical reactions. Digital imaging software (ImageJ) was used to quantify the increase in brownish staining of the lamina propria structures of vocal folds, from superficial to deep layers.

Results. There was an increase of collagen I and III immunoexpression in the elderly larynges, in both layers. Collagens IV and V were immunoexpressed in the vessels endothelium of the lamina propria and in the basement membrane. The immunoexpression of elastin decreased in the elderly larynges, in both lamina propria layers of the vocal folds.

Conclusion. A clear increase of collagens I and III and a decrease of elastic fibers were observed in the lamina propria of vocal folds. The concentration of collagens IV and V was the same across age groups. These findings suggest that as men age, the density of the extracellular matrix increases, brought about by an increase in collagen, while the loss of elastin results in decreased viscoelasticity.

Keywords

collagen, elastin, aged, lamina propria, vocal folds, larynx, immunohistochemistry

Received May 15, 2014; revised October 28, 2014; accepted November 14, 2014.

The physiology of voice production is intimately connected with the characteristics of the vocal fold lamina propria,[1,2] the vocal fold being the main vibrating structure during phonation.[3] In the pursuit to understand the vocal properties typical of age groups, sexes, and diseases, researchers have gone deeper in their morphological studies, trying to elucidate the mechanisms involving the lamina propria microstructures of vocal folds, mainly relating to elastic, collagen, and protein fibers.

Collagen fibers are found across the whole lamina propria, from its superficial to its deepest layers,[4,5] supporting tissue structure through enhanced stability and resistance, relevant to vocal physiology. Gray et al[6] confirmed that collagen and elastic fibers are involved in vocal fold biomechanics, in which the former allow tension and stretching while the latter allow deformation and a quick return to the initial shape.

The composition of the vocal folds extracellular matrix varies significantly between sexes and age groups, implying biomechanical differences that directly influence the vocal properties.[7] The lamina propria trilaminar structure in the adult larynx described by Hirano[8] is not present in the newborn.[3,9,10] It starts to organize only in childhood, from 6 years of age onward. In aging people, anatomical and functional changes occur in the vocal tract, especially in the vocal folds. At this stage of life, structural and functional changes take place in the epithelial covering, the muscle, and the lamina propria of vocal folds, notably, the atrophy of epithelial layers and the vocal muscle, as well as density alterations in collagen and elastic fibers and in hyaluronic acid.[11-13]

Elastic fibers change morphologically in elderly people, relative to adults. Their remodeling with age is reflected by

[1]Ophthalmology, Otorhinolaryngology, Head and Neck Surgery Department, Botucatu Medical School, UNESP–Univ Estadual Paulista, São Paulo, Brazil
[2]Department of Pathology, Botucatu Medical School, UNESP–Univ Estadual Paulista, São Paulo, Brazil

Corresponding Author:
Regina Helena Garcia Martins, SLP, PhD, Botucatu Medical School/UNESP, Ophthalmology, Otorhinolaryngology, Head and Neck Surgery Department, Distrito de Rubião Jr, CEP 18618-970, Botucatu, São Paulo, Brazil.
Email: rmartins@fmb.unesp.br

parameters such as thickness, size, orientation, and location.[14] The ratio of collagen to elastic fibers also evolves, increasing with age, and the predominance of the former explains the rigidity and reduction of the vibratory mucosal wave. Such alterations, combined with loss of elastin and hyaluronic acid, further decrease the viscosity of vocal folds and negatively affect voice emission in the elderly population.[1,4,5,12,15]

Given the complexity and interdependence of vocal fold microstructures, one can expect them to be susceptible to organic changes with aging. Published studies focused more on collagen types I and III, while the rest were little investigated. Thus, the goal of our study was to analyze the concentration of collagen types I, III, IV, and V and of elastin in human larynx senescence by immunohistochemistry.

Methods

The research project was approved by the Ethics Committee for Research on Human Beings of Botucatu Medical School/UNESP, São Paulo, Brazil (reference 3861/2011).

Twenty larynges from elderly men (aged 60-90 years) were included, as well as 10 from male adult controls (aged 30-50 years), collected at autopsy. The elderly group was subdivided into 2 age subgroups: 60 to 75 years (n = 10) and greater than 75 years (n = 10).

The cause of death was retrieved from the autopsy records, and when necessary, additional information was obtained from medical records and family. The exclusion criteria were septicemia; prolonged intubation; systemic infections; persistent dermatologic, autoimmune, or metabolic diseases; smoking habit; and cervical trauma of any kind or other conditions that might compromise the mucosa of vocal folds and invalidate the immunohistochemical analysis.

Fresh larynges were incised at their posterior portion and examined macroscopically to exclude lesions. As a standard procedure, the middle part of the right vocal fold was dissected (**Figure 1**) and immediately embedded in 10% buffered formaldehyde for 24 to 48 hours. Paraffin blocks were prepared for histological slides (hematoxylin and eosin [H&E] stain) and immunohistochemical reactions. The left vocal fold was used in another study.

The following antibodies were used: collagen I (Col1A1, 1:2000 dilution; Dako, Carpinteria, California), collagen III (Col 3A1, dilution 1:1000; Dako), collagen IV (Col4A, 1:40 dilution; Santa Cruz Biotechnology, Santa Cruz, California), collagen V (Col 5A1, 1:100 dilution; Santa Cruz Biotechnology), and elastin (1:200 dilution, Dako). The antigen preparation for immunohistochemical reactions of collagens I, III, and V and of elastin was performed with 1% pepsin at pH 1.8 and incubated for 15 minutes at 60°C and then 30 minutes at 37°C. Blocking was performed with 8% methanol and hydrogen peroxide, followed by 3% Molico milk (Nestlé, Vevey, Switzerland). The secondary complex HRP EnVision (Dako) was added for 1 hour. Diaminobenzidine was added for 5 minutes, and Harris hematoxylin was added for 20 seconds. For collagen IV antigenic sites retrieval, we used a Pascal (Dako) pressure chamber for 3 minutes in the pretreatment solution Trilogy (Cell Marque, Rocklin, California). Following the protocol recommended by our laboratory, the positive control used for collagen I, III, and V and elastin was a fragment of kidney and, for negative control of these same antibodies, only buffered saline, replacing the primary antibody in a series of sections each sampling.

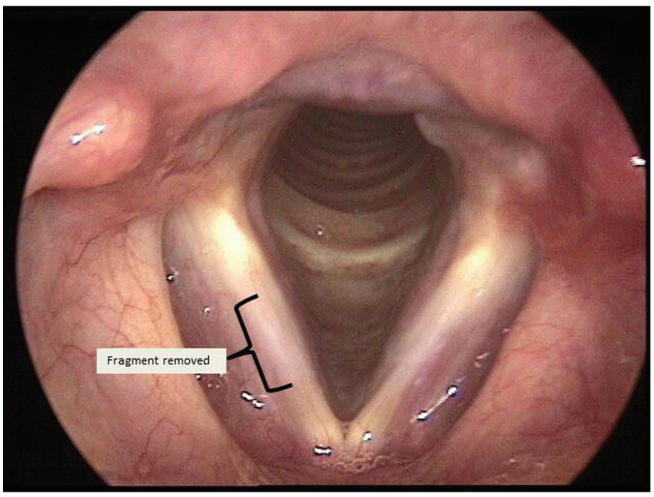

Figure 1. Vocal fold showing the site of the fragment dissection.

Protein Expression Analysis

Reading of histological slides was performed by 3 authors, blindly and randomly. The slides were evaluated at random, without knowledge of the groups and the age of patients, using a light microscope from Zeiss (Axiostar plus, Carl Zeiss do Brasil Ltda), at 40×.

The analyzed locations were the basement membrane, the endothelium vessels, and the lamina propria (superficial and deep layers). Because of the imprecise nature of the boundaries surrounding the intermediate layer of the lamina propria, we decided to divide the lamina propria into 2 layers (superficial and deep) to facilitate presentation of the results. The size of the lamina propria was measured between the basal membrane and muscle fibers. The thickness of the lamina propria was divided into 2 portions subjectively as the superficial layer, corresponding to the upper portion and positioned just below the basement membrane, and the deep layer, corresponding to the lower portion and positioned just above the muscle fibers.

Measurement of antibodies involved quantification of the level of brown staining of the lamina propria and basement membrane structures. The area was set at 2 μm, with percentage used to report results.[16,17] ImageJ software was used in this analysis. We set the polygon as the selection tool for the area of interest, choosing 3 fields at random, and we set the Color Deconvolution plugins to HDAB and Make Binary.

Statistical Analysis

To compare age groups, considering the response profile assessed in 2 layers of the vocal fold (superficial and deep layers), we used a parametric variance analysis combined

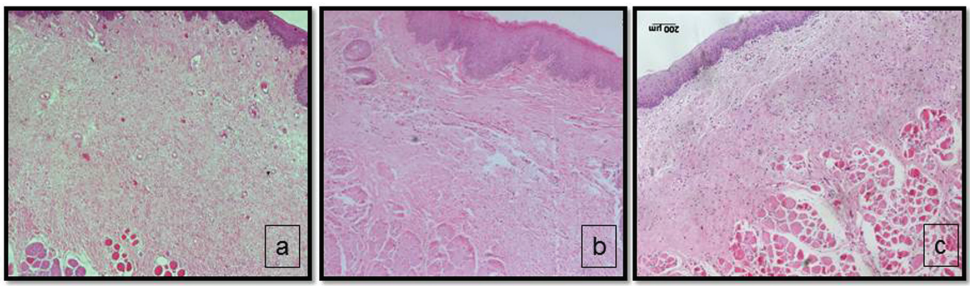

Figure 2. (A) Control group: collagen fibers in the lamina propria, some fibroblasts and vessels. (B, C) Elderly groups: dense collagen in the lamina propria and thickened basement membrane. Hematoxylin and eosin, 10×.

Table 1. Immunoexpression of Antibodies for Collagens I, III, and Elastin, Measured as Area Percentage (%).[a]

	Antibodies								
	Collagen I			Collagen III			Elastin		
Age Group, y	SL	DL	*P* Layers	SL	DL	*P* Layers	SL	DL	*P* Layers
30-50	34.2 (8.5)	3.8 (10.4)	.105	25.6 (21.6-29.4)	23.6 (20.5-29.9)	.046	28.2 (25.7-29.3)	27.5 (23.6- 33.3)	.508
60-75	39.6 (5.1)	44.4 (7.5)	.114	31.2 (23.7-35.6)	24.6 (20.1-29.7)	.0051	20.8 (19.2-23.7)	23.0 (18.7-28.2)	.114
>75	40.4 (5.7)	46.9 (9.3)	.114	32.4 (24.8-44.4)	29.9 (22.1-40.1)	.333	15.9 (10.8-24.3)	18.2 (14.3-25.9)	.285
P ages	.002*			.0061*			.00001*		

[a]Mean (standard deviation) or median (minimum and maximum) according to age group and depth of vocal folds lamina propria. DL, deep layer; SL, superficial layer.
**P* with statistical significance.

with the respective multiple comparisons tests and the nonparametric model when necessary (median, minima, and maxima values). Means and standard deviations of each group were presented when the parametric model was used. For the nonparametric model, medians and their minimum and maximum values were presented, considering the 5% significance level ($P < .05$) within each age range.

Results

The starting point in our study was to analyze H&E slides to check the syntopy of the lamina propria structures in vocal folds and then perform the immunohistochemical study. In these images (**Figure 2**), it was possible to observe in the control group adults that collagen fibers are arranged as a loose fabric between some fibroblasts and vessels in the entire vocal fold lamina propria. In the 60- to 75-year-old group subjects, it was possible to identify a dense collagen thickness deposited on the subepithelial layer (**Figure 2B**). Finally, in the elderly group (greater than 75 years of age), dense collagen uniformly occupied the whole lamina propria (**Figure 2C**).

The immunoexpression of antibodies for collagens I, III, and elastin as well as collagens IV and V is presented in **Tables 1** and **2**, respectively, and shown in **Figures 3** to **7**. We observed that, for collagen I and III, there was an increase of labeling density in the elderly larynges, without any statistical difference between elderly subgroups. Regarding collagens IV and V, a greater labeling density was observed in the endothelium of vessels of the lamina propria and basement membrane. Finally, regarding elastin, we observed a decrease of this antibody as the age increased in both elderly groups.

Table 2. Immunoexpression of Antibodies for Collagens IV and V, Measured as Area Percentage (%).[a]

	Collagen IV		Collagen V	
Age Group, y	SL	DL	SL	DL
30-50	20.9 (2.9)	17.6 (3.7)	38.5 (14.4)	36.6 (8.5)
60-75	20.7 (4.4)	19.3 (4.2)	36.2 (8.9)	35.0 (5.6)
>75	21.3 (3.9)	20.4 (3.5)	31.8 (5.6)	31.3 (6.0)
P ages	.573		.2597	
P Layers	.0034		.397	

[a]Mean (standard deviation) according to age group and depth of vocal folds lamina propria. DL, deep layer; SL, superficial layer.
**P* with statistical significance.

Discussion

Presbyphonia is a physiological effect of the senescence process that takes place in the whole body. Trying to

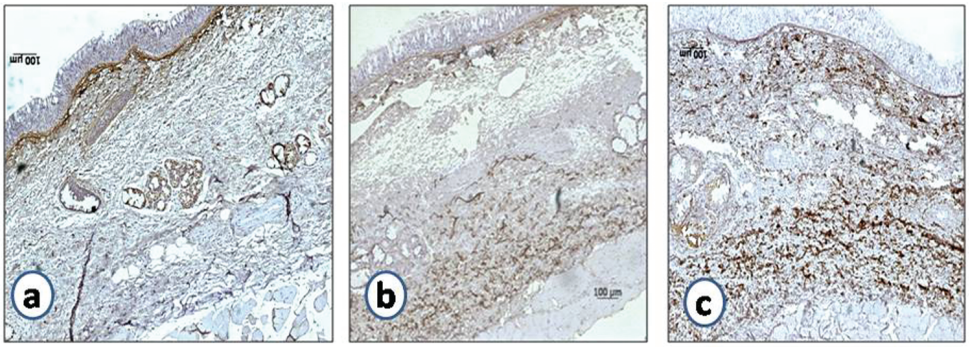

Figure 3. Vocal folds. (A) Control group: collagen I staining the basement membrane and the superficial layer. (B, C) Elderly groups: dense collagen I staining the superficial and deep layers of the lamina propria. Immunohistochemical reaction, 20×.

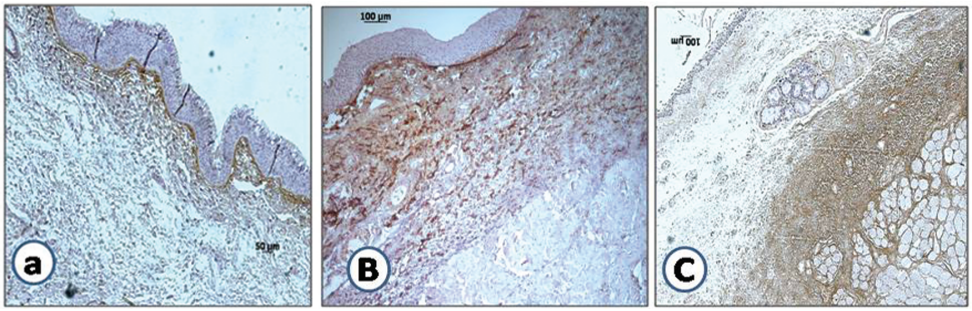

Figure 4. Vocal folds. (A) Control group: collagen III staining the basement membrane and the superficial layer. (B, C) Elderly groups: collagen III staining the superficial (B) and deep layers (C) of the lamina propria. Immunohistochemical reaction, 20×.

correlate the alterations in the vocal folds of the elderly population with the rest of the body's epithelial covering, Ximenes Filho et al[18] performed the simultaneous histomorphometric analysis of vocal folds and inguinal skin from 20 elderly cadavers (10 male and 10 female), finding similar alterations in both locations such as lamina propria and epithelial atrophy.

The immunohistochemical staining revealed that this network is mainly formed by collagen I and III (**Table 1**; **Figures 3** and **4**) with a significant decreased in density of elastin in the elderly larynx (**Table 1**; **Figure 7**). These structural changes in the vocal fold cover are responsible for the hardening of the vocal folds,[6,19,20] clinically manifested by symptoms of hoarseness, vocal fatigue, and vocal range restriction, having a direct impact on speech in higher frequencies.[21,22]

Keeping the proportion of elastic and collagen fibers in the lamina propria of the larynx is important for this organ to retain local resistance, provided by collagen fibers, and distensibility, given by the elastic fibers. Studies of the larynx in animals[23] and in humans[24] have confirmed the role of collagen I in resistance and that of collagen III and elastin in flexibility and elasticity. These parameters work as a ''balance,'' whose equilibrium determines the relative participation of the different vocal fold layers in phonation. When collagen fibers start to predominate in the lamina propria, as seen in our study, the vocal folds become more rigid, which negatively affects voice qualities.[19] According to Ohno et al,[25] the collagen increase in the elderly larynx affects the mucosal wave, resulting in decreases of phonatory intensity and fundamental frequency, especially in women. The voice becomes failing and weak, a vocal pattern known as phonasthenia.[26]

In an attempt to better understand what causes the abnormal production of collagen in the elderly larynx, Kosztyła-Hojna et al[27] studied the ultrastructure of vocal folds retrieved at elderly autopsies or from total laryngectomies due to supraglottic larynx carcinomas, without damage to the actual vocal folds. Using transmission electronic microscopy, the authors observed the destruction of epithelial cells, a vacuolar degeneration of the cell's cytoplasm, a considerable increase of collagen fibers, a vacuolar degeneration of fibroblasts, an increase of the endoplasmic reticulum, and an increase of blood vessels. These authors suggest that the increase of collagen fibers is connected to the cytoplasmatic alterations observed in fibroblasts. This possibility is insufficiently supported by the literature and therefore requires additional studies.

Although the mechanisms are not yet understood, the elderly larynx fibroblasts produce collagen in an excessive and irregular way, as well as a lower amount of hyaluronic acid and elastic fibers. In a study of young and elderly rat larynx fibroblast cultures, Hirano et al[12] observed in the latter lower amounts of hyaluronic acid and higher concentrations of collagen I. They observed the opposite when the fibroblasts were exposed to fibroblast growth factor, which was thus considered by the authors a potential therapeutic tool for

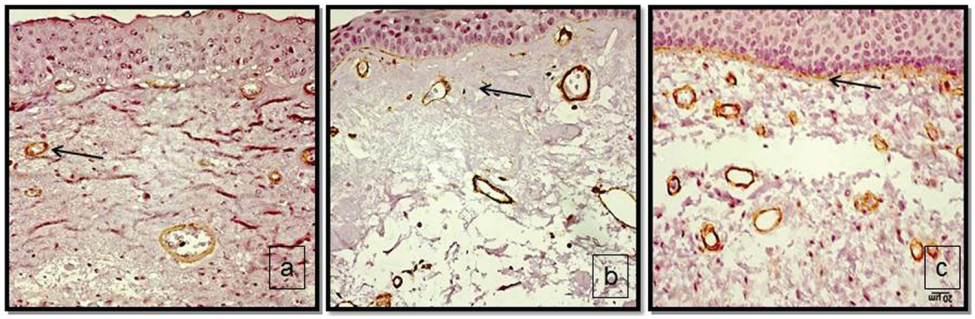

Figure 5. Vocal folds. (A) Control group: collagen IV staining the vessel endothelium (arrow). (B, C) Elderly groups: collagen IV staining the vessel endothelium and the basement membrane (arrows). Immunohistochemical reaction, 20×.

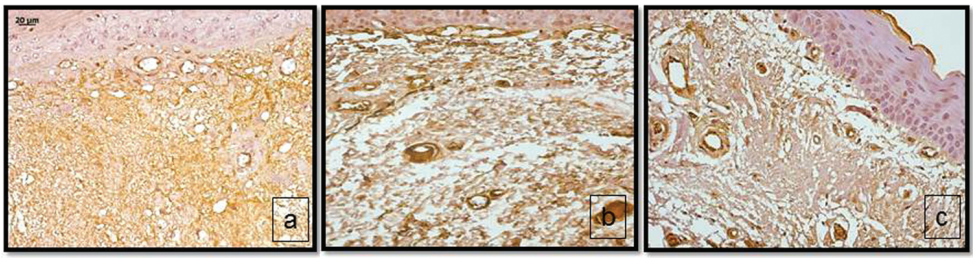

Figure 6. Vocal folds. (A) Control group: collagen V staining the vessel endothelium (arrow). (B, C) Elderly groups: collagen V staining the vessel endothelium and the basement membrane (arrows). Immunohistochemical reaction, 20×.

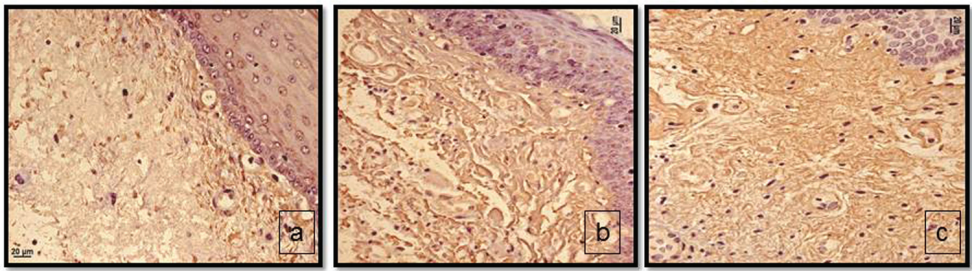

Figure 7. Vocal folds. (A) Control group: uniform elastin antibody staining the lamina propria. (B, C) Elderly groups: sparse distribution of elastin in the lamina propria. Imunohistochemical reaction, 20×.

lamina propria collagen modulation, opening new perspectives for presbyphonia and vocal fold atrophy treatments.[25,28]

The basement membrane and the vessels endothelium having a structural function, they also contain collagen types IV and V, which contribute to support the lamina propria.[29] As in our results, it was observed that these antibodies, when quantified in said locations, also increased with age (**Table 2**). Similar results have been reported in larynges with chronic inflammation in morphological studies.[30] The increase of said collagens is probably related to the simultaneous increase of blood vessels, which is observed with both persistent inflammation and aging.

Elastic fibers in the elderly larynx not only are found in lower amounts but also present structural alterations. Sato and Hirano[15] observed that with aging, these fibers lose elasticity because of an increase in amorphous substance, a decrease in microfibrils, and their metabolic alteration. Thus, we can infer that synchronicity of the phonatory system can depend on biomechanic changes, due to the physiological remodeling of the extracellular matrix.[31]

We can therefore stress that the vibration mechanism of vocal folds is governed by the laryngeal tissue biomechanics. The observed alterations in the elderly larynx lamina propria components are interpreted in the endoscopic examinations as vocal fold atrophy and spindle chink, giving the voice different degrees of hoarseness, asthenia, and breathiness.[32] These ''new'' micro- and macroanatomic configurations present in the elderly larynx may mirror age-dependent physiological remodeling; that is, the structural changes in the vocal folds lamina propria that occur with aging make it necessary for the extracellular matrix and its components to be remodeled to remain functional.

Conclusions

In our methodological conditions, we observed in the vocal folds lamina propria a clear increase of collagens I and III,

as well as a decrease of elastic fibers. The concentration of collagens IV and V did not change according to age group. These findings suggest that as men age, the density of the extracellular matrix increases, brought about by an increase in collagen, while the loss of elastin results in decreased viscoelasticity.

Author Contributions

Anete Branco, interpretation of data for the work, drafting the work, final approval of the work, agreement to be accountable for all aspects of the work; **Alexandre Todorovic Fabro**, interpretation of data for the work, drafting the work, final approval of the work, agreement to be accountable for all aspects of the work, data analysis; **Tatiana Maria Gonçalves**, interpretation of data for the work, drafting the work, final approval of the work, agreement to be accountable for all aspects of the work; **Regina Helena Garcia Martins**, study design, interpretation of the data, writing and approval of the manuscript, critical revision, final approval, agreement to be accountable for all aspects of the work.

Disclosures

Competing interests: None.

Sponsorships: Fapesp and Fundunesp, financial support to immunohistochemical analysis and English version.

Funding source: None

References

1. Butler JE, Hammond TH, Gray SD. Gender-related differences of hyaluronic acid distribution in the human vocal fold. *Laryngoscope*. 2001;111:907-911.
2. Ramos HV, Simões MJ, Pontes PA, et al. Immunohistochemistry as a method to study elastic fibers of human vocal fold. *Braz J Otorhinolaryngol*. 2005;71:486-491.
3. Ishii K, Yamashita K, Akita M, Hirose H. Age-related development of the arrangement of connective tissue fibers in the lamina propria of the human vocal fold. *Ann Otol Rhinol Laryngol*. 2000;109:1055-1064.
4. Hammond T, Gray S, Butler J. Age and gender related collagen distributions in human vocal folds. *Ann Otol Rhinol Laryngol*. 2000;109:913-920.
5. Sato K, Hirano M, Nakashima T. Age-related changes of collagenous. *Ann Otol Rhinol Laryngol*. 2002;111:15-20.
6. Gray SD, Titze IR, Alipour F, Hammond TH. Biomechanical and histologic observations of vocal fold fibrous proteins. *Ann Otol Rhinol Laryngol*. 2000;109:77-85.
7. Kutty JK, Webb K. Tissue engineering therapies for the vocal fold lamina propria. *Tissue Eng Part B Rev*. 2009;15:249-262.
8. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr*. 1974;26:89-94.
9. Hirano M, Kurita S, Nakashima T. Growth, development and aging of human vocal folds. In: Bless DM, Abbs JH, eds. *Vocal Fold Physiology*. San Diego, CA: College-Hill; 1983:22-43.
10. Sato K, Hirano M, Nakashima T. Fine structure of the human newborn and infant vocal fold mucosae. *Ann Otol Rhinol Laryngol*. 2001;110:417-424.
11. Gray SD. Cellular physiology of the vocal folds. *Otolaryngol Clin North Am*. 2000;33:679-697.12
12. Hirano S, Bless DM, Río AM, Connor NP, Ford CN. Therapeutic potential of growth factors for aging voice. *Laryngoscope*. 2004;114:2161-2167.
13. Pontes P, Yamasaki R, Behlau M. Morphological and functional aspects of the senile larynx. *Folia Phoniatr Logop*. 2006;58:151-158.
14. Hammond TH, Gray SD, Butler J, Zhou R, Hammond E. Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngol Head Neck Surg*. 1998;119:314-322.
15. Sato K, Hirano M. Age-related changes of elastic fibers in the superficial layer of the lamina propria of vocal volds. *Ann Otol Rhinol Laryngol*. 1997;106:44-48.
16. Collins TJ. ImageJ for microscopy. *BioTechniques*. 2007;43: S25-S30.
17. Miot HA, Brianezi G. Morphometric analysis of dermal collagen by color clusters segmentation. *An Bras Dermatol*. 2010; 85:361-364.
18. Ximenes Filho JA, Tsuji DH, Nascimento PHS, Sennes LU. Histologic changes in human vocal folds correlated with aging: a histomorphometric study. *Ann Otol Rhinol Laryngol*. 2003;112:894-898.
19. Kendall K. Presbyphonia: a review. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:137-140.
20. Paulsen F, Kimpel M, Lockemann U, Tillmann B. Effects of ageing on the insertion zones of the human vocal fold. *J Anat*. 2000;196:41-54.
21. Gorhan-Rowan MM, Laures-Gore J. Acoustic-perceptual correlates of voice quality in elderly men and women. *J Commun Disord*. 2006;39:171-184.
22. Gregory ND, Chandran S, Lurie D, Sataloff RT. Voice disorders in the elderly. *J Voice*. 2012;26:254-258.
23. Kolachala VL, Torres-Gonzalez E, Mwangi S, et al. A senescence accelerated mouse model to study aging in the larynx. *Otolaryngol Head Neck Surg*. 2010;142:879-885.
24. Tateya T, Tateya I, Bless DM. Collagen subtypes in human vocal folds. *Ann Otol Rhinol Laryngol*. 2006;116:469-476.
25. Ohno T, Hirano S, Rousseau B. Age-associated changes in the expression and deposition of vocal fold collagen and hyaluronan. *Ann Otol Rhinol Laryngol*. 2009;118:735-741.
26. Behlau M, Azevedo R, Pontes P. Conceito de voz normal e classificação das disfonias. In: Behlau M, org. *Voz: o livro do especialista*. Rio de Janeiro, Brazil: Revinter; 2001:53-79.
27. Kosztyła-Hojna B, Andrzejewska A, Rogowski M, Pepinski W. Evaluation of the ultrastructure of the vocal folds mucosa in patients with presbyphonia. *Otolaryngol Pol*. 2002;56:239-243.
28. Ford C. Voice restoration in presbyphonia. *Arch Otolaryngol Head Neck Surg*. 2004;130:1117.
29. Prades JM, Dumollard JM, Duband S, et al. Lamina propria of the human vocal fold: histomorphometric study of collagen fibers. *Surg Radiol Anat*. 2010;32:377-382.
30. Baltaziak M, Kosztyła-Hojna B, Moskal D, Falkowski D. Morphological conditions of dysphonia in patients with vocal folds edema. *Pol Merkur Lekarski*. 2010;29:181-186.
31. Roberts T, Morton R, Al-Ali S. Microstructure of the vocal fold in elderly humans. *Clin Anat*. 2011;24:544-551.
32. Martins RHG, Gonçalvez TM, Pessin AB, Branco A. Aging voice: presbyphonia. *Aging Clin Exp Res*. 2014;26:1-5.

Reprinted by permission of JAMA Otolaryngol Head Neck Surg. 2013; 139(12):1291-1295.

Original Investigation

Influence of Age on Treatment With Proton Pump Inhibitors in Patients With Laryngopharyngeal Reflux Disease
A Prospective Multicenter Study

Young Chan Lee, MD; Jun Seok Lee, MD; Seung Woo Kim, MD; Kee Hwan Kwon, MD, PhD; Young Gyu Eun, MD, PhD

IMPORTANCE Several trials on the predictors of response to proton pump inhibitor (PPI) treatment of laryngopharyngeal reflux (LPR) have shown conflicting results. Furthermore, the influence of age in disease severity and response to PPI therapy is unclear.

OBJECTIVE To assess the difference in disease severity and response to PPI therapy according to age in patients with LPR.

DESIGN, SETTING, AND PARTICIPANTS Prospective multicenter study at 3 tertiary medical centers of 264 consecutive patients with LPR who were referred to the otolaryngology clinic from November 2010 to February 2012.

INTERVENTIONS Participants were prescribed 15 mg of lansoprazole (PPI) twice daily for 3 months.

MAIN OUTCOMES AND MEASURES Reflux Symptom Index (RSI), Reflux Finding Score (RFS), and laryngopharyngeal reflux–health-related quality of life (LPR-HRQOL) were collected at baseline and at 1 and 3 months postbaseline.

RESULTS After 3 months, 35 patients were lost to follow-up and excluded; the remaining 229 patients included 135 men and 94 women. The oldest group (60-79 years; n = 111) showed higher baseline RSI ($P < .001$) and LPR-HRQOL ($P < .001$) scores than the 18- to 39-year-old (n = 35) and 40- to 59-year-old (n = 83) groups. However, baseline RFS scores showed no significant difference among age groups ($P = .44$). Within each age group, the RSI, RFS, and LPR-HRQOL improved significantly with PPI therapy (all $P < .001$); however, no significant difference in improvement of RSI ($P = .59$), RFS ($P = .50$), or LPR-HRQOL ($P = .09$) was seen among the groups. At 3-month follow-up, significantly more responders, defined as those whose RSI score improved by more than 50%, were found in the 18- to 39-year-old and 40- to 59-year-old groups (86% and 75%, respectively) than in the oldest group (57%) ($P = .002$), but there was no significant difference in proportion of responders among age groups at 1-month follow-up ($P = .69$).

CONCLUSIONS AND RELEVANCE In patients with LPR, age seems to affect the subjective symptoms and resulting impact on quality of life but not the laryngeal findings. Furthermore, older patients are more likely not to respond to PPI therapy than younger patients.

Author Affiliations: Department of Otolaryngology–Head and Neck Surgery, Kyung Hee University School of Medicine, Seoul, Korea (Y.C. Lee, J.S. Lee, Eun); Department of Otolaryngology–Head and Neck Surgery, Seoul Veterans Hospital, Seoul, Korea (Kim); Department of Otorhinolaryngology–Head and Neck Surgery, Kangdong Sacred Heart Hospital, Hallym University School of Medicine, Seoul, Korea (Kwon).

Corresponding Author: Young Gyu Eun, MD, PhD, Department of Otolaryngology–Head and Neck Surgery, Kyung Hee University School of Medicine, 1 Hoegi-dong, dongdaemun-gu, Seoul 130-702, Korea (ygeun@hanmail.net).

JAMA Otolaryngol Head Neck Surg. 2013;139(12):1291-1295. doi:10.1001/jamaoto.2013.5556
Published online November 7, 2013.

Copyright 2013 American Medical Association. All rights reserved.

Laryngopharyngeal reflux (LPR) refers to the retrograde flow of stomach contents into the throat and larynx, which leads to symptoms such as chronic dysphonia, throat clearing, cough, globus sensation, and sore throat.[1] In otolaryngologic practice, approximately 10% of patients presenting to outpatient clinics and more than 50% of patients with voice problems receive a diagnosis of LPR.[2] Laryngopharyngeal reflux is a gastrointestinal and otolaryngological condition related to but distinct from gastroesophageal reflux disease (GERD). The reflux of gastric contents is at the core of both LPR and GERD, but the mechanism and symptoms of the 2 disorders are different.[3] The laryngeal mucosa are vulnerable to exposure to acidic substances, so patients often present with laryngopharyngeal symptoms without heartburn or regurgitation.[4] Several studies have suggested that the frequency of GERD complications (esophagitis, Barrett esophagus, stricture) is higher in geriatric patients because of many potential aggravating factors[5-7]; however, the influence of age in LPR is unclear. Although a 3-month empirical trial of proton pump inhibitor (PPI) treatment is generally regarded as a cost-effective modality for the initial management of LPR, the difference in response to PPI therapy according to age is not established. Garrigues et al[8] suggested that response to therapy was associated with younger age and shorter duration of laryngeal symptoms, but the response could not consistently be predicted in patients with chronic posterior laryngitis.

Thus, we conducted a prospective cohort study to assess the difference in symptom severity according to age in patients with LPR through the Reflux Symptom Index (RSI),[9] Reflux Finding Score (RFS),[10] and laryngopharyngeal reflux-health-related quality of life (LPR-HRQOL)[11] score. Furthermore, we evaluated the influence of age on response to PPI therapy.

Methods

Subjects and Study Design

Patients with suspected LPR who were referred to 3 different otolaryngology clinics from November 2010 to February 2012 were assessed for eligibility for the study. All patients underwent otolaryngologic evaluation by 1 of us (S.W.K., K.H.K., Y.G.E.), including laryngoscopy and video strobolaryngoscopy. A diagnosis of LPR was made on the basis of the presence of at least 1 of the following symptoms: hoarseness, chronic cough, throat irritation, laryngospasm, chronic throat clearing, and dysphasia. Diagnosis of LPR was also based on confirmed signs such as erythema, vocal cord edema, subglottic edema, posterior pachydermia, laryngeal edema, ventricular obliteration, and thick endolaryngeal mucus and granuloma from the findings of the laryngoscope. Last, diagnosis required that symptoms not be due to laryngitis caused by upper airway infections and/or allergies. A consensus meeting among the 3 otolaryngologists was conducted to improve interrater reliability. The enrolled participants included all the patients who initially received a diagnosis of LPR and had no history of PPI treatment. Patients younger than 18 years, those experiencing GERD symptoms but not LPR symptoms, and those who had a malignant tumor or major psychosis were excluded. The study protocol was reviewed and approved by the institutional review boards of Kyung Hee University Hospital at Gangdong, Seoul Veterans Hospital, and Samsung Changwon Hospital. Written informed consent was obtained from the participants.

In addition to advice about lifestyle modification (avoidance of caffeine, alcohol, smoking, fatty food, and eating close to bedtime), patients with LPR were prescribed 15 mg of lansoprazole 2 times a day for 3 months. Patients were instructed to take the PPI 30 minutes before meals. The disease severity and changes in subjective symptoms were assessed using 2 surveys, the RSI and LPR-HRQOL, for 3 months.[9,11] The surveys were administered 3 times during this period at outpatient department visits: the first visit and follow-up visits at 1 and 3 months. In addition, to evaluate the objective findings of the laryngeal condition, the RFS by Belafsky et al[10] was conducted by an otolaryngologist. A greater than 50% primary RSI improvement from baseline was considered a response to PPI therapy.

The Questionnaires

The RSI is a high-validity survey that not only assesses the level of severity of LPR but also includes 9 questions to estimate the response to treatment. The questionnaire evaluates the level of symptoms and their severity through a 6-point Likert scale, which ranges from 0 to 5. A high score indicates that patients have more severe symptoms, whereas 0 indicates the absence of symptoms.

The LPR-HRQOL, developed by Carru et al,[11] has been shown to be a reliable and valid rating scale for evaluating the quality of life (QOL) of LPR patients. It consists of a simple questionnaire of 43 questions in the 5 categories of hoarseness, cough, throat clearing, swallowing, and overall impact of acid reflux. The questionnaire consists of basic 7-point Likert scale questions in the first 4 categories and concludes with a 10-point Likert scale question regarding the overall impact of acid reflux. A high score indicates that patients have more severe symptoms, whereas 0 indicates the absence of symptoms.

Examination of the Larynx

All enrolled participants underwent laryngoscopy to evaluate objective signs of LPR on the basis of the RFS. An otolaryngologist performed the examination using a strobolaryngoscope, and 70° rigid endoscopes were generally used. When the vocal cords and surrounding structures were not clearly visible in the rigid endoscope, a flexible nasopharyngoscope was used to increase accuracy. Participants were instructed to verbalize "yee" in a high-pitched tone, low-pitched tone, and regular-pitched tone. Through this procedure, the diagnosis of LPR and the RFS, the rating scale of clinical advanced LPR, were assessed. The RFS ranged from 0 (normal state) to 26; a higher score indicates a deteriorated laryngeal condition. A consensus meeting among the 3 clinics was conducted to improve the interrater and intrarater reliability in RFS scoring. This investigation was performed according to standard protocol and scored by observers blinded to the patient's identity.

Copyright 2013 American Medical Association. All rights reserved.

Table 1. Initial Reflux Symptom Index (RSI), Reflux Finding Score (RFS), and LPR–Health-Related Quality of Life (LPR-HRQOL) According to Age Group

Test	Score, Mean (SD)			*P* Value
	18-39 y (n = 35)	40-59 y (n = 83)	60-79 y (n = 111)	
RSI	13.88 (7.68)	12.20 (8.90)	18.45 (10.43)	<.001
RFS	6.78 (4.86)	7.18 (4.55)	7.75 (3.84)	.44
LPR-HRQOL				
Voice	14.48 (18.65)	18.63 (21.36)	29.55 (21.17)	<.001
Cough	6.42 (7.36)	8.96 (7.36)	14.10 (12.40)	<.001
Throat clearing	6.97 (7.99)	6.15 (7.22)	12.78 (11.50)	<.001
Swallowing	6.05 (7.25)	6.40 (6.38)	11.31 (9.30)	<.001
Overall impact of acid reflux	21.28 (18.46)	21.45 (14.72)	34.81 (23.26)	<.001

Table 2. Improvement in Reflux Symptom Index (RSI), Reflux Finding Score (RFS), and LPR–Health-Related Quality of Life (LPR-HRQOL) After Proton Pump Inhibitor Therapy According to Age

Test	Score, Mean (SD)			*P* Value	
	Baseline	1 Month	3 Months	Within Group	Among Groups
RSI					
18-39 y	13.88 (7.68)	8.77 (6.80)	4.62 (5.35)	<.001	.59
40-59 y	12.20 (8.90)	7.57 (7.31)	5.32 (6.28)	<.001	
60-79 y	18.45 (10.43)	12.89 (9.15)	10.81 (9.86)	<.001	
RFS					
18-39 y	6.21 (4.87)	3.89 (3.10)	2.92 (2.59)	<.001	.50
40-59 y	6.46 (4.70)	4.34 (3.36)	2.98 (2.66)	<.001	
60-79 y	7.74 (3.92)	5.76 (3.39)	4.40 (2.67)	<.001	
LPR-HRQOL					
18-39 y	21.50 (19.29)	14.28 (8.38)	11.93 (4.13)	<.001	.09
40-59 y	20.81 (14.50)	17.41 (12.40)	14.02 (8.03)	<.001	
60-79 y	35.20 (23.29)	24.27 (17.57)	23.08 (19.15)	<.001	

Statistical Analysis

For statistical analysis, SPSS, version 18.0 (SPSS), was used, and all of the data are presented as mean (SD). A *t* test and analysis of variance (ANOVA) were used to compare age differences in RSI, RFS, and LPR-HRQOL data. A repeated measure of ANOVA was used to determine which age group showed a greater response to PPI therapy. An ANOVA model in repeated measures at 3 time points was used with Bonferroni correction for multiple comparisons. Comparison of the proportion of responders according to age was made using a χ^2 analysis. A difference was considered statistically significant when the *P* value was less than .05.

Results

Study Populations

Of 264 consecutive patients considered for the study, 35 were excluded because of loss of follow-up at 3 months. A total of 229 patients with LPR were enrolled and completed the study without loss to follow-up. There were 135 men (59.0%) and 94 women (41.0%). The mean (SD; range) age of the patients was 55.7 (14.0; 18-79) years. Patients were divided into 3 age groups of 18 to 39, 40 to 59, and 60 to 79 years. The number of patients in each group was 35 (15.3%), 83 (36.2%), and 111 (48.5%), respectively.

Difference of RSI, RFS, and LPR-HQOL According to Age

The oldest patient group (60-79 years) with LPR had significantly higher mean (SD) baseline RSI scores than the 18- to 39-year-old and 40- to 59-year-old patient groups (18.45 [10.43] vs 13.88 [7.68] and 12.20 [8.90], respectively; $P < .001$). However, the RFS score showed no significant difference among age groups. The oldest patient group showed significantly worse results on all domains of the LPR-HRQOL (all $P < .001$) (**Table 1**).

Improvement of RSI, RFS, and LPR-HRQOL After PPI Therapy

Within each age group, scores on all 3 tests improved significantly during the period of PPI therapy; however, there was no significant difference among groups in the amount of improvement (RSI, $P = .59$; RFS, $P = .50$; LPR-HRQOL, $P = .09$) (**Table 2**).

Difference in Proportion of Responders on RSI According to Age

Among the age groups, the proportion of responders, as evaluated by RSI score, showed no significant difference at 1 month; however, responders were significantly more plentiful in the 2 younger groups than the oldest group at 3 months ($P = .002$) (**Table 3**).

Copyright 2013 American Medical Association. All rights reserved.

Table 3. Proportion of Responders as Evaluated by Reflux Symptom Index (RSI) According to Age

Follow-up Period, mo	Responders,[a] No. % 18-39 y (n = 35)	40-59 y (n = 83)	60-79 y (n = 111)	*P* Value
1	11 (31)	26 (31)	41 (37)	.70
3	30 (86)	62 (75)	63 (57)	.002

[a] Responders were defined as those whose RSI score improved by more than 50% after proton pump inhibitor therapy.

Discussion

The major finding of this prospective study is that the subjective severity of LPR is significantly greater in older than in younger patients. In addition, the older patients showed lower response rates after PPI therapy.

In a previous cohort study in 100 patients with no history of voice or laryngeal symptoms, 35% were found to have symptoms of LPR and 64% showed 1 or more physical findings of LPR on laryngoscopic examination.[12] Despite the high prevalence of LPR, there are few data on the influence of age on symptom severity or response to PPI treatment. The present prospective study investigated the influence of age on severity and PPI response in LPR. We assessed the subjective severity through the LPR-HRQOL, which evaluated the QOL of patients with LPR, as well as the RSI in groups stratified according to age. To our knowledge, this is the first report of greater severity of disease and negative impact on QOL in geriatric patients with LPR.

It is known that the incidence of GERD symptoms does not increase with age; however, several studies suggest that the frequency of GERD complications such as esophagitis, stricture, or Barrett esophagus is significantly higher in older people.[5,7,13,14] The most likely reason for the increased severity of GERD in older people is the cumulative injury of acid to the esophageal mucosa over time. In addition, a defective antireflux barrier, abnormal esophageal clearance, altered esophageal mucosal resistance, and delayed gastric emptying could contribute to this phenomenon.[15]

It is not known whether the severity of LPR in older patients is greater than in younger patients. Saruç et al[16] demonstrated that age is not a risk factor for the development of LPR. In our results, older patients with LPR showed a higher score on the RSI. Moreover, LPR symptoms had a significantly greater negative impact on the lives of older patients. In a recent study on the QOL impact of LPR, LPR symptoms had a significant correlation with all tested QOL parameters.[17] However, we could not find any difference in RFS, the objective laryngeal finding, among the groups. Our data suggest that age affects the subjective symptoms and resulting impact on the QOL in LPR but not the laryngeal finding. The difference may be the result of a different perspective on their health status among people of different ages.

Many previous studies agree that PPI therapy is the cornerstone of LPR treatment.[18,19] The current management strategy for patients with LPR is empirical therapy with a twice-daily PPI for 3 months[19]; however, the proportion of patients who respond to PPI therapy varies, ranging from 27% to 83% for 1 month of treatment and 41% to 100% for 3 months of treatment.[20-22] Although several randomized clinical trials demonstrated no significant postintervention difference between groups receiving a PPI vs placebo, in a recent open-label observational study, significant improvement in RSI (primary RSI improvement of >50%) was obtained in 75% of patients after 12 weeks.[23] This is similar to the response rate in the 40- to 59-year-old group in our study. Moreover, we were able to find a difference in response among the groups according to age. This is a noteworthy finding in our trial, although there was no placebo group.[21]

Several trials on the predictors of response to PPI treatment have also shown conflicting results. Park et al[24] demonstrated that pretherapy abnormalities in the interarytenoid mucosa and true vocal fold were associated with a 2-fold increase in symptom response to PPI treatment. Williams et al[25] reported that neither baseline GERD symptoms nor endoscopic findings predicted laryngoscopic or symptomatic response. Another study suggested that baseline anxiety levels and heartburn scores and medication dose might be relevant factors in predicting faster response to PPI treatment in carefully selected patients.[26] In our data, different age groups had different proportions of responders as evaluated by the RSI. The response rate in the oldest patients was significantly lower than in other age groups.

Few published articles have investigated PPI resistance in LPR. Amin et al[27] suggested that incomplete suppression might result from a shorter duration of drug action in unresponsive patients, possibly through increased metabolism of the PPI by the liver. Another explanation for poor response to PPI therapy is low bioavailability of the drug. Ashida et al[28] suggested that decreased plasma levels of PPI in patients with resistant gastric ulcers were due to an increase in gastric emptying time. Several authors have showed that older adults have a significant decrease in the amplitude of peristaltic pressures.[29-31] This is associated with a higher prevalence of diabetes mellitus or rheumatological disorders, which may alter esophageal motility in older persons. Therefore, decreased acid clearance in geriatric patients might be a possible cause of decreased response to PPI therapy.

Limitations of the present study include the lack of a placebo group as control. Moreover, we did not demonstrate the reflux events by means of multichannel impedance or pH monitoring studies. Although the gold standard diagnostic method for LPR is dual-probe 24-hour pH monitoring, it is an invasive test with a high false-negative rate.[21] Also, LPR is a fluctuating condition and there can be substantial day-to-day variation of acid exposure in the hypopharynx.[32] However, the response to PPI therapy in patients with suspected LPR is usually so explicit that empirical PPI therapy in LPR is recommended by both gastroenterology and otolaryngology experts and guidelines.[19,33,34] The present study might have meaningful implications for the difference in the effects of PPI therapy according to age.

Copyright 2013 American Medical Association. All rights reserved.

Although there was no significant difference in the objective findings among the different age groups, the subjective severity of LPR in geriatric patients is significantly greater than in younger patients. Furthermore, older patients are more likely not to respond to PPI therapy than younger patients.

ARTICLE INFORMATION

Submitted for Publication: July 2, 2013; final revision received August 18, 2013; accepted September 7, 2013.

Published Online: November 7, 2013. doi:10.1001/jamaoto.2013.5556.

Author Contributions: Dr Eun had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Y. C. Lee, Kim, Kwon, Eun.
Acquisition of data: Y. C. Lee, J. S. Lee, Kim, Kwon, Eun.
Analysis and interpretation of data: Y. C. Lee, J. S. Lee, Eun.
Drafting of the manuscript: Y. C. Lee, J. S. Lee.
Critical revision of the manuscript for important intellectual content: Kim, Kwon, Eun.
Statistical analysis: Y. C. Lee, J. S. Lee.
Study supervision: Kim, Kwon, Eun.

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Wiener GJ, Koufman JA, Wu WC, Cooper JB, Richter JE, Castell DO. Chronic hoarseness secondary to gastroesophageal reflux disease: documentation with 24-h ambulatory pH monitoring. *Am J Gastroenterol*. 1989;84(12):1503-1508.

2. Fraser AG. Review article: gastro-oesophageal reflux and laryngeal symptoms. *Aliment Pharmacol Ther*. 1994;8(3):265-272.

3. Rees CJ, Belafsky PC. Laryngopharyngeal reflux: current concepts in pathophysiology, diagnosis, and treatment. *Int J Speech Lang Pathol*. 2008;10(4):245-253.

4. Koufman J, Sataloff RT, Toohill R. Laryngopharyngeal reflux: consensus conference report. *J Voice*. 1996;10(3):215-216.

5. Collen MJ, Abdulian JD, Chen YK. Gastroesophageal reflux disease in the elderly: more severe disease that requires aggressive therapy. *Am J Gastroenterol*. 1995;90(7):1053-1057.

6. Brunnen PL, Karmody AM, Needham CD. Severe peptic oesophagitis. *Gut*. 1969;10(10):831-837.

7. el-Serag HB, Sonnenberg A. Associations between different forms of gastro-oesophageal reflux disease. *Gut*. 1997;41(5):594-599.

8. Garrigues V, Gisbert L, Bastida G, et al. Manifestations of gastroesophageal reflux and response to omeprazole therapy in patients with chronic posterior laryngitis: an evaluation based on clinical practice. *Dig Dis Sci*. 2003;48(11):2117-2123.

9. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). *J Voice*. 2002;16(2):274-277.

10. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope*. 2001;111(8):1313-1317.

11. Carrau RL, Khidr A, Gold KF, et al. Validation of a quality-of-life instrument for laryngopharyngeal reflux. *Arch Otolaryngol Head Neck Surg*. 2005;131(4):315-320.

12. Reulbach TR, Belafsky PC, Blalock PD, Koufman JA, Postma GN. Occult laryngeal pathology in a community-based cohort. *Otolaryngol Head Neck Surg*. 2001;124(4):448-450.

13. Locke GR III, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ III. Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota. *Gastroenterology*. 1997;112(5):1448-1456.

14. Räihä IJ, Impivaara O, Seppälä M, Sourander LB. Prevalence and characteristics of symptomatic gastroesophageal reflux disease in the elderly. *J Am Geriatr Soc*. 1992;40(12):1209-1211.

15. Castell DO. Esophageal disorders in the elderly. *Gastroenterol Clin North Am*. 1990;19(2):235-254.

16. Saruç M, Aksoy EA, Vardereli E, et al. Risk factors for laryngopharyngeal reflux. *Eur Arch Otorhinolaryngol*. 2012;269(4):1189-1194.

17. Siupsinskiene N, Adamonis K, Toohill RJ. Quality of life in laryngopharyngeal reflux patients. *Laryngoscope*. 2007;117(3):480-484.

18. Ford CN. Evaluation and management of laryngopharyngeal reflux. *JAMA*. 2005;294(12):1534-1540.

19. Koufman JA, Aviv JE, Casiano RR, Shaw GY. Laryngopharyngeal reflux: position statement of the committee on speech, voice, and swallowing disorders of the American Academy of Otolaryngology–Head and Neck Surgery. *Otolaryngol Head Neck Surg*. 2002;127(1):32-35.

20. Vaezi MF. Reflux-induced laryngitis (laryngopharyngeal reflux). *Curr Treat Options Gastroenterol*. 2006;9(1):69-74.

21. Karkos PD, Wilson JA. Empiric treatment of laryngopharyngeal reflux with proton pump inhibitors: a systematic review. *Laryngoscope*. 2006;116(1):144-148.

22. Ulualp SO, Toohill RJ, Shaker R. Outcomes of acid suppressive therapy in patients with posterior laryngitis. *Otolaryngol Head Neck Surg*. 2001;124(1):16-22.

23. Shin MH, Nam SY, Park YH, Son YI. Open-label observational study for evaluating the short-term benefits of rabeprazole medication on laryngopharyngeal reflux. *Clin Exp Otorhinolaryngol*. 2012;5(1):28-33.

24. Park W, Hicks DM, Khandwala F, et al. Laryngopharyngeal reflux: prospective cohort study evaluating optimal dose of proton-pump inhibitor therapy and pretherapy predictors of response. *Laryngoscope*. 2005;115(7):1230-1238.

25. Williams RB, Szczesniak MM, Maclean JC, Brake HM, Cole IE, Cook IJ. Predictors of outcome in an open label, therapeutic trial of high-dose omeprazole in laryngitis. *Am J Gastroenterol*. 2004;99(5):777-785.

26. Siupsinskiene N, Adamonis K, Toohill RJ, Sereika R. Predictors of response to short-term proton pump inhibitor treatment in laryngopharyngeal reflux patients. *J Laryngol Otol*. 2008;122(11):1206-1212.

27. Amin MR, Postma GN, Johnson P, Digges N, Koufman JA. Proton pump inhibitor resistance in the treatment of laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2001;125(4):374-378.

28. Ashida K, Sakaguchi M, Tanaka M, Takiuchi H, Egashira Y, Katsu K. Clinical study on the pathophysiology and treatment of PPI-resistant ulcers. *J Clin Gastroenterol*. 1995;20(suppl 2):S67-S71.

29. Hollis JB, Castell DO. Esophageal function in elderly men: a new look at "presbyesophagus." *Ann Intern Med*. 1974;(3):371-374.

30. Soergel KH, Zboralske FF, Amberg JR. Presbyesophagus: esophageal motility in nonagenarians. *J Clin Invest*. 1964;43:1472-1479.

31. Khan TA, Shragge BW, Crispin JS, Lind JF. Esophageal motility in the elderly. *Am J Dig Dis*. 1977;22(12):1049-1054.

32. Cheung TK, Lam PK, Wei WI, et al. Quality of life in patients with laryngopharyngeal reflux. *Digestion*. 2009;79(1):52-57.

33. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al; American Gastroenterological Association. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. *Gastroenterology*. 2008;135:1383-1391.e5.

34. Altman KW, Prufer N, Vaezi MF. A review of clinical practice guidelines for reflux disease: toward creating a clinical protocol for the otolaryngologist. *Laryngoscope*. 2011;121(4):717-723.

Copyright 2013 American Medical Association. All rights reserved.

Reprinted by permission of J Voice. 2015; 29(6):755-762.

Vocal Fold Vibration in Vocal Fold Atrophy: Quantitative Analysis With High-Speed Digital Imaging

***Akihito Yamauchi, †Hisayuki Yokonishi, *Hiroshi Imagawa, ‡Ken-Ichi Sakakibara, *Takaharu Nito, §Niro Tayama, and *Tatsuya Yamasoba,** *†§*Tokyo and* ‡*Hokkaido, Japan*

Summary: Introduction. Vocal fold vibrations of vocal fold atrophy (VFA), a rapidly increasing voice disorder owing to worldwide societal aging, have not been clarified by high-speed digital imaging (HSDI).
Methods. The HSDI method was performed on 46 patients (33 males and 13 females) with VFA and 20 vocally healthy subjects (8 males and 12 females), and the obtained data were quantitatively evaluated by frame-by-frame analysis, laryngotopography, single- and multi-line kymography, and glottal area waveform.
Results. Overall, patients with VFA revealed larger open quotients, larger lateral phase difference, larger integral glottal width (the average glottal width over a glottal cycle), and smaller speed index than vocally healthy subjects. Some gender difference was noted: in males, lateral phase difference was not significant; and in females, integral glottal width and speed index were not significant. Correlation study revealed moderate correlations between HSDI-derived parameters and conventional acoustic or aerodynamic parameters.
Conclusions. The combination of multiple HSDI analysis methods was effective in documenting the characteristics of vocal fold vibrations in VFA. The knowledge of general vibratory characteristics and gender difference is beneficial for the appropriate clinical care of VFA.
Key Words: Vocal fold atrophy–Presbyphonia–High-speed digital imaging–Aging–Anti-aging.

INTRODUCTION

Vocal fold atrophy (VFA) is a voice disorder resulting from the atrophied muscle and mucosa in the vocal folds.[1] These structural modifications lead to increased glottal air leakage and breathy, rough voice. Aging is considered to be the most major predisposing factor for VFA, although other risk factors have also been proposed (eg, reflux laryngitis, chronic medical conditions, and vocal abuse).[2] The VFA has increased considerably during the past two decades as a result of the worldwide societal aging, and thus, is attracting clinical attention in the world these days.[3,4]

Laryngoscopically, the vocal fold bowing, prominent vocal process, and spindle-shaped glottal gap are usually observed.[1,2] The vibratory characteristics observed with videostroboscopy include normal or decreased amplitude, either complete closure or glottal gap, normal mucosal wave, and small supraglottal area.[1,2,5,6]

The details of vocal fold vibrations in VFA, however, have not yet been documented by high-speed digital imaging (HSDI), although HSDI is considered to be the better choice than videostroboscopy.[7,8] First, HSDI is capable of observing actual vocal fold vibrations with a high frame rate, and guarantees reliable assessment of intra- and intercycle vibratory behaviors, unlike videostroboscopy that only provides reconstructed, averaged, illusory images. Second, HSDI offers wider application to clinical cases than videostroboscopy because HSDI is free from the problem with synchronization and is applicable to severe dysphonia in which videostroboscopy results in desynchronization. Third, various analysis methods for HSDI are now available, and thus HSDI provides more multifaceted information than videostroboscopy that has relatively limited choices of analysis methods. Furthermore, only little is known about the association between vibratory parameters and acoustic or aerodynamic parameters in VFA,[5] and HSDI data have not been reported on this matter. The connection between HSDI parameters and routinely evaluated vocal function parameters in VFA should be beneficial for better understanding the pathophysiological aspects of this clinical entity.

Hence, the purpose of the present study was to quantitatively elucidate the vibratory characteristics in VFA patients using HSDI, and to clarify the relationship between HSDI parameters and aerodynamic/acoustic measures.

MATERIALS AND METHODS

Subjects

Patients who visited the Voice Outpatient Clinic of the Department of Otolaryngology and Head and Neck Surgery at the University of Tokyo Hospital (Tokyo, Japan) and those who were diagnosed with VFA between 2006 and 2013 were included in this study. The diagnosis of VFA was based on careful history taking, acoustic and aerodynamic evaluation, and laryngostroboscopic findings: Patients with objective dysphonia on acoustic or aerodynamic studies; without signs of other laryngeal diseases such as vocal fold paralysis, vocal fold polyp, laryngeal carcinoma, vocal fold scar, or functional dysphonia; and with the prominent vocal process, bowed vocal fold, spindle-shaped or anterior glottal gap, or increased open phase during phonation

Accepted for publication December 15, 2014.
This research was presented in the Seventh East Asian Conference on Phonosurgery; November 26–27, 2010; Tokyo, Japan.
From the *Department of Otolaryngology, The University of Tokyo Hospital, Tokyo, Japan; †Department of Otolaryngology, The Tokyo Metropolitan Police Hospital, Tokyo, Japan; ‡Department of Communication Disorders, The Health Sciences University of Hokkaido, Hokkaido, Japan; and the §Department of Otolaryngology and Tracheo-Esophagology, The National Center for Global Health and Medicine, Tokyo, Japan.
Address correspondence and reprint requests to Akihito Yamauchi, Department of Otolaryngology, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655 Japan. E-mail: drachilles23@yahoo.co.jp
Journal of Voice, Vol. ■, No. ■, pp. 1-8
0892-1997/$36.00
© 2015 The Voice Foundation
http://dx.doi.org/10.1016/j.jvoice.2014.12.008

were diagnosed as those having VFA. The diagnosis was made by three or four certified otorhinolaryngologists specializing in vocal treatment. As for sulcus vocalis, patients with a type 1 sulcus (physiological sulcus) were included in the category of VFA because type 1 sulcus is superficial and generally causes no or minimal functional vocal impairment.[9] Furthermore, type 1 sulcus is considered to be associated with aging.[9] On the other hand, patients with a type 2 sulcus (sulcus vergeture) or a type 3 sulcus (true sulcus vocalis) were excluded from this study.[9]

As a control group, vocally healthy subjects without vocal complaints, history of laryngeal disorders, or signs of laryngeal abnormality with laryngoendoscopy were recruited. As an exception, however, a small glottal gap was permitted for a control group in the present study because vocally healthy elderly population is known to demonstrate a small glottal gap frequently.[10]

All subjects were required to sign a consent form that was approved by our Institutional Review Board. A total of 46 patients with VFA (13 women and 33 men), with the age range between 60 and 91 years, and 20 vocally healthy subjects (12 women and eight men), with the age range between 65 and 81 years, were enrolled in the present study.

TABLE 1.
Clinical Data of All Participants

Parameter (Units)	Control (N = 20)	VFA (N = 46)	*t* Test
Age (yr)	73 ± 5	72 ± 7	0.513
MPT (s)	18.7 ± 6.6	16.4 ± 8.2	0.284
MFR (mL/s)	136 ± 36	210 ± 101	0.002†
AA-F_0 (Hz)	178 ± 50	179 ± 63	0.940
APQ (%)	3.2 ± 1.5	4.1 ± 2.3	0.135
PPQ (%)	0.31 ± 0.46	0.92 ± 0.61	<0.001‡
HNR (dB)	21.5 ± 3.5	12.6 ± 4.9	<0.001‡
Grade	0.80 ± 0.62	1.33 ± 0.52	<0.001‡
Roughness	0.80 ± 0.62	1.17 ± 0.44	0.006†
Breathiness	0.50 ± 0.51	0.65 ± 0.60	0.330

Abbreviations: SD, standard deviation; VFA, vocal fold atrophy; MPT, maximum phonation time; MFR, mean flow rate; AA-F_0, fundamental frequency in acoustic analysis; APQ, amplitude perturbation quotient; PPQ, period perturbation quotient; HNR, harmonics-to-noise ratio.
Notes: Values signify "mean ± SD." The column for *t* test shows the *P* value of Student's *t* test between control and VFA groups.
† $P < 0.01$.
‡ $P < 0.001$.

Background data

Vocal function and voice quality were evaluated by measuring aerodynamic and acoustic parameters. The aerodynamic parameters including the maximum phonation time and mean flow rate were measured with a Nagashima PE-77E Phonatory Function Analyzer (Nagashima Medical, Inc., Tokyo, Japan). Acoustic parameters included the fundamental frequency (AA–F_0), amplitude perturbation quotient, period perturbation quotient, and harmonic-to-noise ratio, which were measured at the University of Tokyo with a dedicated software program, as well as the subjective rating by the GRBAS scale.

Table 1 summarizes the results of aerodynamic and acoustic studies, in which mean flow rate, period perturbation quotient, harmonics-to-noise ratio, and the grade and roughness of the GRBAS scale revealed significant intergroup differences. The scores of Voice Handicap Index-10 and the voice-related quality of life were 13.7 ± 9.4 and 13.8 ± 8.9, respectively, and the rate of synchronization in VFA with videostroboscopy (LS-3A; Nagashima Medical, Inc.) was 67.3%.

High-speed digital imaging

A high-speed digital camera (FASTCAM-1024PCI; Photron, Tokyo, Japan) was connected to a rigid endoscope (#4450.501; Richard Wolf, Vernon Hills, IL) via an attachment lens ($f = 35$ mm; Nagashima Medical, Inc.). Recording was performed under illumination with a 300-W xenon light source at a frame rate of 4500 frames per second and a spatial resolution of 512 × 400 pixels, with an 8-bit grayscale and a recording duration of 1.86 seconds. High-speed digital images of sustained vowel phonation /i/ at a comfortable frequency with a comfortable intensity were recorded. The image sequence of stable vocal fold vibrations were selected for further analyses.

Aerodynamic and acoustic studies were performed approximately 30 minutes before HSDI recording because simultaneous recording was not available at our institution. Both evaluations were done under as similar conditions as possible to allow comparison between the HSDI parameters and the aerodynamic or acoustic parameters.

HSDI analysis methods

The recorded HSDIs were evaluated by frame-by-frame analysis,[11] laryngotopography (LTG),[12] single-/multi-line digital kymography (SLK and MLK, respectively),[13,14] and glottal area waveform (GAW).[15] The details of analysis by these methods are described elsewhere.[11–15]

The size parameters were normalized by the vocal fold length, labeled by "N_L-" (eg, V_L-amplitude mean). The time parameters were normalized by the glottal cycle, labeled by "N_G-" (eg, N_G-lateral phase difference). The size and time parameters were normalized by both glottal cycle and vocal fold length, labeled by "N_{GL}-" (eg, N_{GL}-lateral phase difference).[13]

In the present study, analysis was focused on selected parameters that were considered to be related with the vibratory characteristics of VFA such as amplitude, mucosal wave, lateral/longitudinal phase difference, open quotient, speed index, integral glottal width (the average glottal width over a glottal cycle),[13] maximal/minimal glottal area, glottal area difference, and glottal outlet (normalized supraglottal area).[6]

Frame-by-frame analysis was performed using an assessment form for HSDI developed by the authors, with which vibratory parameters such as symmetry, periodicity, amplitude, mucosal wave, phase difference, glottal closure, and supraglottal hyperactivity were evaluated by two- or four-point scale.[11] For glottal gaps, the incidence (present or absent) and glottal type (incomplete closure, posterior, spindle-shaped, or anterior) were evaluated.

The LTG is a method using a pixel-wise Fourier transform of time-varying brightness curve for each pixel across images and

allows quantitative evaluation of spatial characteristics of frequency and phase. In the present study, the incidence of diplophonia and lateral/longitudinal phase difference normalized by glottal cycle (N_G-lateral/longitudinal phase difference) were evaluated.[12] Image sequence of 512 frames was evaluated.

The SLK analyzes mediolateral vocal fold movements at a midglottal level and allows evaluation of mediolateral and temporal vibratory characteristics such as amplitude (N_L-amplitude mean), mucosal wave (N_L-mucosal wave magnitude mean), phase (N_G-lateral phase difference and N_G-longitudinal phase difference), open quotient (O_q^{SLK}), speed index (SI^{SLK}), and integral glottal width (N_{GL}-integral glottal width).[13] Image sequence of 400 frames was evaluated.

The MLK involves five different longitudinal levels, and can assess temporal and longitudinal oscillatory features such as open quotient (O_q^{MLK}), speed index (SI^{MLK}), and opening/closing longitudinal phase difference (N_G-opening/closing longitudinal phase difference).[14] Image sequence of 400 frames was evaluated.

The GAW provides information on the general dynamics of the glottal area by tracing the vocal fold edges and displaying temporal changes of the glottal area, with which open quotient (O_q^{GAW}), speed index (SI^{GAW}), minimal glottal area (N_L-minimal glottal area), maximal glottal area (N_L-maximal glottal area), glottal outlet (N_L-glottal outlet), and glottal area difference (glottal area difference index = (N_L-maximal glottal area − N_L-minimal glottal area)/N_L-maximal glottal area) can be evaluated.[15] Five consecutive glottal cycles of the most stable segment were evaluated (100–200 frames).

All the HSDI analyses were performed with custom *MATLAB* software (Version 2011a; Mathworks, Inc., Natick, MA) programmed at our institution . An example of HSDI analysis is shown in Figure 1.

Statistics

The difference of clinical and HSDI parameters between VFA patients and vocally healthy subjects were evaluated by Student's *t* test for normally distributed parameters, or either by the Mann-Whitney *U* test or by chi-squared test for other parameters. To assess the correlations with HSDI parameters and reference data (demographic, aerodynamic/acoustic, or stroboscopic data), or the correlations among HSDI parameters, Pearson's correlation analysis for normally distributed parameters and Spearman's correlation analysis for the other parameters were used. In all analyses, *P* value lower than 0.05 was considered significant. Calculations were performed with a custom *MATLAB* software.

RESULTS

HSDI parameters in the VFA

Frame-by-frame analysis. The glottal gap was observed in 45% of the VFA group and 30% of the control group ($P = 0.235$), and the most frequent type was spindle shaped for the VFA group (47.8% of those with glottal gaps) and anterior for the control group (66.7% of those with glottal gaps). Although no parameters revealed significant differences between the two groups, the VFA group tended to demonstrate

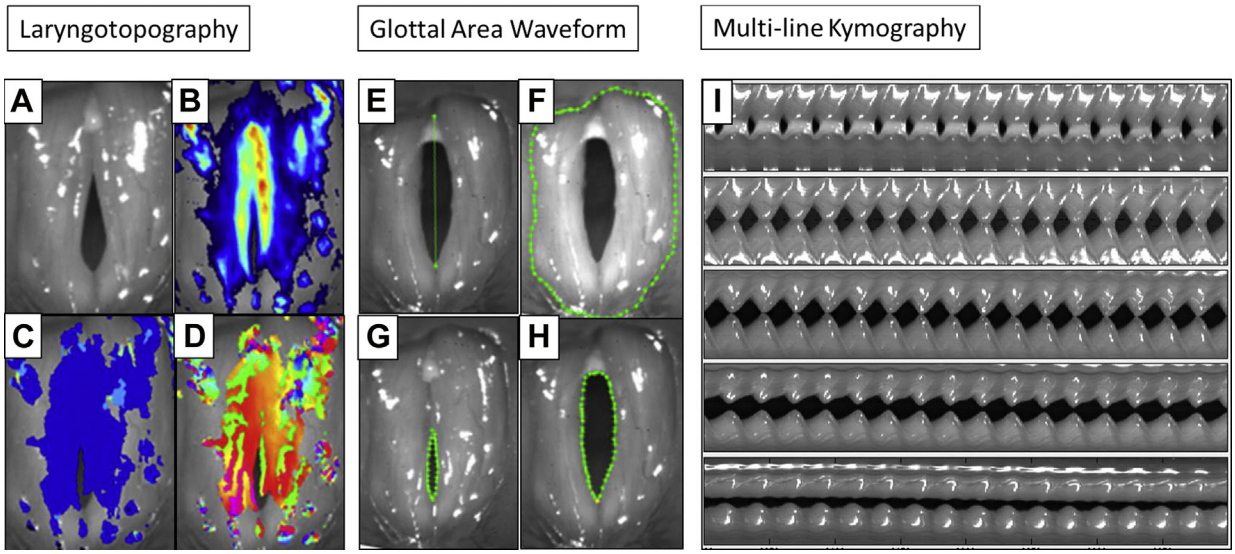

FIGURE 1. An example of the analysis of high-speed digital image is shown. (**A**–**D**) Laryngotopography is shown. (**A**) A static laryngeal image to be superimposed by analyzed topographic data. (**B**–**D**) Spatial distribution of amplitude, frequency, and phase of the maximum amplitude components, respectively. This 71-year-old female patient has a topographic F0 of 225 Hz, left-to-right lateral phase difference (6.3% of a glottal cycle), and anterior-to-posterior longitudinal phase difference (21.9% of a glottal cycle). (**E**–**H**) Glottal area waveform is shown: areas demarcated by a green line in Panels (**E**–**H**) show a minimal glottal area, a maximal glottal area, and a glottal outlet, respectively; and a green line in Panel (**G**) signifies the vocal fold length. This patient has an anterior glottal gap with a N_L-minimal glottal area of 1.2%, a spindle-shaped maximal glottal area with a N_L-maximal glottal area of 11.5%, and no supraglottal hyperfunction with N_L-glottal outlet of 100.2%. (**I**) Multi-line kymography is shown. There are anterior-to-posterior opening and posterior-to-anterior closing longitudinal phase differences (23.8% and 19.0% of a glottal cycle, respectively). O_q^{MLK} and SI^{MLK} are 0.86 and −0.21, respectively.

TABLE 2.
Comparisons of High-Speed Digital Image Parameters Between the Control and VFA Groups

Parameter (Units)	Control (N = 20)	VFA (N = 46)	*t* Test
Laryngotopography			
N_G-lateral phase difference (%)	3.8 ± 4.3	6.3 ± 4.3	0.034*
N_G-longitudinal phase difference (%)	−14.4 ± 16.6	−8.9 ± 13.6	0.175
Single-line digital kymography			
N_L-amplitude mean (%)	8.1 ± 2.7	8.6 ± 3.4	0.544
N_L-mucosal wave magnitude mean (%)	18.6 ± 9.3	16.2 ± 7.9	0.321
O_q^{SLK}	0.57 ± 0.13	0.76 ± 0.17	<0.001†
SI^{SLK}	−0.19 ± 0.18	−0.18 ± 0.18	0.906
N_G-lateral phase difference (%)	9.9 ± 6.5	8.9 ± 10.4	0.689
N_{GL}-integral glottal width (%)	4.5 ± 1.8	7.0 ± 3.5	0.004‡
Multi-line digital kymography			
O_q^{MLK}	0.45 ± 0.11	0.66 ± 0.20	<0.001†
SI^{MLK}	−0.24 ± 0.16	−0.25 ± 0.14	0.908
N_G-opening longitudinal phase difference (%)	−7.4 ± 19.2	−7.3 ± 22.6	0.975
N_G-closing longitudinal phase difference (%)	2.4 ± 19.9	1.6 ± 11.0	0.860
Glottal area waveform			
O_q^{GAW}	0.80 ± 0.18	0.84 ± 0.17	0.908
SI^{GAW}	0.11 ± 0.23	−0.06 ± 0.17	0.002‡
N_L-maximal glottal area (%)	8.8 ± 3.1	16.7 ± 20.3	0.089
N_L-minimal glottal area (%)	0.18 ± 0.53	0.63 ± 1.18	0.104
Glottal area difference index (%)	98.1 ± 5.0	94.1 ± 11.1	0.133
N_L-glottal outlet (%)	61.8 ± 23.6	58.5 ± 21.9	0.592

Abbreviations: VFA, vocal fold atrophy; N_G-, normalized by glottal cycle; N_L-, normalized by vocal fold length; Oq, open quotient; SLK, single-line kymography; SI, speed index; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, multi-line kymography; GAW, glottal area waveform.
Notes: Values for control and VFA columns show "mean ± standard deviation," and the value of *t* test column shows the *P* value of Student's *t* test between all control and VFA groups.
* $P < 0.05$.
† $P < 0.001$.
‡ $P < 0.01$.

smaller amplitude, smaller mucosal wave, greater asymmetry, larger glottal gap, and greater supraglottal hyperactivity than the control group.

Laryngotopography. Three of the patients with VFA (4.9%) revealed diplophonia almost constantly at a comfortable pitch and sound pressure level. The VFA group had larger N_G-lateral phase difference than the control group ($P = 0.034$; Table 2). However, by gender, a significant difference was found only in females (Tables 3 and 4).

Single-line digital kymography. With SLK, patients with VFA revealed significantly greater O_q^{SLK} ($P < 0.001$) and N_{GL}-integral glottal width ($P = 0.004$; Table 2). Although O_q^{SLK}

TABLE 3.
Comparisons of High-Speed Digital Image Parameters Between the Control and VFA Groups in Females

Parameter (Units)	Control (N = 12)	VFA (N = 13)	*t* Test
Laryngotopography			
N_G-lateral phase difference (%)	3.1 ± 4.2	7.3 ± 3.6	0.016*
Single-line digital kymography			
O_q^{SLK}	0.58 ± 0.14	0.77 ± 0.14	0.003†
N_{GL}-integral glottal width (%)	4.4 ± 1.6	5.2 ± 1.8	0.241
Multi-line digital kymography			
O_q^{MLK}	0.46 ± 0.13	0.69 ± 0.21	0.004†
Glottal area waveform			
SI^{GAW}	0.076 ± 0.260	−0.032 ± 0.177	0.261

Abbreviations: VFA, vocal fold atrophy; N_G-, normalized (by glottal cycle); O_q, open quotient; SLK, single-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, multi-line kymography; SI, speed index; GAW, glottal area waveform.
Notes: Values for control and VFA columns show "mean ± standard deviation," and the *t* test column shows the *P* value of Student's *t* test between female control and female VFA groups.
* $P < 0.05$.
† $P < 0.01$.

TABLE 4.
Correlation Coefficients (r) Between High-Speed Digital Image Parameters and Selected Aerodynamic/Acoustic Measures

Parameters	MPT	MFR	AA-F_0	PPQ	HNR	G	R	B
N_L-amplitude mean	−0.06	0.27*	−0.40†	−0.07	0.10	−0.05	−0.05	−0.03
O_q^{SLK}	−0.29†	0.37†	0.02	0.47‡	−0.35†	0.29*	0.29*	0.15
N_{GL}-integral glottal width	−0.20	0.39†	−0.30*	0.19	−0.15	0.18	0.18	0.02
O_q^{MLK}	−0.34†	0.49‡	−0.16	0.34†	−0.31*	0.26*	0.26*	0.13
O_q^{GAW}	−0.28*	0.28*	0.14	0.29*	−0.04	0.24*	0.24*	0.23
SI^{GAW}	0.20	−0.36*	−0.13	−0.28*	0.18	−0.32†	−0.32†	−0.36†
N_L-minimal glottal area	−0.22	0.53‡	−0.43‡	0.12	−0.10	0.20	0.20	0.21
Glottal area difference index	0.17	−0.37†	−0.27*	−0.10	0.12	−0.13	−0.13	−0.09
N_L-glottal outlet	0.37†	−0.11	0.01	−0.09	−0.02	−0.08	−0.08	−0.17

Abbreviations: MPT, maximum phonation time; MFR, mean flow rate; AA-F0, fundamental frequency of acoustic analysis; PPQ, period perturbation quotient; HNR, harmonics-to-noise ratio; G, grade; R, roughness; B, breathiness; NL-, normalized by vocal fold length; O_q, open quotient; SLK, in single-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, in multi-line kymography; GAW, in glottal area waveform; SI, speed index.
* $P < 0.05$.
† $P < 0.01$.
‡ $P < 0.001$.

demonstrated significant differences in both gender, N_{GL}-integral glottal width demonstrated a significant difference only in males (Tables 3 and 5). There were no significant differences in N_L-amplitude mean or N_L-mucosal wave magnitude mean. Although the differences were not statistically significant, N_L-amplitude mean of VFA tended to be greater in males and smaller in females, and N_L-mucosal wave magnitude mean of VFA tended to be smaller in both genders.

Multi-line digital kymography. With MLK, patients with VFA revealed significantly larger O_q^{MLK} ($P < 0.001$; Table 2), and this difference was significant in both genders (Tables 3 and 5).

Glottal area waveform. The VFA group demonstrated significantly smaller SI^{GAW} ($P < 0.001$; Table 2). Although the differences were not significant, the VFA group revealed larger N_L-maximal glottal area and N_L-minimal glottal area (Table 2).

Correlation study

Table 4 summarizes the correlation study between HSDI parameters and background data. Moderate correlations ($0.4 < |r| \leqq 0.7$) were found between mean flow rate and O_q^{MLK}, mean flow rate and N_L-minimal glottal area, AA-F_0 and N_L-amplitude mean, AA-F_0 and N_L-minimal glottal area, and O_q^{SLK} and period perturbation quotient.

Table 6 summarizes the correlation study among the HSDI parameters. A strong correlation ($0.7 < |r|$) was found between N_L-minimal glottal area and glottal area difference index. Otherwise, moderate correlations ($0.4 < |r| \leqq 0.7$) were found in multiple pairs.

DISCUSSION

Videostroboscopy versus HSDI

The use of HSDI offered several advantages to the present study over the previous studies in the literature. First, HSDI enabled the evaluation of time parameters that cannot be assessed by

TABLE 5.
Comparisons of High-Speed Digital Image Parameters Between the Control and VFA Groups in Males were summarized

Parameter (Units)	Control (N = 8)	VFA (N = 33)	*t* Test
Laryngotopography			
N_G-lateral phase difference (%)	4.7 ± 4.4	5.9 ± 4.4	0.511
Single-line digital kymography			
O_q^{SLK}	0.54 ± 0.12	0.75 ± 0.19	0.006*
N_{GL}-integral glottal width (%)	4.6 ± 2.3	7.8 ± 3.8	0.033†
Multi-line digital kymography			
O_q^{MLK}	0.42 ± 0.08	0.65 ± 0.20	0.003*
Glottal area waveform			
SI^{GAW}	0.16 ± 0.17	−0.067 ± 0.166	0.002*

Abbreviations: VFA, vocal fold atrophy; N_G-, normalized (by glottal cycle); Oq, open quotient; SLK, single-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, multi-line kymography; SI, speed index; GAW, glottal area waveform.
Notes: Values for control and VFA columns show "mean ± standard deviation," and the *t* test column shows the *P* value of Student's *t* test between male control and male VFA groups.
* $P < 0.01$.
† $P < 0.05$.

TABLE 6.
Correlation Coefficients (r) Among High-Speed Digital Image Parameters

Parameters	N_L- Minimal GA	SI^{GAW}	$O_q{}^{MLK}$	$O_q{}^{SLK}$	N_L-Amplitude Mean	N_G-O-LPD^{MLK}
GA difference index	−0.90*	−0.46*	−0.65*	−0.54*	−0.34†	−0.15
N_L-minimal GA	1	−0.33†	0.69*	0.47*	−0.24	0.05
SI^{GAW}	—	1	−0.47*	−0.56*	−0.04	−0.39*
$O_q{}^{GAW}$	—	—	0.69*	0.42*	−0.10	0.10
N_G-O-LPD^{LTG}	—	—	0.14	0.19	−0.03	0.50*
$O_q{}^{MLK}$	—	—	1	0.65*	−0.13	0.19
N_{GL}-integral glottal width	—	—	—	0.51*	0.51*	0.28
N_L-MW magnitude mean	—	—	—	—	0.59*	−0.03

Abbreviations: GA, glottal area; N_L-, normalized by vocal fold length; SI, speed index; GAW, in glottal area waveform; O_q, open quotient; N_G-, normalized by glottal cycle; O-LPD, opening longitudinal phase difference; LTG, in laryngotopography; MLK, in multi-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MW, mucosal wave; SLK, in single-line kymography.
* $P < 0.001$.
† $P < 0.01$.

videostroboscopy, which is considered to be meaningful especially because a large part of vibratory characteristics of VFA were reflected in the time parameters (eg, open quotient).

Second, with the use of HSDI, the rate of successful image evaluation increased by 1.5 folds at a rough estimate because videostroboscopic study was successful only in 67.3% because of desynchronization in the present study. The relatively high rate of desynchronization in VFA may be explained by their poor acoustic profile. Patel et al[8] reported that HSDI could be used to augment videostroboscopy for assessment of moderate-to-severe dysphonia, especially in patients with jitter exceeding 0.87%, shimmer exceeding 4.4%, and a signal-to-noise ratio of less than 15.4 dB. In the present study, 21.7% of the VFA group fitted these criteria.

Third, with the application of multiple analysis methods, the present study documented the characteristics of vocal fold vibrations of VFA more extensively and multidirectionally than previous reports, in which vocal fold vibrations were either qualitatively evaluated or quantitatively evaluated with only limited parameters.[5,6]

Although HSDI has disadvantages in comparison with videostroboscopy such as a relatively long time required for analysis (approximately 30 minutes per HSDI at present), a high cost, and the lack of instantaneity (with videostroboscopy, the result of modulation in F_0, sound pressure level, or register can be observed directly and instantaneously),[7,8,16] HSDI is considered to be a good supplementary tool in the assessment of VFA.

Amplitude and integral glottal width

The amplitude mean of VFA was comparable with that of vocally healthy subjects, in this study, which was a consistent result with the previous study.[5] Although not statistically significant, the amplitude of VFA was larger in males and smaller in females than vocally healthy subjects. Various factors can affect amplitude such as amplitude increases as intensity or subglottal pressure increase, or as pitch or stiffness decrease.[6,9,16,17] In female VFA, poor pulmonary function is reported to be frequently associated,[2] which may lead to decreased subglottal pressure and decreased amplitude. In male VFA, a greater glottal flow and lower tension of the thyroarytenoid muscle owing to the muscular atrophy can increase the amplitude.[6,18,19]

On the other hand, N_{GL}-integral glottal width demonstrated a significant difference between the control and VFA groups. The N_{GL}-integral glottal width may be a sensitive parameter than the amplitude *per se* because it has the characteristics of both amplitude and open quotient (Table 6).[13]

Open quotient and speed index

Significant intergroup differences were observed in $O_q{}^{SLK}$ and $O_q{}^{MLK}$ but not in $O_q{}^{GAW}$. This is probably because $O_q{}^{GAW}$ was not a parameter to reflect the size of glottal gap ($O_q{}^{GAW}$ becomes one whether a glottal gap is small or large). Interestingly, the results of $O_q{}^{SLK}$ were comparable with those of $O_q{}^{MLK}$, although $O_q{}^{MLK}$ that assesses the overall glottal area should reflect the pathophysiology of the disease better than $O_q{}^{SLK}$. Perhaps, the midglottal level may represent the vibratory dynamics of overall glottis well enough in VFA, and the information of the glottal ends included in $O_q{}^{MLK}$ may have been less important. Correlation analysis revealed that high $O_q{}^{SLK}$ and $O_q{}^{MLK}$ were associated with poor aerodynamic and acoustic conditions (Table 4). These results seem to stand to reason because weak glottal closure reflected in high open quotient should lead to high glottal flow with high air turbulence.

Speed index of VFA was smaller than that of vocally healthy subjects. Small speed index in the VFA group may originate from the decreased restorative force of the laterally displaced vocal fold toward the medial direction resulting from the disarrangement of collagen fibers or decreased elastin fibers in the lamina propria, the decreased mass or tension of the vocal fold owing to the muscular atrophy.[1,18,19] Contrary to open quotient, SI^{GAW} was more sensitive than SI^{SLK} or SI^{MLK}, probably because SI^{GAW} reflects the general vibratory dynamics than SI^{SLK} or SI^{MLK} (Table 2). Speed index had similar relationships with acoustic and aerodynamic parameters to open quotient (Table 4). These results accord with the findings in the literature, reporting that smaller speed index leads to poorer aerodynamic or acoustic results.[20,21]

Phase difference

With relatively larger N_G-lateral phase difference, vocal fold vibrations in patients with VFA were more asymmetrical than those of vocally healthy subjects. Left-right difference of mass, tension, mucoelasticity of the vocal fold resulting from a different degree of muscular atrophy, and muscular/mucosal degeneration as well as asymmetry of the laryngeal frame may play a role here.[1,18,19]

GAW parameters

The GAW parameters failed to reveal significant intergroup differences although N_L-minimal glottal area and N_L-maximal glottal area were larger, and glottal area difference index was smaller in the VFA group as a trend. This result was consistent with the study of Bloch and Behrman[6] that reported no significant difference in N_L-minimal glottal area between the control and VFA groups. Larger N_L-maximal glottal area found in the present study may be owing to an increased glottal flow in patients with VFA (Table 1), and decreased muscular tension of the vocal fold resulting from the muscular atrophy, leading to a greater lateral excursion of the vocal folds.[18] The smaller glottal area difference index observed in VFA signifies the decreased alternating current of glottal flow, the glottal flow efficiency in other words.

Glottal gap

The result that 30% of elderly vocally healthy subjects had a glottal gap in the present study was consistent with the findings in the literature: Pontes et al[10] reported that the incidence of glottal gap in normal elderly population was 58%, for instance. Strictly speaking, the vocally healthy subjects with a glottal gap in the present study (as well as those in the study of Pontes et al,[10] perhaps) should be termed as "pathological but asymptomatic" rather than "normal," though. Because the preponderant glottal gap was different between the control (anterior) and VFA groups (spindle shaped), the location of the glottal gap may serve as a clue to differentiate VFA from normal aging.

Glottal outlet

No significant intergroup difference of N_L-glottal outlet in the present study was a contradictory result to the report by Bloch and Behrman,[6] who reported significantly smaller N_L-glottal outlet in the VFA group than the normal group. One possible explanation is an interindividual difference of a maladaptive supraglottal hyperactivity as a compensatory strategy for incomplete or decreased glottal closure.[6] Another possible explanation is the posterior displacement of the petiole of epiglottis associated with a descension of the larynx observed in a male low-pitch phonation,[17] or an elevated laryngeal position observed in high-pitch phonation.[22]

Limitations

Overall, the combination of multiple HSDI analysis methods adopted in the present study was effective in the objective documentation of vocal fold vibrations in VFA. Applying the same technique to the evaluation of other laryngeal pathology (eg, vocal fold scar or sulcus vocalis) will be called for in the near future to further validate its utility.

The study design in which the HSDI study and acoustic or aerodynamic studies were performed on separate occasions may be a limitation of the present study, however. Although the effort was made to make the conditions of examination equal as much as possible, there could be a minor variation in F_0 or sound pressure level, leading to relatively low correlations between HSDI parameters and acoustic/aerodynamic parameters. Another limitation may be the use of a rigid endoscope for the HSDI recording, which could yield undesirable laryngeal tension during the study. The short time interval for HSDI analysis as well as the relatively limited subject number (especially of male vocally healthy subjects) may be other limitations. Furthermore, the heterogeneity in the VFA group in the present study may have existed, although the selection of recruited subjects and the diagnosis was based on the agreement of three or four certified otorhinolaryngologists specializing in vocal treatment: Because the differential diagnosis among VFA, sulcus vocalis, and vocal fold scar is not always clear-cut, there is inevitable room for subjectivity.

In the future study, the improvement of the study design by an introduction of simultaneous recording system of HSDI and acoustic signal or aerodynamic data, the introduction of transnasal flexible HSDI, the further refinement of analysis technique with more automation that allows much extended time interval for analysis, and the expansion of subject number will be warranted.

CONCLUSION

The quantitative HSDI analysis of VFA revealed larger open quotients, lateral phase difference and integral glottal width (the average glottal width over a glottal cycle), and smaller speed index than vocally healthy subjects. Gender difference was noted in lateral phase difference, integral glottal width, and speed index. Correlation study revealed mild-to-moderate correlations between HSDI-derived parameters and conventional acoustic or aerodynamic parameters, and moderate-to-strong correlation among HSDI parameters. The combination of multiple HSDI analysis methods was effective in the objective documentation of vocal fold vibrations in VFA.

Acknowledgments

This research was not funded by any organization or grant. There was no conflict of interest to be disclosed.

REFERENCES

1. Kendall K. Presbyphonia: a review. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:137–140.
2. Yamauchi A, Imagawa H, Sakakibara K-I, et al. Vocal fold atrophy in a Japanese tertiary medical institute: status quo of the most aged country. *J Voice*. 2014;28:231–236.
3. Takano S, Kimura M, Nito T, et al. Clinical analysis of presbylarynx-vocal fold atrophy in the elderly individuals. *Auris Nasus Larynx*. 2010;37:461–467.

4. Davids T, Klein AM, Johns MM. Current dysphonia trends in patients over the age of 65: is vocal atrophy becoming more prevalent? *Laryngoscope*. 2012;122:332–335.
5. Omori K, Slavit DH, Matos C, et al. Vocal fold atrophy: quantitative glottic measurement and vocal function. *Ann Otol Rhinol Laryngol*. 1997;106: 544–551.
6. Bloch I, Behrman A. Quantitative analysis of videostroboscopic images in presbylarynges. *Laryngoscope*. 2001;111:2022–2027.
7. Olthoff A, Woywod C, Kruse E. Stroboscopy versus high-speed glottography: a comparative study. *Laryngoscope*. 2007;117: 1123–1126.
8. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008;117:413–424.
9. Ford CN, Inagi K, Khidr A, et al. Sulcus vocalis: a rational analytical approach to diagnosis and management. *Ann Otol Rhinol Laryngol*. 1996;105:189–200.
10. Pontes P, Yamasaki R, Behlau M. Morphological and functional aspects of the senile larynx. *Folia Phoniatr Logop*. 2006;58:151–158.
11. Yamauchi A, Imagawa H, Yokonishi H, et al. Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: comparative study between healthy young and elderly subjects. *J Voice*. 2012;26: 742–750.
12. Yamauchi A, Imagawa H, Sakakibara K-I, et al. Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography. *J Voice*. 2013;27:39–45.
13. Yamauchi A, Yokonishi H, Imagawa H, et al. Quantitative analysis of digital videokymography: a preliminary study on age- and gender-related difference of vocal fold vibration in normal speakers. *J Voice*. 2014; [E-pub ahead of print].
14. Yamauchi A, Imagawa H, Sakakibara K-I, et al. Characteristics of vocal fold vibrations in vocally healthy subjects: analysis with multi-line kymography. *J Speech Lang Hear Res*. 2014;57:648–657.
15. Yamauchi A, Yokonishi H, Imagawa H, et al. Age- and gender-related difference of vocal fold vibration and glottal configuration in normal speakers: analysis with glottal area waveform. *J Voice*. 2014;28:525–531.
16. Woo P. Quantification of videostrobolaryngoscopic findings—measurements of the normal glottal cycle. *Laryngoscope*. 1996;106:1–27.
17. Timcke R, von Leden H, Moore P. Laryngeal vibrations: measurements of the glottic wave. *Arch Otolaryngol*. 1958;68:1–19.
18. Rodeno MT, Sanchez-Fernandez JM, Rivera-Pomar JM. Histochemical and morphometrical aging changes in humans vocal cord muscles. *Acta Otolaryngol*. 1993;113:445–449.
19. Filho JAX, Tsuji DH, do Nascimento PH, Sennes LU. Histologic changes in human vocal folds correlated with aging: a histomorphometric study. *Ann Otol Rhinol Laryngol*. 2003;112:894–898.
20. Titze IR. Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation. *J Speech Lang Hear Res*. 2006;49:439–447.
21. Mehta DD, Zanartu M, Quatieri TF, Deliyski DD, Hillman RE. Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy. *J Acoust Soc Am*. 2011;130:3999–4009.
22. Guzman M, Barros M, Espinoza F, Herrera A, Parra D, Muñoz D, Lloyd A. Laryngoscopic, acoustic, perceptual, and functional assessment of voice in rock singers. *Folia Phoniatr Logop*. 2013;65:248–256.

Reprinted by permission of Laryngoscope. 2014; 124(8):1869-1876.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Preliminary Data on Two Voice Therapy Interventions in the Treatment of Presbyphonia

Aaron Ziegler, MA (ABD), CCC-SLP; Katherine Verdolini Abbott, PhD, CCC-SLP; Michael Johns, MD; Adam Klein, MD; Edie R. Hapner, PhD, CCC-SLP

Objectives/Hypothesis: Presbyphonia is common among elderly individuals, yet few studies have evaluated behavioral treatment approaches for presbyphonia. The primary aim of this study was to assess the short-term efficacy of two types of voice therapy—vocal function exercises (VFE) and phonation resistance training exercise (PhoRTE) therapy—in the treatment of presbyphonia. The secondary aim was to determine if differences in adherence and treatment satisfaction existed between the two therapy approaches.

Study Design: Prospective, randomized, controlled.

Methods: Preliminary data from 16 elderly participants with presbyphonia randomly assigned to VFE, PhoRTE, or a no-treatment control group (CTL) were analyzed. Before and after a 4-week intervention period, participants completed the *Voice-Related Quality of Life* (V-RQOL) questionnaire and a perceived phonatory effort (PPE) task. Additionally, participants receiving treatment completed weekly practice logs and a posttreatment satisfaction questionnaire.

Results: Preliminary data revealed VFE and PhoRTE groups demonstrated a significant improvement in V-RQOL scores. However, only PhoRTE demonstrated a significant reduction in PPE, as suggested by the study's causal model. The CTL group did not demonstrate significant changes. Numerically, VFE registered slightly greater adherence to home practice recommendations than did PhoRTE, but PhoRTE perceived greater treatment satisfaction than VFE.

Conclusions: Findings provide new evidence regarding the efficacy of voice therapy exercises in the treatment of age-related dysphonia and suggest PhoRTE therapy as another treatment method for improved voice-related quality of life and reduced perceived vocal effort in this population.

Key Words: Aging, presbyphonia, voice disorder, treatment.

Level of Evidence: 2b.

***Laryngoscope*, 124:1869–1876, 2014**

INTRODUCTION

Presbyphonia is a common clinical finding among the elderly and poses a significant barrier to life satisfaction.[1,2] This voice disorder results from age-related laryngeal and respiratory degenerative changes, which lead to glottal incompetence[3] and a decline in inspiratory and expiratory pressures.[4] A deterioration in vocal function in the elderly has been putatively linked to a reduced amount and *intensity* of speech.[5] Interestingly, and analogous to findings of senior athletes,[6] the voice of elderly singers sounds younger, clearer, and louder than the elderly nonsinger's voice.[7] Additionally, both elderly male[8] and female[8,9] singers maintain a stable fundamental frequency throughout the lifespan. Those differences suggest the benefit of increased vocal activity for vocal longevity.

Current Evidence for Behavioral Treatment of Presbyphonia

Over the past decade, eight studies have been conducted on voice therapy for presbyphonia.[10–17] In brief, an overwhelming majority of patients with presbyphonia believe voice therapy is beneficial[15] and exhibit a significant improvement in voice-related quality of life,[13,14] a finding not observed in patients who forego voice therapy.[14] Furthermore, patients with presbyphonia report a significant decrease in phonatory effort after completing voice therapy.[13] Most important, patients with presbyphonia who receive voice therapy exhibit a significant improvement in their functional vocal status.[16]

To date, published prospective studies have only investigated the efficacy of voice therapy approaches for treating individuals with presbyphonia,[10–13,17] but none

From the Emory Voice Center, Department of Otolaryngology—Head and Neck Surgery, Emory University (A.Z., M.J., A.K., E.R.H.), Atlanta, Georgia; the Department of Communication Science and Disorders (K.V.); and the McGowan Institute for Regenerative Medicine (K.V.), University of Pittsburgh; the Center for Neural Basis of Cognition (K.V.), Carnegie-Mellon University and University of Pittsburgh, Pennsylvania; and note Aaron Ziegler is now Acting Assistant Professor in Communication Sciences and Disorders, John A. Burns School of Medicine, University of Hawaii at Manoa (A.Z.), Honolulu, Hawaii, U.S.A

Editor's Note: This Manuscript was accepted for publication November 27, 2013.

Presented at the Voice Foundation's 40th Annual Symposium, Philadelphia, PA, U.S.A., 1–5 June 2011.

Financial Disclosures: Aaron Ziegler, Katherine Verdolini Abbott, Michael Johns III, and Adam Klein-None; Edie Hapner—Plural Publishing. Conflict of Interest: Aaron Ziegler and Edie Hapner—Authors for Plural Publishing. The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Aaron Ziegler, MA (ABD), CCC-SLP, Department of Communication Sciences and Disorders, John A. Burns School of Medicine, University of Hawaii at Manoa, 677 Ala Moana Blvd, Suite 625, Honolulu, HI 96813. E-mail: aszieglerr@gmail.com

DOI: 10.1002/lary.24548

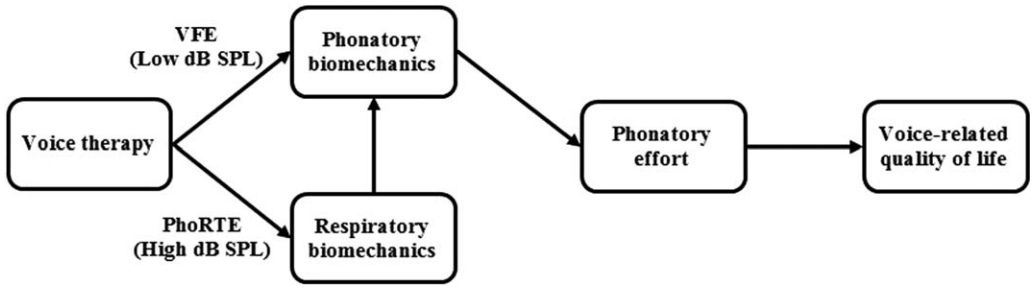

Fig. 1. Proposed flowchart delineating a causal model linking voice therapy to changes in phonatory and respiratory biomechanics, phonatory effort, and voice-related quality of life.

have compared voice therapy techniques to assess the superiority of one approach over another. Furthermore, the literature lacks suggestions for a causal model describing mechanisms of voice change from behavioral treatment of presbyphonia that may assess the potential differential impact of two types of voice therapy. Therefore, a causal model was developed, which stated that targeted voice therapy may affect phonatory biomechanics directly or indirectly through altered respiratory behavior resulting in reduced phonatory effort and lead to an improved voice-related quality of life (Fig. 1).

The causal model suggests that an effective therapeutic approach for presbyphonia will be one that targets the biological bases of the condition, or degenerative respiratory and laryngeal changes as a result of aging. These changes in muscle mass and strength—sarcopenia—are targeted in other parts of the body by engaging in structured exercise that emphasizes an increased level of physical activity to overload the muscle and reverse the sarcopenia process.[18] This type of exercise training—resistance training—has demonstrated positive effects on sarcopenia in older adults by reducing secondary aging effects that occur from muscle atrophy and weakness.[19]

Based on the causal model, it was hypothesized that the intervention groups in this study would result in more positive changes in voice across the experimental period than seen in a no-intervention control group. Furthermore, the causal model suggests that one therapy, a treatment requiring high-vocal intensity phonation and that loads both respiratory and laryngeal musculature, will result in more positive changes than the other therapy, a treatment requiring low vocal intensity phonation.

Study Aims

The purpose of this study was to compare two interventions and no treatment for adults with presbyphonia by using a prospective, randomized, controlled experimental design to assess the short-term efficacy of two voice therapy approaches, as demonstrated by a change in quality of life and perceived phonatory effort. Secondary aims of this study were to examine differences in patient adherence and treatment satisfaction.

MATERIALS AND METHODS

All procedures were approved by the institutional review boards at Emory University and the University of Pittsburgh (IRB #00037045 and #10060268, respectively). The experiment used a prospective, randomized, controlled design.

Participants

Twenty elderly adults aged 60 years and over enrolled in the study (Fig. 2). For this preliminary study, the sample size was selected arbitrarily to generate the necessary results for a power analysis for future studies.

All participants a) reported a current voice problem, including a complaint of reduced vocal loudness or increased vocal effort; b) received a diagnosis of presbyphonia by a fellowship-trained laryngologist[14]; c) received an auditory–perceptual diagnosis of vocal asthenia by a voice-specialized speech-language pathologist (SLP); d) were judged perceptually by a SLP to be free of dysarthria, dysfluency, or language problems; e) passed hearing, cognition, and mood screenings; f) were currently nonsmokers (five years or more); g) reported no progressive neuromuscular diseases affecting voice; h) denied concomitant health problems affecting voice; i) completed menopause, if female; j) reported using current medications for at least one month before participation; k) denied current use of inhaled corticosteroids or prednisone; and l) stated willingness to persist with the 6-week protocol. In addition, participants were included, if stimulable for improved voice quality as assessed by a SLP during the physician's examination visit. Stimulability testing is a routine part of the voice evaluation to determine candidacy for treatment.[20] No participants were excluded based on race, ethnicity, or gender. In accordance with standards on reporting randomized, controlled studies,[21] participant characteristics are provided in Table I.

Procedures

Recruitment, screening, and randomization. Recruitment was performed by a SLP who was part of the multidisciplinary team at the Emory Voice Center. An individual was initially seen for a comprehensive evaluation by a fellowship-trained laryngologist and SLP. Following informed consent, each individual underwent a hearing screening to ensure age-appropriate hearing or adequately managed sensory-neural hearing loss with the use of hearing aids, as evidenced by a response during audiometric testing in a sound-isolated booth at 40 dB HL at 0.5 kHz, 1 kHz, and 2 kHz presented in sound field.[22] Next, each individual underwent a screening to ensure age-appropriate cognitive ability based on results from the *Mini Mental State Examination* (MMSE).[23] A score of ≥ 20 was required for further participation in the study. Then, each individual underwent self-administration of the *Elderly Depression Scale-Short Form* (EDS-SF),[24] and a score of ≤ 5 was required for further participation. Finally, individuals satisfying inclusion criteria were randomized to one of three groups using a

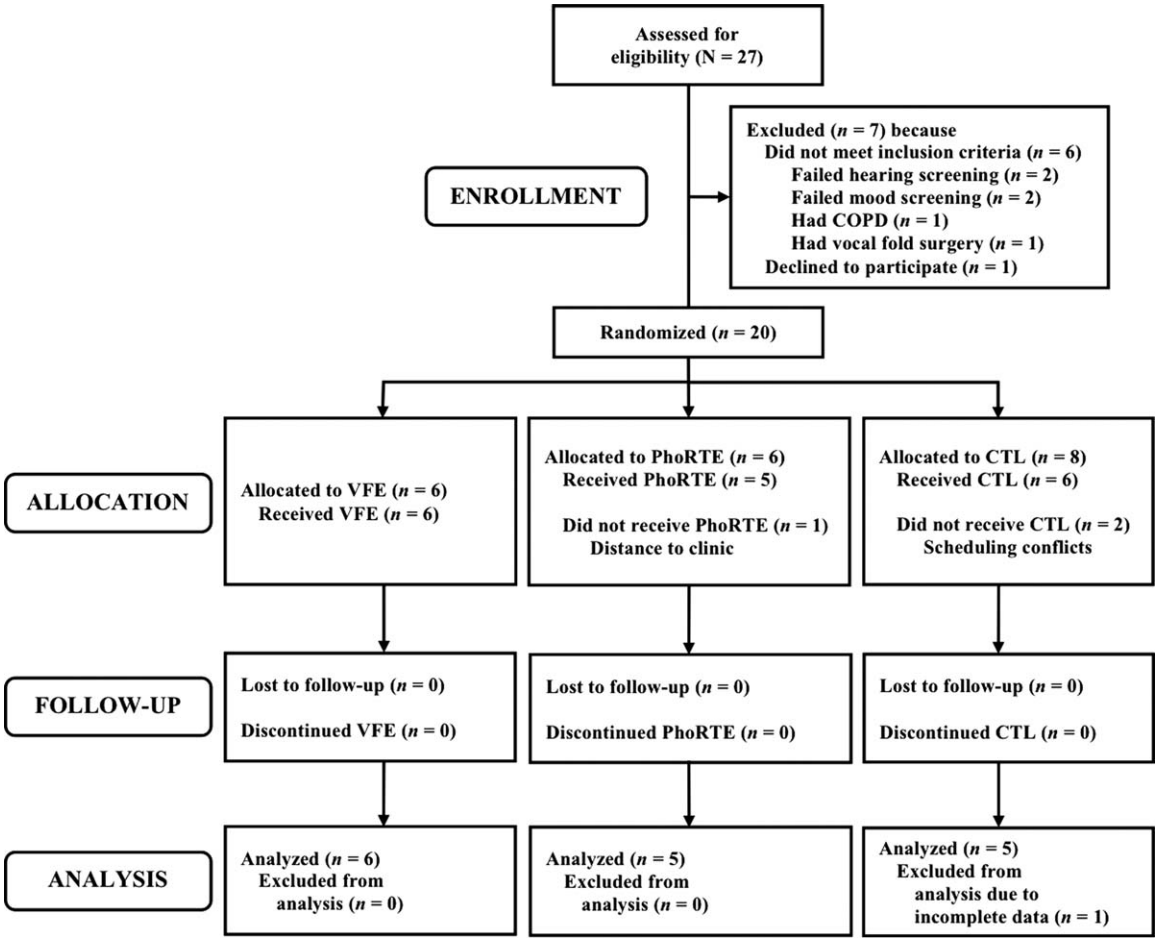

Fig. 2. Flowchart of study procedures.

computer algorithm: vocal function exercises (VFE), phonation resistance training exercise (PhoRTE) therapy, or a no-intervention control group (CTL). All participants were briefly counseled on voice hygiene and given a written copy of a handout that describes vocal hygiene recommendations.

Baseline and follow-up evaluations. At the baseline visit, each participant completed the V-RQOL.[25] Then, the participant was asked to provide an estimation of perceived phonatory effort (PPE). To determine PPE, the participant used a direct magnitude estimation scale[26] on which "100" represented "comfortable effort during phonation," "50" represented "half as much effort as comfortable," "200" represented "two times as much effort as comfortable," and so forth.[27,28]

Participants returned for follow-up measures within one week of completion of the intervention, or 6-weeks postbaseline in the case of the CTL group. At the follow-up visit, each participant completed the V-RQOL[25] and provided a rating of PPE, which were anchored to the participant's baseline ratings to limit drift due to increased awareness of voice. Finally, participants in the VFE and PhoRTE groups completed a post treatment satisfaction questionnaire.[29]

Interventions. Participants receiving an intervention attended four 45-minute treatment sessions—either VFE or PhoRTE—over the course of four weeks, which were provided by one of two participating voice-specialized SLPs. Execution of VFE[30,31] involved four exercises: 1) maximum sustained phonation on /i/ on the pitch F above middle C (males dropped down an octave); 2) an ascending glide over the entire pitch range on /oɬ/; 3) a descending glide over the entire pitch range on /oɬ/; and 4) maximum sustained phonation on the pitches middle C and D, E, F, and G above middle C (males dropped down an octave) on /oɬ/. Participants learned to use low abdominal breathing, a frontal focus with an inverted megaphone mouth shape, and were instructed to complete the exercises as quietly as possible but while maintaining a clear and consistent voice.

PhoRTE[32] (a homophone to the Italian word *forte* meaning loud and strong), adapted from Lee Silverman Voice Treatment (LSVT),[33–35] consisted of four exercises: 1) loud maximum sustained phonation on /a/; 2) loud ascending and descending pitch glides over the entire pitch range on /a/; 3) participant-specific functional phrases using a loud and high voice; and 4) phrases from exercise #3 in a loud and low voice. Low abdominal breathing gestures were encouraged. All feedback thereafter was limited to reminding participants to maintain a "strong" voice. During therapy sessions, participants were expected to maintain a SPL between 80 and 90 dB, as measured by a sound level meter positioned at a microphone-to-mouth distance of 30 cm.

PhoRTE, while derived from the therapeutic studies on LSVT, differed in several ways. First, PhoRTE sessions occurred once weekly as opposed to a more intensive intervention schedule for LSVT (i.e., four days per week for four weeks). Second, PhoRTE incorporated two different manners of producing participant-specific functional phrases (i.e., a loud and high voice and a loud and low voice),[36] Finally, PhoRTE home practice required fewer repetitions than is typically required for

TABLE I.
Summary of Participant Characteristics by Group.

Group/Participant	Sex	Age	Race
VFE			
1	female	83	Caucasian
3	male	66	Caucasian
9	female	74	Caucasian
10	male	78	Caucasian
13	male	78	Caucasian
17	male	60	Caucasian
Mean (SD), $n = 6$	2 females; 4 males	73.2 (8.6)	
PhoRTE			
6	male	79	Caucasian
7	female	78	Caucasian
8	female	72	Caucasian
11	female	80	Caucasian
20	male	71	Asian
Mean (SD), $n = 5$	3 females; 2 males	75.8 (4.0)	
CTL			
2	male	79	Caucasian
4	female	69	Caucasian
5	male	76	African American
14	female	91	Caucasian
15	male	73	Caucasian
Mean (SD), $n = 5$	2 females; 3 males	77.6 (8.4)	
Overall Mean (SD), $N = 16$		75.4 (7.2)1	

CTL = no-treatment control group; PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

patients receiving LSVT (two versus 10 repetitions of each exercise per practice session, respectively).

The PhoRTE exercises were selected because of their high intensity nature that might induce changes to muscle structure and function to reverse the degenerative sarcopenia process.[12] In addition, phonatory–resonatory interaction through a widened mouth and narrow pharynx, as occurs with the use of the vowel /a/, creates an acoustic situation that allows a speaker to shout safely. This megaphone mouth shape at low to medium high pitches raises the first formant frequency to reinforce the fundamental and second harmonic of the source. The resulting phonatory–resonatory interaction helps to recalibrate phonatory effort by assisting vocal fold vibration and maximizing phonatory efficiency. Furthermore, coupling a narrowed epilarynx tube with increased adduction provides maximum power transfer from the glottis to the lips to further increase vocal loudness.[37] Finally, the PhoRTE program subscribes to a task-dependent model of motor control by including functional phrases to help with generalization of voice techniques to conversation.[38]

Home practice program. Participants in both intervention groups were instructed to practice their respective treatments, VFE or PhoRTE, twice daily every day, to perform each exercise twice during each practice session, and to log their practice. Participants were instructed to complete practice logs only for completed exercises. From the practice log, the percent of prescribed exercises completed was computed to measure treatment adherence. The protocols of the two treatments controlled for what was assumed to be equivalent practice durations if the participant was adherent to the twice daily practice sessions. Participants received written instructions on how to complete daily home practice and a compact disc with audio demonstrations of the respective exercises.

RESULTS

Statistical Analysis

Inferential statistical analyses of the preliminary data were used to examine pretreatment to posttreatment changes within groups, and between group differences were examined descriptively for the primary outcome measures (i.e., V-RQOL and PPE). Inferential statistical analyses were also used to investigate between group differences in the secondary outcome measures (i.e., treatment adherence and treatment satisfaction). Due to the preliminary nature of this study and the small sample size, an alpha level of 0.10 was used to minimize the type II error rate in analyzing treatment effects on primary and secondary outcome measures. Of the 20 enrolled participants, only 16 participants were included in the data set for analysis. Of the four who were excluded, three dropped out of the study prior to data collection and one participant in the no-treatment control group had an incomplete data set. Therefore, data from six VFE participants, five PhoRTE participants, and five CTL participants were analyzed.

Participant Characteristics

Participants were seven women (44%) and nine men (56%) aged 60 to 91 years ($M = 75.4$ years, SD = 7.2). Post-hoc analyses using Fisher's exact test and between-subject ANOVAs confirmed the equivalence of groups on gender ($P = .825$, Fisher's Exact Test), age ($F[2, 13] = 0.501$, $P = .617$, $\eta_p^2 = .072$), baseline V-RQOL scores ($F[2, 13] = 0.880$, $P = .438$, $\eta_p^2 = .119$), and baseline PPE ratings ($F[2, 13] = 1.948$, $P = .182$, $\eta_p^2 = .231$) (Tables (I–III)).

V-RQOL

Individual scores, group means and standard deviations, difference scores, and percent change values for the V-RQOL data before and following the 4-week intervention period are displayed in Table II. Results revealed that the VFE and PhoRTE groups experienced a significant improvement in mean pretreatment to posttreatment V-RQOL scores (80.8 to 87.5, $t[5] = 1.964$, $P = .054$, one-tailed, $d = 0.80$ and 88.5 to 95.0, $t[4] = 2.152$, $P = .049$, one-tailed, $d = 0.96$, respectively). The CTL group did not demonstrate a significant change in mean V-RQOL scores (87.5 to 91.5, $t[4] = 1.554$, $P = .195$, $d = 0.70$).

The data were reanalyzed after excluding a PhoRTE participant who commenced therapy without registering quality of life impairment (as evidenced by a score of 100 on the V-RQOL). Removal increased the PhoRTE percent change value (8.03 to 10.66), and it was slightly greater than that of the VFE group (9.30).

TABLE II.
Individual Scores, Mean Pretreatment and Posttreatment Scores, Standard Deviations, Percent Change, and P Values for the VFE, PhoRTE, and CTL Groups on the *Voice-Related Quality of Life*.

Group/Participant	Baseline (Pretreatment)	Follow-Up (Posttreatment)	Absolute Difference	Percent Change	Test Statistic	*P* Value
VFE						
1	80.0	85.0	5.0	6.25		
3	90.0	90.0	0.0	0.00		
9	62.5	85.0	22.5	36.00		
10	90.0	97.5	7.5	8.33		
13	92.5	97.5	5.0	5.41		
17	70.0	70.0	0.0	0.00		
Mean (SD), $n = 6$	80.8 (12.3)	87.5 (10.2)	6.7 (8.3)	9.30 (13.5)	$t = 1.964$**	.054*
PhoRTE						
6	97.5	100.0	2.5	2.56		
7	82.5	97.5	15.0	18.18		
8	75.0	85.0	10.0	13.33		
11	87.5	95.0	7.5	8.57		
20	100.0	97.5	−2.5	−2.50		
Mean (SD), $n = 5$	88.5 (10.4)	95.0 (5.9)	6.5 (6.8)	8.03 (8.25)	$t = 2.152$**	.049*
CTL						
2	90.0	92.5	2.5	2.78		
4	95.0	90.0	−5.0	−5.26		
5	75.0	82.5	7.5	10.00		
14	85.0	95.0	10.0	11.76		
15	92.5	97.5	5.0	5.41		
Mean (SD), $n = 5$	87.5 (7.9)	91.5 (5.8)	4.0 (5.8)	4.94 (6.73)	$t = 1.554$**	.195
Overall Mean (SD), $N = 16$	85.3 (10.4)					

Note. *Significant difference at $P \leq 0.10$ level, one-tailed.
**From repeated-measures *t* test.
CTL = no-treatment control group; PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

PPE

Individual ratings, group means and standard deviations, difference scores and percent change values for PPE ratings before and following the 4-week intervention period are shown in Table III. Results showed that PPE ratings decreased significantly in the PhoRTE group only (144 to 102, $t[4] = -2.370$, $P = .077$, two-tailed, $d = -1.06$). Neither the VFE group nor the CTL group demonstrated a significant difference in PPE ratings (142.5 to 109.2, $t[5] = -1.865$, $P = .121$, two-tailed, $d = -0.76$; 101 to 103, $t[4] = 1.000$, $P = .374$, two-tailed, $d = 0.45$, respectively).

Adherence and Treatment Satisfaction

Participants in the VFE and PhoRTE groups demonstrated adherence to treatment recommendations, and no differences were detected between groups ($P = .411$). One participant in the PhoRTE group practiced significantly less than any other participant and skewed the averaged data for adherence. A post-hoc analysis of the data removing this participant from the PhoRTE data resulted in a more balanced assessment of the practice patterns of the PhoRTE group, 88.2%, nearly equivalent to the average practice of the VFE group (89.3%). Results for treatment satisfaction data revealed no differences in ratings between VFE and PhoRTE on the three questions: extent to which participants a) liked the particular therapy ($P = .285$); b) felt voice changed because of therapy ($P = .227$); and c) felt voice changes were caused by the particular therapy ($P = .550$) (Table IV).

DISCUSSION

The data from this study provide optimism that there may be short-term benefits from two therapy approaches, VFE and PhoRTE, for improvement of voice-related quality of life in elderly individuals with presbyphonia. The causal model tested in this study proposed that therapy-induced changes in laryngeal biomechanics, possibly partly related to changes in respiratory biomechanics, would lead to a reduction in perceived phonatory effort and, ultimately, result in an improvement in voice-related quality of life. Significant pretreatment to posttreatment increases were documented in V-RQOL scores for both intervention groups, in comparison to scores for a no-treatment control group, which did not improve. The magnitude of pretreatment to posttreatment differences on the V-RQOL in each treatment group (VFE and PhoRTE) exceeded changes in an untreated group of elderly individuals with presbyphonia. The improvement of patient-reported outcome measures in a group of elderly individuals with presbyphonia

TABLE III.
Individual and Mean Pretreatment and Posttreatment Ratings, Standard Deviations, Difference Scores, Percent Change, and *P* values for the VFE, PhoRTE, and CTL Groups on Perceived Phonatory Effort.

Group/Participant	Baseline (Pretreatment)	Follow-Up (Posttreatment)	Absolute Difference	Percent Change	Test Statistic	*P* Value
VFE						
1	125	100	−25.0	−20.0		
3	100	100	0.0	00.0		
9	150	100	−50.0	−33.3		
10	200	100	−100.0	−50.0		
13	100	125	−25.0	25.0		
17	180	130	−50.0	−27.8		
Mean (SD), $n = 6$	142.5 (41.7)	109.2 (14.3)	−33.3 (43.8)	−17.7 (26.6)	$t = -1.865$**	.121
PhoRTE						
6	100	100	0.0	00.0		
7	100	50	−50.0	−50.0		
8	200	150	−50.0	−25.0		
11	200	100	−100.0	−50.0		
20	120	110	−10.0	−8.3		
Mean (SD), $n = 5$	144 (51.8)	102 (35.6)	−42.0 (39.6)	−26.7 (23.1)	$t = -2.370$**	.077*
CTL						
2	100	100	0.0	00.0		
4	100	100	0.0	00.0		
5	125	125	0.0	00.0		
14	100	100	0.0	00.0		
15	80	90	10.0	12.5		
Mean (SD), $n = 5$	101 (16.0)	103 (13.0)	2.0 (4.5)	2.5 (5.6)	$t = 1.000$**	.374
Overall Mean (SD), $N = 16$	130 (42.1)					

Note. *Significant difference at $P \leq 0.10$ level, two-tailed.
**From repeated-measures *t* test.
CTL = no-treatment control group; PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

following voice therapy is consistent with results from prior research (Berg et al., 2008; Sauder et al., 2010).

Significant improvement in perceived phonatory effort accompanied voice-related quality of life changes for the PhoRTE group, but not the VFE group, a finding that partially supports the causal model explored in this study, and moreover, that can also be inferred from previous research in a similar cohort.[13] Differences in PPE pretreatment to posttreatment changes between VFE and PhoRTE may be explained by unique vocal tract configurations and their influence on vocal fold vibration. Whereas VFE are characterized by an inverted megaphone-shaped vocal tract, PhoRTE therapy employs a megaphone-shaped vocal tract. Consistent with nonlinear dynamics, rounded vowels such as /o, u/ using a wide open pharynx as in the case of VFE, have been shown to decrease vocal fold adduction. Open vowels such as /a, æ/ using a narrow pharynx and high larynx, as in PhoRTE, have been shown to cause greater vocal fold adduction. In the population of interest, increased adduction is a desired laryngeal target. Perhaps a reduction in the glottal half-width due to increased adduction lowered the required subglottal pressure and resulted in a decrease in perceived phonatory effort.[37,39]

Whereas improvement in V-RQOL scores was accompanied by numerical decreases in PPE in both treatment groups, the no-treatment control group exhibited the opposite finding. For that group, pre- to posttreatment PPE actually increased slightly, even with anchoring the posttreatment estimation of phonatory effort to pretreatment ratings. In light of that finding, elderly individuals who forego therapy seem to employ increased muscle tension at the level of the glottis to achieve phonatory closure during voicing.

Given these preliminary findings, PhoRTE may have a slight advantage over VFE for producing benefit from a physiologic perspective because it demands a higher intensity of effort, which better addresses the overload principle required to induce neuromuscular changes in strength.[40] Increased neuromuscular activity of both the respiratory and laryngeal systems from PhoRTE should lead to even greater improvement in respiratory and laryngeal biomechanics than VFE, ultimately causing a significant reduction in PPE. Furthermore, phonatory efficiency from a megaphone-shaped vocal tract configuration may have also contributed to decreased phonatory effort.[37] Additionally, inclusion of task-specific exercises, as used in PhoRTE, to address the exercise training principle of specificity and promote carryover may result in a greater change in respiratory and laryngeal biomechanics during conversational speech. Consequently, phonatory effort for the PhoRTE group should demonstrate a larger change than VFE.

TABLE IV.
Individual and Group Means, Standard Deviations, and P Values for the VFE and PhoRTE Groups on Weekly Practice Log (% completed) and Posttreatment Satisfaction Questionnaire.

Group/Participant	Adherence Week 1–4	Treatment Satisfaction		
		Like Therapy	Voice Change	Therapy Cause
VFE				
1	78.0	4	4	2
3	79.6	3	3	1
9	100.0	4	5	3
10	95.8	3	4	3
13	87.5	3	4	2
17	94.8	3	4	3
Mean (SD), $n = 6$	89.3 (9.0)	3.3 (.52)	3.9 (.66)	2.3 (.82)
PhoRTE				
6	100.0	3	4	2
7	17.5	3	5	3
8	56.3	4	4	2
11	96.5	5	4	3
20	100.0	4	5	3
Mean (SD), $n = 5$	74.1 (36.6)	3.8 (.84)	4.4 (.55)	2.6 (.55)
Test statistic	t (4.407) = 0.908*	t (9) = −1.137*	t (9) = −1.297*	t (9) = −0.621*
P value, two-tailed	.411	.285	.227	.550

Note. For "like therapy" scale, 1 = not at all; 2 = somewhat; 3 = moderate; 4 = very much; 5 = extremely. For "voice change" scale, 1 = got a lot worse; 2 = got a little worse; 3 = no change; 4 = got a little better; 5 = got a lot better. For "therapy cause" scale, 1 = voice therapy probably irrelevant to voice change; 2 = voice therapy may have caused voice changes; 3 = voice therapy definitely caused voice changes.

*From independent samples *t* test.

PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

In addition to the foregoing results, this study investigated adherence to home treatment recommendations in this population. Participants in both VFE and PhoRTE appeared to exhibit fairly regular practice of their home programs, a finding that is consistent with published literature.[17] Although self-report may be inaccurate, in the absence of any clear difference in mean practice between the VFE and PhoRTE groups, the most straightforward interpretation is that improvements in V-RQOL are not likely strongly related to treatment adherence.

Accordingly, although not significant, PhoRTE practiced less than VFE and yet consistently perceived greater satisfaction with the therapy they received. This finding supports a model of voice therapy in which treatment efficacy is optimized by a combination of biomechanical, learning, and adherence factors.[41] Specifically, the high intensity component of PhoRTE may necessitate less practice time than VFE to generate neuromuscular changes in muscle strength. Furthermore, the inclusion of functional speech tasks may promote fast learning because it addresses task-specificity and generalization to extra-therapy situations. In addition, practice of functional speech tasks for transfer of therapy techniques to unique communication situations, as well as the emphasis on increased vocal intensity to addresses a key patient concern— reduced loudness— may both increase self-efficacy and lead to improved treatment adherence.

Limitations and Future Aims

This study was designed to develop preliminary data to support the use of voice therapy for a subset of people with voice complaints secondary to presbylaryngeus. It was also designed to support the use of an alternative therapy that was based on high-intensity vocal exercise in the treatment of presbyphonia. Accordingly, one of the aims of the study was to develop an effect size for future research into the therapeutic treatment of presbyphonia. A limitation of this study is thus the small number of participants. Yet another limitation, although a no-treatment control group was included in the experimental design to determine the influence of time, was the lack of an experimental *treatment* control group, which would have provided evidence on whether the perceived change was due to a placebo effect. Additionally, a longitudinal study that follows participants for more than six weeks is necessary to assess maintenance of treatment effects. Future studies should include a larger sample size, incorporate a placebo treatment, and follow participants longitudinally. In addition, future studies should assess differences in vocal load between VFE and PhoRTE, as well as pre- to posttreatment changes in acoustic and aerodynamic parameters.

CONCLUSION

Indications from this study on voice therapy in individuals with presbyphonia are that behavioral

approaches are effective in the management of age-related voice problems. The study provides further preliminary evidence that individuals with presbyphonia may benefit from various therapeutic approaches for which patients express treatment satisfaction. Finally, this study contributes additional support to a previous finding that individuals with presbyphonia regularly practice voice exercises and exhibit good adherence to treatment recommendations.

BIBLIOGRAPHY

1. Golub JS, Chen PH, Otto KJ, Hapner E, Johns III MM. Prevalence of perceived dysphonia in a geriatric population. *J Am Geriatr Soc* 2006;54: 1736–1739.
2. Roy N, Stemple J, Merrill RM, Thomas L. Epidemiology of voice disorders in the elderly: preliminary findings. *Laryngoscope* 2007;117:628–633.
3. Linville SE. Vocal Aging. Clifton Park, NY: Delmar Cengage Learning; 2001.
4. Miller MR. Structural and physiological age-associated changes in aging lungs. *Semin Respir Crit Care Med* 2010;31:521–527.
5. Bastian RW, Thomas JP. Talkativeness and vocal loudness: do they correlate with laryngeal pathology? A study of the vocal overdoer/underdoer continuum. 29th Annual Symposium of the Voice Foundation: Care of the Professional Voice, Philadelphia, PA; 2000.
6. Tanaka H, Seals DR. Invited Review: Dynamic exercise performance in Masters athletes: insight into the effects of primary human aging on physiological functional capacity. *J Appl Physiol* 2003;95:2152–2162.
7. Prakup B. Acoustic measures of the voices of older singers and nonsingers. *J Voice* 2012;26:341–350.
8. Brown Jr WS, Morris RJ, Hollien H, Howell E. Speaking fundamental frequency characteristics as a function of age and professional singing. *J Voice*. 1991;5:310–315.
9. Brown Jr WS, Morris RJ, Hicks DM, Howell E. Phonational profiles of female professional singers and nonsingers. *J Voice* 1993;7:219–226.
10. Gorman S, Weinrich B, Lee L, Stemple JC. Aerodynamic changes as a result of vocal function exercises in elderly men. *Laryngoscope* 2008;118:1900–1903.
11. LaGorio LA, Carnaby-Mann GD, Crary MA. Treatment of vocal fold bowing using neuromuscular electrical stimulation. *Arch Otolaryngol Head Neck Surg* 2010;136:398–403.
12. Ramig LO, Gray S, Baker K, et al. The aging voice: a review, treatment data and familial and genetic perspectives. *Folia Phoniatr Logop* 2001; 53:252–265.
13. Sauder C, Roy N, Tanner K, Houtz DR, Smith ME. Vocal function exercises for presbylaryngis: a multidimensional assessment of treatment outcomes. *Ann Otol Rhinol Laryngol* 2010;119:460–467.
14. .Berg EE, Hapner E, Klein A, Johns III MM. Voice therapy improves quality of life in age-related dysphonia: a case-control study. *J Voice* 2008;22: 70–74.
15. Gartner-Schmidt J, Rosen C. Treatment success for age-related vocal fold atrophy. *Laryngoscope* 2011;121:585–589.
16. Mau T, Jacobson BH, Garrett CG. Factors associated with voice therapy outcomes in the treatment of presbyphonia. *Laryngoscope* 2010;120: 1181–1187.
17. Tay EY, Phyland DJ, Oates J. The effect of vocal function exercises on the voices of aging community choral singers. *J Voice* 2012;26:672 e619–627.
18. Freiberger E, Sieber C, Pfeifer K. Physical activity, exercise, and sarcopenia–future challenges. *Wien Med Wochenschr* 2011;161:416–425.
19. American College of Sports M, Chodzko-Zajko WJ, Proctor DN, et al. American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc* 2009;41: 1510–1530.
20. .ASHA. Voice Evaluation Templates. http://www.asha.org/uploadedFiles/slp/healthcare/AATVoiceEvaluation.pdf. Accessed May 31, 2013.
21. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–694.
22. Cruickshanks KJ, Wiley TL, Tweed TS, et al. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. *Am J Epidemiol* 1998;148:879–886.
23. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
24. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Disabil Rehabil* 1986;5:165–173.
25. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice* 1999;13:557–569.
26. Stevens SS. Psychophysics: introduction to its perceptual, neural, and social prospects. New York, NY: Wiley & Sons; 1975.
27. Colton RH, Brown JWS. Some relationships between vocal effort and intraoral air pressure. *J Acoust Soc Am* 1973;53:296–296.
28. Wright HN, Colton RH. Some parameters of vocal effort. *J Acoust Soc Am* 1972;51:141.
29. Nanjundeswaran C, Li NY, Chan KM, Wong RK, Yiu EM, Verdolini-Abbott K. Preliminary data on prevention and treatment of voice problems in student teachers. *J Voice* 2012;26:816 e811–816, e812.
30. Stemple JC, Lee L, D'Amico B, Pickup B. Efficacy of vocal function exercises as a method of improving voice production. *J Voice* 1994;8:271–278.
31. Wrycza Sabol J, Lee L, Stemple JC. The value of vocal function exercises in the practice regimen of singers. *J Voice* 1995;9:27–36.
32. Ziegler A, Hapner E. Phonation Resistance Training Exercise (PhoRTE) Therapy. In: Behrman A, Haskell J, eds. Exercises for Voice Therapy. 2 ed. San Diego, CA: Plural Publishing; 2013.
33. Ramig LO, Countryman S, Thompson LL, Horii Y. Comparison of two forms of intensive speech treatment for Parkinson Disease. *J Speech Hear Res* 1995;38:1232–1251.
34. Smith ME, Ramig LO, Dromey C, Perez KS, Samandari R. Intensive voice treatment in Parkinson Disease: laryngostroboscopic findings. *J Voice* 1995;9:453–459.
35. Ramig LO, Bonitati CM, Lemke JH, Horii Y. Voice treatment for patients with Parkinson Disease: Development of an approach and preliminary efficacy data *J Med Speech Lang Pathol* 1994;2:191–209.
36. de Swart BJ, Willemse SC, Maassen BA, Horstink MW. Improvement of voicing in patients with Parkinson's Disease by speech therapy. *Neurology* 2003;60:498–500.
37. Titze IR, Verdolini Abbott K. V*ocology: The Science and Practice of Voice Habilitation*. Salt Lake City, UT: National Center for Voice and Speech; 2012.
38. Ziegler W. Review. *Aphasiology* 2003;17:3–36.
39. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am* 1988;83:1536–1552.
40. McArdle WD, Katch FI, Katch VL. *Exercise Phsyiology: Nutrition, Energy, and Human Performance*. 7th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2010.
41. Verdolini-Marston K, Burke MK, Lessac A, Glaze L, Caldwell E. Preliminary study of two methods of treatment for laryngeal nodules. *J Voice* 1995;9:74–85.

Reprinted by permission of Laryngoscope. 2016; 126(1):135-141.

The Laryngoscope
© 2015 The American Laryngological, Rhinological and Otological Society, Inc.

Systematic Review

Cricopharyngeal Dysfunction: A Systematic Review Comparing Outcomes of Dilatation, Botulinum Toxin Injection, and Myotomy

Pelin Kocdor, MD; Eric R. Siegel, MS; Ozlem E. Tulunay-Ugur, MD

Objectives: Cricopharyngeal dysfunction may lead to severe dysphagia and aspiration. The objective of this systematic review was to evaluate the existing studies on the effectiveness of myotomy, dilatation, and botulinum toxin (BoT) injection in the management of cricopharyngeal dysphagia.

Methods: PubMed and Web of Science databases were searched to identify eligible studies by using the terms "cricopharyngeal dysfunction," "cricopharyngeal myotomy," "cricopharyngeal botox," "cricopharyngeal dilation," and their combinations from 1990 to 2013. This was supplemented by hand-searching relevant articles. Eligible articles were independently assessed for quality by two authors. Statistical analysis was performed.

Results: The database search revealed 567 articles. Thirty-two articles met eligibility criteria and were further evaluated. The reported success rates of BoT injections was between 43% and 100% (mean = 76%), dilation 58% and 100% (mean = 81%), and myotomy 25% and 100% (mean = 75%). In logistic regression analysis of the patient-weighted averages, the 78% success rate with myotomy was significantly higher than the 69% success rate with BoT injections ($P = .042$), whereas the intermediate success rate of 73% with dilation was not significantly different from that of either myotomy ($P = .37$) or BoT ($P = .42$). There was a statistically significant difference between endoscopic and open myotomy success rates ($P = .0025$). Endoscopic myotomy had a higher success rate, with a 2.2 odds ratio.

Conclusions: The success rate of myotomy is significantly higher than the success rate of BoT injections in cricopharyngeal dysfunction. Moreover, endoscopic myotomy was found to have a higher success rate compared to open myotomy.

Key Words: Cricopharyngeal dysfunction, cricopharyngeal myotomy, cricopharyngeal botox, cricopharyngeal dilation.

Level of Evidence: NA

***Laryngoscope*, 00:000–000, 2015**

INTRODUCTION

Cricopharyngeal (CP) muscle dysfunction can lead to dysphagia, aspiration, and weight loss, causing significant morbidity and reduced quality of life.[1] Etiologies are numerous and include the general categories of anatomic (cricopharyngeal bar), neuromuscular (central, peripheral, or myogenic), iatrogenic, inflammatory, neoplastic, and idiopathic (Table I).[2] The role of the CP muscle in swallowing has been well established. In 1717, Valsalva first described the anatomy of the cricopharyngeus muscle, which was further clarified by Killian in 1907.[3] CP dysfunction has been attributed mainly to the disordered opening of the CP muscle, which is the main component of the upper esophageal sphincter (UES). The opening of the UES necessitates three factors: neural inhibition of tonic intrinsic sphincter muscle contraction, anterior-superior laryngeal elevation that leads to the mechanical distraction of the UES, and passive stretching of the intrinsic sphincter muscles as the bolus passes.[4,5] A heterogeneous spectrum of disorders can lead to CP dysfunction, including failure of neural inhibition of tonic CP contraction, weakness of pharyngeal muscles with reduced laryngeal elevation and UES opening, as well as decreased compliance of the CP muscle, such as due to radiation fibrosis.

Various preoperative techniques can be used for diagnosis (Table II). The most important component has been a thorough history. In most centers this is followed by a videofluoroscopic swallowing study (VFSS) and manometry. These not only demonstrate the dysfunctional UES, but also demonstrate laryngeal elevation, the strength of the pharyngeal muscles, and laryngeal penetration or aspiration. Although some authors find manometry cumbersome and of limited value,[6,7] others strongly advocate the use of it, especially if coupled with fluoroscopy.[8–11] Manofluoroscopy, which ensures improved sensor placement, also allows assessment of pressures at known sensor locations during swallowing.[10,12,13] It is still

From the Department of Otolaryngology–Head and Neck Surgery (P.K., O.E.T.-U.); and the Department of Biostatistics (E.R.S.), University of Arkansas for Medical Sciences, Little Rock, Arkansas, U.S.A.

Editor's Note: This Manuscript was accepted for publication May 21, 2015.

Presented at American Laryngological Association Annual Meeting, Boston, Massachusetts, U.S.A., April 22–23, 2015.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Ozlem E. Tulunay-Ugur, MD, Associate Professor, Director Division of Laryngology, University of Arkansas for Medical Sciences, Department of Otolaryngology–Head and Neck Surgery, 4301 W. Markham, Slot 543, Little Rock, AR 72205.
E-mail: oetulunayugur@uams.edu

DOI: 10.1002/lary.25447

TABLE I.
Causes of Cricopharyngeal Dysfunction.

Central nervous system
Cerebellar infarct
Brain stem infarct
Parkinsonism
Amyotrophic lateral sclerosis
Base of skull neoplasm
Peripheral nervous system
Peripheral neuropathy
Diabetic neuropathy
Bulbar poliomyelitis
Myasthenia gravis
Neoplasm
Cricopharyngeal muscle
Polymyositis
Oculopharyngeal muscular dystrophy
Hyperthyroidism
Hypothyroidism
Cricopharyngeal disruption
Laryngectomy
Supraglottic laryngectomy
Radical oropharyngeal resections
Pulmonary resections
Cricopharyngeal spasm
Hiatal hernia
Gastroesophageal reflux
Idiopathic cricopharyngeal achalasia

Adapted from Halvorson DJ.[30]

not available and not a part of the workup in many institutions. Poirier et al. advocate the use of manometry to assess the physiological abnormalities at the pharyngoesophageal junction, but do not use it as an indication for surgical treatment.[14] Electromyography has been used by some authors to diagnose swallowing disorders.[15,16]

Numerous treatments exist for CP dysfunction, including swallowing therapy, CP dilation, injection of botulinum toxin, and CP myotomy. The traditional surgical treatment for CP dysfunction has been CP myotomy through a transcervical approach. To minimize the complications of an open approach, endoscopic CP myotomy was introduced using the potassium-titanyl-phosphate laser (wavelength, 532 nm) by Halvorson and Kuhn in 1994.[17] Subsequently, carbon dioxide laser (wavelength, 10,600 nm) gained favor because of its ability to coagulate small vessels and minimize thermal damage.[1]

Blitzer and Brin first presented on the use of in-office botulinum toxin (BoT) injections in 1993 as an alternative to surgery for the treatment of UES dysfunction.[18] In most cases, BoT has been injected under endoscopic visualization and general anesthesia, whereas less has been reported on percutaneous BoT injections under electromyographic guidance and local cutaneous anesthesia.[19] The range of BoT doses reported per injection varies from 10 U to 100 U.[20] Bougienage has been used in the treatment of anatomic esophageal strictures for decades.[21] The commonly used approaches are bougies, wire-guided polyvinyl dilators, air-filled pneumatic dilatation, and water-filled balloon dilatation with or without endoscopy guidance.[22]

CP dysfunction can be challenging diagnostically and in regard to the identification of the best treatment modality for a given patient. The scope of this article was to systematically review the literature regarding CP muscle interventions, specifically myotomy, injection of BoT, and dilation of the CP muscle for the treatment of CP dysfunction in adult patients.

MATERIALS AND METHODS

The literature search was performed according to the guidelines of the Cochrane Collaboration for systematic reviews in PubMed and Web of Science using a time frame from January 1990 until March 2013. Only literature published in English was considered. The search included the following keywords: "cricopharyngeal dysfunction," "cricopharyngeal myotomy," "cricopharyngeal botox," "cricopharyngeal dilation," and their combinations. The inclusion criterion for the studies was for the main focus of the article to be on the success rate and complications of the treatment modality. Bibliographies were manually reviewed to obtain additional articles of relevance. Reviews, editorials, case reports with less than four patients, articles with nonhuman data, duplicate publications, and articles on the pediatric patient population were excluded. Articles describing CP dysfunction attributed directly to Zenker's diverticulum and/or requiring diverticulectomy were also excluded. Articles with one specific etiology (except CP achalasia) as the reason for cricopharyngeal dysfunction were excluded; articles with heterogeneous etiology were included in the study.

The eligible articles were assessed for quality using the modified Downs and Black scale,[23] which is a validated checklist for randomized and nonrandomized studies. Any data extraction or assessment disagreements or inconsistencies were resolved through discussion and consensus.

Statistical Analysis

The average success rate of each procedure was calculated two ways: 1) as the crude (unweighted) average of reported success rates across articles and 2) as the patient-weighted average calculated as the total number of reported successes divided by the total number of treated patients. For logistic regression, the events/trials syntax was used, in which "events" and "trials" respectively represented the number of successes and number of patients in each article; this means that the logistic regression was effectively comparing patient-weighted averages between procedures. Additionally, the procedures were scored for invasiveness as botulinum toxin = low, dilation = medium, and myotomy = high, and the trend in success rate with invasiveness was assessed via the Cochran-Armitage trend test. These analyses assessing success rates were also used for complication rates. SAS version 9.3 (SAS Institute, Cary, NC) was employed for all analyses, and a $P < .05$ significance level was employed for all comparisons.

RESULTS

Study selection identified 567 reference articles; of these 42 met eligibility criteria. An additional five potential relevant reports were identified through scanning reference lists. Ultimately, 32 articles were included in the analysis. Thirteen articles were excluded for the

TABLE II.
Selected Studies of Cricopharyngeal Muscle Intervention for Cricopharyngeal Dysfunction

Author	Year	Study (Level of Evidence)	No. of Patients	Outcome Measures	Type of Treatment	Follow-up	Success Rate (%)	Complication
Schneider[24]	1994	PCS (IV)	7	SR, VFSS, M	BoT-A	4.5–5 months	71	None
Blitzer[18]	1997	RCS (IV)	6	C	BoT-A	Not mentioned	100	Not mentioned
Alberty[4]	2000	PCS (IV)	10	SR, VFSS	BoT-A	1–1.5 months	100	Pharyngeal diffusion
Haapaniemi[34]	2001	RCS (IV)	4	C	BoT-A	2–24 months	75	Urine retention
Shaw[35]	2001	RCS (IV)	12	C	BoT-A	1–14 months	83	Pharyngeal diffusion, pharyngeal tear
Parameswaran[33]	2002	RCS (IV)	12	C, VFSS	BoT-A	3–6 months	92	Neck cellulitis
Zaninotto[25]	2004	PCS (IV)	21	SR, VFSS	BoT-A	12–38 months	43	Death through aspiration
Murry[26]	2005	PCS (IV)	13	C, VFSS	BoT-A	1–9 months	92	Not mentioned
Terre[27]	2008	PCS (IV)	10	C, VFSS, M	BoT-A	12 months	80	None
Lee[36]	2009	RCS (IV)	8	VFSS	BoT-A	0.2–1 month	75	Not mentioned
Alfonsi[28]	2010	PCS (IV)	34	SR	BoT-A	2 months	50	None
Woisard-Bassols[29]	2013	PCS (IV)	11	SR, VFSS	BoT-A	12–48 months	45	None
St. Guily[37]	1994	RCS (IV)	11	C	Myotomy (open)	5–53 months	72	None
Herberhold[38]	1995	RCS (IV)	32	C, VFSS	Myotomy (endoscopic)	Up to 7 years	97	Supraglottic edema, imminent mediastinitis
Poirier[14]	1997	RCS (IV)	40	C, VFSS, M	Myotomy (open)	1–255 months	72.5	Retropharyngeal hematoma
Ali[31]	1997	Cohort study (IIb)	8	C	Myotomy (open)	6 weeks	75	Not mentioned
Halvorson[30]	1998	PCS (IV)	18	C	Myotomy (endoscopic)	Not mentioned	78	Not mentioned
Mason[39]	1998	RCS (IV)	31	C, M	Myotomy (open)	2–48 months	77	Neck hematoma, pulmonary edema, pneumonia
Lawson[40]	2003	RCS (IV)	29	C, VFSS, FEES	Myotomy (endoscopic)	1–36 months	88	None
Zaninotto[25]	2004	PCS (IV)	11	SR, VFSS	Myotomy (open)	6–31 months	73	None
Takes[41]	2005	RCS (IV)	10	SR	Myotomy (endoscopic)	2–24 months	60	None
Dauer[2]	2006	RCS (IV)	22	SR	Myotomy (endoscopic + open)	Not mentioned	58	Fever of unknown etiology, chest pain, pharyngocutaneous fistula, tracheotomy
Munoz[5]	2007	RCS (IV)	14	SR, VFSS	Myotomy (open)	6–10 months	25	Not mentioned
Lawson[42]	2008	RCS (IV)	31	SR, VFSS, FEES	Myotomy (endoscopic)	12–23 months	64.5	None
Kos[9]	2010	RCS (IV)	28	VFSS, M	Myotomy (open)	2.5–203 months	79	Fever, aspiration pneumonia, mucosal perforation
Ozgursoy[10]	2010	RCS (IV)	14	SR, VFSS, MF	Myotomy (endoscopic)	6 months	100	Not mentioned
Bachy[32]	2013	PCS (IV	32	SR	Myotomy (endoscopic)	6–99 months	84	Severe bleeding
Lim[6]	1995	RCS (IV)	40	C, VFSS	Myotomy (endoscopic)	2–22 months	90	Esophageal perforation
Ali[31]	1997	Cohort study (IIb)	12	C	Dilatation (Savary)	6 weeks	58	Not mentioned
Hatlebakk[8]	1998	PCS (IV)	10	SR, M	Dilatation (Savary)	6–20 months	90	Not mentioned
Solt[43]	2001	RCS (IV)	5	C, VFSS, M	Dilatation (balloon)	7–33 months	100	Superficial mucosal splitting
Wang[44]	2005	RCS (IV)	6	SR	Dilatation (balloon + French)	8–27 months	100	None
Clary[21]	2011	RCS (IV)	42	SR	Dilatation (French)	Up to 72 months	64	Partial mucosal tears, laryngospasm
Dou[22]	2012	PCS (IV)	38	SR, VFSS	Dilatation (water-inflated balloon)	3–5 months	76.3	None

BoT-A = botulinum toxin serotype A; C = clinical; FEES = flexible endoscopic evaluation of swallowing; M = manometry; MF = manofluorography; PCS = prospective case series; RCS = retrospective case series; SR = self-rating; VFSS = videofluoroscopic swallowing study.

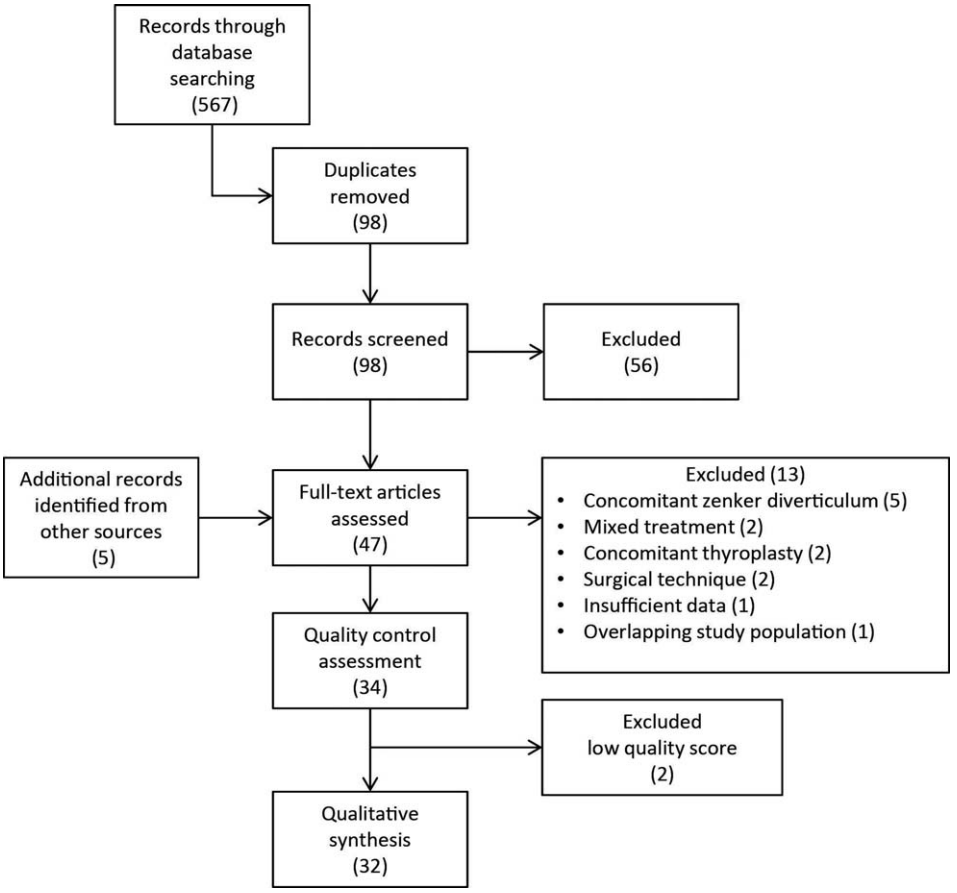

Fig. 1. Flow diagram of the search strategy.

following reasons: 1) surgical technique descriptions (two articles); 2) duplicate and overlapping study populations (one article); 3) insufficient data available to calculate the success rate of the procedure (one article); 4) patients with CP dysfunction besides Zenker's diverticulum over 1.5 cm and/or requiring diverticulectomy (five articles); 5) patients underwent concomitant thyroplasty with BoT injection or myotomy (two articles); 6) patients underwent BoT injection at the same time with myotomy or dilatation (two articles). Studies ranged from 10 to 20 of 25 points on the Downs and Black scale. Two articles receiving a score below 13 were also excluded from the evaluation (Fig. 1).

Twelve studies[4,8,22,24–32] were prospective and 20 [2,5,6,9,10,14,18,21,33–44] were retrospective. All of the publications were observational studies, with a level of evidence of IV; with the exception of one prospective cohort study (IIb) (Table II). All articles except for two dealt with one type of therapeutic procedure; the two exceptions each assessed two procedure types. In these two articles, the authors used one type of procedure for each patient and reported on the success rates and complications of the procedures separately.

Assessment of Success Rates and Complications Between BoT, Dilation, and Myotomy

Of the 32 articles, 12 articles reported on the success rates and complications of BoT injections, six articles on dilation, and 16 articles on myotomy. The range of reported success rates were between 43% and 100% for BoT injections (crude average =76%, patient-weighted average = 69%), between 58% and 100% for dilation (crude average = 81%, patient-weighted average = 73%), and between 25% and 100% for myotomy (crude average = 75%, patient-weighted average = 78%) (Table III).

Patient questionnaires, type of diet tolerated, clinical score of swallowing impairment, and flexible endoscopy had been used for the measurement of success rate in the majority of the articles. In some of the articles, a retrospective review of VFSS had been the choice as an objective tool.

The reported complication rates were between 0% and 25% for BoT injections (crude average = 5%, patient-weighted average = 4%), between 0% and 20% for dilation (crude average = 5%, patient-weighted average = 5%), and between 0% and 39% for myotomy (crude average = 6%, patient-weighted average = 7%) (Table IV). These included pharyngocutaneous fistula, pharyngeal tear, supraglottic edema, imminent mediastinitis, neck cellulitis, retropharyngeal hematoma, neck hematoma, esophageal perforation, laryngospasm, severe bleeding, and death through aspiration.

In logistic regression analysis of the patient-weighted averages, the 78% success rate with myotomy was significantly higher than the 69% success rate with BoT injections ($P = .042$), whereas the success rate of

TABLE III.
Distribution of Success Rates of BoT Injection, Dilation, and Myotomy.

	No. of Articles	Range of Success Rates (Crude Average)	No. of Patients (Sum)	No. of Successes (Sum)	Patient-Weighted Average Success Rate
BoT Injection	12	43%–100% (76%)	148	102	69%
Dilation	6	58%–100% (81%)	113	83	73%
Myotomy	16	25%–100% (75%)	369	286	78%

73% with dilation was not significantly different from that of either myotomy ($P = .37$) or BoT ($P = .42$).

Upon scoring the procedures for invasiveness as BoT injection = low, dilation = medium, and myotomy = high, there was a positive and statistically significant trend favoring increased success rate with increased invasiveness ($P = .039$). In contrast, we found no significant difference in complication rates between procedures via logistic regression, and no significant trend in complication rate with invasiveness via trend analysis.

Subgroup Analysis of Myotomy Procedures

A subgroup analysis was performed to assess the success and complication rates of open versus endoscopic myotomy. For this purpose, one study that used both methods was excluded. There were eight articles reporting outcomes of endoscopic myotomy, whereas seven evaluated open myotomy. Success rates ranged between 60% and 100% with endoscopic myotomy (crude average = 83%, patient-weighted average = 84%) compared to 73% and 79% with open myotomy (crude average = 68%, patient-weighted average = 71%). Comparison of success rates via logistic regression analysis revealed a significant increase in odds of success with the endoscopic procedure (ratio = 2.24, $P = .0025$). Complication rates were reported between 0% and 6% for endoscopic myotomy (crude average = 2%, patient-weighted average = 2%) versus 0% and 39% for open myotomy (crude average = 8%, patient-weighted average = 11%). Comparison of complication rates via logistic regression showed a significant increase in odds of complication with the open procedure (odds ratio = 5.01; $P = .0021$). Brief details of complications were mentioned in Table I.

Subgroup Analysis of BoT Injections

Botulinum toxin units were often reported as a range. We used the midpoint of the BoT unit range in analyzing success and complications rates. Logistic regression analysis indicated that a 20-unit increase in the midpoint BoT dose significantly increased the odds of success (odds ratio = 1.26, $P = .033$) without significantly changing the odds of complication (odds ratio = 0.74, $P = .33$).

DISCUSSION

CP dysfunction can present with various symptoms, often not fitting a common pattern. Patient complaints vary in severity from a lump sensation to complete inability to swallow and life-threatening aspiration. The workup varies among institutions, and there is no agreed on, uniform preoperative or postoperative evaluation technique. Similarly, because outcomes are generally not reported through objective measures, there is continued debate on the best surgical technique and the selection of suitable patients. The aim of this systematic review was to assess the success rates of myotomy, CP dilatation, and botulinum toxin injection in the management of CP dysfunction.

Kaplan is credited for performing the first CP myotomy in 1951 on a patient with bulbar poliomyositis.[45] Varying methods of transcervical myotomy have been described since then as can be seen in Table II. It can be noted that the majority of the articles were on the effectiveness of myotomy (seven papers on open myotomy, eight on endoscopic, and one comparing the two methods) in the management of CP dysfunction. We found the average success rate of myotomy to be 75%, and it was significantly higher than BoT injections ($P = .042$) but not statistically different than dilatation ($P = .37$). The average complication rate of 6% (range = 0%–39%) was not significantly higher than the other methods. Interestingly, myotomy outcomes were significantly better with the endoscopic technique (odds ratio = 2.24), supplemented with the advantage of decreased complication rates ($P = .0021$). Although the risk of mediastinitis and fistula could not be completely excluded by endoscopic laser myotomy, limiting the procedure to the fibers of the cricopharyngeus muscle considerably reduced it.[2,6,10,30,32,38,40–42] Also, any injury to the

TABLE IV.
Distribution of Complications of BoT Injection, Dilation, and Myotomy

	No. of Articles	Range of Complication Rates (Crude Average)	No. of Patients (Sum)	No. of Complications (Sum)	Patient-Weighted Average Complication Rate
BoT Injection	12	0%–25% (5%)	148	6	4%
Dilation	6	0%–20% (5%)	113	6	5%
Myotomy	16	0%–39% (6%)	369	27	7%

recurrent laryngeal nerve is avoided, and the postoperative course is significantly shortened with minimal pain and quick return to swallowing when endoscopic technique can be employed.[30]

The reported articles include patients with various etiologies. Mason et al. reported that the results of myotomy were excellent or good in patients with no discernible (idiopathic) underlying disease, but were not as good in patients with neuropathic or myopathic disease. They also evaluated the role of preoperative manometry and noted that the only factor predicting the success of the procedure, other than the etiology of the disorder, was impaired sphincter opening during manometry (odds ratio = 8.4). They went on to suggest that the most important manometric marker was the absence of the subatmospheric intrasphincteric pressure drop. They concluded that, when combined with an increased intrabolus pressure, the mechanical indicators that the procedure should work are present. Mason et al. also modified the procedure where they divide the sternohyoid and omohyoid muscles (depressors of the hyoid) to improve laryngeal elevation.[39] On the other hand, Poirier et al., in their 40-patient series with a neurogenic origin, reported success if the following criteria were fulfilled: 1) normal voluntary deglutition, 2) adequate tongue movement, 3) intact laryngeal function and phonation, and 4) absence of dysarthria.[14] Kos et al. also report the etiology of the dysphagia to be the most important prognostic factor. The patients with no apparent cause of dysphagia or with non–cancer-related iatrogenic oropharyngeal dysphagia showed 100% improvement. The outcomes in patients with central nervous system damage and extensive head and neck cancer therapy were not as rewarding (25% success rates). Their group also challenged the absence of hypopharyngeal contractions as a contraindication to surgery. In their series, although 71% of the patients with normal constrictor activity showed improvement, 79% with reduced and 71% with absent activity also showed successful outcomes following myotomy.[9] This was also advocated by Ozgursoy and Salassa, and Bammer et al., who reported improved swallowing in patients with weak pharyngeal driving forces.[10,46]

Botulinum toxin injections have been used as a test to determine whether myotomy would be effective.[18] On the other hand, Zaninotto et al. reported success with myotomy even in patients who failed BoT injections, and suggest it should not be used to discriminate between patients who may or may not benefit from surgery.[25]

There is also disagreement between authors on the necessary postoperative studies. Most outcomes are reported on subjective patient improvement. This limits our ability to uniformly compare studies and reported outcomes.

There were fewer studies reporting on the efficacy of CP dilatation. The main advantages include being less invasive and ability to be performed under sedation. This makes it a suitable alternative in patients who cannot undergo general anesthesia along with electromyography-(EMG)-guided in-office BoT injections. Ali et al. performed the only study comparing myotomy and dilatation outcomes. They operated on 20 patients, 12 of whom underwent dilatation and eight myotomy. The patient selection was dictated by clinical circumstances and patient preference, with the exception of patients demonstrating manometric failure of UES relaxations. All of these patients underwent myotomy. They clinically evaluated the patients 6 weeks postoperatively. They had an overall response rate of 65%; 75% of the patients undergoing myotomy and 58% of the patients undergoing dilatation had responded. Unfortunately, when reporting outcomes, they did not differentiate between the two groups.[31] Hatlebakk et al. reported that nine out of the 10 patients remained on an oral diet at 13 months, following dilatation with 18 to 20 mm Savary dilators. On manometry, UES pressures were significantly reduced, and/or the duration and completeness of relaxation increased following dilatation.[8] Solt et al. reported similar improvement in patients without organic stenosis of the UES, with redilatation needed in one patient (out of five) at 21 months.[43] Wang et al. also used dilatation for patients with CP dysfunction that could only be attributed to a CP bar and reported complete response.[44] Clary et al. suggested CP bougie dilatation as a first surgical step. They advocate this two-step approach for two reasons: 1) if dysphagia resolves, the patient can avoid a more morbid myotomy, and 2) if patient experiences no relief, it can suggest a need for further workup to evaluate other causes of dysphagia.[21]

Since the first report of BoT injections for CP dysfunction by Schneider, many have advocated the use of it due to the minimal invasiveness of the procedure, ability to perform in the clinic with EMG guidance, and minimal morbidity.[24] The effective duration varies on the injected site, dosag, and type of disease.[27] Most studies have reported doses between 5 and 50 units[4,18,34] up to 100 units.[20] The maximum duration of the beneficial effects continues to be studied. Terre et al. reported improvement up to a year with a single 100-U injection. They attributed this to the reduction of basal UES pressure, with a subsequent increase in pharyngeal pressure that permitted improvement in sphincter relaxation, as well as the achieved oral diet permitting the strengthening of swallowing musculature.[27] Although Terre et al. recommended BoT injections for patients who had incomplete relaxation of the CP muscle with a certain degree of pharyngeal propulsion, Woisard-Bassols et al. reported good outcomes in patients with CP dysfunction and pharyngolaryngeal weakness.[29] Our review found that BoT injections are not as successful as myotomy, and as the invasiveness of the procedure increased (BoT = low, dilatation = medium, myotomy = high), there was a statistically significant trend favoring increased success rates.

This systematic review has several limitations. Primarily, retrospective chart review studies and prospective cohort studies are subject to selection bias; therefore, the level of evidence provided by this review relies on the strength of the individual articles. The surgeons may select a patient to undergo a particular procedure based on CP dysfunction etiology, patient comorbidities, and surgeon experience. Patients are also allowed to choose the treatment based on recommendations. In CP dysfunction, there is also no universally

agreed on algorithm for management, including preoperative diagnostic testing and patient selection criteria for surgical approach. Due to this, we aimed to only evaluate surgical outcomes. This study is also limited in regard to making recommendations on patient selection for a particular surgical method. Furthermore, the studies reviewed reported outcomes with various methods, relying heavily on self-rating and clinical improvement. Similarly, due to the nature of the disease and infrequency, the largest series in this review included 42 patients. Nevertheless, we believe our data improves our understanding of the surgical management techniques for CP dysfunction and can serve as a starting point for future, well-designed, multicenter prospective trials.

CONCLUSION

In the current systematic review, logistic regression analysis of patient-weighted averages revealed significantly higher success rates with myotomy compared to BoT injections. Although the success rates of dilatation were not found to be significantly different from BoT injections or myotomy, there were also fewer studies assessing myotomy. There was no significant difference in regard to complication rates, and the effectiveness of the procedures improved as the invasiveness increased. As a result, in the well-selected patient, all of these procedures can be employed with good outcomes and minimal morbidity.

BIBLIOGRAPHY

1. Ho AS, Morzaria S, Damrose EJ. Carbon dioxide laser-assisted endoscopic cricopharyngeal myotomy with primary mucosal closure. *Ann Otol Rhinol Laryngol* 2011;120:33–39.
2. Dauer E, Salassa J, Luga L, Kasperbauer J. Endoscopic laser vs open approach for cricopharyngeal myotomy. *Otolaryngol Head Neck Surg* 2006;134:830–835.
3. Lim KG. The mouth of the esophagus. *Laryngoscope* 1907;17:421–428.
4. Alberty J, Oelerich M, Ludwig K, et al. Efficacy of botulinum toxin A for treatment of upper esophageal sphincter dysfunction. *Laryngoscope* 2000;110:1151–1156.
5. Munoz AA, Shapiro J, Cuddy LD, et al. Videofluoroscopic findings in dysphagic patients with cricopharyngeal dysfunction: before and after open cricopharyngeal myotomy. *Ann Otol Rhinol Laryngol* 2007;116:49–56.
6. Lim RY. Endoscopic CO_2 laser cricopharyngeal myotomy. *J Clin Laser Med Surg* 1995;13:241–247.
7. McKenna JA, Dedo HH. Cricopharyngeal myotomy: indications and technique. *Ann Otol Rhinol Laryngol* 1992;101:216–221.
8. Hatlebakk JG, Castell JA, Spiegel J, et al. Dilatation therapy for dysphagia in patients with upper esophageal sphincter dysfunction-manometric and symptomatic response. *Dis Esophagus* 1998;11:254–259.
9. Kos MP, David EF, Klinkenberg-Knol EC, Mahieu HF. Long-term results of external upper esophageal sphincter myotomy for oropharyngeal dysphagia. *Dysphagia* 2010;25:169–176.
10. Ozgursoy OB, Salassa JR. Manofluorographic and functional outcomes after endoscopic laser cricopharyngeal myotomy for cricopharyngeal bar. *Otolaryngol Head Neck Surg* 2010;142:735–740.
11. Olsson R, Nilsson H, Ekberg O. An experimental manometric study simulating upper esophageal sphincter narrowing. *Invest Radiol* 1994; 29:630–635.
12. Salassa JR, DeVault KR, McConnel FM. Proposed catheter standards for pharyngeal manofluorography (videomanometry). *Dysphagia* 1998;3: 105–110.
13. Leonard R, Belafsky PC, Rees CJ. Relationship between fluoroscopic and manometric measures of pharyngeal constriction: the pharyngeal constriction ratio. *Ann Otol Rhinol Laryngol* 2006;115:897–901.
14. Poirier NC, Bonavina L, Taillefer R, et al. Cricopharyngeal myotomy for neurogenic oropharyngeal dysphagia. *J Thorac Cardiovasc Surg* 1997; 113:233–240.
15. Ertekin C, Aydogdu I, Yuceyar N, et al. Electrodiagnostic methods of neurogenic dysphagia. *Electroencephalogr Clin Neurophysiol* 1998;109: 331–340.
16. Elidan J, Shochina M, Gonen B, et al. Electromyography of the inferior constrictor and cricopharyngeal muscles during swallowing. *Ann Otol Rhinol Laryngol* 1990;90:466–469.
17. Halvorson DJ, Kuhn FA. Transmucosal cricopharyngealmyotomy with the potassium-titanyl-phosphate laser in the treatment of cricopharyngeal dysmotility. *Ann Otol Rhinol Laryngol* 1994;103:173–177.
18. Blitzer A, Brin MF. Use of botulinum toxin for diagnosis and management of cricopharyngeal achalasia. *Otolaryngol Head Neck Surg* 1997;116: 328–330.
19. Verhulst J, Ame P, Guatterie M, Barat M, Traissac L. A retrospective study of 91 injections of botulinum toxin into the upper sphincter of the esophagus. *Rev Laryngol Otol Rhinol* 2003;124:315–320.
20. Kim DY, Park CI, Ohn SH, et al. Botulinum toxin type A for poststroke cricopharyngeal muscle dysfunction. *Arch Phys Med Rehabil* 2006;87: 1346–1351.
21. Clary MS, Daniero JJ, Keith SW, et al. Efficacy of large-diameter dilatation in cricopharyngeal dysfunction. *Laryngoscope* 2011;1221:2521–2525.
22. Dou Z, Zu Y, Wen H, et al. The effect of different catheter balloon dilatations modes on cricopharyngeal dysfunction in patients with dysphagia. *Dysphagia* 2012;27:514–520.
23. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377–384.
24. Schneider I, Potatschnig C, Thomfort WF, Eckel HE. Treatment of dysfunction of the CPM with botulinum A toxin: introduction of a new, noninvasive method. *Ann Otol Rhinol Laryngol* 1994;103:31–35.
25. Zaninotto G, Ragona RM, Briani C, et al. The role of botulinum toxin injection and upper esophageal sphincter myotomy in treating oropharyngeal dysphagia. *J Gastrointest Surg* 2004;8:997–1006.
26. Murry T, Wasserman T, Carrau RL, Castillo B. Injection of botulinum toxin A for the treatment of dysfunction of the upper esophageal sphincter. *Am J Otolaryngol* 2005;26:157–162.
27. Terre R, Valles M, Panades A, Mearin F. Long-lasting effect of a single botulinum toxin injection in the treatment of oropharyngeal dysphagia secondary to upper esophageal sphincter dysfunction: a pilot study. *Scand J Gastroenterol* 2008;43:1296–1303.
28. Alfonsi E, Merlo IM, Ponzio M, Montomoli C, et al. An electrophysiological approach to the diagnosis of neurogenic dysphagia: implications for botulinum toxin treatment. *J Neurol Neurosurg Psychiatry* 2010;81:54–60.
29. Woisard-Bassols V, Alshehri S, Simonetta-Moreau. The effects of botulinum toxin injections into the cricopharyngeus muscle of patients with cricopharyngeus dysfunction associated with pharyngo-laryngeal weakness. *Eur Arch Otorhinolaryngol* 2013;270:805–815.
30. Halvorson DJ. The treatment of cricopharyngeal dysmotility with a transmucosal cricopharyngeal myotomy using the potassium-titanly-phosphate (KTP) laser. *Endoscopy* 1998;30:46–50.
31. Ali GN, Wallace KL, Laundl TM, et al. Predictors of outcome following cricopharyngeal disruption for pharyngeal dysphagia. *Dysphagia* 1997; 12:133–139.
32. Bachy V, Matar N, Remacle M, et al. Long-term functional results after endoscopic cricopharyngeal myotomy with CO_2 laser: a retrospective study of 32 cases. *Eur Arch Otorhinolaryngol* 2013;270:965–968.
33. Parameswaran MS, Soliman AMS. Endoscopic botulinum toxin injection for cricopharyngeal dysphagia. *Ann Otol Rhinol Laryngol* 2002;111:871–874.
34. Haapaniemi JJ, Laurikainen EA, Pulkkinen J, Marttila RJ. Botulinum toxin in the treatment of cricopharyngeal dysphagia. *Dysphagia* 2001; 16:171–175.
35. Shaw GY, Searl JP. Botulinum toxin treatment for cricopharyngeal dysfunction. *Dysphagia* 2001;16:161–167.
36. Lee SY, Seo HG, Paik NJ. Botulinum toxin injection for dysphagia: a blinded retrospective videofluoroscopic swallowing study analysis. *Am J Phys Med Rehabil* 2009;88:491–494.
37. St. Guily JL, Perie S, Willig TN, et al. Swallowing disorders in muscular diseases: functional assessment and indications of cricopharyngeal myotomy. *Ear Nose Throat J* 1994;73:34–40.
38. Herberhold C, Walther EK. Endoscopic laser myotomy in cricopharyngeal achalasia. *Adv Otorhinolaryngol* 1995;49:144–147.
39. Mason RJ, Bremner CG, DeMeester TR. Pharyngeal swallowing disorders selection for and outcome after myotomy. *Ann Surg* 1998;228:598–608.
40. Lawson G, Remacle M, Jamart J, Keghian J. Endoscopic CO_2 laser-assisted surgery for cricopharyngeal dysfunction. *Eur Arch Otorhinolaryngol* 2003;260:475–480.
41. Takes RP, Hoogen FJA, Marres HAM. Endoscopic myotomy of the cricopharyngeal muscle with CO_2 laser surgery. *Head Neck* 2005;27: 703–709.
42. Lawson G, Remacle M. Ins and outs of myotomy of the upper esophageal sphincter in swallowing disorders. *B-ENT* 2008;10:83–89.
43. Solt J, Bajor J, Moizs M, et al. Primary cricopharyngeal dysfunction: treatment with balloon catheter dilatation. *Gastrointest Endosc* 2001;6: 767–771.
44. Wang AY, Kadkade R, Kahrilas PJ, Hirano I. Effectiveness of esophageal dilation for symptomatic cricopharyngeal bar. *Gastrointest Endosc* 2005; 61:148–152.
45. Kaplan S. Paralysis of deglutition, a post-poliomyelitis complication treated by section of the cricopharyngeus muscle. *Ann Surg* 1951;133: 572–573.
46. Bammer T, Salassa JR, Klingler PJ. Comparison of methods for determining cricopharyngeal intrabolus pressure in normal patients as possible indicator for cricopharyngeal myotomy. *Otolaryngol Head Neck Surg* 2002;127:299–308.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(3):488-493.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–
Head and Neck Surgery
2015, Vol. 152(3) 488–493
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599814565599
http://otojournal.org
SAGE

Esophageal Visualization as an Adjunct to the Videofluoroscopic Study of Swallowing

Anna Miles, PhD[1], Jessica McMillan, MSc [1], Katie Ward, MSc[2], and Jacqui Allen, MBChB, FRACS[1,2]

No sponsorships or competing interests have been disclosed for this article.

Received August 13, 2014; revised October 23, 2014; accepted December 4, 2014.

Abstract

Objective. Complaints of dysphagia for solids lead to speech-language pathology (SLP) referral. Yet many of these patients are later diagnosed with esophageal rather than oropharyngeal dysphagia. Fluoroscopic screening involving the oropharynx alone fails to identify these patients. The aim of this study was to investigate the prevalence of esophageal abnormalities in an SLP-led videofluoroscopic study of swallowing (VFSS) clinic.

Study Design. Prospective, observational study.

Setting. Radiology suite, public hospital.

Subjects and Methods. In total, 111 consecutive mixed-etiology patients referred to the clinic by otorhinolaryngology (ORL) (59) or by a speech-language pathologist (52) were recruited. A VFSS was performed according to protocol, and at completion, esophageal visualization (in anterior-posterior plane) was performed by administration of a large liquid barium bolus and a barium capsule. All VFSS recordings were analyzed using objective digital measures of timing and displacement.

Results. Sixty-eight percent of patients had an abnormal esophageal transit. One-third of those referred presented exclusively with esophageal abnormalities, while one-third had both oropharyngeal and esophageal abnormalities. Oral abnormalities, reduced pharyngoesophageal segment maximum opening (PESmax), and increasing age were significantly associated with esophageal abnormalities.

Conclusion. Fluoroscopic evaluation of the pharynx alone, without esophageal review, risks incomplete diagnosis of patients with esophageal disorders. Using esophageal visualization allows timely referral for further investigation by appropriate medical specialties, avoiding incomplete management of patients with dysphagia.

Keywords

deglutition, deglutition disorders, dysphagia, esophageal visualization, speech-language pathology, otorhinolaryngology

Associations between oropharyngeal abnormalities and esophageal abnormalities are poorly understood but well documented.[1-3] Oropharyngeal alterations have been reported in patients with gastroesophageal reflux disease.[4] Neurologic diseases such as Parkinson disease[5,6] and systemic conditions such as scleroderma[7] lead to both oropharyngeal and esophageal abnormalities. In a recent study using high-resolution manometry, O'Rourke and colleagues[8] describe a variety of esophageal alterations during voluntary pharyngeal maneuvers (effortful swallow and Mendelsohn), adding to the theory that changing one point in the swallowing system can lead to positive or negative changes elsewhere. In addition, patient accuracy in locating the level of bolus holdup has been shown to be poor, with patients often indicating the cervical region or levels more proximal than the true site of bolus stasis, particularly when this occurs in the esophagus.[1,9,10] Smith and colleagues[10] reported 57% of respondents located a solid bolus impacted at a distal esophageal ring to the level of the sternal notch.

Complaints of dysphagia for solids regularly lead to speech-language pathology (SLP) referral rather than gastroenterology or otorhinolaryngology (ORL). Traditionally, SLP-led videofluoroscopic study of swallowing (VFSS) has assessed the oropharynx exclusively, even when symptoms might suggest esophageal complaints. This results in failure to identify patients with esophageal problems. These patients are sent home with no diagnosis and either continue to manage their symptoms alone or undergo a variety of other diagnostic tests over a prolonged period of time before reaching correct diagnosis and treatment. An esophageal screen was described and validated by Allen and colleagues in 2012.[11] They compared fluid esophageal screens with full esophagrams in 74 mixed-

[1]The University of Auckland, Auckland, New Zealand
[2]Waitemata District Health Board, Auckland, New Zealand

Corresponding Author:
Anna Miles, PhD, Speech Science, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.
Email: a.miles@auckland.ac.nz

etiology patients. Sensitivity of the screen was 63% with 100% specificity. Sensitivity increased to 71% with the inclusion of a barium capsule. In comparison to the esophagram, the esophageal screen subjects patients to approximately 10 times less radiation dose and is relatively quickly completed at the end of a standard VFSS procedure.[11] The aim of this prospective, observational study was to investigate the prevalence of esophageal abnormalities in an SLP-led VFSS clinic.

Methods

This study received appropriate regional ethics approval (University of Auckland Human Participants Ethics Committee 9263).

Participants

Data from 111 inpatients and outpatients consecutively referred to an SLP-led VFSS clinic (between May and December 2013) were collected. Patients were referred by the ORL service (59) or by another speech-language pathologist (52), with a mean (SD) cohort age of 71 (14.95) years (range, 20-95 years). Fifty-one patients were male (46%). Referrals were all for complaints of dysphagia attributed to mixed etiologies: 36 neurological (32%), 37 dysphagia of unknown cause (33%), 28 otorhinolaryngology (eg, head and neck cancer [25%]), and 10 other (eg, unwell elderly [9%]). Participants were excluded if their VFSS recording had no esophageal visualization or had no timer or no calibration ring, since these were required to complete the timing and displacement measures.

Esophageal Videofluoroscopic Assessment

Studies were performed in a radiology suite using a Videofluoroscope (Toshiba, Tokyo, Japan) and recorded at 30 frames per second onto a USB drive. Timing information was superimposed on the fluoroscopic recording in 100ths of a second using a Horita VS-50 Video Stopwatch (Horita, Capistrano Beach, California). A 19-mm diameter radioopaque ring was taped to the patient's chin (in the lateral plane) and shoulder (in the anterior-posterior plane) to allow calibration for displacement measures. A medical radiation technician (MRT) and a speech-language pathologist were present at all procedures.

For the standard VFSS protocol, the patient was screened in the lateral plane. The patient was presented with 1 mL, 3 mL, and then 20 mL of thin barium (E-Z Paque, E-Z-EM Anjou, Canada; 100% w/v) followed by half a cup of thin barium through a straw. The patient was then given 3 mL of barium paste (E-Z paste, E-Z-EM; 60% w/w). The procedure was truncated if required for patient safety.

The esophageal phase involved the introduction of 2 boluses after completion of the standard VFSS protocol: a 20-mL fluid bolus and a 13-mm barium capsule. The patient was positioned in the anterior-posterior plane, standing whenever possible. The patient was asked to ''swallow all in one go'' to avoid deglutitive inhibition. The MRT followed the bolus from the oral cavity through the lower esophageal sphincter (LES) until clearance into the stomach. Screening was continued for up to 15 seconds. If there was still residue in the esophagus, screening was ceased for 15 seconds, then recommenced. If residue was still present, the patient was asked to take a dry swallow to see if clearance occurred. If residue was still present, the patient was then offered a water swallow as well as being asked if he or she could feel any remaining bolus. A screen shot was used to identify complete clearance.

Data Collected

Age, sex, and comorbidities were recorded for each patient. Each VFSS was analyzed using real-time and frame-by-frame viewing (Quicktime Media Player; Apple, Cupertino, California). Videos were scored for the presence of oral, pharyngeal, and esophageal abnormalities (yes/ no). Prolonged bolus manipulation, anterior spillage from lips, premature spillage into the pharynx, and oral residue were all considered oral abnormalities. Nasal regurgitation, reduced epiglottic deflection, delay in swallowing initiation, pharyngeal residue, penetration, aspiration, and prolonged pharyngeal transit time were considered pharyngeal abnormalities. Esophageal bolus stasis, bolus redirection/intraesophageal reflux, gastroesophageal reflux, esophagopharyngeal reflux, hiatal hernia, prolonged esophageal transit, and pill stasis were all considered esophageal abnormalities. Pharyngeal transit time (PTT) was recorded and translated into a binary measure of (1) within normal limits vs (2) more than 2 standard deviations (SD) outside of normal limits.[12] Esophageal transit time (ETT) was also recorded. A conservative cutoff of over 15 seconds was selected for abnormality. Previous published work has defined normal liquid transit through the esophagus as less than 13 seconds.[11,13,14] Maximum penetration-aspiration scale (PAS) scores were recorded, and scores 6 and above were considered an aspiration event.[15] To explore whether esophageal abnormalities can be predicted by objective pharyngeal measures, we calculated the pharyngeal constriction ratio (PCR)[12] and pharyngoesophageal segment maximum opening (PESmax)[12] using the Universal Desktop Ruler (AVPSoft). These measures were also translated into binary measures of (1) within normal limits vs (2) more than 2 SD outside of normal limits.[12] All measures were taken from the largest fluid bolus ingested.

Data Analysis

Swallow studies were reported by an experienced otolaryngologist, specializing in dysphagia management, and by a speech-language pathologist, trained in quantitative analysis of VFSS using the method developed by Leonard and Kendall.[12] Interrater reliability for all measures was calculated on 30% of videos by a third researcher. Videos were randomly selected by a fourth researcher. The third rater was blinded to the first researcher's scoring and patient etiology. Total agreement across measures was 98%, with a κ coefficient of 0.92. Lack of agreement was found for 1 PAS score (1 vs 2), and although there was slight variance in PESmax (maximum variance .08) and PCR (maximum variance .07), this did not change binary measures of within

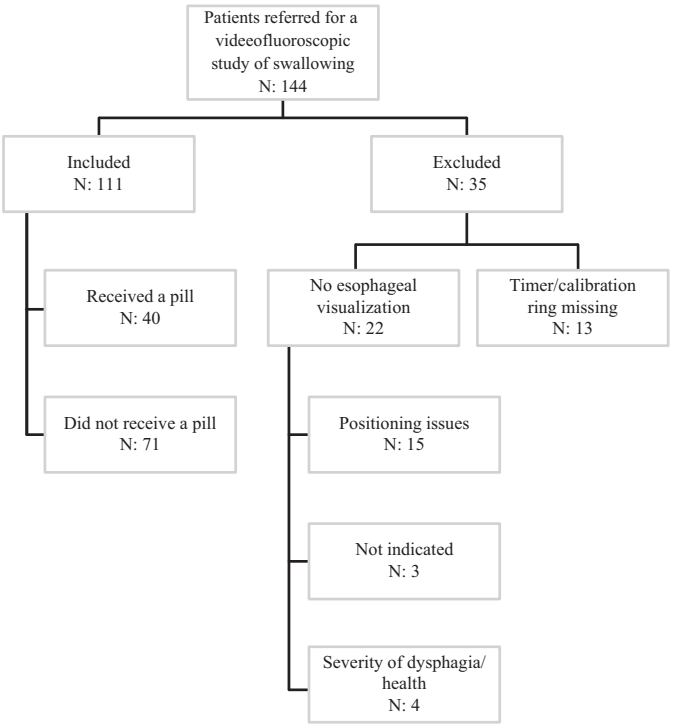

Figure 1. Recruitment inclusion.

Table 1. Frequency of Swallowing Abnormalities in Full Cohort.

Swallowing Indices	Number	% of Cohort
Frequency of oral abnormalities	28	22
Frequency of pharyngeal abnormalities	50	45
Frequency of esophageal abnormalities	76	68
Oral abnormalities alone	4	4
Pharyngeal abnormalities alone	9	8
Esophageal abnormalities alone	34	31
Oral and pharyngeal abnormalities	7	6
Oral and esophageal abnormalities	9	8
Pharyngeal and esophageal abnormalities	15	14
Oral, pharyngeal, and esophageal abnormalities	19	17
No abnormalities	14	13
Frequency of PCR outside 2 SD of norm	10	9
Frequency of PESmax outside 2 SD of norm	20	18
PTT outside 2 SD of norm	19	17
ETT >15 seconds	32	29
ETT pill, >15 seconds (n = 40)	10	25

Abbreviations: ETT, esophageal transit time; PCR, pharyngeal constriction ratio; PESmax, pharyngoesophageal segment maximum opening; PTT, pharyngeal transit time; SD, standard deviation.

normal limits vs outside normal limits. Data were analyzed using SPSS version 20 (SPSS, Inc, an IBM Company, Chicago, Illinois). Descriptive statistics were used to explore the frequency of swallowing abnormalities. Correlation analyses were made using χ^2 for categorical variables and Spearman correlations for continuous variables. Multiple logistic regressions were applied to evaluate the associations between esophageal abnormalities and other clinical indices adjusted for confounding variables based on bivariate analyses (sex, age, etiology, oral abnormalities, pharyngeal abnormalities, PCR, PESmax, PTT, and aspiration event) and 2-way interactions (esophageal abnormalities and PESmax, esophageal abnormalities and oral abnormalities, and esophageal abnormalities and age). First the full model with all confounding factors was fit, and backward selection was used to select the main effect model. The 2-way interactions were then added to the main effect model one by one for the final model.

Excluded Data

Thirty-three videos were excluded from analyses. Reasons for exclusion include inadequate positioning, severe aspiration precluding completion of the study, and missing measurement devices (**Figure 1**). Excluded cases were significantly older ($P < .001$), more likely to be referred by a speech-language pathologist ($P < .001$), and more likely to have a neurologic condition ($P < .01$) than included cases.

Results

Sixty-eight percent of the 111 patients had esophageal abnormalities, with 29% of the total cohort having an esophageal transit time of >15 seconds. Esophageal phase abnormalities were the most common: 68% vs 45% (pharyngeal) and 34% (oral) (**Table 1**). In addition, it was rare for patients to present with isolated oral phase or pharyngeal phase abnormalities (only 4% and 8%, respectively), but one-third of patients demonstrated only esophageal phase abnormalities (**Table 1**).

Referral Sources

There were significant differences in age, sex, etiology, and swallowing indices between referral sources (**Table 2**). However, frequency of esophageal abnormalities was not significantly different (ORL, 73%; SLP, 63%).

Associations with Esophageal Abnormalities

Table 3 displays the associations between esophageal abnormalities and prolonged ETT, with other clinical indices. There was a significant association between esophageal abnormalities and reduced PESmax, with 90% of patients with reduced PESmax opening having concurrent esophageal abnormalities. Esophageal transit time was significantly associated with age ($\rho = .218$, $P < .05$), with an odds ratio of 2.8 for prolonged ETT if a patient was older than 65 years. Oral abnormalities were also significantly associated with prolonged ETT, with 50% of patients with prolonged ETT having oral abnormalities. However, when referral sources were separated, the association was only significant in the SLP cohort (SLP cohort, $P < .05$; ORL cohort, $P = .237$).

Table 2. Comparison of Referral Sources.

Characteristic	Frequency (% within Referral Type)			
	ORL	SLP	χ^2	*P* Value[a]
Number of files	59	52	neurological etiology	neurological etiology
Age (>65 y)	37 (63)	47 (90)	11.50	**.001**
Sex (male)	20 (33)	31 (60)	7.36	**.008**
Etiology—neurological	4 (7)	32 (62)	44.01	**.000**
Dysphagia of unknown cause	27 (46)	10 (19)		
ORL (H&N cancer, GERD)	24 (41)	4 (8)		
Other (unwell elderly, pneumonia)	4 (6)	6 (11)		
Frequency of oral abnormalities	11 (19)	27 (52)	13.60	**.000**
Frequency of pharyngeal abnormalities	10 (17)	40 (77)	40.16	**.000**
Frequency of esophageal abnormalities	43 (73)	33 (63)	1.14	.312
Frequency of PCR outside 2 SD of norm	1 (2)	9 (17)	8.26	**.006**
Frequency of PESmax outside 2 SD of norm	12 (20)	8 (36)	.340	.625
PTT outside 2 SD of norm	7 (12)	12 (23)	2.33	.139
ETT >15 seconds	13 (22)	19 (37)	2.83	.099
Aspiration event occurred (Pen-Asp score 6-8)	2 (3)	12 (23)	9.72	**.003**

Abbreviations: ETT, esophageal transit time; GERD, gastroesophageal reflux disease; H&N, head and neck; ORL, otorhinolaryngology; PCR, pharyngeal constriction ratio; Pen-Asp, penetration-aspiration; PESmax, pharyngoesophageal segment maximum opening; PTT, pharyngeal transit time; SD, standard deviation; SLP, speech-language pathology.
[a]Bolding indicates that the *P* values have reached significance.

Table 3. Associations between Esophageal Abnormalities and Other Clinical Indices.

Characteristic	Esophageal Abnormalities		Esophageal Transit Time >15 Seconds	
	χ^2	*P* Value	χ^2	*P* Value[a]
Age (>65 y)	.000	1.000	3.42	.087
Sex (male)	.006	1.000	.298	.675
Etiology	1.645	.678	2.54	.481
Frequency of oral abnormalities	.089	.826	4.96	**.030**
Frequency of pharyngeal abnormalities	.001	1.000	.446	.504
Frequency of PCR outside 2 SD of norm	.617	.474	.723	.466
Frequency of PESmax outside 2 SD of norm	4.41	**.054**	.029	1.00
PTT outside 2 SD of norm	.047	1.000	1.89	.266
ETT >15 seconds	19.15	**.000**	20.71	**.000**
Aspiration event occurred (Pen-Asp score 6-8)	.009	1.000	.370	.543

Abbreviations: ETT, esophageal transit time; PCR, pharyngeal constriction ratio; Pen-Asp, penetration-aspiration; PESmax, pharyngoesophageal segment maximum opening; PTT, pharyngeal transit time; SD, standard deviation.
[a]Bolding indicates that the *P* values have reached significance.

Discussion

As has been previously reported, there was a high frequency of esophageal abnormalities observed in this cohort of patients referred to an SLP-led VFSS clinic.[1] In fact, esophageal phase abnormalities were the most prevalent finding in all groups regardless of referral diagnosis or source. Esophageal phase abnormalities were not predicted by other phase abnormalities. If fluoroscopic screening had involved the oropharynx alone, one-third of patients would have been sent home with no diagnosis, and one-third of patients would have been treated for oropharyngeal dysphagia without consideration of concurrent esophageal disorders. In agreement with previous manometric data, slower esophageal transit times were found with increasing age.[16] It could, therefore, be hypothesized that, with the inclusion of esophageal visualization, two-thirds of patients potentially had altered clinical recommendations: different diet recommendations, different feeding strategy recommendations, and additional referrals for further investigations and medical specialty input. It may also suggest that esophageal transit time changes with age rather than being a pathologic

finding. Normative data in older adults are required to clarify this and are currently under way.

As one would expect, differences in the primary etiology causing dysphagia were noted between referral sources. However, it is clinically significant that there was no significant difference in the prevalence of esophageal abnormalities between referral sources. These results advocate for esophageal visualization irrespective of the referral source. Esophageal abnormalities appear difficult to predict by etiology. Yet, as has been previously reported, patients with impaired PESmax were more likely to have both pharyngeal[17] and esophageal abnormalities.[1] The association of reduced PESmax and prolonged ETT is hypothesized to represent a compensatory strategy. To limit retrograde flow and minimize the ''threat'' that a retained bolus presents in the esophagus, the upper esophageal sphincter hypertrophies and becomes less compliant.[1,17] This is measured as a reduction in absolute PES opening. Identification of reduction in PES opening therefore should prompt esophageal evaluation.

This study demonstrates that inclusion of esophageal visualization as part of a VFSS protocol can help identify and categorize patients' problems when referred with a symptom of dysphagia. This may allow further investigations to be requested (including formal esophagram) or referral to appropriate medical services to quantitate and characterize the esophageal disorder more thoroughly. Therapeutic recommendations can be refined and targeted to the appropriate service, and the patient receives the most complete information and holistic management.

A short visualization that adds only 2 further swallows (each screened for a maximum of 15 seconds) does not significantly increase radiation exposure or the overall study time. The additional radiation exposure incurred by performing esophageal visualization was recorded as less than 0.1 mSv. Background radiation dose annually exceeds 3 mSv, and therefore the incremental increase in exposure related to esophageal views is very low.[18-21] Compared with formal esophagram, which incorporates several additional views and longer screening time, the simple esophageal visualization may direct referral or management of the patient without exposing him or her to the higher radiation dose of a full esophagram.[11]

Concerns regarding scope of practice with regard to esophageal diagnosis have been raised. Speech-language pathologists are not expected to be esophageal diagnosticians. The visualization provides simple parameters for onward referral, as transit times longer than normal (15 seconds) can be easily measured by automated timer and indicate the need for further review. The referrer, who will decide whether further investigation is warranted and who should perform this, usually directs this. The onus will not fall on the speech-language pathologist performing the test to interpret the clinical significance of any findings. In fact, identification of an esophageal discrepancy likely to produce symptoms, particularly in the absence of other likely causes, may assist the speech-language pathologist in deciding what advice and guidance to give regarding eating strategies, rehabilitative exercises, and body positioning.

Limitations

A proportion of patients (17%) were difficult to screen due to positioning issues, limiting the view of the LES. However, these patients were generally more disabled and often wheelchair bound and may not have tolerated a full esophagram (requiring the ingestion of large quantities of barium in the prone position) either. The barium capsule was used in only 40 procedures, despite evidence that it increases the sensitivity of the screen.[11] This was thought to be due to hesitancy from speech-language pathologist to give patients with dysphagia a capsule. It is likely that capsule use is not safe for all dysphagic patients and that SLP clinical decision making is necessary in evaluating risk in each individual patient. There are no comparative normative data for pill transit times, and this is currently being investigated. PESmax was measured solely in the lateral view. The addition of an anterior-posterior measure would have provided more information regarding the extent of PES opening impairment. A measure of hyoid displacement and/or hyoid-larynx approximation would have added to the study by allowing further analysis of the cause of PES opening impairment. This was not a validation study, and no formal esophagram was performed routinely for comparison. Accuracy of esophageal abnormality detection therefore cannot be confirmed.

Conclusion

Esophageal abnormalities are highly prevalent in patients referred to a VFSS clinic with a symptom of dysphagia. One-third of patients present only with esophageal phase abnormalities. Traditional fluoroscopic screening of the oropharynx alone fails to identify these patients. Esophageal visualization is a useful adjunct to VFSS as it provides preliminary information regarding the esophageal phase of swallowing. It enables appropriate referrals to radiology, ORL, and/or gastroenterology to be made and avoids patients being falsely reassured, misdiagnosed, and mismanaged.

Acknowledgments

The authors thank the Waitemata District Health Board–Speech-Language Therapy Department for access to their clinical database.

Author Contributions

Anna Miles, study design, data collection, analysis and manuscript preparation and final approval; **Jessica McMillan**, study design, data collection, analysis and manuscript preparation and final approval; **Katie Ward**, data collection and analysis and final approval; **Jacqui Allen**, study design, analysis and manuscript preparation and final approval.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Jones B, Ravich W, Donner M, Kramer S, Hendrix T. Pharyngoesophageal interrelationships: observations and working concepts. *Gastrointest Radiol*. 1985;10:225-233.
2. Triadafilopoulos G, Hallstone A, Nelson-Abbott H, Bedinger K. Oropharyngeal and esophageal interrelationships in patients with nonobstructive dysphagia. *Dig Dis Sci*. 1992;37:551-557.
3. Gullung J, Hill E, Castell D, Martin-Harris B. Oropharyngeal and esophageal swallowing impairments: their association and the predictive value of the modified barium swallow impairment profile and combined multichannel intraluminal impedance-esophageal manometry. *Ann Otol Rhinol Laryngol*. 2012;121:738-745.
4. Mendell D, Logemann J. A retrospective analysis of the pharyngeal swallow in patients with a clinical diagnosis of GERD compared with normal controls: a pilot study. *Dysphagia*. 2002;17:220-226.
5. Castell J, Johnston BT, Colcher A, et al. Manometric abnormalities of the oesophagus in patients with Parkinson's disease. *Neurogastroenterol Motil*. 2001;13:361-364.
6. Leopold N, Kagel M. Pharyngo-esophageal dysphagia in Parkinson's disease. *Dysphagia*. 1997;12:11-18.
7. Montesi A, Pesaresi A, Cavalli M, et al. Oropharyngeal and esophageal function in scleroderma. *Dysphagia*. 1991;6:219-223.
8. O'Rourke A, Morgan L, Coss-Adame E, et al. The effect of voluntary pharnygeal swallowing maneuvers on esophageal swallowing physiology. *Dysphagia*. 2014;29:262-268.
9. Edwards D. Discriminatory value of symptoms in the differential diagnosis of dysphagia. *J Clin Gastroenterol*. 1976;5: 49-57.
10. Smith D, Ott D, Gelfand D, Chen M. Lower esophageal mucosal ring: correlation of referred symptoms with radiologic findings using a marshmallow bolus. *AJR Am J Roentgenol*. 1998;171:1361-1365.
11. Allen J, White C, Leonard R, Belafsky P. Comparison of esophageal screen findings on videofluoroscopy with full esophagram results. *Head Neck*. 2012;34:264-269.
12. Leonard R, Kendall K. *Dysphagia Assessment and Treatment Planning: A Team Approach*. 2nd ed.London, UK: Singular Publishing; 2008.
13. Torrico S, Corazziari E, Habib F. Barium studies for detecting esophagopharyngeal reflux events. *Am J Med*. 2003;115:124S-129S.
14. Barloon T, Bergus G, Lu C. Diagnostic imaging in the evaluation of dysphagia. *Am Fam Physician*. 1996;53:535-546.
15. Rosenbek J, Robbins J, Roecker E, Coyle J, Woods J. A penetration-aspiration scale. *Dysphagia*. 1996;11:93-98.
16. Grand L, Lacima G, Ros E, et al. Deterioration of esophageal motility with age: a manometric study of 79 healthy subjects. *Am J Gastroenterol*. 1999;94:1795-1801.
17. Belafsky P, Rees C, Allen J, Leonard R. Pharyngeal dilation in cricopharyngeus muscle dysfunction and zenker diverticulum. *Laryngoscope*. 2010;120:889-894.
18. Herfarth H, Palmer L. Risk of radiation and choice of imaging. *Dig Dis*. 2009;27:278-284.
19. Raelson C, Kanal K, Vavilaia M, et al. Radiation dose and excess risk of cancer in children undergoing neuroangiography. *AJR Am J Roentgenol*. 2009;193:1621-1628.
20. Fazel R, Krumholz H, Wang Y, et al. Exposure to low-dose ionizing radiation from medical imaging procedures. *N Engl J Med*. 2009;361:849-857.
21. Smith-Bindman R, Lipson J, Marcus R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Arch Intern Med*. 2009;169:2078-2086.

Reprinted by permission of Am J Gastroenterol. 2013; 108(3):366-372.

Randomized Controlled Trial Comparing Aerosolized Swallowed Fluticasone to Esomeprazole for Esophageal Eosinophilia

Fouad J. Moawad, MD[1,2], Ganesh R. Veerappan, MD[1,2], Johnny A. Dias, DO[1], Thomas P. Baker, MD[3], Corinne L. Maydonovitch, BS[1] and Roy K.H. Wong, MD[1,2]

OBJECTIVES: **Patients with clinical symptoms of esophageal dysfunction and dense eosinophilic infiltration of the esophageal mucosa are suspected to have eosinophilic esophagitis (EoE). Topical steroids are often used as first-line therapy for EoE, although some patients respond clinically to proton pump inhibitors (PPIs). The purpose of this study was to compare the histological and clinical response of patients with esophageal eosinophilia treated with aerosolized swallowed fluticasone propionate vs. esomeprazole.**

METHODS: **This prospective single-blinded randomized controlled trial enrolled newly diagnosed patients with suspected EoE, defined as having clinical symptoms related to esophageal dysfunction with at least 15 eosinophils/high power field (hpf). Patients underwent 24-h pH/impedance monitoring to establish gastroesophageal reflux disease (GERD). Patients were stratified by the presence of GERD and randomized to receive fluticasone 440 mcg twice daily or esomeprazole 40 mg once daily for 8 weeks followed by repeat endoscopy with biopsies. The primary outcome was histological response of esophageal eosinophilia, defined as <7 eosinophils/hpf. Secondary outcomes included clinical change in symptoms using the validated Mayo dysphagia questionnaire (MDQ) and interval change in endoscopic findings following treatment.**

RESULTS: **Forty-two patients (90% male, 81% white, mean age 38±10 years) were randomized into fluticasone (n=21) and esomeprazole (n=21) treatment arms. In all, 19% (8/42) of patients had coexisting GERD and were equally stratified into each arm (n=4). Overall, there was no significant difference in resolution of esophageal eosinophilia between fluticasone and esomeprazole (19 vs. 33%, P=0.484). In patients with established GERD, resolution of esophageal eosinophilia was noted in 0% (0/4) of the fluticasone group compared with 100% (4/4) of the esomeprazole group (P=0.029). In GERD-negative patients, there was no significant difference in resolution of esophageal eosinophilia between treatment arms with fluticasone and esomeprazole (24 vs.18%, P=1.00). The MDQ score significantly decreased after treatment with esomeprazole (19±21 vs. 1.4±4.5, P<0.001), but not with fluticasone (17±18 vs. 12±16, P=0.162). Improvement in endoscopic findings and other histological markers were similar between treatment groups.**

CONCLUSIONS: **Fluticasone and esomeprazole provide a similar histological response for esophageal eosinophilia. With regard to clinical response, esomeprazole was superior to fluticasone, particularly in patients with established GERD.**

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

Am J Gastroenterol 2013; 108:366–372; doi:10.1038/ajg.2012.443; published online 12 February 2013

INTRODUCTION

Eosinophilic esophagitis (EoE) is a chronic inflammatory immune-mediated condition characterized by symptoms of esophageal dysfunction and the presence of dense eosinophilia on esophageal biopsies (1). Management, most often with topical steroids, is aimed at improving clinical symptoms and reversing the inflammatory changes within the esophagus to prevent tissue remodeling and formation of fibrosis (2).

[1]Department of Medicine, Walter Reed National Military Medical Center, Gastroenterology Service, Bethesda, Maryland, USA; [2]Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA; [3]Department of Pathology, Walter Reed National Military Medical Center, Bethesda, Maryland, USA.
Correspondence: Fouad J. Moawad, MD, Department of Medicine, Walter Reed National Military Medical Center, Gastroenterology Service, 8901 Rockville Pike, Bethesda, Maryland 20889, USA. E-mail: Fouad.Moawad@us.army.mil
Received 16 May 2012; accepted 16 September 2012

Proton pump inhibitors (PPIs) are also used to control symptoms associated with EoE and to treat gastroesophageal reflux disease (GERD), the main differential diagnosis of EoE, which can present with similar clinical symptoms and histopathology. According to the AGA consensus statement published in 2007, administration of PPIs to presumed EoE patients was part of the diagnostic evaluation, primarily to exclude GERD as a cause of esophageal eosinophilia (3). If dense eosinophilia persisted following therapy, then a diagnosis of EoE is made. However, some patients with a phenotypic appearance more suggestive of EoE rather than GERD (i.e. young atopic patient presenting with food impaction with concentric rings on endoscopy and having elevated eosinophils on esophageal biopsies) can respond to PPI therapy (4,5). This phenomenon now recognized by the more recent and updated EoE consensus statement has been termed PPI-responsive esophageal eosinophilia or PPI-responsive EoE (2,6).

The aim of this study was to perform a randomized controlled trial to compare the efficacy of fluticasone propionate (FP) to esomeprazole (ESO) in patients with esophageal eosinophilia. A secondary aim of this study was to determine whether the presence of GERD impacted the response to therapy in each treatment group.

METHODS

Study design and patient population

This is a prospective investigator-blinded randomized study. Adult patients (age≥18 years) seen at Walter Reed Army Medical Center (WRAMC) with esophageal eosinophilia were enrolled from April 2008 to October 2010. All patients had at least one clinical symptom of esophageal dysfunction (dysphagia, food impaction, heartburn) with ≥15 eosinophils/hpf (eos/hpf; high power field) on index endoscopy. Patients who had a history of secondary hypereosinophilic disorders, severe coagulopathy, or who were pregnant were excluded from the study. Patients who were dilated at index endoscopy were not excluded from the study. Baseline demographic data, history of coexisting atopic diathesis (seasonal allergies, food allergies, asthma, and eczema), and data from index endoscopy (concentric rings, longitudinal furrows, white plaques, mucosal tearing/friability, strictures, Schatzki rings, erosive esophagitis) were collected. All patients completed a validated dysphagia questionnaire, known as the 2-week Mayo Dysphagia Questionnaire (MDQ), following index endoscopy once eosinophilic infiltration was established on biopsies (7). This 29-item instrument is scored from 0 to 100 based on the presence and severity of dysphagia and whether patients avoided or had trouble swallowing different foods (oatmeal, banana, apple, ground meat, bread, and fibrous meat) (**Supplementary file**). Informed consent was obtained from all patients. This study was approved by the WRAMC Institutional Review Board (Work Unit number: 08-14045) and the study was registered at www.clinicaltrials.gov (NCT00895817).

GERD diagnosis

Upon enrollment into the study, all patients underwent 24-h pH with impedance studies. Location of the lower esophageal sphincter was determined by esophageal manometry utilizing a Solar Stationary GI motility system (Medical Measurement Systems USA, Dover, NH) and an electrically powered water perfusion pump (Mui Scientific, Ontario, CA). A 24-h pH/impedance catheter was then placed 5 cm above the proximal location of the lower esophageal sphincter. The catheter was connected to a ZepHr-reflux recording system (Sandhill Scientific, Highlands Ranch, CO) to capture pH/impedance, as well as symptom data. Subjects returned to our clinic the following day for analysis of the study. Data was analyzed with Bioview Analysis software (Sandhill Scientific). GERD was defined by the validated Johnson-DeMeester score (8,9). This scoring method takes into account six parameters, which include: total % time pH below 4, % time pH below 4 in the upright position, % time pH below 4 in the supine position, the total number of reflux episodes within a 24-h period, the number of reflux episodes longer than 5 min, and the longest reflux episodes in minutes. A composite score is then calculated with a score of greater than 22 being indicative of GERD. The pH drops without accompanying reflux events on impedance and reflux events during meals were excluded from analysis.

Randomization and drug administration

A computer-generated list of random numbers was used to separate patients into two equal treatment groups (esomeprazole and fluticasone proprionate). Concealed allocation using a sealed opaque envelope containing data on the sequence of randomization was maintained by a research pharmacist. Following data from the 24-h pH study, patients were stratified into GERD-negative or GERD-positive groups. Within each group, subjects were randomized to receive either 40 mg of ESO once daily or 440 mcg of FP twice daily. Patients randomized to ESO were instructed to take the medicine 30–60 min before their first meal. Patients randomized to FP were educated by the research pharmacist on correct delivery of the medication using an inhaler without the use of a spacer and instructed not to drink or eat 30 min following administration. The research pharmacist observed the patients priming the metered dose inhaler and administering at least one puff to ensure correct delivery. Adherence was assessed in the ESO arm by counting the number of pills at the end of the treatment study. For the FP arm, the number of puffs was counted using a specially designed metered dose inhaler, which recorded the number of puffs administered. Patients were considered adherent to treatment if ≥80% of the medication was taken during the study period.

Follow-up

Following 8 weeks of treatment, patients underwent repeat upper endoscopy with esophageal biopsies. A total of eight samples using standard biopsy forceps (Boston Scientific, Natick, MA) were taken from all patients, four from the proximal esophagus, ~15 cm from the gastroesophageal junction, and four from the distal esophagus, ~3 cm above the gastroesophageal junction. All endoscopies were performed with Olympus P160 or 180 endoscopes (Olympus, Tokyo, Japan). Endoscopic data was collected including concentric rings, longitudinal furrows,

© 2013 by the American College of Gastroenterology

white plaques, mucosal tearing/friability, strictures, Schatzki rings, and erosive esophagitis. The MDQ was completed by all patients at the end of the study.

Histopathology

All biopsy samples were embedded in formalin and stained with hematoxylin and eosin. Samples to include slides from index endoscopy and following treatment were reviewed and read by a single-blinded expert gastrointestinal pathologist. Each slide had three separate sections and all were reviewed. Eosinophils were counted in all available fields limited by the size of the biopsy specimens. A high-powered field was considered ×40 magnification on our microscopes, which measured 0.19 mm^2. The total number of eosinophils in all fields was counted and the peak eosinophil count per hpf was reported. Findings of eosinophilic microabscesses, intercellular edema, evidence of eosinophilic degranulation, epithelial basilar hyperplasia, and whether eosinophils were confined to the epithelial basal layer or extended to the epithelial surface (full thickness involvement) were also noted.

End points

The primary end point measured was histological response defined as achieving <7 eos/hpf in both proximal and distal esophageal biopsies following 8 weeks of treatment. Secondary outcomes measured included symptomatic change in dysphagia on the basis of the score from the MDQ and interval change in endoscopic and other histological findings.

Statistical analysis

Sample size estimation was based on the following assumptions: 10% of patients will be GERD positive and respond to PPIs compared with 55% of the patients treated with topical steroids. Controlling the probability of a Type I error at $\alpha = 0.05$, a sample of 38 patients in the treatment groups (19 in each arm) will have 80% power to detect a difference in treatment response of 45%.

Data were collated and analyzed with SPSS 15.0 statistical analysis package (SPSS Inc, Chicago, IL). Categorical data are expressed as frequency and percentage, and continuous data as means and standard deviation (s.d.). Histological response, comparison of histological markers, and endoscopic features between the two treatment groups were analyzed using Fisher's exact test. Within each treatment arm, the change in MDQ score after treatment was compared with Wilcoxon signed rank test. Adherence to treatment was compared using Mann–Whitney *U* test. Spearman rank correlation coefficient (r_s) was used to assess the relationship between reflux impedance episodes in the proximal and distal esophagus to eosinophil counts, as well as the relationship of change in MDQ and eosinophils count post treatment. A *P* value of <0.05 was considered statistically significant. Analysis was performed as intention to treat.

RESULTS

Forty-two patients with esophageal eosinophilia were enrolled into the study. The mean age ± s.d. was 38 years old ±10, 81% were Caucasian, 10% Hispanic, 7% African American, and 1% other. Sixty-two percent of patients had a history of coexisting atopic diathesis (33% seasonal allergies, 29% food allergies, 10% asthma, and 5% eczema). The primary indication for endoscopy was dysphagia (69%), followed by food impaction (19%), heartburn (12%), and other (2%). Nineteen percent (8/42) had GERD by Johnson-DeMeester score and were equally stratified into each treatment arm. Endoscopy revealed Los Angeles Grade A erosive esophagitis in seven patients, all of whom had GERD by Johnson-DeMeester score. There were significantly more acid reflux episodes on 24-h pH monitor in the FP arm compared with the ESO arm (45.8±40.6 vs. 25.5±19.3, $P = 0.045$), as well as impedance reflux episodes (63.6±23.1, 44.8±21.4, $P = 0.012$) (**Table 1**).

There was no significant change in mean eosinophil counts before and after treatment in either arm (FP: 55.9±25 vs. 39.2±29.4, $P = 0.102$; ESO: 42.9±18.9 vs. 30.5±33.7, $P = 0.174$) (**Figure 1**). Histological response was achieved in 33% (7/21) of ESO patients vs. 19% (4/21) of FP patients, $P = 0.484$. Among the eight patients with GERD, all four patients randomized to ESO achieved histological response, whereas none of the four patients randomized to FP achieved a histological response, $P = 0.029$. In the 34 GERD-negative patients, response was achieved in 18% (3/17) of ESO patients vs. 24% (4/17) of FP patients, $P = 1.000$. Among patients with coexisting allergies, response was similar between the two treatment groups, FP: 27% (3/11) vs. ESO: 33% (5/15), $P = 1.000$.

The histological response for FP vs. ESO was similar in the proximal esophagus (29 vs. 55%, $P = 0.118$) and in the distal esophagus (19 vs. 40%, $P = 0.181$). In two patients of the FP arm and four patients of the ESO arm, histological response was achieved in the proximal but not distal esophagus. Improvement in other histological markers of EoE following treatment (basal cell hyperplasia, intercellular edema, eosinophilic microabscess, eosinophilic degranulation, and eosinophilic distribution within the epithelium) was similar between the two treatment arms (**Table 2**).

Frequency and severity of dysphagia were similar at baseline between the two treatment arms. The majority of patients reported moderate to severe dysphagia on question no. 3 of the MDQ with no significant difference between the two groups, (FP 77% vs. ESO 83%, $P = 0.512$). In terms of frequency of dysphagia (question no. 4), the majority of patients reported symptoms from less than once per week to several times per week with no difference between the two treatment arms (FP 89% vs. ESO 83%, $P = 0.646$). Only three patients indicated symptoms with every meal. On baseline questionnaire, 30% of patients randomized to FP avoided fibrous foods (meat, chicken, bread, celery, salad) compared with 42% of patients randomized to ESO, $P = 0.381$.

The MDQ score before and after therapy significantly improved in the ESO group (19±21 vs. 1.4±4.5 $P = 0.001$), but not in the FP group (17±18 vs. 12±16, $P = 0.162$) (**Figure 2**). A similar finding was noted among GERD-negative patients: there was a significant difference in MDQ score before and after treatment with ESO (16±14 to 1.7±5.0, $P = 0.001$) but not with FP (18±19 to 10±16, $P = 0.086$). Overall, there was no significant correlation between the change in symptoms by MDQ and the change in eosinophil count in the proximal ($r_s = 0.001$, $P = 0.996$) and distal ($r_s = 0.101$, $P = 0.558$) esophagus.

Table 1. Baseline patient characteristics

Variables	Fluticasone (*n*=21)	Esomeprazole (*n*=21)	*P* value
Mean age±s.d., years	37.0±11.1	38.0±8.8	0.771
Male, *n* (%)	19 (90.5)	19 (90.5)	1.000
Race, n *(%)*			0.766
Caucasian	17 (81.0)	17 (81.0)	—
Hispanic	3 (14.3)	2 (9.5)	—
African American	1 (4.7)	2 (9.5)	—
Coexisting allergies, n *(%)*			
Any atopic disease	11 (52.4)	15 (71.4)	0.341
Seasonal allergies	7 (33.3)	7 (33.3)	1.000
Asthma	2 (9.5)	2 (9.5)	1.000
Eczema	2 (9.5)	0 (0.0)	0.488
Food allergies	3 (14.3)	9 (42.9)	0.085
Concomitant use of allergy medications, n *(%)*			
Antihistamines	4 (19.0)	4 (19.0)	1.000
Nasal steroid spray	3 (14.3)	1 (4.7)	0.606
Leukotriene antagonist	1 (4.7)	1 (4.7)	1.000
GERD by pH score	4 (19.0)	4 (19.0)	1.000
Erosive esophagitis[a]	3 (14.3)	4 (19.0)	1.000
pH reflux episodes			
No. of episodes	45.8±40.6	25.5±19.3	0.045
% (*x*/*n*) abnormal patients	28.5% (6/21)	9.5% (2/21)	0.238
Impedance reflux episodes			
No. of episodes	63.6±23.1	44.8±21.4	0.012
% (*x*/*n*) abnormal patients	38% (8/21)	14.2% (3/21)	0.159
Mayo dysphagia score	17.1±17.8	19.5±20.7	0.691
Pretreatment eosinophil count			
Proximal biopsy	39.1±33.2	32.9±19.4	0.473
Distal biopsy	38.4±22.3	34.2±25.2	0.593

GERD, gastroesophageal reflux disease.
[a]All cases of erosive esophagitis were Los Angeles Grade A.

With regard to resolution of endoscopic findings, no significant difference was seen between FP and ESO (**Table 2**). Dilation was performed on 15 patients on index endoscopy primarily for a dominant stricture seen or a coexisting Schatzki ring (8 patients taking FP and 7 patients taking ESO). No dilations were performed during the treatment period or on follow-up endoscopy post therapy. There was significant improvement in clinical symptoms based on a mean decrease in the MDQ score seen in both treatment groups who underwent dilation (FP = −10.6±10.5, P=0.027 and ESO = −14.3±14.0, P=0.027). Among patients who did not undergo dilation (n=27) on index endoscopy, there was a significant decrease in MDQ score in the ESO group (−20±24, P=0.005) but not in the FP group (−1.9±21.5, P=0.721).

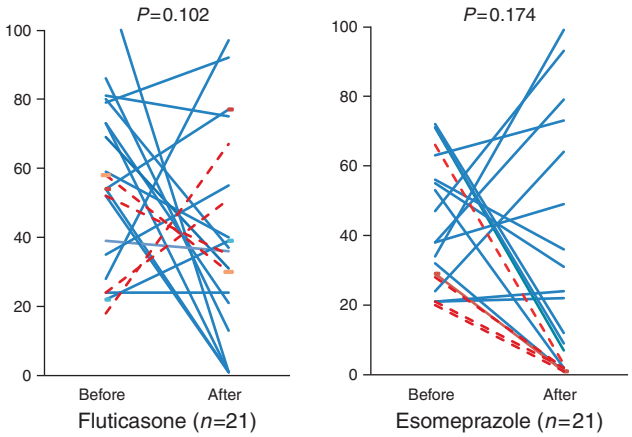

Figure 1. There was no significant change in eosinophil count post treatment in either arm regardless of coexisting gastroesophageal reflux disease (GERD). Dashed lines indicate GERD-positive patients.

To further explore the relationship between reflux and eosinophilia, we examined the association between impedance reflux episodes and eosinophil counts. There was no significant association between impedance reflux episodes and eosinophil counts in the proximal (r_s=0.263, P=0.116) or distal esophagus (r_s=0.162, P=0.359).

In both treatment arms, adherence to therapy was very good. There was no significant difference in adherence to treatment in both groups (FP = 86±24% vs. ESO = 92±10%, P=0.977).

Two patients randomized to the FP arm discontinued treatment during the study period. One patient had worsening of migraine headaches, which he attributed to FP. Another patient had bothersome GERD-related symptoms and discontinued the steroid, and began treatment with a PPIs. Both patients were analyzed as intention to treat. One patient in the fluticasone arm developed esophageal candidiasis. He was asymptomatic during the study period and this was discovered on follow-up endoscopy and confirmed on esophageal biopsies. He was treated with a course of oral fluconazole. No adverse events occurred in the PPI arm.

DISCUSSION

This randomized controlled single-blinded study demonstrated a similar histological response between esomeprazole and fluticasone treatment groups in patients with esophageal eosinophilia. With regard to clinical improvement, based on a validated symptom questionnaire, ESO was significantly better than FP regardless of a concomitant GERD diagnosis.

We defined our patient population who had clinical symptoms of esophageal dysfunction and elevated eosinophil counts on biopsies as having esophageal eosinophilia rather than EoE, based on the most recent updated consensus statement (2). Although swallowed FP is commonly accepted as the first-line treatment for EoE, in this study we examined its efficacy in patients who had phenotypic appearance of EoE and elevated eosinophil counts yet were PPI naive. Our data demonstrated a 19% histological response in patients treated with FP.

© 2013 by the American College of Gastroenterology

Table 2. Data presented for percent (number) of patients with histological and endoscopic findings in pretreatment and posttreatment

	Fluticasone			Esomeprazole		
	Pre	Post	% Improve[a]	Pre	Post	% Improve[a]
Histological findings						
Basal cell hyperplasia	100% (21)	81% (17)	19% (4/21)	100% (21)	52% (11)	43% (9/21)
Intercellular edema	100% (21)	76% (16)	24% (5/21)	86% (18)	57%(12)	35% (7/18)
Eosinophilic microabscess	86% (18)	52% (11)	44% (8/18)	71% (15)	38% (8)	67% (10/15)
Eosinophilic degranulation	95% (20)	52% (11)	50% (10/20)	76% (16)	43% (9)	50% (8/16)
Eosinophilic distribution[b]	95% (20)	76% (16)	25% (5/20)	90% (19)	48%(10)	47% (9/19)
Endoscopic findings						
Stenosis on index endoscopy	24% (5)	14% (3)	80% (4/5)	24% (5)	10% (2)	80% (4/5)
Concentric rings	76% (16)	76% (16)	13% (2/16)	76% (16)	52% (11)	44% (7/16)
Longitudinal furrows	81% (17)	76% (16)	18% (3/17)	81% (17)	52% (11)	41% (7/17)
White plaques	19% (4)	29% (6)	50% (2/4)	24% (5)	0% (0)	100% (5/5)

% improve, % improvement; pre, pretreatment; post, posttreatment.
Percent improvement is among a subgroup of patients who had a pretreatment finding. There was no statistically significant difference in improvement between treatment arms.
[a]Percent improvement is among patient who had a pretreatment finding.
[b]Involving entire thickness of epithelium.

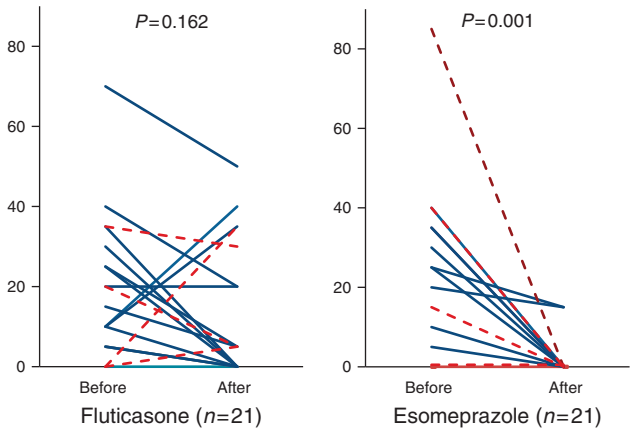

Figure 2. Change in Mayo dysphagia questionnaire (MDQ) score following treatment. There was significant clinical improvement in ESO but not in the FP treatment arm. Dashed lines indicate gastroesophageal reflux disease (GERD)-positive patients.

Response to topical steroids for EoE has varied in the literature with small case series and retrospective studies reporting response rates in up to 80% (1,10). In contrast, data from prospective controlled studies have demonstrated lower response rates. Konikoff *et al.* (11) reported a 50% efficacy (10/20) in a pediatric population with histological response defined as ≤1 eos/hpf. Again, complete histological response was seen in 50% (18/36) of patients after 4 weeks of therapy with FP (12). In another study comparing the efficacy of topical steroids to PPI therapy in adults with EoE, only 15% achieved complete histological response with FP, defined as 5 eos/hpf (13). A recent study by Alexander *et al.* (14) reported a 62% histological response to fluticasone, defined as a more than 90% decrease in mean levels of eosinophils following 6 weeks of therapy. Interestingly, no significant clinical improvement was noted when compared with placebo.

The relatively poor response seen in many randomized studies may at least partially be due to a lack of uniformity in medication dosage, duration of therapy, the definition of response, and delivery method. Additionally, variability in the definition of EoE exists in the literature (15) and in many studies EoE is often synonymous with esophageal eosinophilia. Alexander *et al.* (14) used a higher dose of swallowed fluticasone (880 mcg twice daily), which may reflect the higher response rate seen in their study compared with other randomized controlled studies using topical fluticasone for EoE. Five patients developed esophageal candidiasis, compared with one patient in our study, suggesting this may be dose dependent. Another reason our response rate may have been different when compared with other randomized studies may be due to the definition of response. Variability exists in the literature with some studies using a change in histological grade (12) and even a decrease in percent eosinophils in comparison with baseline (14,16). The majority of the studies have used absolute changes in eosinophil count as an end point to include 1 eos/hpf (11), 5 eos/hpf (13,17), and <7 eos/hpf (18). We chose to use <7 eos/hpf as a target end point. However, we also examined our data using <5 eos/hpf and obtained identical results (19% fluticasone and 33% for esomeprazole).

Another explanation for the low response to FP may be related to the mechanism of drug delivery. FP is an aerosolized medication administered through a metered dose inhaler, which is intended for the airways, and it is unclear how much drug is actually delivered to the esophagus. One would assume that response may be superior in the proximal esophagus than the distal esophagus with this type of delivery. However, when we compared the histological response between the proximal and distal esophagus, there was no

significant difference in eosinophilia between these two locations. It is noteworthy that studies using oral viscous budesonide, which can be administered as a slurry mixed with a sugar substitute or as a swallowed nebulizer, demonstrated higher response rates (72–80%) (17,18). This suggests that the delivery system, and not the steroid type, may be the reason for low response rates. To further study and explore this point, a recent study demonstrated that oral viscous budesonide was more effective in reducing eosinophil counts when compared with nebulized budesonide (19).

Interestingly, there was a greater number of acid and impedance reflux episodes in the FP arm compared to the ESO arm. However, the percentage of patients with abnormal acid (>50) and impedance episodes (>73) was similar between both treatment arms. Therefore, it is unclear whether an increased number of reflux episodes had any role in the response to topical steroids.

The rationale for using PPIs for the treatment of esophageal eosinophilia is that it may help treat underlying GERD, which may contribute to EoE (20). It may be difficult to completely separate these two entities and it appears that many adult EoE patients respond at least clinically to PPI therapy. In one study, 8 of 17 (47%) adult patients presenting with food impaction and dense esophageal eosinophilia (>20 eos/hpf) responded clinically to PPI therapy (5). The mechanism for this is not clear, although PPIs may help heal a disrupted epithelial barrier and therefore reduce further immune activation (2). In our study, the four patients with esophageal eosinophilia and GERD had a complete response to PPI therapy.

The histological response achieved in three patients with PPI therapy despite having a negative 24-h pH study suggests these patients may have PPI-responsive EoE, which stimulates an interesting discussion regarding the pathophysiology of this entity. This described response to esomeprazole in GERD-negative esophageal eosinophilia patients can be attributed to anti-inflammatory properties of PPI independent of acid inhibition. Interestingly, PPIs have been shown to exhibit anti-inflammatory properties by acting directly on principal cytokines (IL-4 and IL-13) involved in the recruitment of eosinophils in the esophagus (21). Additionally, PPI therapy was shown to block the release of eotaxin-3, which has an integral role in the pathogenesis of EoE (22). Other studies have also reported a histological response to PPI therapy in EoE patients even following a negative 24-h pH study. Peterson *et al.* (13) demonstrated a 33% (4/12) histological response to PPI therapy. Similar results have been reported in children as well. In 43 pediatric EoE patients, a subset of patients responded histologically to PPI therapy following a normal pH study (23). These studies coupled with our data suggest that a subset of patients whose clinical presentation suggests EoE and have dense eosinophilia on esophageal biopsies indeed appear to be PPI responsive.

Significant clinical improvement was seen in the majority of patients with esophageal eosinophilia taking PPI even though only a few had histological resolution. Other studies have also reported a poor correlation between clinical remission and histological response. In a randomized placebo controlled study, vomiting was the only symptom that significantly improved in EoE children treated with topical steroids and did not correlate with histological response (11). Similar findings have been described by Molina *et al.* (24) in which clinical remission was achieved in the majority of patients with food impaction or dysphagia despite persistence of eosinophilia. Interestingly, a dissociation was recently reported between clinical response and histological severity in EoE children with the majority of patients reporting continued symptoms despite histological remission (25). These data combined with ours suggest one cannot use clinical response as a surrogate for histological change. However, whether achieving histological response in EoE is an important outcome is yet to be determined. In theory, normalization of the eosinophila may prevent further tissue remodeling, fibrosis, and possibly stricture formation.

Esophageal dilation is one of the most effective treatments for alleviating the symptoms of EoE even though it does not address the underlying disease pathophysiology (26). In our cohort, all of the 15 patients (8 in FP arm and 7 in ESO arm) with dysphagia and esophageal eosinophilia who underwent dilation at index endoscopy had significant improvement in symptoms. Similar clinical improvements have been noted in other studies in EoE (26,27). In contrast, among the 27 patients who did not undergo dilation, only those randomized to ESO had significant improvement in symptoms. As a similar number of patients underwent dilation in each treatment arm and dilation was not performed during the 8-week treatment course or at follow-up endoscopy, it is unlikely that dilation had an effect on the change in MDQ scores before and after therapy.

There are some limitations of our study. A placebo arm would have helped better define the natural history of esophageal eosinophilia. As we were comparing one medication administered as a pill to another given as an inhaler, we found adding a placebo arm to be challenging. Additionally the delivery system of aerosolized fluticasone may have hampered the true assessment of steroid response rates, and using another agent such as oral viscous budesonide may give us a different understanding of the true steroid response. Another limitation of our study was the relatively small sample size, which limited the subgroup analyses. Additionally, we used the 2-week dysphagia score, which may not have accurately reflected the symptoms in some EoE patients, particularly in those with intermittent dysphagia. However, the 2-week MDQ was the only validated questionnaire available at the time of this study. To date, no dysphagia questionnaire developed exclusively for EoE has been published.

In conclusion, histological response between ESO and FP were similar in the treatment of esophageal eosinophilia, with neither drug having overwhelming treatment success. On the other hand, significant improvement in clinical symptoms was demonstrated with PPI therapy. Further larger studies are needed to better define the optimal treatment for patients with esophageal eosinophilia and to better describe the subgroup and natural history of such patients who respond to PPI therapy.

CONFLICT OF INTEREST

Guarantor of the article: Fouad J. Moawad, MD.

Specific author contributions: Study concept and design, subject enrollment, acquisition of data, analysis and interpretation of data,

© 2013 by the American College of Gastroenterology

drafting of the manuscript, critical revision of the manuscript for important intellectual content: Fouad J. Moawad; study concept and design, study enrollment, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content: Ganesh R. Veerappan; data collection, verification and analysis, drafting of the manuscript: Johnny A. Dias; histological analysis, drafting of the manuscript: Thomas P. Baker; study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content: Corinne L. Maydonovitch; study concept and design, drafting of the manuscript, critical revision of the manuscript for important intellectual content: Roy K.H. Wong.

Financial support: None.

Potential competing interests: None.

Disclaimer

The opinions are solely those of the authors and do not represent an endorsement by the Department of Defense. This is US Government work. There are no restrictions on its use.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Topical steroids are commonly used as first line treatment for patients with a clinical presentation suggestive of EoE and who have dense eosinophilic infiltration.
- ✓ PPI therapy is often prescribed in esophageal eosinophilia patients to help treat coexisting GERD and establish diagnosis of EoE.

WHAT IS NEW HERE

- ✓ Topical fluticasone had a lower than expected response rate which may be dose and delivery related.
- ✓ PPIs induce histological response in some patients with esophageal eosinophilia regardless of the presence of GERD.
- ✓ PPIs significantly improved clinical symptoms in patients with esophageal eosinophilia even in the absence of GERD.

REFERENCES

1. Moawad FJ, Veerappan GR, Wong RK. Eosinophilic esophagitis. Dig Dis Sci 2009;54:1818–28.
2. Liacouras CA, Furuta GT, Hirano I *et al.* Eosinophilic esophagitis: updated consensus recommendations for children and adults. J Allergy Clin Immunol 2011;128:3–20.
3. Furuta GT, Liacouras CA, Collins MH *et al.* Eosinophilic esophagitis in children and adults: a systematic review and consensus recommendations for diagnosis and treatment. Gastroenterology 2007;133:1342–63.
4. Ngo P, Furuta GT, Antonioli DA *et al.* Eosinophils in the esophagus-peptic or allergic eosinophilic esophagitis? Case series of three patients with esophageal eosinophilia. Am J Gastroenterol 2006;101:1666–70.
5. Desai TK, Stecevic V, Chang CH *et al.* Association of eosinophilic inflammation with esophageal food impaction in adults. Gastrointest Endosc 2005;61:795–801.
6. Hirano I. Eosinophilic esophagitis and gastroesophageal reflux disease: there and back again. Clin Gastroenterol Hepatol 2011;9:99–101.
7. Grudell AB, Alexander JA, Enders FB *et al.* Validation of the Mayo Dysphagia Questionnaire. Dis Esophagus 2007;20:202–5.
8. Johnson LF, Demeester TR. Twenty-four-hour pH monitoring of the distal esophagus. A quantitative measure of gastroesophageal reflux. Am J Gastroenterol 1974;62:325–32.
9. Johnson LF, DeMeester TR. Development of the 24-hour intraesophageal pH monitoring composite scoring system. J Clin Gastroenterol 1986;8 (Suppl 1): 52–8.
10. Noel RJ, Putnam PE, Collins MH *et al.* Clinical and immunopathologic effects of swallowed fluticasone for eosinophilic esophagitis. Clin Gastroenterol Hepatol 2004;2:568–75.
11. Konikoff MR, Noel RJ, Blanchard C *et al.* A randomized, double-blind, placebo-controlled trial of fluticasone propionate for pediatric eosinophilic esophagitis. Gastroenterology 2006;131:1381–91.
12. Schaefer ET, Fitzgerald JF, Molleston JP *et al.* Comparison of oral prednisone and topical fluticasone in the treatment of eosinophilic esophagitis: a randomized trial in children. Clin Gastroenterol Hepatol 2008;6:165–73.
13. Peterson KA, Thomas KL, Hilden K *et al.* Comparison of esomeprazole to aerosolized, swallowed fluticasone for eosinophilic esophagitis. Dig Dis Sci 2010;55:1313–9.
14. Alexander JA, Jung KW, Arora AS *et al.* Swallowed fluticasone improves histologic but not symptomatic response of adults with eosinophilic esophagitis. Clin Gastroenterol Hepatol 2012;10:742–9.
15. Sperry SL, Shaheen NJ, Dellon ES. Toward uniformity in the diagnosis of eosinophilic esophagitis (EoE): the effect of guidelines on variability of diagnostic criteria for EoE. Am J Gastroenterol 2011; 106:824–32.
16. Straumann A, Conus S, Degen L *et al.* Budesonide is effective in adolescent and adult patients with active eosinophilic esophagitis. Gastroenterology 2010;139:1526–37.
17. Straumann A, Conus S, Degen L *et al.* Long-term budesonide maintenance treatment is partially effective for patients with eosinophilic esophagitis. Clin Gastroenterol Hepatol 2011;9:400–9.
18. Aceves SS, Bastian JF, Newbury RO *et al.* Oral viscous budesonide: a potential new therapy for eosinophilic esophagitis in children. Am J Gastroenterol 2007;102:2271–9.
19. Dellon ES, Sheikh A, Speck O *et al.* Viscous Topical Is More Effective Than Nebulized Steroid Therapy for Patients With Eosinophilic Esophagitis. Gastroenterology 2012;143:321–4.
20. Spechler SJ, Genta RM, Souza RF. Thoughts on the complex relationship between gastroesophageal reflux disease and eosinophilic esophagitis. Am J Gastroenterol 2007;102:1301–6.
21. Cortes JR, Rivas MD, Molina-Infante J *et al.* Omeprazole inhibits IL-4 and IL-13 signaling signal transducer and activator of transcription 6 activation and reduces lung inflammation in murine asthma. J Allergy Clin Immunol 2009;124:607–10.
22. Zhang X CE, Hou X *et al.* In esophageal epithelial cell lines from patients with eosinophilic esophagitis (EoE), omeprazole blocks the stimulated secretion of eotaxin-3: a potential anti-inflammatory effect of omeprazole in EoE that is independent of acid inhibition. Gastroenterology 2010;138 (suppl 1): S–122.
23. Dranove JE, Horn DS, Davis MA *et al.* Predictors of response to proton pump inhibitor therapy among children with significant esophageal eosinophilia. J Pediatr 2009;154:96–100.
24. Molina-Infante J, Ferrando-Lamana L, Ripoll C *et al.* Esophageal eosinophilic infiltration responds to proton pump inhibition in most adults. Clin Gastroenterol Hepatol 2011;9:110–7.
25. Pentiuk S, Putnam PE, Collins MH *et al.* Dissociation between symptoms and histological severity in pediatric eosinophilic esophagitis. J Pediatr Gastroenterol Nutr 2009;48:152–60.
26. Schoepfer AM, Gonsalves N, Bussmann C *et al.* Esophageal dilation in eosinophilic esophagitis: effectiveness, safety, and impact on the underlying inflammation. Am J Gastroenterol 2010;105:1062–70.
27. Straumann A, Spichtin HP, Grize L *et al.* Natural history of primary eosinophilic esophagitis: a follow-up of 30 adult patients for up to 11.5 years. Gastroenterology 2003;125:1660–9.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2015; 124(3):221-226.

Article

Utility of the Transnasal Esophagoscope in the Management of Chemoradiation-Induced Esophageal Stenosis

Annals of Otology, Rhinology & Laryngology
2015, Vol. 124(3) 221–226
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489414550858
aor.sagepub.com
SAGE

Kevin A. Peng, MD[1], Aaron J. Feinstein, MD, MHS[1], Jonathan B. Salinas, MD[1], and Dinesh K. Chhetri, MD[1]

Abstract

Objective: This study aimed to describe management of esophageal stenosis after chemoradiation therapy for head and neck squamous cell carcinoma (HNSCC), with particular emphasis on techniques and outcomes with the use of the transnasal esophagoscope (TNE) in the office as well as operating room settings.

Methods: Retrospective analysis of all patients with esophageal stenosis following head and neck cancer radiation, with or without chemotherapy, and managed with TNE-assisted esophageal dilation over a 5-year period. Preoperative and postoperative swallowing function were assessed objectively with the Functional Outcome Swallowing Scale (FOSS; ranging from score 0, a normal diet, to score 5, complete dependence on nonoral nutrition).

Results: Twenty-five patients met inclusion criteria. The mean pretreatment FOSS score was 4.4, whereas the mean posttreatment FOSS score was 2.7 (Wilcoxon signed-rank test, $P < .001$). Prior to dilation, 16 patients were completely gastrostomy-tube dependent (FOSS 5), of whom 12 (75%) were able to tolerate oral nutrition for a majority of their diet following treatment according to our protocol. No complications were noted.

Conclusion: Dysphagia following chemoradiation therapy for HNSCC is often related to esophageal stenosis. With the aid of TNE, we have developed a successful treatment strategy for esophageal stenosis with improved success rates.

Keywords

chemoradiation, esophageal dilation, esophageal stenosis, head and neck squamous cell carcinoma, transnasal esophagoscopy

Introduction

Squamous cell carcinomas of the head and neck occur frequently, with more than 500000 cases diagnosed worldwide annually.[1] Radiation with concurrent chemotherapy (CRT) is an increasingly used treatment modality for these cancers. As survival rates improve with advances in care, organ preservation—that is, the maintenance of normal mechanisms of breathing, deglutition, and communication—becomes of paramount importance. Following successful treatment of head and neck squamous cell cancer (HNSCC), dysphagia is the most common symptom decreasing quality of life, affecting 50% to 64% of patients after CRT.[2,3]

Whereas early dysphagia is usually temporary, late dysphagia often results from chronic inflammation and fibrosis and is much more difficult to manage.[4-7] This fibrosis may progress to hypopharyngeal or esophageal strictures, which occur in approximately 21% of patients undergoing CRT.[8,9] Risk factors implicated in stricture formation in the general population include reflux, older age, and caustic ingestion; among head and neck cancer patients with HNSCC, additional factors include hypopharyngeal primary site, combined chemoradiation (vs radiation alone), radiation dose, prior neck dissection, female sex, and treatment-induced mucositis.[10]

Objective assessment of dysphagia is essential and comprises 2 complementary tests: the videofluoroscopic swallow study, also known as a modified barium swallow study (MBSS), and the functional endoscopic evaluation of swallow (FEES).[11] Whereas the advantages of FEES include rapidity of the test in an office setting, direct observation of native secretions and swallow anatomy, and lack of radiation for the procedure, MBSS is superior in evaluating the oral and upper esophageal phases.[12] In addition to these tests, flexible transnasal esophagoscopy has seen increasing use in the otolaryngology dysphagia clinic, particularly in

[1]Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

Corresponding Author:
Dinesh K. Chhetri, MD, Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA, 10833 Le Conte Avenue, CHS 62-132, Los Angeles, CA 90095, USA.
Email: dchhetri@mednet.ucla.edu

Table 1. Functional Outcome Swallowing Scale.[a]

Stage	Description
0	Normal function; asymptomatic
1	Normal function; episodic or daily symptoms of dysphagia
2	Compensated abnormal function manifested by significant dietary modifications or prolonged mealtime, without weight loss or aspiration
3	Decompensated abnormal function, with weight loss of 10% or less of body weight over 6 months due to dysphagia, or daily cough, gagging, or aspiration during meals
4	Severely decompensated abnormal function, with weight loss of more than 10% of body weight over 6 months due to dysphagia, or severe aspiration with bronchopulmonary complications; nonoral feeding recommended for most of nutrition
5	Nonoral feeding for all nutrition

[a]Adapted from Salassa.[15]

evaluating the presence, severity, and length of esophageal stenoses.[13]

Following assessment of the stenosis, esophageal dilations with Savary-Gilliard dilators or controlled radial expansion (CRE) balloons can be performed in both operating room and office-based settings with modifications of prior techniques.[14] In this study, we review our management of esophageal stenosis after CRT for HNSCC. We place special emphasis on the use of the transnasal esophagoscope (TNE) to demonstrate that this method has a high success rate with minimal potential for complications. The Functional Outcome Swallowing Scale (FOSS), described by Salassa[15] in 1999, was used to quantify dysphagia prior to and following treatment (Table 1).

Methods

Institutional review board approval was obtained for this study. Inclusion criteria were as follows: history of HNSCC treated with radiation and/or chemotherapy, presence of esophageal stenosis, management of esophageal stenosis by esophagoscopy and dilation, and documentation of swallowing function with instrumental swallow tests (MBSS and/or FEES) both prior to and following dysphagia treatment. Patients with multilevel esophageal stenosis were excluded. Patients who underwent surgery, including tracheostomy, neck dissection, or resection of the primary tumor, were also excluded, with the following exceptions: gastrostomy tube (G-tube) placement, tonsillectomy, or panendoscopy with biopsies.

We managed esophageal stenosis with the following algorithm. After a history and physical examination were performed, FEES was performed in office to assess the current safe diet, and transnasal esophagoscopy was performed if esophageal stenosis was suspected based on MBSS, dysphagia to solid foods, or severe piriform sinus residue. Following a definitive office diagnosis of esophageal stenosis, patients were scheduled for surgery.

In the operating room, suspension direct laryngoscopy was performed under general anesthesia and the rigid operating laryngoscope was placed in the postcricoid space. The TNE was then passed through the laryngoscope into the hypopharynx and advanced into the esophagus. Several dilation scenarios were possible at this point. (1) In a majority of cases, the stricture was seen, and the TNE could be passed atraumatically beyond the stenosis and into the distal esophagus. A CRE balloon was then passed through the stricture under direct visualization and dilation was performed to 18 mm. Alternatively, a Savary-Gilliard dilator guidewire could be passed through the working port of the scope, and dilation could then be performed over the guidewire after retracting the scope completely. (2) The stricture was seen, but the TNE would not pass through the stricture. This indicated that the stenosis diameter was smaller than the diameter of our scope (5.1 mm). At this point, gentle passage of a Savary-Gilliard dilator guidewire was attempted while directly visualizing its passage through the stenotic opening. If this was possible without resistance, dilation was then performed using Savary-Gilliard dilators up to 9 to 10 mm just past the stenosis. Then, the TNE was passed through the stenosis to ensure normal esophageal lumen, after which CRE balloon dilation was performed, typically to 15 mm. (3) A complete stricture was encountered, and the TNE could not pass. In this case, the G-tube was removed and retrograde esophagoscopy was performed. The TNE could be inserted through the G-tube site without dilation of the G-tube tract (Figure 1) and was advanced through the lower esophageal sphincter to the upper esophagus to the stricture site. Anterograde palpation of the esophageal stricture using a blunt instrument such as a rigid esophageal suction tube assisted in identifying the stricture, and under direct retrograde visualization with the TNE, the stricture was punctured. A Savary-Gilliard guidewire was passed through the stenosis with both retrograde and anterograde visualization. This guidewire could be inserted in an anterograde manner under direct visualization of the TNE and dilation performed as in situation 2 above. Topical mitomycin-C (MMC), which inhibits fibroblast proliferation, was applied to the affected region in all cases, using cotton pledgets at a concentration of 0.4 mg/mL for 4 minutes, a technique that has been successfully applied in the treatment of upper aerodigestive tract stenosis.[16-19]

A second dilation was scheduled for 1 to 2 weeks after the first dilation. The second dilation allows an assessment of efficacy of the first dilation, which helps to counsel patients on the anticipated treatment course in regard to repeat dilations and provides an opportunity for the second

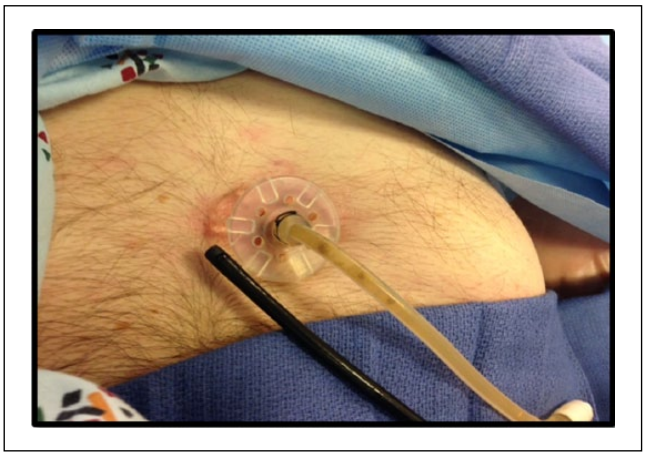

Figure 1. The transnasal esophagoscope (TNE) used in this study pictured adjacent to a standard gastrostomy tube (G-tube), showing similarity of diameters. With the G-tube removed, the TNE can be passed for retrograde esophagoscopy without further dilation of the G-tube site.

application of MMC.[19] Dilation was typically performed to 18 to 20 mm diameter using the CRE balloon. All patients were referred for swallow therapy after the second dilation. Some patients were scheduled for office dilation depending on the degree of stenosis and residual dysphagia. Office dilation was undertaken preferentially as many patients had significant trismus and were high anesthetic risks regarding intubation. Office esophageal dilations were accomplished as follows: bilateral nasal cavities were anesthetized and decongested with topical lidocaine and oxymetazoline. Thereafter, transnasal esophagoscopy was performed via the more patent nasal cavity; once the stenosis was identified, a CRE balloon dilator was passed via the contralateral nasal cavity and to the level of stenosis under direct visualization. Passage of the balloon was sometimes aided by bending the tip slightly to traverse the nasopharyngeal curvature. Dilation was then performed using the CRE balloon, typically to 18 mm. Patients received proton-pump inhibitors for the first 3 months after initial dilation, with further prescriptions based on the presence of reflux symptoms. Esophagoscopy and dilations were performed until the patient's symptoms were alleviated satisfactorily.

To analyze outcomes of our esophageal stenosis treatment algorithm, pretreatment and posttreatment FOSS scores were compared with a Wilcoxon signed-rank test. The number of dilations undergone by each patient was noted, as was the elapsed time between dilations.

Results

Among the 115 patients identified, 81 were excluded due to use of surgery (eg, neck dissection, tumor resection) during their initial treatments and 9 were excluded due to requiring additional surgical procedures at the time of esophageal dilation. There were 25 patients, 21 male and 4 female, who met inclusion criteria (Table 2). The median age was 63 years (range, 40-84 years). The most common primary site was oropharynx (n = 13, 52%), followed by an unknown primary (n = 5, 20%). All patients received combined chemoradiation therapy. Median time from completion of CRT to initiation of esophageal stenosis management was 6.0 months (range, 2 months to 30 years). All patients in the study had single-level stenosis.

Table 2. Patient Characteristics.

Age, y	
Median	63
Range	40-84
Sex	
Male	21
Female	4
Primary site	
Oropharynx	13 (52%)
Unknown primary	5 (20%)
Hypopharynx	3 (12%)
Nasopharynx	2 (8%)
Larynx	1 (4%)
Oral cavity	1 (4%)
Elapsed time between termination of CRT and initiation of esophageal stenosis treatment	
Median	6.0 months
Range	2 months to 30 years
Elapsed time between esophageal dilations	
Median	21 days
Range	6 days to 1.8 years
Number of dilations performed	
Median	2
Range	1-16
Functional Outcome Swallowing Scale score	
Prior to treatment of esophageal stenosis	
Mean	4.36
Range	2-5
Following treatment of esophageal stenosis	
Mean	2.40
Range	1-5

The median number of dilations performed on each patient was 2 (range, 1-16). For patients undergoing multiple dilations, the median time between procedures was 21 days (range, 6 days to 21 months). In 3 patients (12%) who were completely G-tube dependent, retrograde esophagoscopy was performed via the G-tube with the TNE to delineate the esophageal lumen. Mean pretreatment FOSS score for all patients was 4.4 (median, 5; range, 2-5); mean posttreatment FOSS score was 2.7 (median, 3; range, 1-5). A Wilcoxon signed-rank test confirmed a

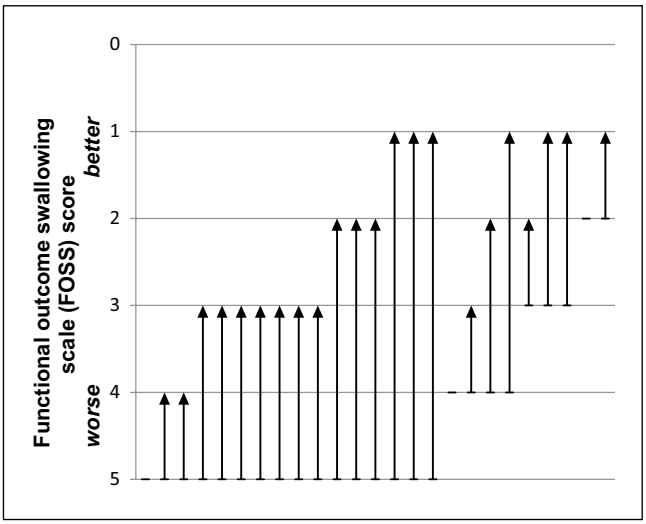

Figure 2. Improvement in Functional Outcome Swallowing Scale (FOSS) score was seen in all but 3 of 25 patients following our esophageal dilation protocol; no patients worsened after therapy. Arrows depict change in FOSS scores following therapy.

statistically significant difference between FOSS scores prior to and following esophageal stenosis treatment ($P <$ .001). The FOSS score did not worsen in any patients (Figure 2).

Prior to treatment, 16 patients (64%) were completely dependent on nonoral nutrition, primarily via G-tube (FOSS score of 5); following treatment, only 2 patients (8%) were completely dependent on nonoral nutrition. Of the 16 patients completely dependent on nonoral nutrition prior to treatment, 12 (75%) transitioned to oral intake for a majority of their nutrition following therapy (FOSS score of 3 or better). Out of all patients studied, 6 (24%) were ultimately on a normal diet following therapy (FOSS score of 0 or 1).

Only 3 patients required 10 or more dilations. Two of these had required initial combined anterograde-retrograde dilations via the gastrostomy, whereas the third received numerous maintenance office dilations. They were typically treated about 3 months apart as they subjectively felt improvements with each office dilation.

Patients who were treated within 6 months after completion of CRT (early dilation) had improved results relative to those treated beyond 6 months (late dilation). Among the 13 patients with early dilation, the mean pretreatment and posttreatment FOSS scores were 4.5 and 2.2, respectively, whereas the 12 patients with late treatment had mean pretreatment and posttreatment FOSS scores of 4.2 and 2.7, respectively. Only 1 of 13 early patients had a posttreatment FOSS score of 4 or 5, as compared to 3 of 12 patients in the late group. There were no documented complications, including zero occurrences of esophageal perforation or mediastinitis.

Discussion

Dysphagia resulting from esophageal stenosis following successful chemoradiation therapy for HNSCC has a significant effect on quality of life.[20] In this setting, optimal treatment is accomplished with the use of serial dilation.[6,21,22] At our institution, we have developed an algorithm to manage esophageal stenosis in the setting of prior CRT, where initial evaluation includes the complementary studies of MBSS, FEES, and transnasal esophagoscopy.

The first dilation occurs in a controlled, operative setting under general anesthesia. The flexible scope is preferred because many of these patients have trismus, friable pharyngeal mucosa, and/or lack of extension precluding rigid esophagoscopy. The otolaryngologist is also more familiar with use of this scope, which has improved maneuverability compared to the regular or even the "ultrathin" but long scope that is typically used in gastroenterology. Following visualization of the stenosis, dilation is performed with CRE balloon or Savary-Gilliard dilators. When using the latter, a guidewire is first passed atraumatically through the stenosis—either parallel to the scope or through the working port of the scope—before the dilator is introduced, thus minimizing the risk of mucosal trauma or extraluminal passage. Retrograde esophagoscopy via the gastrostomy site remains a safe option for patients with complete stenosis. Mitomycin-C can also be applied at this time. The complication risk is very low, and all patients could be discharged to home after recovery from anesthesia. Depending on the severity of stenosis, the timing and the setting of future dilations (office vs operative) are determined.

In our series of patients, we have demonstrated excellent outcomes with our structured management of esophageal stenosis. On Wilcoxon signed-rank test, there was a statistically significant improvement (ie, decrease) in FOSS score, with 6 patients (24%) ultimately tolerating a normal diet (FOSS score of 1). Sixteen patients (64%) were initially G-tube dependent (FOSS score of 5); 12 of these patients (75%) tolerated the oral route for the majority of nutrition (FOSS score of 3 or better) following our therapy.

This compares favorably to previous series: Silvain et al[6] described an early series of 11 patients with esophageal stricture, 9 of whom underwent dilation. This series noted complications in 4 patients, including 1 death, and 4 patients were described to have a semisolid diet after treatment. Dhir et al[23] performed dilations on 21 patients who had undergone radiation with or without surgery and achieved dysphagia relief in 15 of 20 (75%) patients for a median of 14 weeks; however, long-term follow-up was not available. Laurell et al[7] described a similar group who developed moderate to severe esophageal stenosis; their management included both endoscopic dilation and microvascular free flap esophageal reconstruction. In this study, a "nearly normal" diet was achieved in 17 of 22 (78%) patients, although

there was no report of preintervention or postintervention G-tube status. Ahlawat et al[24] performed dilation on 24 patients and reported technical success (endoscopic dilation to 14 mm) in 19 patients and functional success (occasional dysphagia to solid foods) in 18 patients. Again, G-tube status was not available. Our technique improves on these outcomes, however, as the rate of conversion from G-tube dependence to predominantly oral nutrition—75% in our study population—greatly exceeds the success rates reported previously.[6,7,23,24] Furthermore, whereas others have demonstrated good results (81% of patients maintaining weight with oral diet) from dilation of the hypopharynx and upper esophagus,[25] we have achieved these results without complications and with serial dilations in the clinic setting without general anesthesia.

Our use of the TNE accomplishes both diagnostic and therapeutic purposes. Transnasal esophagoscopy is well tolerated in awake patients in the office setting, and we employ the same scope in the operating room, which is beneficial for consistency in assessing the degree of stenosis. Exposure for rigid esophagoscopy may be quite difficult or impossible following CRT, and thus use of the flexible TNE improves our ability to treat challenging cases. Some of the residual esophageal lumens are quite small, and using the 5.1-mm TNE allows successful passage through the stenosis that is not always achieved with the larger gastroscopes. Similarly, the small size allows retrograde passage through the gastrostomy without requiring dilation, thus minimizing morbidity; our results compare very favorably to another series of 45 patients using the retrograde approach reporting G-tube site morbidities in 7 of 63 (11%) procedures.[26] The ability to perform transnasal esophagoscopy and dilation in the office setting confers additional advantages, not in the least that general anesthesia and its concomitant risks are avoided.

Conclusion

Patients with esophageal stenosis after CRT can be successfully managed, with the majority achieving a full oral diet. Transnasal esophagoscopy is an important tool in our armamentarium of management of esophageal stenosis following chemoradiation for head and neck cancer. The versatility of transnasal esophagoscopy as an adjunct to esophageal dilation, with either guidewire or balloon dilators, allows for its use in both operative and office settings. As demonstrated here, our algorithm is well tolerated, highly effective, and associated with little morbidity.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005;55(2):74-108.
2. Francis DO, Weymuller EA Jr, Parvathaneni U, Merati AL, Yueh B. Dysphagia, stricture, and pneumonia in head and neck cancer patients: does treatment modality matter? *Ann Otol Rhinol Laryngol*. 2010;119(6):391-397.
3. Nguyen NP, Sallah S, Karlsson U, Antoine JE. Combined chemotherapy and radiation therapy for head and neck malignancies: quality of life issues. *Cancer*. 2002;94(4):1131-1141.
4. Lazarus CL, Logemann JA, Pauloski BR, et al. Swallowing disorders in head and neck cancer patients treated with radiotherapy and adjuvant chemotherapy. *Laryngoscope*. 1996;106(9, pt 1):1157-1166.
5. Smith RV, Kotz T, Beitler JJ, Wadler S. Long-term swallowing problems after organ preservation therapy with concomitant radiation therapy and intravenous hydroxyurea: initial results. *Arch Otolaryngol Head Neck Surg*. 2000;126(3):384-389.
6. Silvain C, Barrioz T, Besson I, et al. Treatment and long-term outcome of chronic radiation esophagitis after radiation therapy for head and neck tumors. A report of 13 cases. *Dig Dis Sci*. 1993;38(5):927-931.
7. Laurell G, Kraepelien T, Mavroidis P, et al. Stricture of the proximal esophagus in head and neck carcinoma patients after radiotherapy. *Cancer*. 2003;97(7):1693-1700.
8. Hutcheson KA, Lewin JS. Functional outcomes after chemoradiotherapy of laryngeal and pharyngeal cancers. *Curr Oncol Rep*. 2012;14(2):158-165.
9. Lee WT, Akst LM, Adelstein DJ, et al. Risk factors for hypopharyngeal/upper esophageal stricture formation after concurrent chemoradiation. *Head Neck*. 2006;28(9):808-812.
10. Best SR, Ha PK, Blanco RG, et al. Factors associated with pharyngoesophageal stricture in patients treated with concurrent chemotherapy and radiation therapy for oropharyngeal squamous cell carcinoma. *Head Neck*. 2011;33(12):1727-1734.
11. Nguyen NP, Smith HJ, Sallah S. Evaluation and management of swallowing dysfunction following chemoradiation for head and neck cancer. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15(2):130-133.
12. Platteaux N, Dirix P, Dejaeger E, Nuyts S. Dysphagia in head and neck cancer patients treated with chemoradiotherapy. *Dysphagia*. 2010;25(2):139-152.
13. Postma GN, Bach KK, Belafsky PC, Koufman JA. The role of transnasal esophagoscopy in head and neck oncology. *Laryngoscope*. 2002;112(12):2242-2243.
14. Lew RJ, Shah JN, Chalian A, Weber RS, Williams NN, Kochman ML. Technique of endoscopic retrograde puncture and dilatation of total esophageal stenosis in patients with radiation-induced strictures. *Head Neck*. 2004;26(2):179-183.
15. Salassa JR. A functional outcome swallowing scale for staging oropharyngeal dysphagia. *Dig Dis*. 1999;17(4):230-234.

16. Veen EJ, Dikkers FG. Topical use of MMC in the upper aerodigestive tract: a review on the side effects. *Eur Arch Otorhinolaryngol*. 2010;267(3):327-334.
17. Gillespie MB, Day TA, Sharma AK, Brodsky MB, Martin-Harris B. Role of mitomycin in upper digestive tract stricture. *Head Neck*. 2007;29(1):12-17.
18. Rosseneu S, Afzal N, Yerushalmi B, et al. Topical application of mitomycin-C in oesophageal strictures. *J Pediatr Gastroenterol Nutr*. 2007;44(3):336-341.
19. Smith ME, Elstad M. Mitomycin C and the endoscopic treatment of laryngotracheal stenosis: are two applications better than one? *Laryngoscope*. 2009;119(2):272-283.
20. Nguyen NP, Frank C, Moltz CC, et al. Impact of dysphagia on quality of life after treatment of head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2005;61(3):772-778.
21. Ellenhorn JD, Lambroza A, Lindsley KL, LaQuaglia MP. Treatment-related esophageal stricture in pediatric patients with cancer. *Cancer*. 1993;71(12):4084-4090.
22. Moghissi K, Pender D. Management of proximal oesophageal stricture. *Eur J Cardiothorac Surg*. 1989;3(2):93-97.
23. Dhir V, Vege SS, Mohandas KM, Desai DC. Dilation of proximal esophageal strictures following therapy for head and neck cancer: experience with Savary Gilliard dilators. *J Surg Oncol*. 1996;63(3):187-190.
24. Ahlawat SK, Davidson BJ, Al-Kawas FH. Successful use of biliary accessories in antegrade dilation of complex upper esophageal stricture due to chemoradiation and surgery. *Dis Esophagus*. 2008;21(1):86-89.
25. Sullivan CA, Jaklitsch MT, Haddad R, et al. Endoscopic management of hypopharyngeal stenosis after organ sparing therapy for head and neck cancer. *Laryngoscope*. 2004;114(11):1924-1931.
26. Goguen LA, Norris CM, Jaklitsch MT, et al. Combined antegrade and retrograde esophageal dilation for head and neck cancer-related complete esophageal stenosis. *Laryngoscope*. 2010;120(2):261-266.

Reprinted by permission of Laryngoscope. 2014; 124(10):2345-2351.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Correlation Between Reflux and Multichannel Intraluminal Impedance pH Monitoring in Untreated Volunteers

Marie E. Jetté, MS; Eric A. Gaumnitz, MD; Martin A. Birchall, MD; Nathan V. Welham, PhD; Susan L. Thibeault, PhD

Objectives/Hypothesis: Although probable causative agents have been identified (e.g., refluxate components, tobacco smoke), the definitive mechanism for inflammation-related laryngeal mucosal damage remains elusive. Multichannel intraluminal impedance combined with pH monitoring (MII/pH) has emerged as a sensitive tool for diagnosis and characterization of gastroesophageal reflux disease with laryngopharyngeal manifestations. To determine the relationship between laryngeal signs and MII/pH, we examined correlations between Reflux Finding Score (RFS) ratings of videostroboscopic laryngeal examinations and findings from MII/pH.

Study Design: Correlational study.

Methods: Healthy, untreated volunteers (n = 142) underwent reflux diagnosis using data acquired from MII/pH testing. Eight trained clinicians performed RFS ratings of corresponding laryngeal examinations. Averaged RFS ratings were compared to MII/pH data using Pearson correlation coefficients. The relationship between RFS and MII/pH findings and demographic/clinical information (age, sex, smoking status, reflux) was assessed using general linear modeling. Rater reliability was evaluated.

Results: Posterior commissure hypertrophy was negatively correlated with minutes of nonacid refluxate ($R = -0.21$, $P = .0115$). General linear modeling revealed that 28% to 40% of the variance in ratings of ventricular obliteration, erythema/hyperemia, vocal fold edema, diffuse laryngeal edema, posterior commissure hypertrophy, and granulation/granuloma could be explained by main and interaction effects of age, sex, smoking status, and reflux. Intra- and inter-rater reliability for RFS were poor–fair.

Conclusions: These results support the theory that the RFS is not specific for reflux in healthy, untreated volunteers, suggesting there may be alternate explanations for inflammatory clinical signs commonly ascribed to reflux in this population.

Key Words: Impedance monitoring, pH monitoring, gastroesophageal reflux, laryngopharyngeal reflux, laryngopharyngeal reflux diagnosis, Reflux Finding Score.

Level of Evidence: 1b

***Laryngoscope*, 124:2345–2351, 2014**

INTRODUCTION

Chronic laryngitis, one of the most commonly diagnosed dysphonias among healthcare professionals,[1] is characterized by a variety of inflammatory changes observed in patients with an array of symptoms. Gastroesophageal reflux disease (GERD) has been implicated as a probable etiologic factor for chronic laryngitis,[2–4] though treatment with proton pump inhibitors (PPIs), the current standard of care for GERD, demonstrates a nonsignificant benefit over placebo.[5] In spite of lack of efficacy data supporting the use of PPIs, 46.2% of patients with a diagnosis of chronic laryngitis receive medication.[6] Although reflux with laryngeal manifestations (laryngopharyngeal reflux [LPR]) may be an activator of laryngeal inflammation, the extent to which the effects of LPR alone contribute to the clinical picture of chronic laryngitis is unknown.

The Reflux Finding Score (RFS) was developed by Belafsky et al.[7] to document physical LPR findings on a standardized scale, with scores ranging from 0 (no evidence of reflux) to 26 (severe evidence of reflux). To validate this scale, RFS scores from 40 patients with clinically diagnosed LPR documented by esophageal-pharyngeal pH monitoring were compared to scores from 40 age-matched, asymptomatic controls who had not undergone confirmatory pH monitoring, and a statistically significant difference in scores was found.[7] Based on these results, the authors concluded with 95% certainty that a person with RFS >7 has LPR. Other researchers have determined that findings and symptoms ascribed to LPR are not specific to LPR.[8] Milstein et al.[9] found at

From the Department of Surgery, Department of Communication Sciences and Disorders (M.E.J.), Department of Medicine (E.A.G.), Department of Surgery (N.V.W., S.L.T.), University of Wisconsin–Madison, Madison, Wisconsin, U.S.A; and University College London Ear Institute (M.A.B.), The Royal National Throat, Nose and Ear Hospital, London, United Kingdom.

Editor's Note: This Manuscript was accepted for publication April 22, 2014.

Presented at the 2014 Combined Otolaryngological Spring Meetings, American Laryngological Society, Las Vegas, Nevada, U.S.A., May 15, 2014.

This work was supported by grants T32 DC-9401 and R01 DC 9600 from the NIH/NIDCD.

The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Susan L. Thibeault, PhD, 5107 Wisconsin Institutes for Medical Research, 1111 Highland Avenue, Madison, WI 53705. E-mail: thibeaul@surgery.wisc.edu

DOI: 10.1002/lary.24737

least one sign of laryngeal tissue irritation in the majority of volunteers undergoing laryngoscopy with no history of ear-nose-throat complaints or diagnosis of reflux. Similarly, Hicks et al.[10] demonstrated that 86% of normal, healthy, adult volunteers had findings commonly associated with reflux. Moreover, studies examining reliability of subjective laryngoscopic ratings of LPR have revealed mixed results ranging from poor to good.[11]

Ambulatory pH monitoring has been lauded as the gold standard for diagnosis of acid reflux; however, its role in diagnosing LPR remains controversial. In a review of multiple studies, Vaezi et al.[12] revealed that only 54% of patients with laryngoscopic signs of reflux have abnormal esophageal acid exposure on pH probe. They suggest that such low accuracy demonstrates either overdiagnosis of reflux as the cause of laryngeal pathology or lack of sensitivity of pH monitoring in documenting LPR.[12] Other diagnostic tools developed more recently include multichannel intraluminal impedance (MII), pharyngeal pH monitoring,[13] and hypopharyngeal MII (HMII).[14] Impedance monitoring (including MII and HMII) measures both acid and nonacid reflux in liquid and gaseous forms by measuring electrical resistance between different points along the esophagus. Combined with pH monitoring, impedance may offer improved detection of reflux events associated with LPR, though its role in LPR diagnosis has not been established.

The primary goal of this study was to examine correlations between endoscopic findings using RFS and measures acquired from MII with pH (MII/pH) monitoring in healthy, untreated volunteers. Given that the pathophysiology of laryngeal inflammation has not yet been defined and concerns have been published in the literature regarding the specificity of the RFS, we hypothesized that there would be poor correspondence between these sets of variables.

MATERIALS AND METHODS

Participant Selection

Participants aged 21 to 65 years were recruited with newspaper and email advertisements and signs in the clinic and around the University of Wisconsin–Madison. Participants underwent videolaryngostroboscopic examination and 24-hour MII/pH, with each procedure performed on separate dates. The protocol was approved by the institutional review board of University of Wisconsin–Madison, and informed consent was obtained from all participants.

Participants were excluded from the study if they had a history of radiation therapy to the head and neck within the past 5 years, lung or gastroesophageal surgery, chronic sinusitis or rhinitis in the last year, an acute traumatic event near the larynx in the last year, tracheostomy or other significant laryngeal or tracheal surgery, and substance or alcohol abuse in the past year. Consumption of more than 10 (women) and 17 (men) units of alcohol per week (means of United Kingdom and United States recommended weekly limits) excluded participants.[15] Further exclusion criteria included malignancy (except superficial basal cell carcinoma) within the past 5 years; presence of an infectious cause of laryngitis in the past 3 months; need for continuous therapy with diazepam, phenytoin, mephenytoin, warfarin, anticholinergics, antineoplastics, prostaglandin analogs, H2-receptor antagonists, steroids (inhaled, oral, or intravenous), promotility drugs, and sucralfate; use of any PPI or H2 blockers in the past year; theophylline or any other investigational compound or participation in an investigational drug study in the previous 60 days. Women were excluded if pregnant or lactating. Nonsmokers had not smoked during the previous year. Smokers were defined by consumption of a minimum of five cigarettes/5 g of tobacco per day for the duration of 1 or more years, thereby distinguishing them from light smokers.[16,17]

Laryngoscopy

Participants underwent videolaryngostroboscopic examination using rigid or flexible endoscope (Pentax Medical, Lincoln Park, NJ). Topical anesthetic was avoided unless the participant exhibited extreme gag reflex and was unable to tolerate examination. The larynx was visualized during sustained phonation on /i/ and quiet breathing. Digital recordings of laryngoscopic examinations were edited, randomized by clip number (List Randomizer, random.org), and organized into two video montages (iMovie; Apple, Cupertino, CA) representing two randomizations. Sixteen video clips were chosen randomly (List Randomizer, random.org) and included at the end of each video montage to assess intrarater reliability.

Reflux Finding Score

Eight raters provided ratings for this analysis using an adapted RFS (Table I). Raters included clinicians with 55 combined years of experience in voice disorders. A 45-minute training presentation was developed demonstrating published photographic examples of each RFS item[7,18,19] as well as their descriptions. Following training, raters were presented with still images from five examinations and performed group consensus ratings. Notes from the presentation and consensus ratings were saved and raters were able to access these while completing the RFS. Six of eight raters completed the training session with consensus. Two raters that did not attend reviewed the presentation and consensus notes before completing ratings. No demographic or MII/pH data were provided to raters. Raters were also blinded to the purpose of investigation and participant classification.

Combined Multichannel Intraluminal Impedance and 24-Hour pH Probe

After a four-hour fast, participants underwent conventional esophageal manometry (circumferential probe; Medtronic, Shoreview, MN) to locate lower and upper esophageal sphincters (LES and UES, respectively). The MII/pH catheter had two antimony electrodes placed such that proximal sensor was positioned 1 cm below and distal sensor 15 cm below the UES. Impedance was measured through seven sensors placed along a 2.3-mm polyurethane catheter. This catheter was placed transnasally immediately following manometry. Configuration of the catheter allowed recording of changes in intraluminal impedance at 3, 5, 7, 9, 15, and 17 cm above the LES. Data from impedance channels and pH electrodes were transmitted at 50 Hz and stored together on a portable data recorder (Sleuth; Sandhill Scientific Inc., Highlands Ranch, CO) for later synchronization. Participants were monitored for 18–24 hours and encouraged to eat regular meals and participate in routine activities. Change in position (upright and supine) and symptomatic events including heartburn or regurgitation were documented by using buttons on the data recorder. Data were uploaded and analyzed using commercially available software (Bioview Analysis; Sandhill Scientific Inc.).

TABLE I.
Reflux Finding Score Rating Rubric Adapted From Belafsky, Postma, and Koufman.[7]

Reflux Finding Score	
Subglottic edema (pseudosulcus; aka "infraglottic edema")	2 = present, 0 = absent
Ventricular obliteration (false vocal fold edge is indistinct; "complete" refers to the true and false folds appearing to touch)	2 = partial, 4 = complete
Erythema/hyperemia (redness)	2 = arytenoids only, 4 = diffuse
Vocal fold edema (mild is slight swelling, moderate is more perceptible, severe is sessile)	1 = mild, 2 = moderate, 3 = severe, 4 = polypoid
Diffuse laryngeal edema (size of airway relative to size of larynx)	1 = mild, 2 = moderate, 3 = severe, 4 = obstructing
Posterior commissure hypertrophy (pachydermia; mild is mustache-like appearance, moderate is straight line across back of larynx, severe is bulging into airway, and obstructing is airway obliterated)	1 = mild, 2 = moderate, 3 = severe, 4 = obstructing
Granuloma/granulation	2 = present, 0 = absent
Thick endolaryngeal mucus	2 = present, 0 = absent
Total =	

Analysis of pH data. Acid reflux episodes were defined as drops in pH to <4 for at least 5 seconds. Total acid exposure time (%) was calculated as total time of acid reflux episodes divided by monitoring time. Johnson/DeMeester score[20] was obtained using six parameters: 1) total percentage time pH <4.0, 2) percentage time pH <4.0 in an upright position, 3) percentage time pH <4.0 in a recumbent position, 4) total number acid reflux episodes, 5) total number acid reflux episodes longer than 5 minutes, and 6) duration of longest acid reflux episode.

Analysis of MII data. Recorded meal periods were excluded from analysis. On impedance, gas reflux was defined as rapid (>3,000 Ω/s) retrograde moving increase in impedance in at least two impedance sites. Liquid reflux was defined as retrograde moving 40% fall in impedance in two distal impedance sites. Proximal reflux was considered when refluxate reached the 15-cm impedance sensor. Total bolus exposure time (%) was defined as the combination of durations of gas and liquid reflux events divided by total time monitored.

Interpretation of combined dual-channel MII/pH data. Participants were assigned to cohorts—GERD, LPR, normal—based on MII/pH data. GERD was defined by acid exposure percent time of the distal pH probe >4.0, DeMeester score >14.7, and/or bolus exposure percent time of more than 1.4%.[21] LPR was defined by >31 proximal reflux events.[22,23] Normal was defined by the following criteria: acid exposure percent time of the distal pH probe <4.0, DeMeester score <14.7, and <31 proximal reflux events.[22]

Statistical Analysis

To determine inter-rater reliability, intraclass correlation coefficients (ICC) were calculated. Pearson correlation coefficients were used to evaluate intrarater reliability. Average within rater agreement across all eight raters was computed for each RFS item. RFS ratings for each videostroboscopic examination were averaged across all ratings from eight individual raters. Pearson correlation coefficients were used to determine correlations between average RFS ratings and findings on MII/pH and correlations between age and average RFS ratings. General linear models, including repeated measures analysis of variance and analysis of covariance, were fitted to assess main effects of age, cohort, sex, and smoking status, as well as the two-, three-, and four-way interaction effects of age*sex, age*cohort, age*smoking status, cohort*sex, cohort*smoking status, sex*smoking status, age*sex*smoking status, age*cohort*smoking status, age*sex*cohort, cohort*sex*smoking status, and age*cohort*sex*smoker for all RFS ratings. *t*-tests were used to examine differences in variables that could not be accounted for by linear modeling. All analyses were performed with SAS software (SAS Institute Inc., Cary, NC) with type I error set at 0.05.

RESULTS

Clinical and Demographic Characteristics

Of 155 original video clips included in the montages provided to raters, 13 were excluded from rating and analysis due to insufficient views from anterior commissure to posterior pharyngeal wall. Data from 142 participants including videolaryngostroboscopic recordings, MII/pH variables (Table II), and averaged RFS ratings (Table II) were therefore included in the final analysis. Analysis of MII/pH data revealed 38 participants with GERD (27%), 44 with LPR (31%), and 60 normal (42%). Of 142 participants, 116 (82%) had total RFS >7, and 55 (39%) had total RFS >11. Age, sex, smoking, reflux cohort, and total RFS characteristics of these participants are summarized in Table III. Videostroboscopic examination and MII/pH testing were completed with an average of 61 days between each procedure.

RFS Rater Reliability and Agreement

ICC for intrarater reliability ranged from 0.05 to 0.45 (Table IV). Results demonstrate poor to fair reliability for all RFS rating items. Inter-rater reliability was assessed on 256 observations from eight raters. ICC ranged from 0.21 to 0.48 (Table IV), indicating poor to fair inter-rater reliability for all RFS rating items. Average intrarater agreement examines overall levels of rater self-consistency, for each rater and RFS rating. Results are based on repeated ratings of 16 video clips, and indicate that individual raters were 54.8% to 71.7% reliable across all ratings, and that they produced the same value for any individual variable 48.75% to 78.75% of the time (Table V).

TABLE II.
Summary of Reflux Finding Score and Multichannel Intraluminal Impedance pH Monitoring Variables.

	Mean	SD	Minimum	Maximum
RFS variables				
Subglottic edema	0.78	0.74	0	2
Ventricular obliteration	1.77	0.83	0	3.75
Erythema/hyperemia	2.91	0.87	0.5	4
Vocal fold edema	1.25	0.75	0	3.75
Diffuse laryngeal edema	1.05	0.57	0	3
Posterior commissure hypertrophy	1.63	0.66	0	3
Granulation/granuloma	0.38	0.45	0	2
Thick endolaryngeal mucus	1.18	0.72	0	2.25
Total	10.38	3.63	1.75	19.875
MII/pH variables				
Measured by pH monitoring				
% total time pH <4	3.34	7.95	0	80.5
% upright time pH <4	3.69	7.70	0	62.8
% supine time pH <4	2.56	9.41	0	94.9
No. of reflux episodes	19.50	14.20	0	76
No. of reflux episodes ≥5 minutes	1.09	2.64	0	18
Longest reflux episode (min)	13.84	43.69	0	444.8
Johnson/DeMeester score	14.78	31.85	0.8	256.0
Measured by multichannel intraluminal impedance				
Acid refluxate (min)	13.67	15.63	0	102.9
Nonacid refluxate (min)	7.08	9.78	0	102.3
Total % time reflux (min)	1.79	1.66	0.1	12
No. of reflux events	44.3	21.01	7	105
No. of acid reflux events	24.42	17.28	0	91
No. of nonacid reflux events	19.78	11.56	0	62
No. of reflux events that reached the proximal esophagus	23.75	13.97	3	72
No. of acid reflux events that reached the proximal esophagus	14.69	11.32	0	52
No. of nonacid reflux events that reached the proximal esophagus	9.06	6.65	0	38

MII = multichannel intraluminal impedance; RFS = Reflux Finding Score; SD = standard deviation.

Correlations Between RFS and MII/pH

Average RFS ratings for each videostroboscopic examination were compared to individual MII/pH variables resulting in 144 analyzed correlations across 142 participants. There was a single significant correlation between posterior commissure hypertrophy and minutes of nonacid refluxate ($R = -0.21$, $P = .0115$). No other correlations were significant (data not shown).

Effect of Clinical and Demographic Characteristics on RFS

Average RFS ratings for each variable were analyzed relative to clinical and demographic data including cohort, sex, and smoking status. Age was analyzed as a main effect and also included in a separate interaction effects model (Table VI). Interaction effects of cohort, sex, smoking status, and age influenced averaged RFS ratings. General linear modeling, including all variables and their interactions (Table VI, model 2), explained 25% to 40% of the variance observed in many RFS ratings. Although both models tested could not account for variance in ratings of subglottic edema and thick endolaryngeal mucus, further analysis revealed the main effects of sex on both of these variables ($P = .025$, $P = .049$, respectively).

DISCUSSION

The major finding of this study was a single statistically significant correlation between RFS and MII/pH variables in a group of healthy, non–treatment-seeking, untreated volunteers. We found a negative correlation between posterior commissure hypertrophy and duration (minutes) of nonacid reflux ($R = -0.21$, $P = .0115$), suggesting that posterior commissure hypertrophy is decreased with greater duration of nonacid reflux. This result is supported by biological evidence demonstrating less proinflammatory cytokine gene expression with greater acid exposure in biopsies taken from the posterior commissure.[24] Though this correlation coefficient is statistically significant, it is meaningless unless properly interpreted for clinical relevance. Calculating

TABLE III.
Participant Characteristics.

Characteristic	No. (%)	Mean Age, yr
Sex		
Male	64 (45)	40.1
Female	78 (55)	43.5
Cigarette smoking		
Nonsmoker	107 (75)	40.6
Smoker	35 (25)	42.9
Reflux cohort		
GERD	38 (27)	43.4
LPR	44 (31)	37.5
Normal	60 (42)	42.5
Total reflux finding score		
<7	26 (18)	42.0
≥7	116 (82)	41.3
<11	87 (61)	41.4
≥11	55 (39)	41.2

GERD = gastroesophageal reflux disease; LPR = laryngopharyngeal reflux.

coefficient of determination (R^2) yields 0.044, meaning that 4.4% of variation in ratings of posterior commissure hypertrophy can be explained or accounted for by variation in duration of nonacid reflux. This interpretation of the data suggests there are other factors (e.g., demographic characteristics) aside from reflux findings measured by MII/pH that may explain variability in RFS ratings. It is also possible that there is an inherent lack of RFS validity for specific reflux diagnosis.

The primary outcome measures of our study were eight RFS ratings in addition to total RFS averaged across eight trained clinician raters and 16 MII/pH variables. Though averaged RFS ratings were used for analysis, it is worth noting that inter- and intrarater reliability for RFS was poor–fair. In a review of the literature examining reliability for laryngopharyngeal findings in LPR, Powell and Cocks[11] presented a summary from nine publications demonstrating variable reliability ranging from poor–good. They suggested variability might be related to methods of assessment or statistical tests used. Potential explanations for poor intrarater reliability observed in our study relate to the inherent limits of human raters' visual-perceptual systems and the RFS scale itself. Rosen[25] has suggested several limitations and possible errors associated with visual-perceptual ratings of videostroboscopy, including rater fatigue and lack of variability of videos. Additionally, whereas some variables (e.g., subglottic edema) can be scored as 0 (absent) or 2 (present), other variables (e.g., vocal fold edema) are scored on a 5-point scale (0, 1, 2, 3, 4). When data are pooled for statistical calculation of intrarater reliability, the difference between ratings of 0 and 2 is given greater weight than the difference between ratings on a five-point scale. Examining agreement in conjunction with reliability gives an indication of statistical penalties resulting from limits of the scale. For example, upon repeat rating of thick endolaryngeal mucus, clinicians on average agreed with their initial rating 72.4% of the time, whereas intrarater reliability was calculated at $R = 0.12$ ($P = .0001$) indicating poor reliability. Agreement implies that two raters assign identical meanings to each score for each variable, whereas reliability indicates that raters rate variables in parallel fashion, without implying that score values have the same meaning. If the range of scores is restricted (e.g., raters consistently avoid extremes of a scale or scores vary little with respect to variable rated), reliability coefficients may be low, even if raters agree. In this study, it is possible raters avoided severe extremes of the RFS given they were rating images from non–treatment-seeking volunteers as opposed to a pathologic population.

To bolster the clinical relevance of our findings, we used combined MII/pH variables semidiagnostically to categorize our study population into cohorts including LPR, GERD, and normal based on normative data.[21–23] Our study is the first to report on the incidence of

TABLE IV.
Intrarater and Inter-rater Reliability.

	Intrarater Reliability		Inter-rater Reliability	
RFS Variable	*R*	*P* Value	*R*	*P* Value
Subglottic edema	0.05	.06	0.48	<.0001
Ventricular obliteration	0.45	<.0001	0.24	<.0001
Erythema/hyperemia	0.10	.001	0.34	<.0001
Vocal fold edema	0.29	<.0001	0.39	<.0001
Diffuse laryngeal edema	0.17	<.0001	0.29	<.0001
Posterior commissure hypertrophy	0.021	038	0.34	<.0001
Granulation/granuloma	0.20	<.0001	0.21	<.0001
Thick endolaryngeal mucus	0.12	.0001	0.43	<.0001
Total	0.21	.0001	0.48	<.0001

Interpretation of correlation coefficients: <0.40 = poor; 0.40–0.59 = fair; 0.60–0.74 = good; >0.74 = excellent.
RFS = Reflux Finding Score.

TABLE V.
Percent Intrarater Agreement for Each Individual Rater and RFS Rating as Well as Averages Across Raters and RFS Ratings.

	Intrarater % Agreement								Average % Agreement
RFS Variable	R1	R2	R3	R4	R5	R6*	R7	R8*	All Raters
Subglottic edema	70.0	81.8	81.3	72.7	50.0	60.0	90.9	53.8	70.1
Ventricular obliteration	81.8	53.9	100.0	71.4	75.0	88.9	72.7	56.3	75.0
Erythema/hyperemia	73.3	57.2	62.5	66.7	61.5	87.5	61.5	66.7	67.1
Vocal fold edema	76.9	71.4	68.8	58.3	63.6	25.0	30.8	46.7	55.2
Diffuse laryngeal edema	75.0	76.9	18.8	50.0	54.5	50.0	27.3	37.5	48.7
Posterior commissure hypertrophy	42.9	58.3	56.3	58.3	46.2	62.5	66.7	57.2	56.0
Granulation/granuloma	75.0	100.0	81.3	90.0	75.0	77.8	66.7	64.3	78.7
Thick endolaryngeal mucus	78.6	57.1	68.8	68.7	92.9	66.7	92.3	56.3	72.6
Average % agreement	71.7	69.6	67.2	67.0	64.8	64.8	63.6	54.8	

*Indicates rater did not attend training.
RFS = Reflux Finding Score.

GERD and LPR based on MII/pH in untreated, non–treatment-seeking healthy volunteers. Within our participant group, more than half (58%) was categorized as either LPR or GERD, whereas 42% demonstrated normal findings on MII/pH. Similarly, categorization of participants using published thresholds for total RFS of 7[7] and 11[26] yielded 82% and 39%, respectively, categorized as LPR, supporting Hicks et al.'s finding that 86% of normal, healthy, adult volunteers had signs associated with reflux.[10] In a study investigating the diagnostic usefulness of MII/pH in 98 patients with suspected LPR off PPI therapy for at least 2 weeks, Lee et al. found that 54% demonstrated pathologic GERD,[27] a finding consistent with our data in spite of the difference in study populations. It should be noted that in our study, LPR was determined based on impedance and pH findings in the proximal esophagus, not in the hypopharynx, which may have resulted in overestimation of incidence of LPR. Supporting this possibility, an investigation of 34 asymptomatic, untreated research participants using hypopharyngeal MII/pH revealed a single LPR event recorded from one participant (3%), whereas in symptomatic, untreated patients, 24/184 (13%) had at least one LPR event documented.[14] In clinical practice, gastroenterologists use MII/pH to diagnose reflux in patients with persistent symptoms despite acid-suppressive therapy. Diagnosis includes examining symptom association[28] (i.e., determining whether episodes recorded by MII/pH are associated with a corresponding symptom) and comparing MII/pH variables in patients on therapy to normative values.[29] As we were attempting to use MII/pH as the sole objective measure of reflux in a non–treatment-seeking population, symptom association and treatment response were not evaluated within the present research design.

The clinical/demographic interaction and main effects observed within our dataset provide insight into factors that explain some variance in RFS ratings. General linear modeling including main and interaction effects of age, reflux cohort, sex, and smoking status could explain 25% to 40% of the variance observed in all RFS variables except subglottic edema and thick endolaryngeal mucus, suggesting that RFS ratings are influenced by clinical and demographic factors. Inflammatory

TABLE VI.
Summary of Results of Generalized Linear Modeling.*

	Age		Model 1		Model 2 (With Age)	
RFS Variable	R^2	*P*	R^2	*P*	R^2	*P*
Subglottic edema	0.001	0.69	0.11	.19	0.16	.51
Ventricular obliteration	0.08	0.0006	0.29	<.0001	0.40	<.0001
Erythema/hyperemia	0.01	0.19	0.35	<.0001	0.39	<.0001
Vocal fold edema	0.04	0.03	0.33	<.0001	0.39	<.0001
Diffuse laryngeal edema	0.03	0.04	0.28	<.0001	0.37	<.0001
Posterior commissure hypertrophy	0.03	0.04	0.17	.01	0.25	.03
Granulation/granuloma	0.03	0.03	0.05	.80	0.28	.01
Thick endolaryngeal mucus	0.03	0.70	0.10	.20	0.16	.47
Total	0.02	0.09	0.35	<.0001	0.39	<.0001

*Generalized linear modeling demonstrating the total variance (R^2) accounted for by: 1) the main effect of age; 2) the main and interaction effects of cohort, sex, smoking status; and 3) the main and interaction effects of cohort, sex, smoking status, and age for each RFS variable.
RFS = Reflux Finding Score.

signs measured with RFS are in part related to the combinations of sex, smoking status, and age of the larynx being rated as opposed to reflux alone. Subglottic edema, also referred to as pseudosulcus and infraglottic edema, has long been thought to be predictive of,[30] and specific for,[19] LPR; however, our results demonstrate that males receive greater ratings than females on this variable regardless of reflux cohort, smoking status, and age. It seems possible that this finding so commonly ascribed to inflammation from reflux may be a result of anatomic differences between males and females. Males also received greater ratings than females for thick endolaryngeal mucus, suggesting that this finding provides more information about the sex of the person being examined than it does about reflux.

Although attempts were made to eliminate bias, we recognize limitations in our study design that may have prejudiced our results. Of primary consideration is that we examined data from non–treatment-seeking volunteers, a population not representative of a typical clinical population. It would be ideal to repeat the study in treatment-seeking patients for whom laryngeal inflammation impacts vocal function, thereby addressing the role of reflux specific to diagnosis of chronic laryngitis. We also recognize that we persisted in analyzing averaged RFS ratings in spite of poor reliability, though we attempted to avoid this issue by providing raters with training. Finally, we acknowledge that reflux status may have changed in the time between videostroboscopic examination and MII/pH testing. This could be avoided in future studies by completing videostroboscopic examination immediately prior to MII/pH.

CONCLUSION

Our data demonstrate an overall lack of correlation between RFS and MII/pH, supporting the hypothesis that RFS is not specific for reflux in non–treatment-seeking, untreated volunteers. Our findings also illustrate that in spite of training, raters demonstrated poor–fair inter- and intrarater reliability on RFS, consistent with results from other studies. Finally, we suggest that clinical and demographic characteristics, including sex, smoking status, and age, contribute to differences in RFS ratings.

Acknowledgments

The authors thank Dr. Glen Leverson for statistical support.

BIBLIOGRAPHY

1. Cohen SM, Kim J, Roy N, Asche C, Courey M. Prevalence and causes of dysphonia in a large treatment-seeking population. *Laryngoscope* 2012; 122:343–348.
2. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope* 1991;101:1–78.
3. Koufman J, Sataloff RT, Toohill R. Laryngopharyngeal reflux: consensus conference report. *J Voice* 1996;10:215–216.
4. Hanson DG, Jiang JJ. Diagnosis and management of chronic laryngitis associated with reflux. *Am J Med* 2000;108(suppl 4a):112S–119S.
5. Qadeer MA, Phillips CO, Lopez AR, et al. Proton pump inhibitor therapy for suspected GERD-related chronic laryngitis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol* 2006;101:2646–2654.
6. Cohen SM, Kim J, Roy N, Courey M. Assessing factors related to the pharmacologic management of laryngeal diseases and disorders. *Laryngoscope* 2013;123:1763–1769.
7. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope* 2001;111:1313–1317.
8. Barry DW, Vaezi MF. Laryngopharyngeal reflux: More questions than answers. *Cleve Clin J Med* 2010;77:327–334.
9. Milstein CF, Charbel S, Hicks DM, Abelson TI, Richter JE, Vaezi MF. Prevalence of laryngeal irritation signs associated with reflux in asymptomatic volunteers: impact of endoscopic technique (rigid vs. flexible laryngoscope). *Laryngoscope* 2005;115:2256–2261.
10. Hicks DM, Ours TM, Abelson TI, Vaezi MF, Richter JE. The prevalence of hypopharynx findings associated with gastroesophageal reflux in normal volunteers. *J Voice* 2002;16:564–579.
11. Powell J, Cocks HC. Mucosal changes in laryngopharyngeal reflux–prevalence, sensitivity, specificity and assessment. *Laryngoscope* 2013;123: 985–991.
12. Vaezi MF, Hicks DM, Abelson TI, Richter JE. Laryngeal signs and symptoms and gastroesophageal reflux disease (GERD): a critical assessment of cause and effect association. *Clin Gastroenterol Hepatol* 2003;1:333–344.
13. Becker V, Graf S, Schlag C, et al. First agreement analysis and day-to-day comparison of pharyngeal pH monitoring with pH/impedance monitoring in patients with suspected laryngopharyngeal reflux. *J Gastrointest Surg* 2012;16:1096–1101.
14. Hoppo T, Sanz AF, Nason KS, et al. How much pharyngeal exposure is "normal"? Normative data for laryngopharyngeal reflux events using hypopharyngeal multichannel intraluminal impedance (HMII). *J Gastrointest Surg* 2012;16:16–24; discussion 24–25.
15. Furtwaengler NA, de Visser RO. Lack of international consensus in low-risk drinking guidelines. *Drug Alcohol Rev* 2013;32:11–18.
16. Kenford SL, Wetter DW, Welsch SK, Smith SS, Fiore MC, Baker TB. Progression of college-age cigarette samplers: what influences outcome. *Addict Behav* 2005;30:285–294.
17. Husten CG. How should we define light or intermittent smoking? Does it matter? *Nicotine Tob Res* 2009;11:111–121.
18. Belafsky PC, Postma GN, Amin MR, Koufman JA. Symptoms and findings of laryngopharyngeal reflux. *Ear Nose Throat J* 2002;81:10–13.
19. Belafsky PC, Postma GN, Koufman JA. The association between laryngeal pseudosulcus and laryngopharyngeal reflux. *Otolaryngol Head Neck Surg* 2002;126:649–652.
20. Johnson LF, DeMeester TR. Development of the 24-hour intraesophageal pH monitoring composite scoring system. *J Clin Gastroenterol* 1986; 8(suppl 1):52–58.
21. Lee JH, Park SY, Cho SB, et al. Reflux episode reaching the proximal esophagus are associated with chronic cough. *Gut Liver* 2012;6:197–202.
22. Shay S, Tutuian R, Sifrim D, et al. Twenty-four hour ambulatory simultaneous impedance and pH monitoring: a multicenter report of normal values from 60 healthy volunteers. *Am J Gastroenterol* 2004;99:1037–1043.
23. Cho YK. How to interpret esophageal impedance pH monitoring. *J Neurogastroenterol Motil* 2010;16:327–330.
24. Thibeault SL, Smith ME, Peterson K, Ylitalo-Moller R. Gene expression changes of inflammatory mediators in posterior laryngitis due to laryngopharyngeal reflux and evolution with PPI treatment: a preliminary study. *Laryngoscope* 2007;117:2050–2056.
25. Rosen CA. Stroboscopy as a research instrument: development of a perceptual evaluation tool. *Laryngoscope* 2005;115:423–428.
26. Postma GN, Halum SL. Laryngeal and pharyngeal complications of gastroesophageal reflux disease. GI Motility Online. 2006. Available at: http://www.nature.com/gimo/contents/pt1/full/gimo46.html. Accessed on January 31, 2014.
27. Lee BE, Kim GH, Ryu DY, et al. Combined dual channel impedance/pH-metry in patients with suspected laryngopharyngeal reflux. *J Neurogastroenterol Motil* 2010;16:157–165.
28. Weusten BL, Roelofs JM, Akkermans LM, Van Berge-Henegouwen GP, Smout AJ. The symptom-association probability: an improved method for symptom analysis of 24-hour esophageal pH data. *Gastroenterology* 1994;107:1741–1745.
29. Tutuian R, Castell DO. Review article: complete gastro-oesophageal reflux monitoring—combined pH and impedance. *Aliment Pharmacol Ther* 2006;24(suppl 2):27–37.
30. Hickson C, Simpson CB, Falcon R. Laryngeal pseudosulcus as a predictor of laryngopharyngeal reflux. *Laryngoscope* 2001;111:1742–1745.

Reprinted by permission of Laryngoscope. 2013; 123(10):2463-2468.

The Laryngoscope
© 2013 The American Laryngological,
Rhinological and Otological Society, Inc.

The Role of Impedance Monitoring in Patients With Extraesophageal Symptoms

Robert T. Kavitt, MD, MPH; Elif Saritas Yuksel, MD; James C. Slaughter, DrPH; C. Gaelyn Garrett, MD; David Hagaman, MD; Tina Higginbotham, MPA; Michael F. Vaezi, MD, PhD, MSc (Epi)

Objectives/Hypothesis: Ambulatory esophageal impedance monitoring is commonly employed to assess for nonacid reflux in patients with extraesophageal reflux. We aimed to determine if *on* therapy impedance data can be predicted from *off* therapy upper endoscopy, manometry, or pH parameters.

Study Design: Prospective Cohort Study.

Methods: Patients with extraesophageal reflux symptoms and either partial- or nonresponders to twice-daily PPI underwent impedance monitoring *on* twice-daily PPI, as well as manometry, upper endoscopy, and 48-hour wireless pH monitoring *off* acid-suppressive medications for 1 week. Percent time pH < 4 and number of reflux episodes were obtained. Multivariable linear regression was used to determine association between the impedance data *on* therapy and upper endoscopy, manometry, and pH parameters measured *off* therapy.

Results: Seventy-five patients (77% female, median BMI 29, 38% with hiatal hernia, and 19% with esophagitis) were studied both *on* and *off* therapy. Thirty-five percent had abnormal impedance monitoring *on* therapy and 84% had abnormal pH testing *off* therapy. There was no significant ($P = 0.184$) overall correlation between total number of impedance events and the baseline physiologic parameters of hiatal hernia, degree of acid reflux, or manometric findings, with only weak correlation ($r = 0.54$, $P = 0.045$) with % time pH < 4 among patients with esophagitis.

Conclusions: In patients with suspected extraesophageal reflux refractory to PPI therapy, impedance measures *on* therapy cannot be predicted from traditional baseline esophageal physiologic parameters. We recommend caution regarding overinterpretation of impedance data. Laryngoscope, 000:000–000, 2013

Key Words: Impedance, GERD, refractory reflux.

Level of Evidence: 2b.

***Laryngoscope,* 123:2463–2468, 2013**

INTRODUCTION

Gastroesophageal reflux disease (GERD) is a common disorder with increasing prevalence in the Western world.[1] Approximately 40% of adults frequently complain of heartburn,[2] and GERD remains the leading outpatient physician diagnosis for gastrointestinal disorders in the United States.[3]

Extraesophageal reflux (EER) is widely implicated in the etiology of laryngeal, pharyngeal, and pulmonary symptoms, and controversy exists regarding the diagnosis and management of this condition.[4] Currently, most patients with signs and symptoms attributed to EER are empirically treated with proton pump inhibitors (PPIs). However, symptomatic improvement on PPIs is not as consistent compared to those with classic GERD.[5] Patients with presumed EER refractory to initial empiric medical therapy are often referred for further testing. Current guidelines recommend diagnostic testing, which can include the use of upper endoscopy and pH and/or impedance monitoring.[6]

Combined impedance-pH monitoring can detect various types of esophageal refluxate: gas, liquid, acid, or nonacid, and is used to clarify the mechanisms of PPI-refractory symptoms.[7–11] Multicenter studies utilizing impedance-pH testing in patients with PPI-refractory symptoms suggest that approximately one-third of patients exhibit weakly acid or nonacid reflux. In the background of potent acid suppression, the clinical significance of these findings currently remains controversial. While some advocate for the clinical utility of impedance-pH monitoring in assessing the impact of weakly acidic material on patients' persistent symptoms, others are not as enthusiastic. Studies have suggested that 40% to 50% of patients with persistent symptoms on acid-suppressive therapy have no temporal correlation between their symptoms and any type of reflux.[12,13]

While there is no doubt that impedance-pH testing is currently the most accurate and detailed method for detecting reflux of all kinds, the clinical indications for its use are still evolving. Its role in the management of GERD patients awaits further definition, mainly due to

From the Division of Gastroenterology, Hepatology and Nutrition (R.T.K., E.S.Y., T.H., M.F.V.), the Department of Biostatistics (J.C.S.), the Vanderbilt Voice Center; Department of Otolaryngology (C.G.G.), and the Allergy, Sinus and Asthma Program (D.H.), Vanderbilt University Medical Center, Nashville, Tennessee, U.S.A.

Editor's Note: This Manuscript was accepted for publication August 22, 2012.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Michael F. Vaezi, MD, PhD, MSc (Epi), Division of Gastroenterology, Hepatology and Nutrition, Vanderbilt University Medical Center, 1660 TVC, 1301 22nd Ave. South, Nashville, TN 37232-5280. E-mail: Michael.vaezi@vanderbilt.edu

DOI: 10.1002/lary.23734

the paucity of high-quality studies examining the clinical relevance of impedance findings. Thus, the aim of our study was to determine if *on* therapy impedance values assessing for nonacid reflux can be predicted from the traditionally recognized and commonly employed clinical markers of reflux based on *off* therapy upper endoscopy, manometry, or pH monitoring.

MATERIALS AND METHODS

The study was performed in accordance with the Declaration of Helsinki, Good Clinical Practice, and applicable regulatory requirements. The Vanderbilt Institutional Review Board approved this clinical trial (IRB# 090872).

STUDY DESIGN AND PATIENT POPULATION

The study population consisted of patients with suspected reflux-associated extraesophageal symptoms refractory to PPI therapy referred to the Esophageal Motility Center at Vanderbilt University Medical Center for evaluation and treatment. Refractory symptoms were defined as less than 50% improvement in the chief complaint after at least 12 weeks of twice-daily PPI therapy. This group was chosen since they represent the patient population for whom pH and/or impedance monitoring is currently indicated.[14,15] In order to assess the severity and frequency of their extraesophageal symptoms, patients were asked to complete a questionnaire previously described in the literature.[16] The following information were collected for all patients: presence, severity, and frequency of GERD symptoms (heartburn ± regurgitation) and extraesophageal symptoms (cough, hoarseness, throat clearing, sore throat, globus sensation, postnasal drip symptoms, chest pain), current medications, demographic data (age, sex, race, body mass index), history of alcohol and tobacco use, and presence of voice/laryngeal and nasal symptoms.

Patients underwent esophagogastroduodenoscopy (EGD), wireless 48-hour pH monitoring, and esophageal manometry *off* acid suppression to assess the baseline esophageal mucosal integrity, motility, and acid exposure. They also underwent 24-hour impedance/pH monitoring while *on* twice daily PPI therapy to determine the presence of acid and nonacid reflux in the setting of acid suppression. The presence and size of a hiatal hernia were determined at endoscopy; the presence and severity of esophagitis was graded by the Los Angeles Classification; and the presence of Barrett's esophagus was also noted.[17] Inclusion criteria were age greater than 18 years and chronic EER symptoms refractory to PPI therapy. Patients were excluded from the study if they were unwilling to undergo testing, were pregnant, had undergone surgery for reflux or peptic ulcer disease, or if they had a serious illness that would interfere with study participation.

Wireless pH Monitoring

Ambulatory pH monitoring was performed for 48 hours using a wireless monitoring device (Given Imaging, Duluth, GA). Patients were instructed to stop taking all PPIs and H2-receptor antagonists (H2RAs) for at least 7 days prior to undergoing testing. Wireless capsules were calibrated by submersion in buffer solutions at pH 7.0 and pH 1.0, and then activated by magnet removal. Patients underwent EGD with conscious sedation for visual anatomic inspection and distance measurements from the incisors to the squamocolumnar junction (SCJ). Capsules were then placed using the manufacturer's delivery system at 6 cm above the SCJ and attached with vacuum suction. Capsule placement was confirmed with endoscopy. After successful placement, patients were given wireless pH recorders to wear about their waists, or to keep within 3 feet to 5 feet at all times. Recording devices receive pH data sampling transmitted by the capsule at 433 Hz with 6 second sampling intervals. Patients were instructed to perform their normal daily activities and dietary practices. Distal esophageal pH recording was conducted for a total of 48 hours. During this time patients kept diaries of meal times, symptoms, and supine positioning.

After completion of the 48-hour study, data were downloaded from recording devices to dedicated computers. Patient diary information was manually entered into the computer-based record. Measurements of the total, upright, and supine percentage time when esophageal pH was below 4 were determined over day 1 and day 2 of the wireless study. Total acid exposure time (% total time pH less than 4) greater than 5.5% was considered abnormal, while greater than 8.2% was considered abnormal for the upright state and greater than 3.0% was considered abnormal for the supine state.[18]

Esophageal Motility Testing

High-resolution manometry (Given Imaging, Duluth, GA) was used to measure the location of the lower esophageal sphincter prior to placement of the impedance-pH catheter. A solid-state assembly with 36 circumferential sensors spaced at 1-cm intervals (outer diameter 4.2 mm) was used. This device detects pressure over a length of 2.5 mm in each of the 12 radially dispersed sectors of the 36 pressure-sensing elements. The sector pressures are averaged, making the sensors a circumferential pressure detector. Prior to recording the transducers were calibrated at 0 and 100 mm Hg using externally applied pressure. Using this device the lower esophageal sphincter was measured and the proximal location noted for placement of the impedance-pH catheter.

Combined Impedance-pH Monitoring

Patients underwent impedance-pH monitoring while *on* at least twice daily PPI therapy for 1 month prior to evaluation. They were instructed to fast for 4 hours before testing. Each patient's primary symptom complaint was recorded as part of the preprocedure evaluation. Patients were given diaries to record the timing of initiation and completion of meals and position changes (upright or supine) during 24-hour impedance and pH monitoring.

Impedance testing was performed using a combined impedance-pH monitoring device (Sandhill Scientific Inc; Highlands Ranch, CO) comprising a data recorder (Sleuth System; Sandhill Scientific Inc.) and a 2.1-mm diameter

polyvinyl catheter embedded by one pH and six impedance sensors at predefined positions. The pH sensors were calibrated before placement using standardized buffer solutions at pH 4.0 and 7.0, as recommended by the manufacturer. The catheter was placed intranasally so that the esophageal pH sensor was positioned 5 cm above the manometrically defined upper border of the lower esophageal sphincter. Intraluminal impedance was measured at 3, 5, 7, 9, 15, and 17 cm above the lower esophageal sphincter. Data sampling frequency for both impedance and pH sensors was 50 Hz. Studies were performed for 24 hours, after which patients returned to the lab for catheter removal and data review.

Data were downloaded from the recorder and analyzed using BioView Analysis software (Sandhill Scientific Inc). Reflux episodes were identified by computerized detection (Autoscan; Sandhill Scientific Inc.) of proximally directed decreases in impedance. Tracings were then manually reviewed by an experienced investigator (MFV) to confirm accuracy and correct any errors. Total, upright, and supine reflux events were recorded. Acid reflux events were defined as those occurring with pH less than 4.0, and nonacid reflux events were defined as those occurring at pH of 4.0 or greater. For impedance parameters, total number of reflux events greater than 48 was considered abnormal.[19–22]

Statistical Analysis

Data were collected and stored at the secure Web-based Vanderbilt Digestive Disease Center REDCap (Research Electronic Data Capture) (1 UL1 RR024975 NCRR/NIH). REDCap is an application designed to support data capture for research studies providing: 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources. There was strict control and supervision of the data entry and access for this study.

Continuous variables were summarized using the median, 25th, and 75th percentiles. Categorical variables were summarized using percentages. We used separate Kruskal-Wallis tests to determine if the total number of impedance events and the percent total time pH less than four differed by esophagitis and pH groups. Spearman's rank correlation was used to estimate the overall association between impedance events (on therapy) and percent total time pH less than 4 (off therapy). Multivariable linear regression was used to assess the relationship between *on*-therapy impedance parameters and *off*-therapy pH parameters, manometry findings, and upper endoscopy findings. All analyses were conducted using the R statistical program.

RESULTS

Demographics and Endoscopic Findings

Seventy-five patients with suspected extraesophageal reflux underwent testing with 48-hour wireless pH monitoring *off* PPI therapy and 24-hour impedance/pH monitoring *on* PPI therapy between 2005 and 2012. Their chief complaints included: cough (53%), asthma (12%), hoarseness (7%), throat clearing (6%), pulmonary fibrosis (6%), sore throat (6%), postnasal drip (5%), and sinusitis (5%). A total of 58/75 (77%) of patients were female, 71/75 (95%) were Caucasian, with a median (IQR) age of 55 (45–64) years, and median (IQR) body mass index of 29 (25–33). Hiatal hernias greater than 1 cm in size were present in 29/75 (38%), of which 61% were 2 to 3 cm in size, and 39% were 4 cm or greater. Esophagitis was present in 14/75 (19%) of patients, of which 90% were grade A or B by Los Angeles Classification. None of the patients were found to have endoscopic evidence of possible Barrett's esophagus.

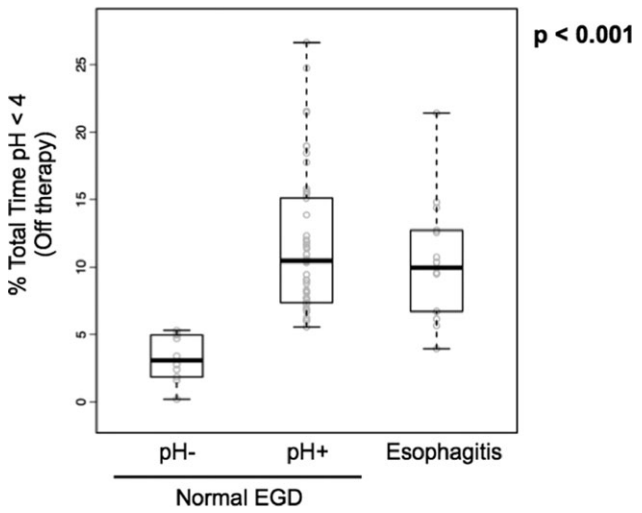

Fig. 1. Individual and group median (IQR) % total time pH < 4 for the three patient subgroups. As expected, acid reflux was significantly less in those with normal endoscopy and pH than those with abnormal pH or esophagitis.

Impedance and pH Parameters

Overall. Median (IQR) total, upright, and supine impedance events *on PPI therapy* were 39 (20–54), 29 (6–12), and 4 (1–8), respectively. The impedance events were predominantly mixed gas and liquid of nonacidic nature (pH > 4) and were abnormal (greater than 48 reflux events) in 35% of tested subjects. Abnormal acid reflux was not present in any of the subjects when tested *on PPI* therapy. Wireless pH testing *off PPI therapy* showed a median (IQR) % time pH < 4 of 9% (7%–13%) in total, 11% (8%–16%) in the upright state, and 2% (1%–8%) in the supine state. Eighty-four percent of patients had abnormal wireless pH testing results *off* therapy.

Subgroups. Patients were divided into three subgroups based on endoscopic and pH findings: Esophagitis, normal endoscopy but abnormal pH (referred to as "pH+" in Figures 1 and 2), and normal endoscopy with normal pH (referred to as "pH−" in Figures 1 and 2). Figures 1 and 2 depict individual and group results on the degree of

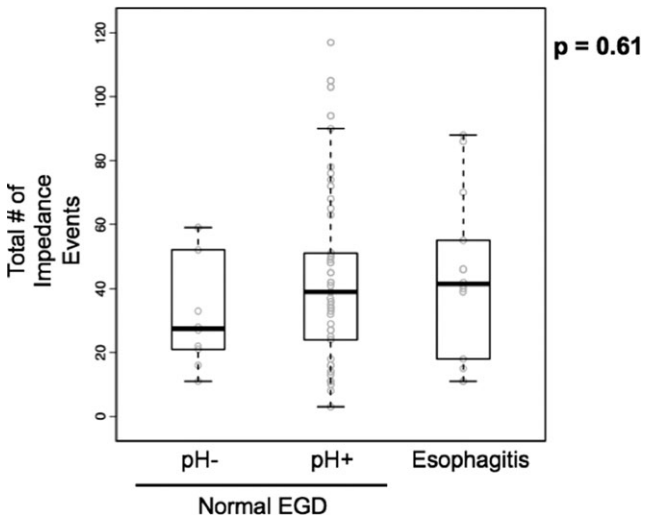

Fig. 2. Individual and group median (IQR) total number of impedance events for the three patient subgroups. Impedance events were similar between the groups.

acid reflux (% total time pH < 4) *off* therapy and total number of impedance events *on* therapy by patient subgroup. As expected, patients with a normal upper endoscopy and normal pH testing exhibited a significantly ($P < 0.001$) lower % total time pH < 4 (3.08% [1.99%–4.89%]), compared to the patients with a normal upper endoscopy and abnormal pH testing (10.45% [7.35%–15.10%]) and those with esophagitis (9.98% [6.70%–12.66%]) (Fig. 1). However, there was no significant difference among the three groups with respect to median (IQR) total number of impedance events on therapy ($P = 0.61$) (27.50 [21.25–47.25];39 [24–51]; 42 [23.25–52.75], respectively) (Fig. 2). Median (IQR) lower esophageal sphincter pressure for the three groups were similar ($P = 0.78$) between the three groups (3.0 [1.0–7.0], 5.5 [0.4–15.5], and 8.0 [1.3–29], respectively).

Predictors of Impedance Findings

No significant overall relationship was observed between the total number of impedance events and the traditionally recognized physiologic parameters for reflux. Specifically, there was no correlation (r = 0.15, $P = 0.2$) between the total number of impedance events *on* therapy and the total % time pH < 4 *off* therapy (Fig. 3A). Presence or absence of hiatal hernia did not change this relationship (Fig. 3B). However, among patients with endoscopic evidence of esophagitis, there was a weak significant correlation (r = 0.54, $P = 0.045$) between the *on* therapy impedance and *off* therapy pH parameters (Fig. 3C). However, the impedance parameters *on* therapy could not be predicted based on severity of reflux parameters at baseline. Number of impedance events *on* therapy were similar ($P = 0.99$) between patients who had no or mild reflux (% time pH < 4 of less than 10%) compared to those with moderate to severe reflux (% time pH < 4 of more than 10%) (Fig. 4).

DISCUSSION

In a unique cohort of patients who had both *off* and *on* therapy testing, we sought to determine if *on* therapy impedance data can be predicted from the more traditionally recognized and commonly employed *off* therapy upper endoscopy, manometry, or pH parameters. In this stepwise diagnostic approach, we found that among patients with extraesophageal reflux refractory to PPI

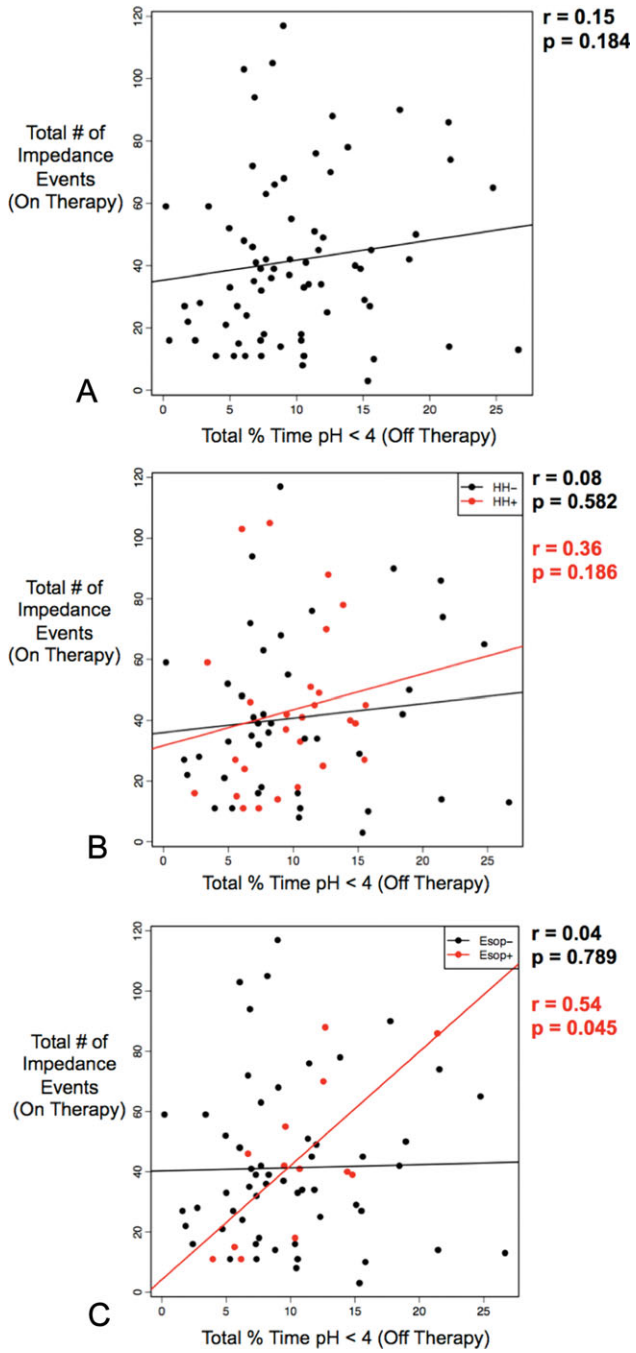

Fig. 3. Correlation between % total time pH < 4 *off therapy* and total number of impedance events *on therapy* for Overall (A), stratified by hiatal hernia (B), and stratified by endoscopic presence of esophagitis (C). HH+ = hiatal hernia present; HH– = no hiatal hernia; Esop– = no esophagitis; Esop+ = esophagitis. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

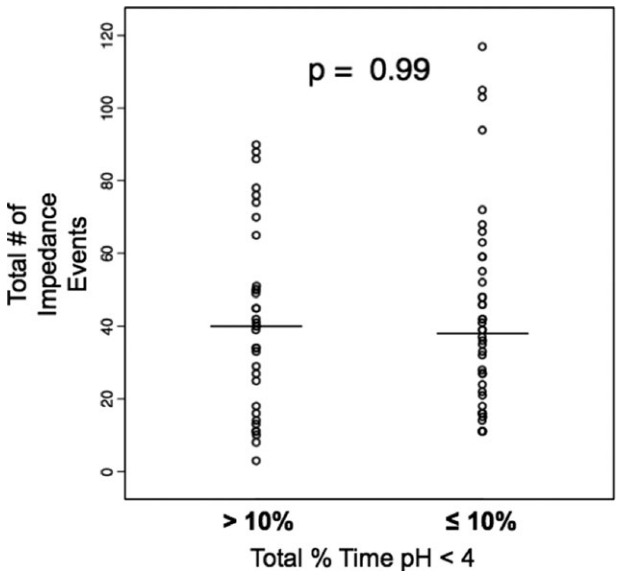

Fig. 4. Impedance parameters dichotomized by severity of baseline acid reflux. The impedance parameters were no different in those with no or mild reflux (defined as % total time pH $<$ 4 of less than or equal to 10%) compared to those with moderate to severe reflux ($>$10% acid reflux).

therapy, impedance testing results *on* therapy cannot be predicted from customary baseline parameters, except among patients with esophagitis in whom the correlation was weak at best, with only 19% of the population having esophagitis. Our data suggests that impedance parameters *on* therapy do not correlate well with any reflux parameters previously employed to assess disease severity. Thus, we urge caution regarding the over-interpretation of impedance parameters, as the clinical relevance of impedance testing remains unclear at this time.

The presence of non- or weakly acidic reflux in patients on PPI therapy is suggested to imply continued reflux and the need for additional therapies.[10,11] In a group of 19 patients who had positive symptom association with acid or nonacid reflux on impedance testing, a retrospective phone interview study suggested 94% fundoplication success.[23] However, two recent prospective trials have questioned the clinical reliability of symptom indices in reflux disease.[24,25] Furthermore, the most recent surgical trial in patients with extraesophageal syndrome showed that impedance parameters *on* therapy did not predict symptom response postfundoplication.[26] In this study, the traditional parameters of increased acid exposure, presence of hiatal hernia, and typical reflux symptoms (heartburn and regurgitation) were predictive of extraesophageal symptom response to fundoplication.

Important controversy in patients with continued symptoms, despite aggressive PPI therapy, is whether to conduct testing *on* or *off* PPI therapy. Employing both impedance-pH monitoring *on* therapy and wireless pH monitoring *off* therapy in the same group of patients with PPI-refractory symptoms, we confirmed that non- or weakly acid reflux may be present in up to 35% of patients; however, continued acid reflux was not seen in any patient. Our data are in agreement with two prior studies; one showing that continued acid reflux is a rarity on twice daily PPI therapy,[27] and the other showing continued nonacid reflux by impedance testing in 37% of patients refractory to PPI therapy.[12] More important, we could not identify any *off* therapy traditionally employed physiologic parameter that could predict the *on* therapy impedance findings. Furthermore, patients with more severe reflux by pH testing defined as % time pH $<$ 4 of greater than 10% had similar impedance parameters than those with no or mild reflux at baseline (Fig. 4). Thus, it appears that the impedance parameters do not correlate with any of the traditionally employed tools in assessing reflux severity. For example, it has been shown that patients with hiatal hernia typically have higher reflux scores compared to those without hiatal hernia,[28] esophagitis severity is expected to correlate with hiatal hernia size and esophageal acid exposure,[29] and % time pH $<$ 4 increases in a graded fashion across the GERD spectrum.[30] Thus, given the lack of any correlation between impedance results and these traditional markers, we urge caution regarding the clinical relevance of impedance testing.

Our study is unique in that the same patient population underwent physiologic testing off and on PPI therapy. However, some limitations of our study should also be highlighted. First, the results from our study underscore the need for larger outcome studies among patients with refractory symptoms and abnormal impedance testing. Second, our present analysis discusses the impedance findings with respect to abnormal number of reflux events in the distal esophagus. We did not evaluate proximal extent and liquid, gas, or mixed nature of the refluxate, as some believe may be important in a subgroup of treatment-resistant patients.[31] Additionally, we have used number of reflux events as the primary measure as opposed to SI or SAP. However, the use of SI and SAP is problematic in this group since patients have already declared lack of clinical response to aggressive acid suppression, and recent studies suggest that these metrics may not be reliable or reproducible.[24,25]

CONCLUSION

In a unique group of patients who had both *off* therapy traditional esophageal physiologic testing and *on* therapy impedance monitoring, our study shows limited correlation between the latter results with the former previously recognized and employed methodologies. There remains uncertainty regarding the clinical utility of impedance testing among patients with extraesophageal symptoms, and we recommend caution in over-interpretation of impedance pH monitoring data.

BIBLIOGRAPHY

1. Dent J, El-Serag HB, Wallander MA, Johansson S. Epidemiology of gastro-oesophageal reflux disease: a systematic review. *Gut* 2005;54: 710–717.
2. Locke GR 3rd, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ 3rd. Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota. *Gastroenterology* 1997;112: 1448–1456.

3. Shaheen NJ, Hansen RA, Morgan DR, et al. The burden of gastrointestinal and liver diseases, 2006. *Am J Gastroenterol* 2006;101: 2128–2138.
4. Frye JW, Vaezi MF. Extraesophageal GERD. *Gastroenterol Clin North Am* 2008;37:845–858, ix.
5. Qadeer MA, Phillips CO, Lopez AR, et al. Proton pump inhibitor therapy for suspected GERD-related chronic laryngitis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol* 2006;101:2646–2654.
6. Kahrilas PJ, Shaheen NJ, Vaezi MF. American Gastroenterological Association Institute technical review on the management of gastroesophageal reflux disease. *Gastroenterology* 2008;135:1392–413, 413 e1–5.
7. Park W, Vaezi MF. Esophageal impedance recording: clinical utility and limitations. *Curr Gastroenterol Rep* 2005;7:182–189.
8. Sifrim D, Blondeau K. Technology insight: the role of impedance testing for esophageal disorders. *Nat Clin Pract Gastroenterol Hepatol* 2006;3: 210–219.
9. Tutuian R, Castell DO. Review article: complete gastro-oesophageal reflux monitoring—combined pH and impedance. *Aliment Pharmacol Ther* 2006;24 (suppl 2):27–37.
10. Agrawal A, Castell DO. Clinical importance of impedance measurements. *J Clin Gastroenterol* 2008;42:579–583.
11. Bredenoord AJ, Tutuian R, Smout AJ, Castell DO. Technology review: esophageal impedance monitoring. *Am J Gastroenterol* 2007;102: 187–194.
12. Mainie I, Tutuian R, Shay S, et al. Acid and non-acid reflux in patients with persistent symptoms despite acid suppressive therapy: a multicentre study using combined ambulatory impedance-pH monitoring. *Gut* 2006;55:1398–1402.
13. Zerbib F, Roman S, Ropert A, et al. Esophageal pH-impedance monitoring and symptom analysis in GERD: a study in patients off and on therapy. *Am J Gastroenterol* 2006;101:1956–1963.
14. DeVault KR, Castell DO. Updated guidelines for the diagnosis and treatment of gastroesophageal reflux disease. *Am J Gastroenterol* 2005;100: 190–200.
15. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. *Gastroenterology* 2008;135:1383–1391, 91 e1–5.
16. Park W, Hicks DM, Khandwala F, et al. Laryngopharyngeal reflux: prospective cohort study evaluating optimal dose of proton-pump inhibitor therapy and pretherapy predictors of response. *Laryngoscope* 2005;115: 1230–1238.
17. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut* 1999;45:172–180.
18. Pandolfino JE, Richter JE, Ours T, Guardino JM, Chapman J, Kahrilas PJ. Ambulatory esophageal pH monitoring using a wireless system. *Am J Gastroenterol* 2003;98:740–749.
19. Zerbib F, des Varannes SB, Roman S, et al. Normal values and day-to-day variability of 24-h ambulatory oesophageal impedance-pH monitoring in a Belgian-French cohort of healthy subjects. *Aliment Pharmacol Ther* 2005;22:1011–1021.
20. Shay S, Tutuian R, Sifrim D, et al. Twenty-four hour ambulatory simultaneous impedance and pH monitoring: a multicenter report of normal values from 60 healthy volunteers. *Am J Gastroenterol* 2004;99:1037–1043.
21. Pritchett JM, Aslam M, Slaughter JC, Ness RM, Garrett CG, Vaezi MF. Efficacy of esophageal impedance/pH monitoring in patients with refractory gastroesophageal reflux disease, on and off therapy. *Clin Gastroenterol Hepatol* 2009;7:743–748.
22. Tutuian R, Mainie I, Agrawal A, et al. Normal values for ambulatory 24-hour combined impedance-pH monitoring on acid suppressive therapy. *Gastroenterology* 2006;130:A171.
23. Mainie I, Tutuian R, Agrawal A, Adams D, Castell DO. Combined multichannel intraluminal impedance-pH monitoring to select patients with persistent gastro-oesophageal reflux for laparoscopic Nissen fundoplication. *Br J Surg* 2006;93:1483–1487.
24. Slaughter JC, Goutte M, Rymer JA, et al. Caution about overinterpretation of symptom indexes in reflux monitoring for refractory gastroesophageal reflux disease. *Clin Gastroenterol Hepatol* 2011;9:868–874.
25. Vaezi MF, Higginbotham T, Kavitt RT, et al. Do patients report symptoms when they occur? Fundamental consequences for interpretation of ambulatory reflux monitoring. *Gastroenterology* 2012;142:S–415.
26. Francis DO, Goutte M, Slaughter JC, et al. Traditional reflux parameters and not impedance monitoring predict outcome after fundoplication in extraesophageal reflux. *Laryngoscope* 2011;121:1902–1909.
27. Charbel S, Khandwala F, Vaezi MF. The role of esophageal pH monitoring in symptomatic patients on PPI therapy. *Am J Gastroenterol* 2005;100: 283–289.
28. Sloan S, Rademaker AW, Kahrilas PJ. Determinants of gastroesophageal junction incompetence: hiatal hernia, lower esophageal sphincter, or both? *Ann Intern Med* 1992;117:977–982.
29. Jones MP, Sloan SS, Rabine JC, Ebert CC, Huang CF, Kahrilas PJ. Hiatal hernia size is the dominant determinant of esophagitis presence and severity in gastroesophageal reflux disease. *Am J Gastroenterol* 2001;96: 1711–1717.
30. Vaezi MF, Richter JE. Role of acid and duodenogastroesophageal reflux in gastroesophageal reflux disease. *Gastroenterology* 1996;111:1192– 1199.
31. Zerbib F, Duriez A, Roman S, Capdepont M, Mion F. Determinants of gastro-oesophageal reflux perception in patients with persistent symptoms despite proton pump inhibitors. *Gut* 2008;57:156–160.

Reprinted by permission of Otolaryngol Head Neck Surg. 2008; 139(3):414-420.

Otolaryngology–Head and Neck Surgery (2008) 139, 414-420

ORIGINAL RESEARCH—LARYNGOLOGY AND NEUROLARYNGOLOGY

Double-blind, placebo-controlled trial with esomeprazole for symptoms and signs associated with laryngopharyngeal reflux

Oliver Reichel, MD, Holger Dressel, MD, Katrin Wiederänders, and **Wolfgang J. Issing, MD, PhD,** Munich, Germany; and Newcastle Upon Tyne, UK

OBJECTIVE: To determine the efficacy of proton pump inhibitor (PPI) therapy with esomeprazole on symptoms and signs associated with laryngopharyngeal reflux (LPR).

STUDY DESIGN AND METHODS: Prospective, double-blind, randomized, placebo-controlled study. Sixty-two patients with a reflux finding score (RFS) > 7 and a reflux symptom index (RSI) > 13 were enrolled and received either esomeprazole 20 mg twice daily or placebo for three months. RSI and RFS were assessed at baseline, after six weeks, and after three months.

RESULTS: Reductions of total RSI and RFS as well as of several subscores were significantly higher in the treatment group compared to placebo after three months ($P < 0.05$ each). The difference between study groups was most pronounced for posterior commissure hypertrophy ($P < 0.01$).

CONCLUSION: In the treatment of LPR-related symptoms a high placebo effect can be observed. However, compared to control, twice-daily PPI treatment for three months demonstrated a significantly greater improvement in laryngeal appearance and LPR symptoms.

© 2008 American Academy of Otolaryngology–Head and Neck Surgery Foundation. All rights reserved.

Laryngopharyngeal reflux (LPR) has had a significantly increasing impact on otolaryngologist office visits in the last decade. In the three-year period 1990-1993, annual visits to otolaryngologists by patients suffering from reflux-related problems averaged 89,000; this average increased to 421,000 annual visits during the three-year time span 1998-2001.[1] The generally recommended treatment in patients with LPR is twice-daily-dose proton pump inhibitor (PPI) therapy for three to six months.[2,3] However, there is some controversy regarding the adequate PPI dosage and even the efficacy of these drugs for LPR treatment.[3,4] Some authors criticize these recommendations for LPR treatment as being based on poor levels of evidence from uncontrolled studies.[5] In addition, several studies in the past could not demonstrate superiority of PPIs over placebo for treatment of suspected LPR.[6-9] Therefore, the primary objective of this prospective, double-blind, randomized, and placebo-controlled study was to evaluate the effect of a three-month treatment with esomeprazole 20 mg twice daily on symptoms and laryngeal signs in patients with suspected LPR. We chose esomeprazole for this study as we could show adequate measurable acid suppression with a once-daily dose of this PPI in a large number of LPR patients in a former trial[4] and because esomeprazole provides greater 24-hour control of intragastric acid than all other available PPIs at standard doses used to treat erosive esophagitis.[10]

MATERIALS AND METHODS

The following procedures were performed in accordance with the Declaration of Helsinki, Good Clinical Practice, and applicable regulatory requirements. The study was approved by the Ethics Committee of the Medical Faculty of Ludwig Maximilians University Munich (project number 265-05). Before initiation of any procedure, signed informed consent was obtained from all patients.

Participants

Between February 2006 and July 2007, 62 consecutive patients (30 women and 32 men; age range, 21-77 years; mean, 48.7 years) were enrolled. They presented to the Department of Otorhinolaryngology–Head and Neck Surgery of Ludwig Maximilians University, Munich, with unspecific otolaryngologic and respiratory disorders such as chronic cough, dysphagia, throat clearing, globus sensation, hoarseness, sore throat, and heartburn. All patients were examined by one otolaryngologist and rigid or flexible laryngoscopy had to reveal mucosal abnormalities consistent with LPR reflected by a reflux finding score (RFS) > 7.[11] In addition, the reflux symptom index (RSI),[12] a self-administered nine-item outcomes instrument for the diagnosis of LPR, had to exceed the value of 13 for inclusion. None of the study patients were treated with PPIs or any other antireflux medication for at least three months. Further

Received April 9, 2008; revised May 27, 2008; accepted June 3, 2008.

0194-5998/$34.00 © 2008 American Academy of Otolaryngology–Head and Neck Surgery Foundation. All rights reserved.
doi:10.1016/j.otohns.2008.06.003

exclusion criteria included age younger than 18 years, a history of laryngeal malignancy or gastrointestinal surgery, and the need for continuous therapy with warfarin, coumarin, or acetylsalicylacid. No patient with a clinically significant condition that could put the patient at risk, affect the patient's ability to participate in the study, influence the trial results, or necessitate surgery during the study was included. Patients with a contraindication to esomeprazole (eg, known or suspected hypersensitivity or allergy to esomeprazole or other PPIs) or participants of another investigational drug study in the past 30 days were excluded from participation. No pregnant or lactating women were enrolled, and in case of childbearing potential, effective contraception had to be maintained. Finally, patients with drug or alcohol addiction or with any psychiatric disease were not allowed to participate in the study.

Study Procedures

Patients passing all inclusion and exclusion criteria were then sequentially randomized in a 1:1 ratio to receive either coated tablets of esomeprazole 20 mg (twice daily, 30 minutes before meals) or a placebo tablet indistinguishable from the esomeprazole tablet (Fig 1). Both esomeprazole and placebo were given for a total of three months and were supplied in four small plastic bottles, each containing 48 tablets. Esomeprazole and placebo were provided by Astra-Zeneca (Wedel, Germany), and Astra-Zeneca also performed the randomization. Patients and investigators were blinded as to the medication randomization. Enrolled patients did not receive any instructions regarding lifestyle modifications to reduce acid reflux like avoidance of fatty meals or caffeine. However, all included subjects were explicitly instructed to take the medication with water at least 30 minutes before morning and evening meals. The RFS was readministered six weeks and three months after the start of treatment by the same otolaryngologist by laryngoscopy and each patient filled in the RSI questionnaire on these two follow-up visits. The primary objective of the study was the comparison of the total RSI and RFS after a treatment period of three months. At the final visit all patients completing the trial were additionally asked if they thought that therapy had completely resolved their complaints (possible answers were "yes" or "no"). In order to check for compliance with treatment, pill counting was performed at each visit. It was considered adequate if 75% or more of the treatment medication was taken. If one or more study medication bottles were not returned, compliance was not calculated.

Statistical Analysis

For sample size calculation we assumed that a typical LPR-associated subscore (eg, posterior commissure hypertrophy) would improve by at least one point in 65% of the PPI group and in 30% of the placebo group. For comparison of those two proportions in independent samples, a sample size of 31 patients per treatment group was calculated (two-tailed z test with alpha = 0.05, power = 80%, and accounting for a 10% drop-out rate). Baseline characteristics of both groups were reported as mean values and standard deviations (SD) or proportions and compared using the Mann-Whitney U test for continuous variables and the χ^2 test for categorical variables (Table 1). Treatment or placebo effects after six weeks and three months were reported as mean differences and standard errors (SEM) from baseline and were tested using Wilcoxon signed rank tests (Tables 2 and 3). Differences in mean changes between esomeprazole and placebo were additionally tested using Mann-Whitney U test (Tables 2 and 3). In order to compare the subjective estimation of the therapeutic drug effect between the two study groups, again the χ^2 test was used. A P value < 0.05 was considered statistically significant. The statistical analyses were done with SPSS 14.0 for windows (version 14.0.1; SPSS Inc, Chicago, IL).

Role of the Funding Source

Astra-Zeneca had no involvement in the study design, collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

RESULTS

Fifty-eight patients completed the study; the drop-out rate was low with 6.5%. Thirty patients received esomeprazole and the control group consisted of 28 individuals (Fig 1, Table 1). Baseline characteristics between the two study groups were comparable. This was also true for all respective subscores of both RSI and RFS.

Compared to baseline, the total RSI and RFS were significantly reduced in both study groups after a treatment period of six weeks (Table 2). However, differences in total scores between the esomeprazole and placebo group were not statistically significant. The only symptom with a statistically significant decrease for the esomeprazole group was heartburn ($P < 0.05$). When the differences in laryngeal findings found after six weeks of treatment were compared for the two study groups, only diffuse laryngeal edema showed a significantly stronger improvement with esomeprazole ($P < 0.05$). For ventricular obliteration and laryngeal erythema a nonsignificant trend for better improvement in the esomeprazole group was found.

At the final visit, patients of both study groups again reported on an improvement of symptoms, reflected by a highly significant reduction from baseline of the total RSI ($P < 0.001$) (Table 3). However, the total improvement was significantly stronger in the esomeprazole group ($P < 0.05$). While in esomeprazole recipients a highly significant reduction for each single item of the RSI was evident, no significant improvement from baseline for swallowing difficulties, coughing after meals, or heartburn was found in the

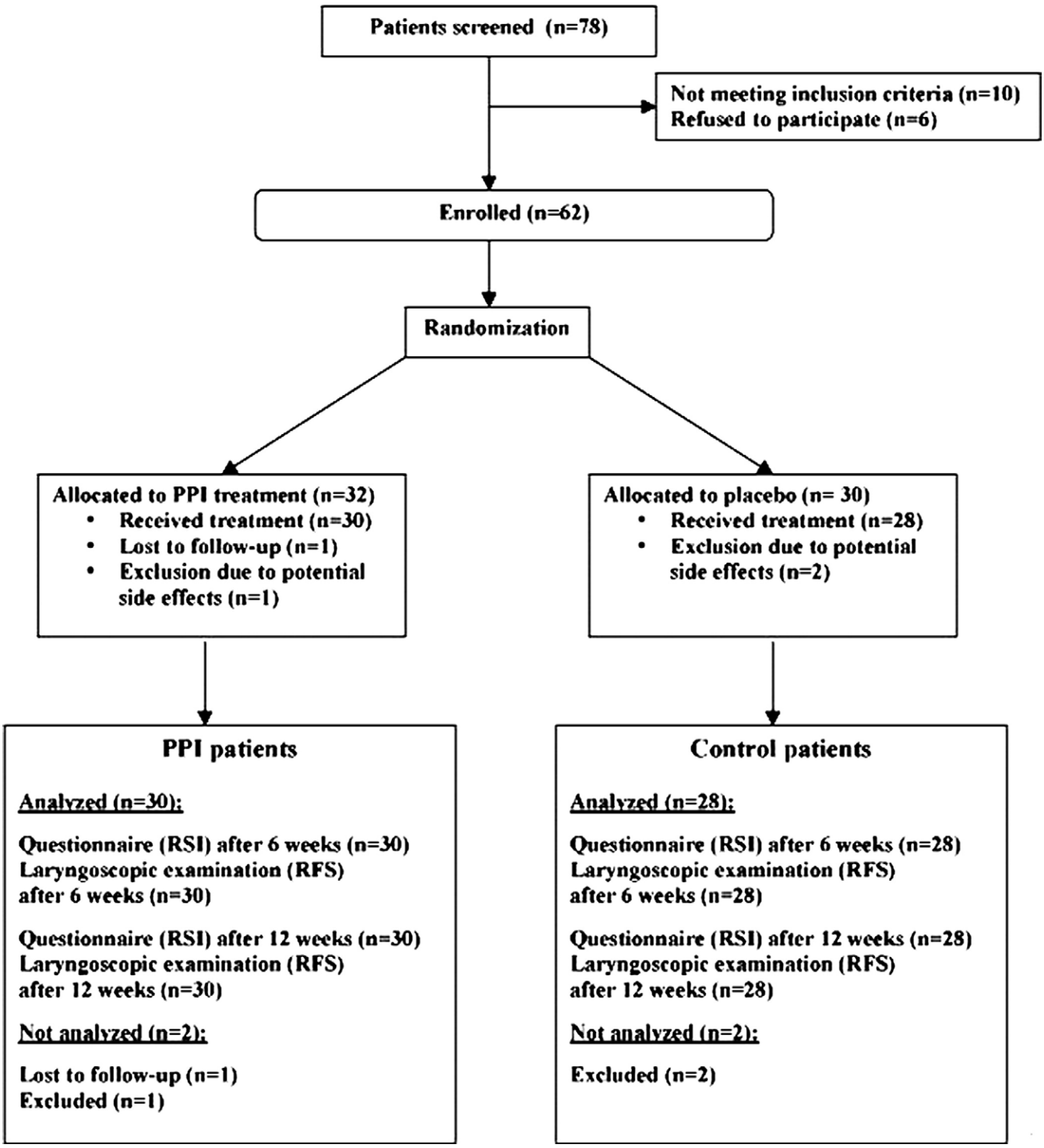

Figure 1 Summary of patient flow throughout study. PPI treatment = esomeprazole 20 mg twice daily.

placebo group. Similar to the result of the first follow-up visit, there was a significantly stronger improvement of the symptom heartburn in patients receiving esomeprazole. The comparison of the total RFS reduction between the two study groups reveals a statistically significant difference in favor of the esomeprazole group ($P < 0.05$). In contrast to esomeprazole recipients, patients receiving placebo did not show any significant improvement of ventricular obliteration or laryngeal erythema. After a treatment period of three months, the final examination revealed a statistically significant stronger improvement of erythema and diffuse laryngeal edema in the esomeprazole group ($P < 0.05$). The most significant difference between the study groups was the much stronger reduction of posterior commissure hypertrophy in the esomeprazole group after three months of therapy ($P < 0.01$).

Table 1
Baseline characteristics*

Variable	Control (n = 28)	PPI (n = 30)	*P* value†
Sex: male	13 (46.4%)	17 (56.7%)	0.160
Smoker	5 (17.9 %)	6 (20.0%)	0.871
Total RSI before therapy	21.79 ± 6.69	23.1 ± 7.45	0.503
Total RFS before therapy	14.89 ± 2.5	14.9 ± 2.75	0.969
Age	47.6 ± 16	49 ± 13.9	0.663

n, number of subjects in each group; *PPI,* proton pump inhibitor (esomeprazole); *RFS,* reflux finding score; *RSI,* reflux symptom index.
*Comparison of baseline characteristics between groups, reported as mean ± standard deviation (SD) or proportion of subjects (%).
†*P* value = Mann-Whitney *U* test or χ^2 test, all nonsignificant.

When asked about persisting complaints at the end of the trial, 11 of 26 placebo recipients (42%) felt themselves to be free of symptoms compared to 22 of 28 patients (>78%) from the esomeprazole group (two patients of each study group could not clearly answer the question with "yes" or "no"). This difference is statistically significant ($P = 0.006$).

All documented adverse events were considered mild and not treatment-related. The most common ones were bronchitis (n = 7), pharyngitis (n = 5), and headache (n = 5). One patient receiving esomeprazole was lost to follow-up after the initial visit. Two female patients complained of abdominal pain two and three weeks after receiving the study medication and were withdrawn from treatment. Their treatment was unblinded, and as they were part of the placebo group their complaints could not be caused by esomeprazole. Another female patient developed a skin rash six weeks after the beginning of therapy. Although this drop-out patient was not a placebo recipient, allergologic testing excluded drug hypersensitivity to esomeprazole (Fig 1).

DISCUSSION

Our study is the second largest to assess the efficacy of a PPI therapy for symptoms and signs associated with LPR compared to placebo and the second one choosing esomeprazole for a controlled trial. The results of this study demonstrate a significant placebo effect in the therapy of LPR-related symptoms as reported by patients especially within the first 6 weeks of PPI treatment. This phenomenon was already observed in former placebo-controlled studies investigating the efficacy of PPI therapy in patients with

Table 2
Change from baseline within each group after six weeks

	Visit I vs visit II		
	Esomeprazole	Placebo	Difference
RSI			
Total	9.87 ± 1.61‡	6.93 ± 1.61‡	2.94 ± 2.27
Hoarseness	1.03 ± 0.31†	1.07 ± 0.35†	−0.04 ± 0.47
Throat clear	1.20 ± 0.31†	1.32 ± 0.32†	−0.12 ± 0.45
Throat mucus	1.10 ± 0.28†	0.75 ± 0.22†	0.35 ± 0.35
Difficulty swallowing	1.23 ± 0.30†	0.57 ± 0.27*	0.66 ± 0.41
Coughing after meals	0.53 ± 0.34	−0.11 ± 0.32	0.64 ± 0.47
Breathing difficulties	0.50 ± 0.29	0.57 ± 0.25*	−0.07 ± 0.39
Annoying cough	0.87 ± 0.39*	0.68 ± 0.27*	0.19 ± 0.48
Throat sensations	1.60 ± 0.32‡	1.25 ± 0.34†	0.35 ± 0.46
Heartburn	1.53 ± 0.32‡	0.75 ± 0.27†	0.78 ± 0.42*
RFS			
Total	3.47 ± 0.47‡	2.46 ± 0.71†	1.00 ± 0.85
Subglottic edema	−0.07 ± 0.07	0.21 ± 0.19	−0.28 ± 0.12
Ventricular	0.60 ± 0.17†	0.36 ± 0.18	0.24 ± 0.25
Erythema	0.77 ± 0.21†	0.43 ± 0.26	0.34 ± 0.33
Vocal fold edema	0.33 ± 0.09†	0.46 ± 0.13†	−0.13 ± 0.16
Diffuse laryngeal edema	0.60 ± 0.13‡	0.21 ± 0.11	0.39 ± 0.16*
Posterior commissure hypertrophy	0.80 ± 0.15‡	0.54 ± 0.13†	0.26 ± 0.20
Granuloma/granulation tissue	0.07 ± 0.07	0.00 ± 0.18	0.07 ± 0.19
Thick endolaryngeal mucus	0.40 ± 0.23	0.29 ± 0.25	0.11 ± 0.33

Data are given as mean differences ± standard error of the mean (SEM). The first two data columns show differences from baseline and the third demonstrates the difference between study groups. *P* values are from Wilcoxon signed rank test and Mann-Whitney *U* test as appropriate.
*$P < 0.05$; †$P < 0.01$; ‡$P < 0.001$.

Table 3
Change from baseline within each group after three months

	Visit I vs visit III		
	Esomeprazole	Placebo	Difference
RSI			
Total	14.27 ± 1.58‡	7.79 ± 1.74‡	6.48 ± 2.34*
Hoarseness	1.37 ± 0.33‡	1.18 ± 0.36†	0.19 ± 0.48
Throat clear	1.80 ± 0.26‡	1.18 ± 0.31†	0.62 ± 0.41
Throat mucus	1.43 ± 0.33†	0.96 ± 0.27†	0.47 ± 0.43
Difficulty swallowing	1.40 ± 0.31‡	0.57 ± 0.28	0.83 ± 0.42
Coughing after meals	1.17 ± 0.37†	0.21 ± 0.33	0.95 ± 0.50
Breathing difficulties	0.87 ± 0.27†	0.57 ± 0.26*	0.30 ± 0.38
Annoying cough	1.60 ± 0.37†	1.07 ± 0.34†	0.53 ± 0.51
Throat sensations	2.27 ± 0.28‡	1.43 ± 0.33†	0.84 ± 0.43
Heartburn	1.97 ± 0.31‡	0.64 ± 0.42	1.32 ± 0.52*
RFS			
Total	4.60 ± 0.63‡	2.32 ± 0.76†	2.28 ± 0.98*
Subglottic edema	0.00 ± 0.10	0.21 ± 0.21	−0.21 ± 0.24
Ventricular	0.53 ± 0.20*	0.36 ± 0.21	0.18 ± 0.28
Erythema	0.77 ± 0.20†	0.07 ± 0.19	0.70 ± 0.27*
Vocal fold edema	0.67 ± 0.13‡	0.46 ± 0.16†	0.20 ± 0.20
Diffuse laryngeal edema	0.83 ± 0.16‡	0.36 ± 0.13*	0.48 ± 0.21*
Posterior commissure hypertrophy	0.97 ± 0.15‡	0.32 ± 0.12*	0.65 ± 0.19**
Granuloma/granulation tissue	0.07 ± 0.07	0.00 ± 0.15	0.07 ± 0.16
Thick endolaryngeal mucus	0.67 ± 0.20†	0.58 ± 0.25*	0.10 ± 0.32

Data are given as mean differences ± standard error of the mean (SEM). The first two data columns show differences from baseline and the third demonstrates the difference between study groups. *P* values are from Wilcoxon signed rank test and Mann-Whitney *U* test as appropriate.
*$P < 0.05$; †$P < 0.01$; ‡$P < 0.001$.

LPR.[7,8,13] After three months of treatment, however, the differences in symptom improvement between the two study groups became more obvious, reflected by a significantly stronger decrease in the total RSI after esomeprazole therapy. As a consequence, even in patients with suspected LPR reporting early symptom relief under PPI medication, the treatment should continue for at least three months.

In contrast to Vaezi et al,[9] who had performed the largest double-blind and placebo-controlled trial evaluating a PPI effect on LPR symptoms and laryngeal signs, we could demonstrate that twice-daily esomeprazole was superior to placebo in improving both LPR symptoms and laryngeal findings. In the mentioned study, Vaezi et al included 145 patients receiving either 40 mg of esomeprazole (n = 95) or placebo (n = 50) twice daily for 16 weeks. From their findings, Vaezi et al concluded that compared with placebo the PPI therapy was of no therapeutic benefit on signs and symptoms associated with LPR.[9] However, they excluded patients with moderate to severe heartburn from their study. Thus, patients with a symptom typical for gastroesophageal reflux disease (GERD) but also relevant for LPR (6% to 43% of LPR patients suffer from heartburn[14,15]) were not part of the large study population. This could have affected the study results. Another reason for this study result differing from our findings might be the fact that Vaezi et al did not use the RFS for control of laryngeal changes due to its lack of external validation.

Four further studies in the past also revealed no statistically significant benefit of a PPI therapy on characteristic LPR symptoms and signs compared to placebo. Havas et al performed a double-blind, placebo-controlled, and randomized study with 15 LPR patients evaluating the therapeutic efficacy of 30 mg lansoprazole twice daily for 12 weeks.[6] From their findings the authors concluded that lansoprazole was not more effective than placebo in the treatment of cervical symptoms of LPR and posterior pharyngolaryngitis. As the study population was small and no regular statistical analysis of the results was performed, the significance of this conclusion remains unclear. In a randomized, double-blind, crossover study by Eherer et al, pantoprazole 40 mg twice daily for three months did not significantly affect symptom or laryngeal scores compared with placebo in 14 LPR patients.[16]

Nevertheless, the authors hypothesized that pantoprazole compared to placebo may have resulted in faster improvement of LPR symptoms. Again, the limitation of this study is a small sample size, which makes it difficult to draw meaningful conclusions from this data. Another double-blind, randomized trial by Steward et al, comparing two-month rabeprazole (20 mg twice daily) to placebo-control PPI treatment, also failed to demonstrate significantly greater improvement in reflux symptoms, health status, or laryngeal appearance.[7] However, for laryngeal symptoms such as hoarseness, dry cough, and throat clearing a statis-

tically significant better improvement was noted for PPI-treated patients. In our opinion, it is not surprising that in this study laryngeal signs of LPR showed no stronger improvement in PPI recipients compared to control within two months, as the physical findings of LPR improve more slowly than the symptoms.[17] In our study we could also find no statistically significant difference in the laryngeal appearance (reflected by the total RFS) between the esomeprazole and placebo group after a short treatment period of six weeks (Table 2). Another critical point for the mentioned study is the fact that patients were not instructed on when to take the medication. This could have seriously affected the results of this study. In a randomized, double-blind, placebo-controlled trial, Wo et al evaluated the efficacy of single-dose pantoprazole 40 mg for 12 weeks in newly diagnosed LPR.[8] The response was similar between the pantoprazole and placebo group. As 60% of the study subjects had additional abnormal distal esophageal reflux and the single-dose PPI was not sufficient to reach a pH-documented measurable suppression of hypopharyngeal acid reflux, the results of this study are not adequately comparable to those of our study with a double-dose PPI design. Noordzij et al performed another prospective, placebo-controlled, randomized, and double-blind study to determine the efficacy of 40 mg omeprazole twice daily for two months in the treatment of LPR.[13] The authors could demonstrate a significant improvement of a composite laryngeal symptom score in the omeprazole group but not the placebo group. Again, the endoscopic laryngeal signs did not change significantly over the course of the study for either treatment group. As mentioned above, the two-month follow-up period may not have been sufficient to detect changes in laryngeal appearance. Another study investigating the therapeutic benefit of lansoprazole 30 mg twice daily for treating LPR, by El-Serag et al, provided evidence that lansoprazole therapy for three months achieved significantly better symptomatic response than placebo.[18] Due to a selected referral study population with a high likelihood of GERD, the authors suggested not to generalize their results to patients with LPR in a primary care setting.

The analysis of the respective RFS subscores in our study revealed a highly significant reduction of posterior commissure hypertrophy in the esomeprazole but not in the placebo group after a treatment period of 12 weeks ($P < 0.01$). This laryngeal sign is supposed to be one of the mucosal alterations most related to LPR.[19] The fact that the most significant difference in laryngeal appearance between the two study groups after 12 weeks could be detected in this special area in our opinion strongly indicates the efficacy of a PPI treatment in patients with symptoms and signs associated with LPR. The improvement of posterior commissure hypertrophy was not significantly better in the esomeprazole group compared to control at the first follow-up. This result underlines the importance of a PPI treatment for at least three months in patients with suspected LPR. Another striking result was that we found heartburn to be the only RSI subscore with a statistically significant difference in improvement between the two study groups after both six weeks and three months. This finding, in our opinion, clearly demonstrates that PPI therapy has achieved an earlier and stronger reduction of distal reflux episodes potentially causing heartburn compared to placebo. This result also reflects the well-known fact that esophagitis caused by reflux shows an earlier resonse to PPI treatment than reflux-induced laryngitis.[3,17]

Several issues should be addressed. As we did not randomize our otolaryngologic evaluations performed by only one examiner, the term double-blind can only be used to describe the medication randomization. Another critical aspect of our trial may be the fact that we did not perform 24-hour pH monitoring to diagnose LPR objectively before patient inclusion. According to reports of other authors, patient tolerance is poor for this procedure.[7] As such, we decided to assess LPR-related symptoms and signs alone with the use of RSI and RFS. Moreover, Vaezi et al, in the above-mentioned study, found that only a small proportion of their patients undergoing pH monitoring had documented pharyngeal acid exposure despite typical LPR symptoms and laryngoscopic signs.[9] From this finding they concluded that sensitivity of pH monitoring for detection of proximal esophageal or hypopharyngeal reflux episodes might not be more than 50%.[9] Another argument against performing pH monitoring before inclusion was the fact that more and more authors doubt that 24-hour pH monitoring, although supposed to be the gold-standard test for LPR, is the preferable initial step in the work-up of most patients with LPR.[9,18,20] A third limitation of our study might be the short follow-up of 12 weeks. The significance of our study results probably would have been higher after a treatment period of six months. However, only Vaezi et al performed a trial with a follow-up of more than three months.[9] Another critical aspect might be the dose of esomeprazole used (20 mg). It can be hypothesized that the differences in RSI and RFS improvement between the study groups would have been even more significant with a dose of 40 mg twice daily. This is the dose generally recommended for treating LPR.[2] However, we chose a dose of 20 mg esomeprazole twice daily as many institutions and general practitioners in Germany prefer to start with the lower dose. Nevertheless, we suppose our results clearly demonstrate a therapeutic effect of PPI treatment for LPR-related symptoms and signs. This estimation is confirmed by the subjective opinion of our study patients concerning the drug effect, with only 42% of placebo recipients and more than 78% of the esomeprazole group being free of symptoms ($P < 0.006$).

CONCLUSION

Especially during the first weeks of PPI therapy, a significant placebo effect appears to exist in the treatment of LPR-related symptoms. However, compared to placebo,

twice-daily esomeprazole treatment with a dose of 20 mg for three months demonstrated a significantly greater improvement in laryngeal appearance and LPR-related symptoms, as reflected by the total score differences of RFS and RSI. The most striking difference between the study groups was the significant reduction of posterior commissure hypertrophy in the PPI group.

AUTHOR INFORMATION

From the Department of Otorhinolaryngology–Head and Neck Surgery (Dr Reichel and Ms Wiederänders) and the Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine (Dr Dressel), Ludwig Maximilians University, Munich, Germany; and the Department of Otolaryngology (Dr Issing), The Freeman Hospital, Newcastle Upon Tyne, UK.

Corresponding author: Oliver Reichel, MD, Department of Otolaryngology–Head and Neck Surgery, Klinikum Großhadern, Marchioninistraße 15, 81377 Munich, Germany.

E-mail address: Oliver.Reichel@med.uni-muenchen.de.

AUTHOR CONTRIBUTIONS

Oliver Reichel, study design, data collection; **Holger Dressel**, data analysis, statistical analysis; **Katrin Wiederänders**, data collection; **Wolfgang J. Issing**, study design.

FINANCIAL DISCLOSURE

This study was partially supported by a grant from Astra-Zeneca, Wedel, Germany. AstraZeneca had no involvement in the study design, collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

REFERENCES

1. Altman KW, Stephens RM, Lyttle CS, et al. Changing impact of gastroesophageal reflux in medical and otolaryngology practice. Laryngoscope 2005;115:1145–53.
2. Koufman JA, Aviv JE, Casiano RR, et al. Laryngopharyngeal reflux: position statement of the Committee on Speech, Voice, and Swallowing Disorders of the American Academy of Otolaryngology–Head and Neck Surgery. Otolaryngol Head Neck Surg 2002;127:32–5.
3. Ford CN. Evaluation and management of laryngopharyngeal reflux. JAMA 2005;294:1534–40.
4. Reichel O, Keller J, Rasp G, et al. Efficacy of once-daily esomeprazole treatment in patients with laryngopharyngeal reflux evaluated by 24-hour pH monitoring. Otolaryngol Head Neck Surg 2007;136:205–10.
5. Karkos PD, Wilson JA. Empiric treatment of laryngopharyngeal reflux with proton pump inhibitors: a systemativ review. Laryngoscope 2006; 116:144–8.
6. Havas T, Huang S, Levy M, et al. Posterior pharyngolaryngitis: double-blind, randomised placebo-controlled trial of proton pump inhibitor therapy. Aust J Otolaryngol 1999;3:243–6.
7. Steward DL, Wilson KM, Kelly DH, et al. Proton pump inhibitor therapy for chronic laryngo-pharyngitis: a randomized placebo-control trial. Otolaryngol Head Neck Surg 2004;131:342–50.
8. Wo JM, Koopman J, Harrell SP, et al. Double-blind, placebo-controlled trial with single-dose pantoprazole for laryngopharyngeal reflux. Am J Gastroenterol 2006;101:1972–8.
9. Vaezi MF, Richter JE, Stasney CR, et al. Treatment of chronic posterior laryngitis with esomeprazole. Laryngoscope 2006;116:254–60.
10. Miner P Jr, Katz PO, Chen Y, et al. Gastric acid control with esomeprazole, lansoprazole, omeprazole, pantoprazole, and rabeprazole: a five-way crossover study. Am J Gastroenterol 2003;98:2616–20.
11. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). Laryngoscope 2001;111:1313–7.
12. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the Reflux Symptom Index (RSI). J Voice 2002;16:274–7.
13. Noordzij JP, Khidr A, Evans BA, et al. Evaluation of omeprazole in the treatment of reflux laryngitis: a prospective, placebo-controlled, randomized, double-blind study. Laryngoscope 2001;111:2147–51.
14. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. Laryngoscope 1991;101(suppl 53):1–78.
15. Kamel PL, Hanson D, Kahrilas PJ. Omeprazole for the treatment of posterior laryngitis. Am J Med 1994;96:321–6.
16. Eherer AJ, Habermann W, Hammer HF, et al. Effect of pantoprazole on the course of reflux-associated laryngitis: a placebo-controlled double-blind crossover study. Scand J Gastroenterol 2003;38:462–7.
17. Belafsky PC, Postma GN, Koufman JA. Laryngopharyngeal reflux symptoms improve before changes in physical findings. Laryngoscope 2001;111:979–81.
18. El-Serag HB, Lee P, Buchner A, et al. Lansoprazole treatment of patients with chronic idiopathic laryngitis: a placebo-controlled trial. Am J Gastroenterol 2001;96:979–83.
19. Book DT, Rhee JS, Toohill RJ, et al. Perspectives in laryngopharyngeal reflux: an international survey. Laryngoscope 2002;112:1399–406.
20. DelGaudio JM, Waring JP. Empiric esomeprazole in the treatment of laryngopharyngeal reflux. Laryngoscope 2003;113:598–601.

Reprinted by permission of Arch Otolaryngol Head Neck Surg. 2010; 136(11):1089-1092.

ORIGINAL ARTICLE

ONLINE FIRST

Treatment of Clinically Diagnosed Laryngopharyngeal Reflux Disease

Tarek Fouad Youssef, MD; Mohamed Rifaat Ahmed, MD

Objectives: To determine the incidence of *Helicobacter pylori* (HP) stool antigen (HPSA) in patients with laryngopharyngeal reflux disease (LPRD), and to make a comparison of 2 treatment regimens that have been used based on the presence or absence of HPSA positivity in patients with LPRD.

Design: Randomized controlled study.

Setting: Suez Canal University Hospital, Ismalia, Egypt.

Patients: A total of 212 patients with symptoms of LPRD.

Intervention: Patients were evaluated by laryngoscopy, ambulatory pH monitoring for 24 hours, and HPSA testing. Esomeprazole magnesium as a monotherapy was evaluated vs triple therapy in patients with HP infection.

Main Outcome Measures: To determine the incidence of HPSA in patients with LPRD, and to make a comparison of 2 treatment regimens that have been used based on the presence or absence of HPSA positivity in patients with LPRD.

Results: Persistent dry cough and a feeling of a lump in the throat (globus sensation) were the most frequent symptoms of LPRD, while posterior laryngeal inflammation was the main laryngoscopic finding. Results from the HPSA test were positive in 57% of the studied group. Patients with negative HPSA were treated with esomeprazole as single modality with a reported improvement score of 96.6%. Patients with positive HPSA test results were divided into 2 groups: 1 received only esomeprazole, with reported improvement in 40%, whereas the second group was treated with esomeprazole, plus amoxicillin sodium and clarithromycin (triple therapy) and reported a 90% incidence of symptom improvement.

Conclusion: The incidence of HP infection in patients with LPRD in our study was 57%. Triple therapy showed a higher cure rate in patients with HPSA-positive test results.

Arch Otolaryngol Head Neck Surg. 2010;136(11):1089-1092.
Published online September 20, 2010.
doi:10.1001/archoto.2010.165

GASTROESOPHAGEAL REflux disease (GERD) is defined as a backward flow of gastric contents into the esophagus.[1] Beaver et al[2] suggested that laryngopharyngeal reflux disease (LPRD) means a backward flow of the stomach contents up to the throat. The clinical symptoms usually occur secondary to a refluxate of hydrochloric acid and pepsin.[3] The gastric refluxate in the larynx might be the causative factor in posterior laryngeal inflammation, laryngeal contact ulcers, and laryngeal granuloma formation.[3,4] It is associated with many otolaryngology disorders, such as reflux laryngitis, cervical dysphagia, globus pharyngeus, chronic cough, laryngeal or tracheal stenosis, and laryngeal carcinoma.[4] The incidence of laryngopharyngeal symptoms is greater than expected.[5]

There is a complex multifactor set of pathophysiologic characteristics of LPRD besides simple acid reflux.[6] *Helicobacter pylori* (HP) is a gram-negative, microaerophilic bacterium that can cause infection of the stomach and is also strongly linked to the development of duodenal and gastric ulcers.[7,8] A relationship between the rates and degree of reflux esophagitis with HP infection has been reported, but to our knowledge, no relationship with reflux laryngitis has been reported.[9,10] The HP stool antigen (HPSA) test is a rapid, noninvasive diagnostic method based on a sandwich enzyme immunoassay with antigen detection, which has a high sensitivity and specificity.[11,12]

No standard guidelines are available for treatment of LPRD; proton pump inhibitors, twice daily for 8 weeks, have been recommended[13] if HP is present. However, clinical guidelines may consider revision to add a triple therapy regimen.

Author Affiliations:
Department of Otolaryngology–Head and Neck Surgery, Faculty of Medicine, Suez Canal University, Ismalia, Egypt.

Table 1. Relationship Between Laryngopharyngeal Reflux Disease (LPRD) Symptoms With HPSA Test Results[a]

LPRD Symptom	Patients With HPSA Test Result, No.	
	Positive (n=122)	Negative (n=90)
Chronic cough	59	46
Feeling of lump in throat	52	46
Frequent throat clearing	40	38
Bad/bitter taste	37	32
Hoarseness	24	30

Abbreviation: HPSA, *Helicobacter pylori* stool antigen.
[a] All comparisons were nonsignificant ($P<.05$ was considered statistically significant).

Table 2. Relationship Between Laryngoscopic Findings and HPSA Test Results[a]

Laryngoscopic Findings	Patients With HPSA Test Result, No.	
	Positive (n=122)	Negative (n=90)
Red, irritated arytenoids	61	55
Swelling of the vocal cords	32	27
Small laryngeal ulcers	11	13
Granulomas in the larynx	2	2

Abbreviation: HPSA, *Helicobacter pylori* stool antigen.
[a] All comparisons were nonsignificant ($P<.05$ was considered statistically significant).

We determined the incidence of HPSA-positive findings among patients diagnosed as having LPRD and compared the efficacy of 2 treatment regimens.

METHODS

We performed a randomized controlled study at Suez Canal University Hospital, Ismalia, Egypt. A total of 212 patients with LPRD symptoms (hoarseness, chronic unexplained cough, frequent throat clearing, a feeling of a lump in the throat [globus sensation], and a bad or bitter taste in the mouth[3]) were included in this study, but we excluded patients with a history of smoking, alcohol intake, chronic rhinosinusitis, or treatment for LPRD.

All patients underwent laryngoscopic examination to confirm reflux signs, then 24 hours of pH monitoring was ordered (the pH test was considered to be positive for LPRD when the pH was lower than 4; HPSA testing was ordered when a fresh stool sample was obtained).[11,12]

Patients with negative HPSA test results received once-daily esomeprazole magnesium, 40 mg, for 4 weeks.[14] Patients with positive HPSA test results were divided into 2 equal randomized groups: one was a control group that received only esomeprazole magnesium, 40 mg, for 4 weeks,[14] and the other was a study group that received triple therapy comprising esomeprazole magnesium, 40 mg, plus amoxicillin sodium, 1 g, and clarithromycin, 500 mg, for the same period.[14]

A senior otolaryngologist (T.F.Y.) who was blind to the treatment protocol performed follow-up evaluation for all patients after the end of medical treatment.

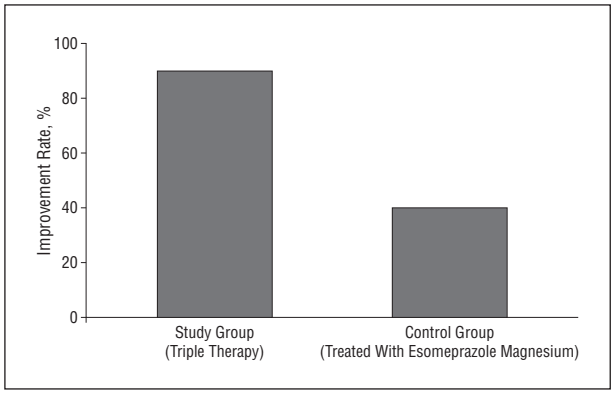

Figure. The clinical improvement in both controls and study patients with laryngopharyngeal reflux disease is seen. Triple therapy comprised esomeprazole plus amoxicillin sodium and clarithromycin. In both groups, treatment was daily for 4 weeks.

RESULTS

The mean age of the 212 patients in the study was 32.4 years. Cough, the main LPRD symptom, was found in 105 patients (49%), followed by globus sensation in 98 patients (46%), frequent throat clearing in 78 patients (36%), a bad or bitter taste in the mouth in 69 patients (32%), and hoarseness in 54 (25%).

Red, irritated arytenoids was the main laryngoscopic finding in 116 patients (54%), followed by swelling of the vocal cords in 59 patients (27%), small laryngeal ulcers in 24 patients (11%), and laryngeal granulomas in 4 patients (2%).

Among the patients in the study, the HPSA test results were positive in 57% of cases, and we found them to be statistically nonsignificant in relation to patient symptoms (**Table 1**) ($P<.05$ was considered statistically significant). Also, they were found to be nonsignificant in relation to the laryngoscopic findings (**Table 2**).

Marked improvement in symptoms occurred in 87 of the 90 patients with negative HPSA test results who received once-daily esomeprazole magnesium, 40 mg, for 4 weeks.[14]

The 122 patients with positive HPSA test results were randomized into 2 equal groups (61 patients each). The control group (61 patients) received only esomeprazole magnesium, 40 mg, for 4 weeks; 23 patients (40%) showed marked improvement in symptoms, partial improvement occurred in 9 patients (16%), while 25 patients (44%) reported no improvement. Four patients discontinued follow-up.

The second study group (61 patients) received triple therapy comprising esomeprazole magnesium, 40 mg, plus amoxicillin sodium, 1 g, and clarithromycin, 500 mg,[14] for the same period. Two patients discontinued follow-up. Fifty-three patients (90%) showed marked improvement in symptoms, partial improvement occurred in 3 patients (5%), and 3 patients (5%) showed no improvement (**Figure**).

COMMENT

In the practice of otolaryngology, it is now common to encounter patients with LPRD symptoms. Most of these pa-

©2010 American Medical Association. All rights reserved.

tients have been seen in thoracic and gastroenterology departments with atypical GERD symptoms. Laryngopharyngeal reflux disease is a diagnostic dilemma given the lack of solid guidelines for diagnosis and management.

In a recent report, Barry and Vaezi state,[15] "more questions than answers" were given, which best describes the current state of knowledge of LPRD. Our current study contributed several more questions.

In 1 limb of the study, a trial was made to associate HP infection with the degree or severity of symptoms and laryngoscopic findings. It was shown clearly based on statistical analysis that HP has no relation with any of the symptoms or signs of HPSA-positive or HP-negative individuals.

The second limb of the study compared the efficacy of proton pump inhibitor monotherapy vs triple therapy, and we have shown in our results that triple therapy gave better results in patients with positive HPSA test results. This study presents as much raw data as possible in compliance with the most recent guidelines to enable future evidence-based meta-analysis.

Gastroesophageal reflux disease is a common acid-related disorder presenting with a broad spectrum of symptoms with or without complications.[3] The incidence of laryngopharyngeal symptoms is greater than expected.[5] There are more complex multifactorial pathophysiologic characteristics of LPRD than simply acid reflux.[6] Laryngopharyngeal reflux disease is considered to be a variant of GERD in which the incidence of throat and laryngeal symptoms is more evident and encountered in practice more often than expected.[16]

A large number of studies have raised the issue of the role of HP infection and its role in the pathophysiologic mechanism of GERD, but the interest in its role in LPRD has not been adequately studied.[17] An estimated prevalence rate of HP infection of 30% among the general population has been given and shows that it is quite common.[18] Various theories and mechanisms have been proposed to clarify its role in GERD.

In our study, 212 patients with symptoms of LPRD and positive results from 24 hours of pH monitoring were evaluated clinically. The most common symptoms were dry, persistent cough (49%) followed by a globus sensation (46%); other studies have also reported a globus sensation or throat-clearing, voice change, persistent sore throat, dysphagia, and cough as the predominant symptoms.[19-21]

The common reported findings of LPRD are in the domain of posterior laryngitis; we reported red, irritated arytenoids in 54% and swollen vocal folds (27%); other reports[20,21] found endoscopic abnormalities in up to 98% of patients with LPRD, including nonspecific hyperemia, usually of the posterior larynx.

In our study, the 57% incidence rate of positive HPSA test scores is higher than that reported by Haruma et al,[21] who mentioned that in Japan there is a relationship between HP infection and LPRD with a reported incidence of 31% to 41%. *Helicobacter pylori* stool antigen testing is a relatively new, noninvasive diagnostic technique with high sensitivity and specificity[11,22]

Several authors suggested a correlation of HP infection and the degree of GERD,[9,10,19,23,24] while others[25] did not find any association between HP positivity and symptoms; the latter is in agreement with our data, which failed to demonstrate such a connection, and this variable report adds more to the dilemma of diagnosing LPRD.

As mentioned in the introductory paragraphs, we did not aim to point to a specific treatment regimen, a task better left for meta-analysis trials, but our raw data showed that patients with LPRD and with negative HPSA test results benefit from esomeprazole magnesium, 40 mg, for 4 weeks, with marked symptom improvement in most cases. While the patients with positive HPSA test results who received only esomeprazole magnesium, 40 mg, for 4 weeks showed a 40% rate of improvement, the second study group of patients with positive HPSA test results receiving triple therapy showed a 90% rate of improvement. Reports of a more successful triple therapy in GERD[26] are in agreement with our results, but still, no clear guidelines for treatment of LPRD are available.

In conclusion, the incidence of the HP infection in patients with LPRD in our study is 57%. Second, HP infection should be considered when treatment is prescribed to patients with LPRD because the standard therapy for GERD might be insufficient. Finally, the use of triple therapy (esomeprazole magnesium, 40 mg, plus amoxicillin sodium, 1 g, and clarithromycin, 500 mg) in the treatment of LPRD with HP infection might result in a higher cure rate.

Submitted for Publication: March 9, 2010; final revision received June 20, 2010; accepted July 22, 2010.
Published Online: September 20, 2010. doi:10.1001/archoto.2010.165
Correspondence: Mohamed Rifaat Ahmed, MD, Department of Otolaryngology–Head and Neck Surgery, Faculty of Medicine, Suez Canal University, Ismalia, Egypt (m_rifaat@hotmail.com).
Author Contributions: Both authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Analysis and interpretation of data:* Youssef and Ahmed. *Critical revision of the manuscript for important intellectual content:* Youssef and Ahmed. *Statistical analysis:* Youssef and Ahmed. *Obtained funding:* Ahmed. *Administrative, technical, and material support:* Ahmed.
Financial Disclosure: None reported.

REFERENCES

1. Koufman JA. Gastroesophageal reflux disease. *Otolaryngol Head Neck Surg*. 1993; 81(9)(suppl 2):2351.
2. Beaver ME, Stasney CR, Weitzel E, et al. Diagnosis of laryngopharyngeal reflux disease with digital imaging. *Otolaryngol Head Neck Surg*. 2003;128(1):103-108.
3. Bough ID Jr, Sataloff RT, Castell DO, Hills JR, Gideon RM, Spiegel JR. Gastroesophageal reflux laryngitis resistant to omeprazole therapy. *J Voice*. 1995; 9(2):205-211.
4. Hanson DG, Kamel PL, Kahrilas PJ. Outcomes of antireflux therapy for the treatment of chronic laryngitis. *Ann Otol Rhinol Laryngol*. 1995;104(7):550-555.
5. Dinis PB, Subtil J. *Helicobacter pylori* and laryngopharyngeal reflux in chronic rhinosinusitis. *Otolarygol Head Neck Surg*. 2006;134(1):67-72.
6. Reimer C, Bytzer P. Management of laryngopharyngeal reflux with proton pump inhibitors. *Ther Clin Risk Manag*. 2008;4(1):225-233.
7. Kim JG. Treatment of *Helicobacter pylori* infection [in Korean]. *Korean J Gastroenterol*. 2005;46(3):172-180.

©2010 American Medical Association. All rights reserved.

8. Yamaoka Y, ed. *Helicobacter pylori: Molecular Genetics and Cellular Biology*. Norwich, England: Caister Academic Press; 2008.
9. Oridate N, Takeda H, Yamamoto J, et al. *Helicobacter pylori* seropositivity predicts outcomes of acid suppression therapy for laryngopharyngeal reflux symptoms. *Laryngoscope*. 2006;116(4):547-553.
10. Tezer MS, Kockar MC, Koçkar O, Celik A. Laryngopharyngeal reflux finding scores correlate with gastroesophageal reflux disease and *Helicobacter pylori* expression. *Acta Otolaryngol*. 2006;126(9):958-961.
11. Gulcan EM, Varol A, Kutlu T, et al. *Helicobacter pylori* stool antigen test. *Indian J Pediatr*. 2005;72(8):675-678.
12. Oderda G, Rapa A, Ronchi B, et al. Detection of *Helicobacter pylori* in stool specimens by non-invasive antigen enzyme immunoassay in children: multicentre Italian study. *BMJ*. 2000;320(7231):347-348.
13. Scott LJ, Dunn CJ, Mallarkey G, Sharpe M. Esomeprazole: a review of its use in the management of acid-related disorders. *Drugs*. 2002;62(10):1503-1538.
14. Ni YH, Lin JT, Huang SF, Yang JC, Chang MH. Accurate diagnosis of *Helicobacter pylori* infection by stool antigen test and 6 other currently available tests in children. *J Pediatr*. 2000;136(6):823-827.
15. Barry DW, Vaezi MF. Laryngopharyngeal reflux: more questions than answers. *Cleve Clin J Med*. 2010;77(5):327-334.
16. Thomson ABR, Barkun AN, Armstrong D, et al. The prevalence of clinically significant endoscopic findings in primary care patients with uninvestigated dyspepsia: the Canadian Adult Dyspepsia Empiric Treatment–Prompt Endoscopy (CADET-PE) study. *Aliment Pharmacol Ther*. 2003;17(12):1481-1491.
17. Issing WJ, Karkos PD, Perreas K, Folwaczny C, Reichel O. Dual-probe 24-hour ambulatory pH monitoring for diagnosis of laryngopharyngeal reflux. *J Laryngol Otol*. 2004;118(11):845-848.
18. Ahmad I, Batch AJ. Acid reflux management: ENT perspective. *J Laryngol Otol*. 2004;118(1):25-30.
19. Pinar E, Oncel IS, Calli C, Atalay M. Detection of gastroesophageal reflux by scintigraphy in patients with laryngopharyngeal symptoms and findings [in Turkish]. *Kulak Burun Bogaz Ihtis Derg*. 2003;10(4):153-158.
20. Yorulmaz I, Ozlugedik S, Kucuk B. Gastroesophageal reflux disease: symptoms versus pH monitoring results. *Otolaryngol Head Neck Surg*. 2003;129(5):582-586.
21. Haruma K, Hamada H, Mihara M, et al. Negative association between *H. pylori* infection and reflux esophagitis in older patients: case-control study in Japan. *Helicobacter*. 2000;5(1):24-29.
22. Rouev P, Chakarski I, Doskov D, Dimov G, Staykova E. Laryngopharyngeal symptoms and gastroesophageal reflux disease. *J Voice*. 2005;19(3):476-480.
23. Tauber S, Gross M, Issing WJ. Association of laryngopharyngeal symptoms with gastroesophageal reflux disease. *Laryngoscope*. 2002;112(5):879-886.
24. Ercan I, Cakir BO, Uzel TS, Sakiz D, Karaca C, Turgut S. The role of gastric Helicobacter pylori infection in laryngopharyngeal reflux disease. *Otolaryngol Head Neck Surg*. 2006;135(1):52-55.
25. Postma GN, Johnson LF, Koufman JA. Treatment of laryngopharyngeal reflux. *Ear Nose Throat J*. 2002;81(9)(suppl 2):24-26.
26. Ching SS, Sabanathan S, Jenkinson LR. Treatment of Helicobacter pylori in surgical practice: a randomised trial of triple versus quadruple therapy in a rural district general hospital. *World J Gastroenterol*. 2008;14(24):3855-3860.

Reprinted by permission of Laryngoscope. 2015; 125(5):1093-1097.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Does Body Mass Index Predict Tracheal Airway Size?

Brian D'Anza, MD; Jesse Knight, MD; J. Scott Greene, MD, FACS

Objectives/Hypothesis: To determine the relationship between body mass index along with other anthropomorphic variables as they relate to tracheal airway dimensions.

Study Design: Retrospective case series.

Methods: This was a radiographic study of 123 consecutive hospitalized patients undergoing tracheotomy over a 4-year period (2007–2011). We measured airway dimensions in axial computed tomography imaging and made comparisons with height, weight, body mass index, gender, and age. Measurements were taken at the first tracheal ring level including anterior-posterior length, width, and calculated area. We expected higher body mass index not to be a good predictor of larger airway dimensions.

Results: The linear regression model showed body mass index was significantly inversely related to tracheal width after controlling for gender and age ($P = .0389$). For every 1 kg/m^2 increase in body mass index, the tracheal width decreased by 0.05 mm. There was a trend for airway area to diminish with increasing body mass index.

Conclusions: These results are consistent with the hypothesis that obese patients do not have larger airways. Our study indicated a trend toward smaller airways as body mass index increased. Specifically, as body mass index increases, tracheal width appears to decrease. This information should help medical professionals avoid the tendency to use a larger tube to secure the airway of an obese patient. Hopefully, this will result in further research into the field and may prevent future airway injuries in a society where obesity has become epidemic.

Key Words: Tracheal airway size, endotracheal tube size, obesity.

Level of Evidence: 4

***Laryngoscope*, 125:1093–1097, 2015**

INTRODUCTION

General guidelines exist for endotracheal tube (ETT) size selection in adults.[1,2] Variations in the choice of tube selection are influenced by factors such as patient age, gender, and body habitus. Generally speaking, a larger diameter tube is used for adult males and those with larger body habitus when compared to females or smaller individuals. Tracheostomy tube selection follows similar decision making algorithms in the adult population.[3] Due to normal anatomical variability, it is difficult to standardize recommendations for endotracheal and tracheostomy tube sizes for adults.[4] On a case-by-case basis, considerations for choosing a tube size might include those mentioned previously. Body dimensions that could be used to predict airway sizes include body mass index (BMI). A literature review shows the lack of a study comparing BMI with airway dimensions.

Larger-than-necessary ETTs are known to cause laryngeal or tracheal trauma and are to be avoided. Common problems can include laryngeal webs, vocal cord ulcerations, vocal cord paralysis or paresis, subglottic stenosis, and tracheal stenosis among others.[5,6] However, our experience has shown the tendency of emergency medical providers and critical care physicians to place a larger tube in a larger patient. Based on our observations at time of tracheotomy, we have found that it is common to find a smaller than anticipated trachea in an obese patient. Based on this experience, we hypothesize that higher BMI is not a good predictor of larger airway size. The purpose of our study was to evaluate the airway dimensions and identify anatomical concerns for the use of relatively oversized ETTs in an obese population. We measured airway dimensions in axial computed tomography (CT) imaging of 123 patients who underwent tracheotomy and made comparisons with their height, weight, BMI, gender, and age.

MATERIALS AND METHODS

The Geisinger Medical Center Office of Research Compliance and Institutional Review Board approved the retrospective electronic chart case series titled Predictors of Airway Size. The study was performed reviewing the information on 123 patients who underwent tracheotomy surgery by the otolaryngology department over a 4-year period (2007–2011) and who also had CT imaging of the trachea within the previous 3 months. Anthropomorphic measurements were taken from the time of

From the Department of Otolaryngology, Head and Neck Surgery, Facial Plastic Surgery (B.D., J.S.G), Geisinger Medical Center, Danville, Pennsylvania; and the ENT Associates of Los Alamos, PLLC (J.K.), Los Alamos, New Mexico, U.S.A.

Editor's Note: This Manuscript was accepted for publication September 3, 2014.

Presented at the 2013 American Academy of Otolaryngology–Head and Neck Surgery National Meeting, Vancouver, British Columbia, Canada, September 30, 2013.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Brian D'Anza, MD, Department of Otolaryngology, Head and Neck Surgery, Facial Plastic Surgery, Geisinger Medical Center, 100 N. Academy Ave., Danville, PA 17822. E-mail: bddanza@geisinger.edu

DOI: 10.1002/lary.24943

TABLE I.
Descriptive Statistics for the Entire Sample Stratified by Gender.

Variable	Entire Sample, N = 123	Female, n = 47	Male, n = 76	*P* Value
Age, yr	63.1 (15.3)	65.1 (14.2)	61.9 (15.9)	.2636
BMI, kg/m^2	32.4 (11.8)	33.5 (12.5)	31.7 (11.3)	.3919
CT AP, mm	20.8 (4.4)	18.8 (4.4)	22.1 (4.0)	<.0001
CT width, mm	17.2 (3.5)	15.6 (2.8)	18.3 (3.6)	.7164
Airway area, mm^2	294.8 (112.8)	240.0 (87.4)	328.8 (113.7)	<.0001

Values for age, BMI, CT AP, CT width, and airway area reflect the average (mean) calculation. Standard deviation is noted in parentheses

AP = anterior-posterior diameter; BMI = body mass index; CT = computed tomography.

tracheotomy in 76 males and 47 females to include height, weight, calculated BMI, age, and gender. Two hundred eighty patients were noted to have undergone tracheotomy over a 4-year period, but 157 patients were excluded due to malignancy involving the tracheal airway or with suboptimal imaging precluding accurate assessment of the tracheal dimensions. This excluded any patient who required placement of an ETT at the time of the CT study, or any other intervention such as a nasogastric tube that might impinge on the membranous tracheal wall dimensions. None of the patients included in this study had a prior tracheostomy or airway procedure. All of the included patients were spontaneously breathing and in a supine position at the time of the CT study. In all cases, the reason for ETT and tracheostomy placement was ventilator-dependent respiratory failure.

We noted the size of the ETT in place at the time of the tracheotomy procedure and documented this for each patient. Axial CT measurements were analyzed at the level of the first tracheal ring, as this was the area most commonly associated with the presence of a balloon cuff and thus the most likely area for tracheal injury.[7–9] We did not measure the level of the cricoid due to the circumferential cartilage ring structure at this level, making the airway less likely to be compressed or narrowed.

Measurements included the anterior-posterior diameter (CT AP) as well as the width (CT width) of the trachea at this point. Cross-sectional airway areas were also calculated and compared. Data are described using means and standard deviation for continuous variables and frequency and percentages for categorical variables. Data are described for the full sample and stratified by gender. Comparison across gender was accomplished using the two-sample t test and Pearson χ^2 tests, as appropriate. Pearson correlation was estimated between the continuous variables. Finally, linear regression was used to identify predictors of airway size. BMI classification was taken from current World Health Organization standards.[10]

TABLE II.
ETT Sizes for the Entire Sample.

ETT Size, mm	Entire Sample	Female	Male
6.0	12 (9.8%)	8 (17.0%)	4 (5.3%)
6.5	3 (2.4%)	2 (4.3%)	1 (1.3%)
7.0	12 (9.8%)	8 (17.0%)	4 (5.3%)
7.5	42 (34.2%)	21 (44.7%)	21 (27.6%)
8.0	52 (42.3%)	8 (17.0%)	44 (57.9%)
8.5	1 (0.8%)	0 (0%)	1 (1.3%)
9.0	1 (0.8%)	0 (0%)	1 (1.3%)

ETT sizes are all standardized numbers based on internal diameter in millimeters. Numbers in columns designate number of patients. Percentages designate the percentage of patients who had the specified size of tube placed.

ETT = endotracheal tube.

TABLE III.
BMI Classification and Mean ETT Sizes.

			Mean ETT Sizes, mm		
BMI, kg/m^2	Males	Females	Males	Females	All
Underweight	4	1	7.5	N/A	7.5
Normal	18	9	7.69	6.88	7.42
Overweight	15	10	7.70	6.95	7.40
Obese	39	27	7.83	7.39	7.65
	Range 14–70	Range 17–75			

The number in each table row for the columns titled Males and Females reflects the number of patients in each category.

Underweight: BMI = <18.50, Normal: BMI 18.50–24.99, Overweight: BMI 25–29.99, Obese: BMI ≥30. Mean ETT sizes are all based on internal diameter in millimeters.

BMI = body mass index; ETT = endotracheal tube.

RESULTS

Descriptive statistics for the entire sample size including age, BMI, CT AP, and CT width are shown in Table I, and the ETT sizes placed at the time of tracheostomy are shown in Table II. These statistics are stratified by gender. The BMI range for males was 14 to 70 and was 17 to 75 for females. Using these outlier patients as examples, the male with a BMI of 14 had an anterior-posterior tracheal airway measurement of 26 mm and a width of 21 mm. The male with a BMI of 70 had an anterior-posterior tracheal airway measurement of 24 mm and a width of 20 mm for comparison. The female with the BMI of 17 had an anterior-posterior tracheal airway measurement of 16 mm and a width of 15 mm. The female with the BMI of 75 had an anterior-posterior tracheal airway measurement of 14 mm and a width of 10 mm.

We have also included the mean ETT sizes calculated for the entire sample stratified by BMI classification. We found that for both obese men and women, the higher the BMI, the higher the average ETT size. Specifically, compared with normal sized patients, men and women who were obese had on average 0.14 and 0.51 mm-larger ETTs placed, respectively. This is demonstrated in Table III.

Most importantly, findings from the linear regression models yielded significant associations between BMI and tracheal airway dimensions. Specifically, BMI was inversely related to tracheal width after controlling for gender and age ($P = .0389$). For every 1 kg/m^2 increase in BMI, the tracheal width decreased by 0.05 mm. In addition, although not statistically significant, on linear regression analysis there was a trend for the airway area to decrease as the BMI increased.

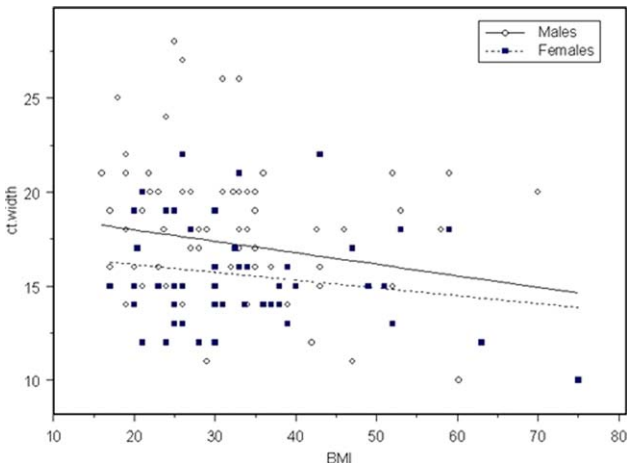

Fig. 1. Airway computed tomography (CT) width (mm) versus body mass index (BMI). The CT width on the y-axis is measured in millimeters. BMI on the x-axis is measured using weight in kilograms and height in meters and is calculated by dividing the subject's weight by the square of his/her height (kg/m^2). The white circle represents one individual male subject. The solid square represents one individual female subject. The solid line represents the linear regression analysis for all male subjects. The dotted line represents the linear regression analysis for all female subjects. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

Figures 1 and 2 provide graphical representations of the linear regression analyses.

Linear regression models also yielded significant associations for gender, with males having increased size on both tracheal CT AP diameter and width measurements on CT imaging compared to females. After controlling for BMI and age, there was an increase in CT AP diameter and CT width for males of 3.34 and 2.59 mm, respectively (Table IV).

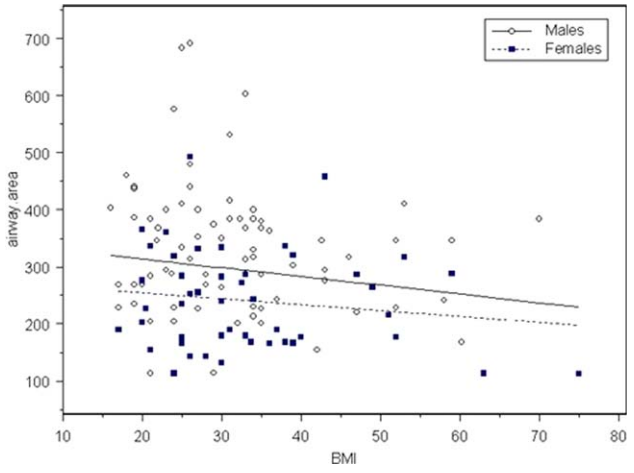

Fig. 2. Airway area (mm^2) versus body mass index (BMI). The airway area on the y-axis is measured in mm^2. BMI on the x-axis is measured using weight in kilograms and height in meters and is calculated by dividing the subject's weight by the square of his/her height (kg/m^2). The white circle represents one individual male subject. The solid square represents one individual female subject. The solid line represents the linear regression analysis for all male subjects. The dotted line represents the linear regression analysis for all female subjects. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

TABLE IV.
Linear Regression Results.

	CT AP, mm		CT Width, mm	
Variable	Estimate (SE)	*P* Value	Estimate (SE)	*P* Value
Male	3.34 (0.78)	<.0001	2.59 (0.61)	<.0001
BMI	−0.02 (0.03)	.5040	−0.05 (0.025)	.0389
Age	0.03 (0.02)	.2208	−0.002 (0.019)	.9359

Estimate is the score on change in millimeters for a given variable based on the linear regression best fit line from the scatter plot.
AP = anterior-posterior diameter; BMI = body mass index; SE = standard error; CT = computed tomography.

We compared each anthropomorphic variable for the strongest effect by using standardized regression coefficients. These were estimated from a linear regression model after the risk factors had been rescaled. Among males, the strongest predictor of CT AP diameter and airway area was height, and for CT width it was BMI. Among females, the strongest predictor of CT AP diameter and airway area was BMI, and for tracheal width it was height. However, none of the variables was significant. Therefore, no further modeling was performed.

Among both males and females, airway area correlated directly with height. Those findings are shown in Figure 3.

DISCUSSION

Our results support BMI to be inversely related to tracheal width on CT imaging. As shown in Figure 1, for every 1 kg/m^2 increase in BMI, the CT width decreased by 0.05 mm ($P = .0389$). All anthropomorphic measurements showed a trend for decreased airway dimensions (area, AP diameter, width) with increasing BMI based on linear regression of scatter plots. This would suggest consistency with the trend we have clinically observed in smaller tracheal airway sizes in patients noted to have much larger body habitus. We speculate that this could be a secondary effect from numerous factors, one being increased pressure on the trachea due to increased adiposity in these patients.

Animal studies have revealed a relationship between the natural caudal traction of the trachea by the thoracic contents and airway patency. Prior research has shown there is an influence of thoracic volumes on upper airway obstruction and compression.[11–13] Studies have shown that obese patients have problems with lower and upper airway compression due to increased weight and adiposity.[14,15] Specifically, abdominal obesity is suggested to negatively influence upper airway function during sleep.[14] It is believed that increased abdominal adiposity causes diaphragmatic compression of intrathoracic contents, which results in their cephalic deviation. As a result, the natural caudal traction of intrathoracic contents via the trachea is reduced and thus increases the distensibility of the airway. This is

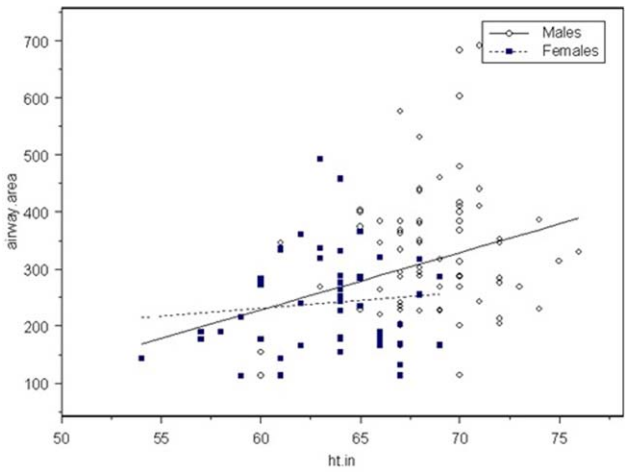

Fig. 3. Airway area (mm^2) versus height (inches). Airway area on the y-axis is measured in $mm.^2$ Height (ht) in on x-axis is measured in inches. The white circle represents one individual male subject. The solid square represents one individual female subject. The solid line represents the linear regression analysis for all male subjects. The dotted line represents the linear regression analysis for all female subjects. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

thought to compromise upper airway patency, especially when obese patients are supine.

Obese subjects have been shown to have increased positive end expiratory pressures due to extrinsic compression of surrounding tissues.[15] Based on the current study results, these relationships may also extend to the tracheal airway. The cartilage framework is more resistant to external thoracic pressures than the soft tissues of the lungs, oropharynx, or hypopharynx; however, the lack of caudal traction may be playing a role in the increased collapsibility and resultant decreased caliber of the airway. This may be especially pronounced as it relates to the membranous portion of the trachea as evidenced in the significantly decreased airway width of the obese study subjects. Other factors, such as genetic differences in cartilage composition, strength of surrounding musculature, inherent conditions of the trachea, such as tracheomalacia, could all be contributing. Further studies are needed to better examine these relationships in the trachea.

It is worth mentioning that when compared to females, males were found to have larger tracheal dimensions at the first tracheal ring. This is consistent with previous reports.[3]

As demonstrated in Table I, the study population had average tracheal airway dimensions that were slightly smaller than published normative values.[16,17] A large radiographic study performed by Breatnach et al. showed that the average tracheal airway width and AP diameter were 25 and 27 mm in men and 21 and 23 mm in women, respectively. These are generally accepted to be average airway dimensions for a presumably normal-sized population. Conversely, our patient sample had a much higher proportion of obese patients compared to the general population, comprising well over half the cohort. According to the Centers for Disease Control, the percentage of the US population that was obese was 35.5% and 35.8% for men and women, respectively.[18] This would seem to further corroborate the evidence that there may be an inverse correlation between obesity and airway size.

A patient with a higher body mass index might be presumed to have a larger tracheal airway. Anecdotally, we have found it is not uncommon to see an inappropriately sized ETT chosen for use in such a patient. Our data corroborate this contention, as obese patients were found on average to have a larger ETT in place, as shown in Table III. More importantly, the results of our study suggest that further research is needed in this area, as it is incorrect to assume a larger patient deserves a larger tube based on body size.

There are several limitations to our data. First, our study represents a retrospective case series of a limited population without randomization. Other limitations of this study include the inability to account for any chronic airway disease, such as tracheomalacia, which could lead to dynamic airway changes. Patients with acquired tracheomalacia can have more dynamic changes of the trachea, which has been noted on CT imaging studies previously.[19] Even cross-sectional changes during breathing and coughing as a result of changes in head and neck position and intrathoracic pressure can be noted.[4] One could make the argument that a larger ETT is required for obese patients to counteract the compressive forces on the airway. This must be balanced against the concern for further airway trauma and resultant complications such as stenosis, webbing, and ulcerations. The study population is also nearly exclusively from the intensive care unit, as the overwhelming majority of the tracheostomies performed at our institution are due to ventilator-dependent respiratory failure. This may bias the results as it relates to the type and size of tubes chosen. However, it is precisely this population of severely ill patients who require long-term intubation and who we are concerned about developing complications of tracheal trauma.

CONCLUSION

The population in this study demonstrated significant decreases in airway size with increasing BMI. Obese patients demonstrated radiographic evidence of a significant decrease in width and area of the airway. There was a trend for larger ETTs being placed in patients with a higher BMI. These findings have implications for airway management in obese populations. This study does not confirm the need to place a smaller tube in obese patients, but does suggest that due to smaller tracheal airway sizes, there may be a higher risk of airway injury in obese populations. Larger tubes are not anatomically indicated simply because of a greater BMI. More research in this population is needed to address airway management of obese patients. Currently, there is a lack of literature addressing this topic. This becomes more important, as population rates of obesity have become epidemic.

Acknowledgments
The authors thank H. Lester Kirchner, PhD, for his invaluable contributions to the data analysis.

BIBLIOGRAPHY

1. Allen MS. Surgical anatomy of the trachea. *Chest Surg Clin N Am* 1996;6: 627–635.
2. Miller RD. *Miller's Anesthesia*. 7th ed. London, United Kingdom: Churchill Livingstone; 2009: Chapter 50: Airway Management in Adult: 1573-1610.
3. Randestad A, Lindholm CE, Fabian P. Dimensions of the cricoid cartilage and the trachea. *Laryngoscope* 2001;110:1957–1961.
4. Gray H, Williams PL, Warwick R. In: *Gray's Anatomy*. 36th ed. Philadelphia, PA: Saunders; 2000:1246.
5. Deeb ZE, Williams JB, Campbell TE. Early diagnosis and treatment of laryngeal injuries from prolonged intubation in adults. *Otolaryngol Head Neck Surg* 1999;120:25–29.
6. Gaynor EB, Greenberg SB. Untoward sequelae of prolonged intubation. *Laryngoscope* 1985;95:1461–1467.
7. Sarper S, Ayten A, Eser I, Ozbudak O, Demircan A. Tracheal stenosis after tracheostomy or intubation. *Tex Heart Inst J* 2005;32:154–158.
8. Stauffer JL, Olson PE, Petty TL. Complications and consequences of endotracheal intubation and tracheostomy. A prospective study of 150 critically ill adult patients. *A J Med* 1981;70:65–76.
9. Pearson FG, Andrews MJ. Detection and management of tracheal stenosis following cuffed tube tracheostomy. *Annals Thoracic Surg* 1971:12:359–374.
10. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. WHO Technical Report Series 894. Geneva, Switzerland: World Health Organization; 2000.
11. Van de Graaff WB. Thoracic influence on upper airway patency. *J Appl Physiol* 1988;65:2124–2131.
12. Series F, Marc I. Influence of lung volume dependence of upper airway resistance during continuous negative airway pressure. *J Appl Physiol* 1994;77:840–844.
13. Series F, Cormier Y, Desmeules M. Influence of passive changes of lung volume on upper airways. *J Appl Physiol* 1990;68:2159–2164.
14. Pankow W, Podszus T, Gutheil T, Penzel T, Peter JH, Von Wichert P. Expiratory flow limitation and intrinsic positive end expiratory pressure in obesity. *J Appl Physiol* 1998:85:1236–1243.
15. Ferretti A, Giampiccolo P, Cavalli A, Milic-Emili J, Tantucci C. Expiratory flow limitation and orthopnea in massively obese subjects. *Chest* 2001; 119:1401–1408.
16. Brodsky JB, Macario A, Mark JBD. Tracheal diameter predicts double-lumen tube size: a method for selecting left double lumen tubes. *Anesthesia Analg* 1996;82:861–864.
17. Breatnach E, Abbott GC, Fraser RG. Dimensions of the normal human trachea. *Am J Roentgenol* 1984;142:903–906.
18. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of obesity in the United States, 2009–2010. NCHS Data Brief. 2012;(82):1–8.
19. Aquino S, Shepard J, Ginns L, Moore R, Halpern E, Grillo H. Acquired tracheomalacia: detection by expiratory CT Scan. *J Comput Assist Tomogr* 2001;25:3:394–399.

Reprinted by permission of Laryngoscope. 2015; 125(5):1137-1143.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Causes and Consequences of Adult Laryngotracheal Stenosis

Alexander Gelbard, MD; David O. Francis, MD, MS; Vlad C. Sandulache, MD, PhD; John C. Simmons, MD; Donald T. Donovan, MD; Julina Ongkasuwan, MD

Objectives/Hypothesis: Laryngotracheal stenosis (LTS) is largely considered a structural entity, defined on anatomic terms (i.e., percent stenosis, distance from vocal folds, overall length). This has significant implications for identifying at-risk populations, devising systems-based preventive strategies, and promoting patient-centered treatment. The present study was undertaken to test the hypothesis that LTS is heterogeneous with regard to etiology, natural history, and clinical outcome.

Study Design: Retrospective cohort study of consecutive adult tracheal stenosis patients from 1998 to 2013.

Methods: Subjects diagnosed with laryngotracheal stenosis (ICD-9: 478.74, 519.19) between January 1, 1998, and January 1, 2013, were identified. Patient characteristics (age, gender, race, follow-up duration) and comorbidities were extracted. Records were reviewed for etiology of stenosis, treatment approach, and surgical dates. Stenosis morphology was derived from intraoperative measurements. The presence of tracheostomy at last follow-up was recorded.

Results: One hundred and fifty patients met inclusion criteria. A total of 54.7% had an iatrogenic etiology, followed by idiopathic (18.5%), autoimmune (18.5%), and traumatic (8%). Tracheostomy dependence differed based on etiology ($P < 0.001$). Significantly more patients with iatrogenic (66%) and autoimmune (54%) etiologies remained tracheostomy-dependent compared to traumatic (33%) or idiopathic (0%) groups. On multivariate regression analysis, each additional point on Charlson Comorbidity Index was associated with a 67% increased odds of tracheostomy dependence (odds ratio 1.67; 95% confidence interval 1.04–2.69; $P = 0.04$).

Conclusions: Laryngotracheal stenosis is not a homogeneous clinical entity. It has multiple distinct etiologies that demonstrate disparate rates of long-term tracheostomy dependence. Understanding the mechanism of injury and contribution of comorbid illnesses is critical to systems-based preventive strategies and patient-centered treatment.

Key Words: Tracheal stenosis, subglottic, laryngotracheal stenosis, intubation, tracheostomy.

Level of Evidence: 4.

***Laryngoscope*, 125:1137–1143, 2015**

INTRODUCTION

Laryngotracheal stenosis (LTS) is a life-threatening, fixed, extrathoracic restriction in pulmonary ventilation. LTS is an umbrella term, encompassing luminal compromise at the level of the larynx, subglottis, or trachea, which exists in a watershed of specialty care. Diagnosis is frequently delayed as patients rapidly transition from acute inpatient care to outpatient facilities. The majority of patients are not definitively diagnosed until outpatient specialty evaluation.[1] Many specialists (e.g., intensivists, otolaryngologists, interventional pulmonologists, thoracic surgeons) initially interact with this population, which makes it difficult to establish the natural history of the disease, define universal predictors of disease outcome, and create cogent personalized plans of care. Additionally, long-term sequelae of intensive respiratory support (endotracheal intubation and elective tracheostomy) do not develop on a timescale necessary for recognition by practitioners providing acute care, impeding quality-driven improvement efforts.[2]

Moreover, LTS is generally described in terms of its structural characteristics, defined in anatomic terms (i.e., percent stenosis, distance from vocal folds, overall length). This neglects the unique biology driving luminal compromise in heterogeneous patient populations and has significant implications for identifying at-risk populations, devising systems-based preventive strategies, and promoting patient-centered treatment directed at the diverse pathophysiology driving airway injury. The present study was undertaken to test the hypothesis that LTS is heterogeneous with regard to etiology, natural history, and clinical outcome.

PATIENTS AND METHODS

This study was performed in accordance with the Declaration of Helsinki, Good Clinical Practice, and was approved by the Baylor College of Medicine Institutional Review Board (IRB No. H33195).

Patients

Subjects diagnosed with laryngotracheal stenosis (ICD-9: 478.74, 519.19) between January 1, 1998, and January 1, 2013, were

From the Department of Otolaryngology, Vanderbilt School of Medicine (A.G., D.O.F.), Nashville, Tennessee; and the Bobby R. Alford Department of Otolaryngology–Head and Neck Surgery Baylor College of Medicine (V.C.S., J.C.S., D.T.D., J.O.), Houston, Texas, U.S.A.

Editor's Note: This Manuscript was accepted for publication September 8, 2014.

All authors have completed and submitted the International Committee of Medical Journal Editors Form for Disclosure of Potential Conflicts of Interest. The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Alexander Gelbard, MD, Assistant Professor, Department of Otolaryngology, Vanderbilt School of Medicine, Medical Center East, S. Tower, 1215 21st Ave. South, Suite 7302, Nashville, TN 37232-8783. E-mail: alexander.gelbard@vanderbilt.edu

DOI: 10.1002/lary.24956

TABLE I.
Definitions of LTS Etiology of Injury Utilized in This Study.

Idiopathic	No history of significant laryngotracheal injury. No significant history of endotracheal intubation or tracheotomy within 2 years of the presentation. No thyroid or major anterior neck surgery. No neck irradiation. No caustic or thermal injuries to the laryngotracheal complex. No history of vasculitis. Negative titers for angiotensin-converting enzyme and antinuclear cytoplasmic antibody. The lesion must involve the subglottis.
Autoimmune	Patients with documented clinical, along with serologic and/or histologic, diagnosis of Wegener's granulomatosis, relapsing polychondritis, systemic lupus erythematous, rheumatoid arthritis, epidermolysis bullosa, sarcoidosis, or amyloidosis
Polytrauma	Patients presenting with laryngotracheal stenosis following documented traumatic injuries involving multiple organ systems
Iatrogenic	Patients who developed subglottic or tracheal stenosis following tracheostomy *or* subglottic or tracheal stenosis developing within 2 years of intubation

identified. Those with a history of tracheal malignancy or isolated laryngeal stenosis were excluded. Laryngeal and tracheal stenosis both share an association with prolonged endotracheal intubation, as well as many of the same comorbid medical risk factors. However, isolated laryngeal stenosis remains a distinct anatomic and structural injury with a unique treatment algorithm that merits dedicated independent study and is not discussed in the present work. Patients meeting inclusion were grouped into four categories based on stenosis etiology: 1) idiopathic, 2) iatrogenic, 3) autoimmune, and 4) polytrauma (Table I).

Data Collected

Patient characteristics (age, gender, race, follow-up duration) and comorbidities were extracted. Records were reviewed for etiology of stenosis, treatment approach (i.e., endoscopic, open), and surgical dates. Stenosis morphology (% luminal obstruction, distance from glottis [cm], and overall length [cm]) and tracheomalacia were derived from intraoperative findings. Patients were staged with the established Cotton-Myer, Lano, and McCaffrey classification systems, as previously described[3–5] (Table II). The number and frequency between repeat procedures were captured.

Procedures

Treatments for tracheal stenosis included: 1) endoscopic dilations of the stenotic trachea,[6] 2) open surgical resection of the diseased tracheal segment with end-to-end anastomosis,[7] and 3) permanent tracheostomy. The treatment algorithm consisted of initial endoscopic dilation for all patients. In patients who required multiple dilation procedures, rigorous selection criteria were applied for consideration of open surgical reconstruction. Patients less than 45 years old, without type 2 diabetes or connective tissue disease, and with stenosis 2 cm or more below the glottis and less than 2 cm in length were offered open surgical reconstruction.

TABLE II.
Definitions of Clinical LTS Classification Systems.

Cotton-Myer	I	<70% obstruction
	II	70%–90% obstruction
	III	>90% obstruction
	IV	Complete obstruction
Lano	I	One subsite* involvement
	II	Two subsite involvement
	III	Three subsite involvement
McCaffrey	I	Subglottis or trachea < 1 cm
	II	Subglottis > 1 cm
	III	Subglottis and trachea >1 cm
	IV	Any lesion involving glottis

*Subsites defined as glottis, subglottis, and trachea.

Outcomes

Presence of a tracheostomy at last follow-up was the primary outcome. This represented failure of surgical management to correct airway narrowing.

Statistical Analysis

All data management and analysis were done using STATA/MP version 12.1 software (STATACorp, College Station, Texas). Univariate analyses were performed using analysis of variance, Pearson's chi-squared tests, and Fisher's exact tests, as appropriate. Stepwise multivariate logistic regression analysis was used to identify independent risk factors for tracheostomy. A significance level of $P < 0.20$ on univariate analysis was used as the criterion for inclusion in the multivariate model. As per convention, $P < 0.05$ was required for statistical significance in the model.

RESULTS

A total of 340 patients with a diagnosis of tracheal or laryngeal stenosis were identified. Excluded were those with tracheal malignancy (N = 9) and isolated bilateral vocal-fold immobility (N = 181). In all, 150 patients met inclusion criteria. The most common etiology was iatrogenic (54.7%), followed by idiopathic (18.5%), autoimmune (18.5%), and traumatic (8%: Table III). Mean follow-up was 39.3 months (95% confidence interval [CI], 31.9–46.6), but varied significantly by etiology ($P < 0.001$; Table III).

Univariate Analysis

Patient Characteristics. Age at presentation differed significantly by strata ($P = 0.002$) with those in the traumatic group being significantly younger than all others (34.4 years, CI 23.5–45.3; Table III). Gender distribution also differed based on etiology ($P < 0.002$; Table III). In order, the idiopathic group had a significantly higher percentage of females (93%) than autoimmune (68%), iatrogenic (62%), or traumatic (33%) LTS patients. Charlson Comorbidity Index (CCI) varied between groups ($P < 0.001$). Iatrogenic and autoimmune strata had significantly higher indices than either idiopathic or traumatic strata (Table III).

TABLE III.
Demographics, Stenosis Characteristics, Comorbidities, and Treatment of LTS Grouped by Etiology of Injury.

Patient Characteristics	Idiopathic (n = 28)	Polytrauma (n = 12)	Autoimmune (n = 28)	Iatrogenic (n = 82)	Significance (*P*)
Demographics					
Follow-up (Mean months, 95% CI)	56.07	12.3	69.1	27.05	<0.001
	(41.5–70.6)	(7.2–17.5)	(39.7–98.6)	(20.9–33.1)	
Age (Mean years, 95% CI)	50.35	35.7	45.1	51	0.002
	(45.9–54.8)	(24.1–47.4)	(39.7–50.4)	(48.0–54.7)	
Sex (% female)	93	33	68	62	0.002
Race (%)					
Caucasian	89	50	71	63	0.330
African American	7	17	14	16	
Asian	0	8	0	2	
Hispanic	4	17	14	17	
Disease Morphology					
% Stenosis (Mean %, 95% CI)	57.86	69.6	68.5	72.8	0.010
	(52.3–63.4)	(55.1– 84.1)	(60.6–76.4)	(68.1–77.6)	
Distance below glottis (Mean cm, 95% CI)	1.289	2.17	1.94	1.77	0.110
	(1.0–1.6)	(1.29–3.05)	(1.38–2.51)	(1.5–2.02)	
Stenosis Length (Mean cm, 95% CI)	1.657	1.95	2.12	2.167	0.440
	(1.3−1.99)	(0.99–2.9)	(1.62–2.62)	(1.91–2.42)	
Comorbidities					
Charlson Index (Mean, 95% CI)	0.07	0.00	1.28	1.32	<0.001
	(0–0.16)	(0)	(0.99 −1.58)	(0.94–1.7)	
DMII (%)	0	0	11	39	<0.001
MI (%)	0	0	3.6	28	<0.001
CHF (%)	0	0	0	13	0.027
CVA (%)	0	0	0	7	0.008
COPD (%)	4	0	7	13	0.390
Connective tissue (%)	0	0	100	0	<0.001
GERD (%)	18	8	21	20	0.859
Treatment					
No. procedures/year (Mean, 95% CI)	1.75	3.41	1.8	2.65	0.490
	(0.8–2.6)	(1.6–5.2)	(0.9–2.7)	(1.7–3.6)	

CHF = congestive heart failure; CI = confidence interval; COPD = chronic obstructive pulmonary disease; CVA = cerebrovascular accident; DMII = diabetes mellitus type 2; GERD = gastroesophageal reflux disease; MI = myocardial infarction.

Examination of the individual components of the CCI showed cardiovascular comorbidities (i.e., myocardial infarction, congestive heart failure, peripheral vascular disease, and cerebrovascular disease) and diabetes mellitus type 2 (DMII) were significantly more prevalent in the iatrogenic strata than in other etiologies (Fig. 1A; Table III). There were not significant differences in the rate of gastroesophageal reflux disease (GERD) between strata (Table III).

Disease Morphology. Degree of stenosis differed between etiologic strata ($P = 0.01$). Idiopathic LTS involved less of the tracheal lumen (mean 57%; CI 52%–63%) than those in the autoimmune or iatrogenic groups (Table III). There were no differences in the mean distance from the glottis ($P = 0.11$) or the length of stenoses between strata ($P = 0.44$). In the iatrogenic group, LTS occurred in the subglottis (1.5 cm from the glottis) in 59% of patients (49/82) (Fig. 1B). Even in those patients presenting with iatrogenic LTS following tracheostomy, 41% (16/39) had subglottic injuries on intraoperative examination.

Treatment. There was no difference in number of surgeries per year of follow-up ($P = 0.49$) or the types of surgeries performed by etiologic strata ($P = 0.14$; Table III). Most patients were treated with tracheal dilation (84%), followed by T-tube placement (8%), resection (6%), and no treatment (2%).

Tracheostomy Dependence. Tracheostomy dependence differed based on etiologic strata ($P < 0.001$; Fig. 1C). Significantly more patients in the iatrogenic (66%) and autoimmune (54%) groups were tracheostomy-dependent at last follow-up compared to those in either the traumatic (33%) or idiopathic (0%) groups. Tracheostomy dependence also differed based on established staging systems (Fig. 2A). When stratified via Cotton-Myer staging (based on the degree of luminal

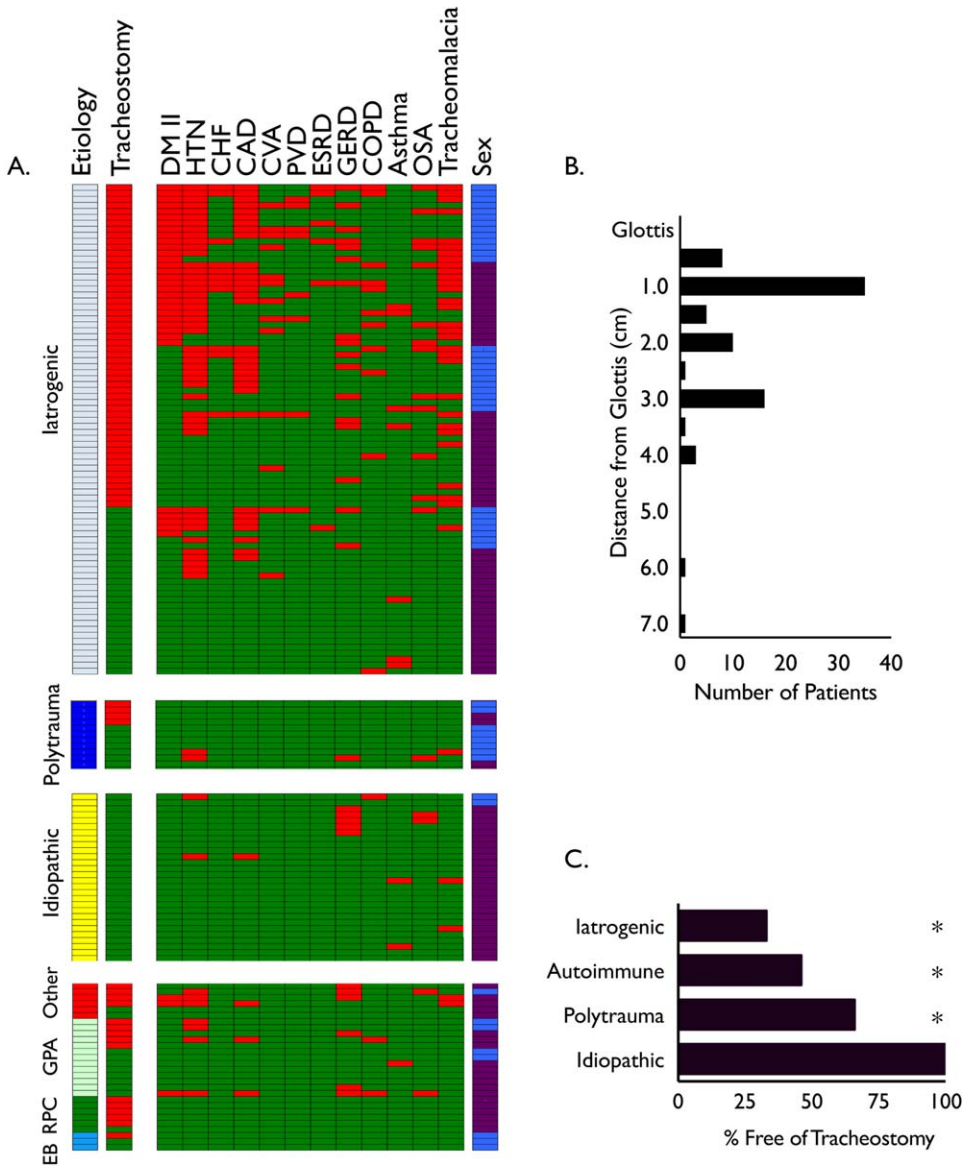

Fig. 1. Heatmap grouped by different etiologies of stenosis. Each line represents an individual patient. Tracheostomy status (red indicating tracheostomy), medical comorbidities (presence highlighted in red), and sex (blue indicating male, purple indicating female). In autoimmune subgroup: GPA (granulomatosis with polyangitis, i.e., Wegener's granulomatosis), RPC (relapsing polychondritis), EB (epidermolysis bullosa) (A). Location of tracheal stenosis in iatrogenic injuries. Histogram showing location of stenotic lesion in iatrogenic subgroup in relation to distance from glottis (B). Tracheostomy status of different etiologies at last follow-up. Asterisk denotes statistical significance from idiopathic group (C).

obstruction), significantly more patients with grade III (90%) and grade IV (90%) lesions were tracheostomy-dependent at last follow-up compared to those in either the grade II (38%) or grade I (32%) groups ($P < 0.001$; Fig. 2A). When stratified by the Lano classification (based on the stenosis location), increasing subsite involvement was significantly associated with a higher rate of tracheostomy ($P < 0.001$; Fig. 2A). When staged according to the McCaffrey classification system (based on both stenosis location and length), increased stage was associated with progressively increased risk of tracheostomy ($P < 0.01$; Fig. 2A).

All three of the established, adult LTS staging systems accurately stratified patients' outcomes based on the severity of their structural injury. Overall (consistent with prior reports), patients in our series with more severe luminal compromise, those with longer stenosis, and those with lesions spanning multiple subsites (glottis, subglottis, and/or trachea) had a much higher incidence of tracheostomy. However, this observation did not hold when patients were stratified by etiology of injury (Table IV.) No patients in the idiopathic group required tracheostomy (even those with lengthy, severe stenosis involving multiple subsites). Conversely, patients with iatrogenic injuries had a significantly higher rate of tracheostomy, even when matched at lower stenosis grades when compared with the other etiologic strata. The non-uniform rate of tracheostomy observed in different

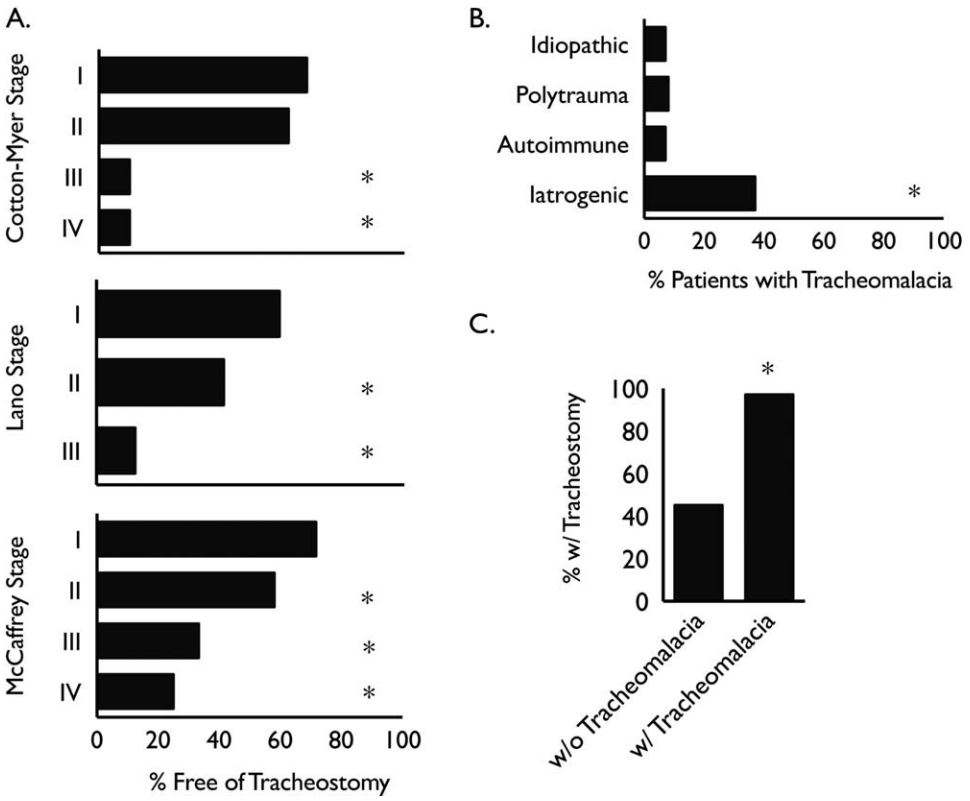

Fig. 2. Tracheostomy status of different Cotton-Myer, Lano, and McCaffrey stages at last follow-up. For Cotton-Myer staging, asterisk denotes statistical significance between grade I and II vs. grade III and IV (A). Diagnosis of tracheomalacia stratified by etiology. Asterisk denotes statistical significance between iatrogenic etiology and all other groups (B). Rate of tracheostomy in iatrogenic etiology patients with and without a diagnosis of tracheomalacia. Asterisk denotes statistical significance (C).

etiologic groups was seen in all three established LTS classification systems (Table IV).

Tracheal Structural Instability. Patients with iatrogenic injuries had a significantly higher rate of tracheomalacia observed on bronchoscopic evaluation (37% vs. 8%; $P < 0.001$; Fig. 2B). Given the retrospective nature of this work, it is not possible to establish a causative relationship between the initial injury and the loss of structural integrity associated with tracheomalacia. However, it is interesting that among the iatrogenic group, 45% of patients without malacia required tracheostomy, whereas 97% of those with malacia necessitated long-term tracheostomy ($P < 0.001$; Fig. 2C).

Multivariate Analysis

Multivariate regression analysis was performed to determine independent predictors of ultimate tracheostomy dependence. Each additional point on CCI was associated with a 67% increased odds of tracheostomy dependence (odds ratio [OR] 1.67; 95% CI 1.04–2.69; $P = 0.04$). Moreover, there was a 3% increased odds of tracheostomy dependence for each additional percentage of airway compromise (OR 1.03, 95% CI 1.01–1.06; $P = 0.001$). LTS patient characteristics (etiology, age, sex, race) were not significantly associated with odds of tracheostomy dependency.

DISCUSSION

Although most airway stenosis appears similar on anatomic imaging and clinical examination, we present data supporting the hypothesis that different mechanisms of injury are associated with differing rates of long-term tracheostomy dependence. The relationships between the anatomic stenosis characteristics (% stenosis, location, and length) and endoscopic or open surgical "success" have been established through pioneering work in children[8,9] and adults.[7] In advanced centers, procedural intervention for LTS offers a high rate of long-term tracheostomy free survival.[4,10,11] However, success in these large published series remains critically dependent on patient selection. With our consecutive series of both inpatient and outpatient consultations, we believe that this study captured a more representative cross-section of symptomatic LTS patients than many prior adult surgical case series. In the "real world," those patients deemed poor operative candidates (e.g., sicker patients) are often left with limited therapeutic options regardless of the structural morphology of their stenosis.

Endotracheal intubation and tracheostomy can be lifesaving but should not be considered benign procedures. They harbor significant long-term risks to communication,[12] swallowing,[13] and breathing,[14] particularly in the subset of patients with comorbid illness.[15] Ironically,

TABLE IV.
Percentage of LTS Patients With Tracheostomy by Cotton-Myer, Lano, and McCaffrey Stage, Grouped by Etiology of Injury.

	Cotton-Myer			
	I	II	III	IV
Idiopathic	0	0	0	n/a
Polytrauma	0	25	100	100
Autoimmune	36	50	100	100
Iatrogenic	57	44	92	88

	Lano		
	I	II	III
Idiopathic	0	0	n/a
Polytrauma	27	100	n/a
Autoimmune	50	42	83
Iatrogenic	60	75	100

	McCaffrey			
	I	II	III	IV
Idiopathic	0	0	0	0
Polytrauma	0	60	50	n/a
Autoimmune	63	50	50	n/a
Iatrogenic	36	65	82	80

n/a refers to an absence of patients within a given stage.

this is also the population that more frequently requires intensive respiratory support. In our series, each additional point on CCI was associated with a 67% increased odds of tracheostomy dependence. Although this association does not appear surprising, we believe that it is powerful. It demonstrates the suitability of the CCI to serve as a systems-based protocol to identify patients who mandate a heightened awareness of complications from these procedures.

Consistent with previously published series,[4,16,17] despite many risk factors for iatrogenic injury being clarified over the past 40 years,[15,18–20] more than half the LTS burden in our cohort was potentially preventable. Overall, 59% of iatrogenic injuries occurred within the subglottis; therefore, they are attributable to intubation. In a post hoc analysis, 83% (15/18) of the "healthy" patients (those without DMII or cardiovascular disease) with iatrogenic LTS were women. This previously reported observation[21] suggests that endotracheal tube size may contribute to tracheal injury and should be carefully considered in the smaller female trachea.[22]

As has been consistently shown across other large series,[15] patients with DMII are particularly vulnerable to airway injury and have a higher likelihood of long-term tracheostomy dependence when injury occurs. Interestingly, the rate of GERD was not significantly different between the etiologic subgroups. Although other investigators have suggested a tight relationship between GERD and adult idiopathic LTS, this was not seen in our patient population. The limits of retrospective review prevent us from direct comparison of the objective data on the frequency and severity of reflux episodes between individuals and subgroups. Increased body mass index also has a suggested association, with increased risk of tracheal injury with intubation and worse response to procedural intervention. Our series lacked the biometric data to address this concept. Additionally, the limits of our tertiary care referral center (with limited out-of-network medical records) prevented us from exploring the relationship between the length of intubation or type of tracheostomy procedure (open vs. percutaneous) and the ultimate injury severity or treatment outcome.

A strong association between the degree of stenosis and ultimate decannulation has previously been reported in children.[23] Our series supports these prior observations in the pediatric population and now extends them to adults. As previously reported in adults, the location of injury and the length of stenosis are also essential components to predict long-term tracheostomy dependence. Critically, we now also offer data supporting an additional relationship between the cause of upper airway injury and its ultimate response to therapy. This relationship had been assumed; we offer the first formal demonstration.

Anatomic staging systems are numerous,[3–5,24–28] yet the ideal system in adult LTS remains unresolved. The most established allow some degree of prognosis, promote individualized treatment planning, and facilitate multi-institutional comparison. In this work, we utilized three separate, established LTS classification systems. As expected, they all effectively stratify the patient's risk of long-term tracheostomy. Of interest, however, in adult LTS it appears that the McCaffrey and Lano systems offer more precision than does the Cotton-Myer scale.

In general, although those patients in our series with more severe luminal compromise, longer stenoses, and lesions spanning multiple subsites had a much higher incidence of tracheostomy, this observation did not hold in the idiopathic group (patients who never required tracheostomy), suggesting a unique injury. Conversely, whereas lower LTS stages (in all 3 systems) overall had a lower rate of tracheostomy, patients with iatrogenic injuries had a significantly higher rate, even when matched at lower stenosis grades (identically in all 3 systems). Grouping LTS patients solely by an anatomic classification of their injury neglects a critical component of the heterogeneous biology responsible for tracheal scar.

Patients with iatrogenic stenosis appear to possess unique medical comorbidities. Their disease ultimately behaves differently, as evidenced by their disparate rate of long-term tracheostomy dependence, even when matched for similar degree of luminal compromise. These separate subgroups likely merit tailored treatment strategies.

The finding of the high rate of tracheomalacia in the subgroup with iatrogenic injuries, and the significant association between tracheomalacia and long-term tracheostomy dependence in this subgroup, raises questions regarding the relative contributions of mucosal injury

and cartilaginous injury in LTS. Ultimately, we believe the degree of tracheal wall injury (what we term "superstructure instability") may have significant prognostic power for overall response to therapy. However, this is difficult to quantify at present with our current diagnostic modalities and remains an area of active research.

Our study represents one of the largest published adult LTS series in the scientific literature. The data supports the hypothesis that laryngotracheal stenosis is a common endpoint to multiple pathophysiologic processes. Although different mechanisms of airway injury physiologically affect the patient in similar ways, we show that they occur in unique populations and have divergent responses to therapy. Management and prevention strategies should carefully consider this heterogeneous pathophysiology. This difference is not reflected in staging systems limiting themselves to an anatomic description of the tracheal scar.

CONCLUSION

Relief through endoscopic dilation, or open tracheal resection, is attainable in some cases of LTS; however, treatment is not universally successful. It is incumbent on the scientific community to move beyond viewing LTS as a purely anatomic problem, remedied only through surgical reconstruction. Rather, the management of airway stenosis should transition to increasingly personalized plans of care based on early recognition of at-risk populations, and an understanding of the divergent pathophysiology affecting the unique subgroups with LTS.

Acknowledgments

Study concept and design: Drs. A. Gelbard and V.C. Sandulache; acquisition of data: Drs. A. Gelbard and J.C. Simmons; analysis and interpretation of data: Drs. A. Gelbard and D.O. Francis; drafting of the manuscript: Drs. A. Gelbard and D.O. Francis; critical revision of the manuscript: Drs. J. Ongkasuwan and D.O. Francis; statistical analysis: Dr. D.O. Francis.

Administrative, technical, or material support: Dr. D.T. Donovan; study supervision: Drs. D.T. Donovan and J. Ongkasuwan.

BIBLIOGRAPHY

1. Nouraei SA, Singh A, Patel A, Ferguson C, Howard DJ, Sandhu GS. Early endoscopic treatment of acute inflammatory airway lesions improves the outcome of postintubation airway stenosis. *Laryngoscope* 2006;116:1417–1421.
2. Institute of Medicine (U.S.). Committee on Quality of Health Care in America. *Crossing the Quality Chasm : A New Health System for the 21st century*. Washington, DC: National Academy Press; 2001.
3. Cotton RT. Pediatric laryngotracheal stenosis. *J Pediatr Surg* 1984;19: 699–704.
4. Lano CF Jr, Duncavage JA, Reinisch L, Ossoff RH, Courey MS, Netterville JL. Laryngotracheal reconstruction in the adult: a ten year experience. *Ann Otol Rhinol Laryngol* 1998;107:92–97.
5. McCaffrey TV. Classification of laryngotracheal stenosis. *Laryngoscope* 1992;102:1335–1340.
6. Baugnee PE, Marquette CH, Ramon P, Darras J, Wurtz A. [Endoscopic treatment of post-intubation tracheal stenosis. *Apropos of 58 cases]. Rev Mal Respir* 1995;12:585–592.
7. Grillo HC, Donahue DM, Mathisen DJ, Wain JC, Wright CD. Postintubation tracheal stenosis: treatment and results. *J Thorac Cardiovasc Surg* 1995;109:486–493.
8. Hartley BE, Rutter MJ, Cotton RT. Cricotracheal resection as a primary procedure for laryngotracheal stenosis in children. *Int J Pediatr Otorhinolaryngol* 2000;54:133–136.
9. Manning PB, Rutter MJ, Border WL. Slide tracheoplasty in infants and children: risk factors for prolonged postoperative ventilatory support. *Ann Thorac Surg* 2008;85:1187–1191; discussion 1191–1192.
10. Grillo HC. Primary reconstruction of airway after resection of subglottic laryngeal and upper tracheal stenosis. *Ann Thorac Surg* 1982; 33: 3–18.
11. Laccourreye O, Brasnu D, Seckin S, Hans S, Biacabe B, Laccourreye H. Cricotracheal anastomosis for assisted ventilation-induced stenosis. *Arch Otolaryngol Head Neck Surg* 1997;123:1074–1077.
12. Whited RE. Posterior commissure stenosis post long-term intubation. *Laryngoscope* 1983;93:1314–1318.
13. Macht M, Wimbish T, Clark BJ, et al. Postextubation dysphagia is persistent and associated with poor outcomes in survivors of critical illness. *Crit Care* 2011;15:R231.
14. Streitz JM Jr, Shapshay SM. Airway injury after tracheotomy and endotracheal intubation. *Surg Clin North Am* 1991;71:1211–1230.
15. Wu X, Su ZZ, Hu LJ, et al. [Analysis of the risk factors causing tracheal stenosis after tracheotomy for mechanical ventilation in 560 patients]. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 2007;42:839–842.
16. Kleiss IJ, Verhagen AF, Honings J, Schuurbiers OC, van der Heijden HF, Marres HA. Tracheal surgery for benign tracheal stenosis: our experience in sixty three patients. *Clin Otolaryngol* 2013;38:343–347.
17. Mutrie CJ, Eldaif SM, Rutledge CW, et al. Cervical tracheal resection: new lessons learned. *Ann Thorac Surg* 2011;91:1101–1106; discussion 1106.
18. Mathias DB, Wedley JR. The effects of cuffed endotracheal tubes on the tracheal wall. *Br J Anaesth* 1974;46:849–852.
19. Seegobin RD, van Hasselt GL. Endotracheal cuff pressure and tracheal mucosal blood flow: endoscopic study of effects of four large volume cuffs. *Br Med J (Clin Res Ed)* 1984;288:965–968.
20. Andrews MJ. The incidence and pathogenesis of tracheal injury following tracheostomy with cuffed tube and assisted ventilation. Analysis of a 3-year prospective study. *Br J Surg* 1971;58:749–755.
21. Zias N, Chroneou A, Tabba MK, et al. Post tracheostomy and post intubation tracheal stenosis: report of 31 cases and review of the literature. *BMC Pulm Med* 2008;8:18.
22. Griscom NT, Wohl ME. Dimensions of the growing trachea related to age and gender. *AJR Am J Roentgenol* 1986;146:233–237.
23. Hartnick CJ, Hartley BE, Lacy PD, et al. Surgery for pediatric subglottic stenosis: disease-specific outcomes. *Ann Otol Rhinol Laryngol* 2001;110: 1109–1113.
24. Anand VK, Alemar G, Warren ET. Surgical considerations in tracheal stenosis. *Laryngoscope* 1992;102:237–243.
25. Brichet A, Verkindre C, Dupont J, et al. Multidisciplinary approach to management of postintubation tracheal stenoses. *Eur Respir J* 1999;13: 888–893.
26. Freitag L, Ernst A, Unger M, Kovitz K, Marquette CH. A proposed classification system of central airway stenosis. *Eur Respir J* 2007;30: 7–12.
27. Grillo HC, Mark EJ, Mathisen DJ, Wain JC. Idiopathic laryngotracheal stenosis and its management. *Ann Thorac Surg* 1993;56:80–87.
28. Grundfast KM, Morris MS, Bernsley C. Subglottic stenosis: retrospective analysis and proposal for standard reporting system. *Ann Otol Rhinol Laryngol* 1987;96:101–105.

Reprinted by permission of Am J Surg. 2014; 208(5):770-774.

The American Journal of Surgery (2014) 208, 770-774

The American Journal of Surgery®

Clinical Science

Incidence of overall complications and symptomatic tracheal stenosis is equivalent following open and percutaneous tracheostomy in the trauma patient

William W. Kettunen, M.D.[a], Stephen D. Helmer, Ph.D.[a], James M. Haan, M.D.[a,b,*]

[a]*Department of Surgery, The University of Kansas School of Medicine – Wichita, 929 N. Saint Francis Street, Room 3082, Wichita, KS 67214, USA;* [b]*Department of Trauma Services, Via Christi Hospital, Saint Francis Campus, Wichita, KS, USA*

KEYWORDS:
Tracheal stenosis;
Percutaneous tracheostomy;
Open tracheostomy

Abstract

BACKGROUND: While percutaneous tracheostomy (PT) is becoming the procedure of choice for elective tracheostomy, there is little late complication data. This study compared incidence of, and factors contributing to, tracheal stenosis following PT or open tracheostomy (OT).

METHODS: A 10-year review was conducted of trauma patients undergoing tracheostomy. Data on demographics, injury severity, tracheostomy type, complications, and outcomes were compared between patients receiving PT or OT and for those with or without tracheal stenosis.

RESULTS: Of 616 patients, 265 underwent OT and 351 underwent PT. Median injury severity score was higher for PT (26 vs 24, $P = .010$). Overall complication rate was not different (PT = 2.3% vs OT = 2.6%, $P = .773$). There were 9 tracheal stenosis, 4 (1.1%) from the PT group and 5 (1.9%) from the OT group ($P = .509$). Mortality was higher in OT patients (15.5% vs 9.7%, $P = .030$). Patients developing tracheal stenosis were younger (29.8 vs 45.2 years, $P = .021$) and had a longer intensive care unit length of stay (28.3 vs 18.9 days, $P = .036$).

CONCLUSION: Risk of tracheal stenosis should not impact the decision to perform an OT or PT.
© 2014 Elsevier Inc. All rights reserved.

Percutaneous tracheostomy (PT) is becoming the procedure of choice for elective tracheostomy in trauma patients. Many studies have proven this more prevalent technique to be safe, and possibly more cost-effective, than the traditional open tracheostomy (OT).[1–3] Most of the literature consists of observational data or small prospective studies, therefore debate still continues as to which method is preferred.

The literature is less clear on late complications, specifically tracheal stenosis. The exact incidence of tracheal stenosis following tracheostomy is difficult to quantify because many patients are critically ill and may die before decannulation, are lost to follow-up after being dismissed from a level-I trauma center, or are asymptomatic.[4–6] With a shortage of evidence, some postulate that the percutaneous technique predisposes patients to tracheal stenosis, more so than the open technique. They cite that the

No current or previous support was received from industry or organizations that might have influenced this work.

Presented at the 71st Annual Meeting of the American Association for the Surgery of Trauma, September 12 to 15, 2012, Kauai, Hawaii.

* Corresponding author. Tel.: +1-316-268-5538; fax: +1-316-291-7892.

E-mail address: James.Haan.Research@viachristi.org

Manuscript received May 22, 2013; revised manuscript December 4, 2013

0002-9610/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amjsurg.2013.12.036

ostomy is often times placed higher on the trachea percutaneously than it is when performed open and there is more trauma and granulation tissue to the trachea when passing dilators percutaneously.[7]

The purpose of this investigation was to compare outcomes and complications between OT and PT. All major complications, including tracheal stenosis, were recorded to determine the incidence of, and any risk factors for, tracheal stenosis.

Patients and Methods

A retrospective review of all trauma patients who received a tracheostomy from August 1, 2001 to July 31, 2011 was conducted. Patients were identified using the trauma registry of an established American College of Surgeons-verified level-1 trauma center. Patient demographics, mechanism of injury, injury severity score (ISS), Glasgow coma scale (GCS) score, time from injury to tracheostomy creation, method of performing tracheostomy (open vs percutaneous), complications associated with tracheostomy (tracheo-innominate artery fistula, tracheal stenosis, scar and excess granulation tissue requiring surgical scar revision, loss of airway requiring conversion to open, and bleeding requiring conversion to open), intensive care unit (ICU) length of stay (LOS), mechanical ventilator days, overall LOS, and patient disposition were collected using the trauma registry and patient records. Tracheal stenosis was identified based on clinical symptoms (ie, difficulty with decannulation or shortness of breath with exertion). Complications were defined as being early, those occurring within the first 48 hours of tracheostomy, or late, those occurring more than 48 hours post-tracheostomy. Outcomes and complication data were collected from the in-hospital stay and from rehospitalizations. Study subjects were not contacted for long-term follow-up.

Analyses were conducted using IBM SPSS Statistics for Windows, Version 19.0. (IBM Corp, Armonk, NY). Data were initially summarized. Primary analyses were conducted comparing outcomes between patients based on the method of tracheostomy creation (OT vs PT). Secondary analyses were conducted comparing outcomes between patients who developed tracheal stenosis and patients who did not develop tracheal stenosis. Quantitative data were analyzed using the Student t test. If heterogeneity of variance was identified, the Mann–Whitney test was used. Comparisons of ordinal data were analyzed with the Mann–Whitney test. Qualitative data were analyzed with chi-square analysis or the Fisher's exact test in instances where cell size was 5 or less observations. All analyses were conducted as 2-tailed tests and statistical significance was defined as $P < .05$.

This study was reviewed and approved for implementation by the Institutional Review Board of Via Christi Hospitals Wichita, Inc.

Results

During the 10-year study period, 629 tracheostomies were performed on trauma patients. We excluded 13 patients who had an emergency cricothyroidotomy or whose LOS was for more than 1 day. Of the remaining 616 patients, the average age was 45.0 ± 20.6 years, the majority were male (n = 458, 74.4%), white (n = 534, 86.7%), and median ISS and GCS scores were 25 (25th and 75th percentiles = 17 and 33) and 5 (25th and 75th percentiles = 3 and 14), respectively. Forty-three percent (n = 265) had an OT and 57% (n = 351) had a PT. There were no significant differences in age, sex, GCS score, mechanism of injury, interval from admission to tracheostomy formation, ICU LOS, ventilator days, or hospital LOS between the 2 groups (Table 1). There was a significant

Table 1 Comparison of demographics, injury severity, mechanism of injury, and hospitalization details for patients who received a tracheostomy through an open or percutaneous procedure

Parameter	Open procedure	Percutaneous procedure	P value
No. of subjects (%)	265 (43.0%)	351 (57.0%)	
Age (years)*	45.0 ± 21.3	44.9 ± 20.1	.932
Sex (male)	204 (77.0%)	254 (72.4%)	.194
Injury severity score†	24.0 (17.0, 30.0)	26.0 (18.0, 34.0)	.010
Glasgow coma scale score†	6.0 (3.0, 15.0)	3.0 (3.0, 14.0)	.116
Mechanism of injury			.068
Blunt	244 (92.1%)	337 (96.0%)	
Penetrating	18 (6.8%)	13 (3.7%)	
Drowning	1 (.4%)	1 (.3%)	
Burn	2 (.8%)	0 (.0%)	
Admission to tracheostomy interval (days)	7.0 ± 5.4	7.0 ± 4.7	.988
Intensive care unit days*	19.3 ± 15.2	18.9 ± 11.8	.223
Mechanical ventilation days*	16.7 ± 12.9	15.8 ± 11.2	.945
Hospital length of stay (days)*	27.6 ± 19.9	26.7 ± 29.2	.643

*Mean ± standard deviation.
†Median (25th and 75th percentile).

Table 2 Comparison of complication, disposition, and death data for patients who received a tracheostomy through an open or percutaneous procedure

Parameter	Open procedure Number (%)	Percutaneous procedure Number (%)	*P* value
Complication	7 (2.6%)	8 (2.3%)	.773
Tracheal stenosis	5 (1.9%)	4 (1.1%)	.509
Other major complications	2 (.8%)	4 (1.1%)	.704
Disposition			.007
Home/home with home health care/jail/mental health facility	44 (16.6%)	35 (10.0%)	
Rehabilitation center/select specialty hospital acute care/other acute hospitals	165 (62.3%)	262 (75.1%)	
Skilled nursing unit/nursing home	14 (5.3%)	14 (4.0%)	
Hospice/death	42 (15.8%)	38 (10.9%)	
Death	41 (15.5%)	34 (9.7%)	.030

difference in ISS between the 2 groups with the percutaneous group having a higher median score than the open group (24 vs 26, $P = .007$).

The overall complication rate was similar between the OT and PT groups (Table 2). The incidence of tracheal stenosis was also similar when comparing the open group with the percutaneous group (1.9% vs 1.1%, $P = .509$). The open group had an incidence of major complications other than tracheal stenosis of .8%. These complications involved scar and excess granulation tissue requiring surgical scar revision ($n = 2$). The percutaneous group had an incidence of other major complications of 1.1% ($n = 4$). These included tracheo-innominate artery fistula ($n = 1$), loss of airway requiring conversion to open ($n = 2$), and bleeding requiring conversion to open ($n = 1$). The patient with a tracheo-innominate artery fistula hemorrhaged while on the floor. The hemorrhage was occluded manually while the patient was taken to the operating room for repair, but the patient exsanguinated before repair could be accomplished. All conversions to an open procedure occurred during the initial hospitalization.

Four of the 9 patients were immediately diagnosed with tracheal stenosis after a failed decannulation attempt; however, the other 5 patients presented in a delayed fashion after being decannulated. The delay ranged from 3 to 12 months, with patients presenting with shortness of breath with exertion ($n = 4$) and with trouble extubating after elective laparoscopic cholecystectomy ($n = 1$). All of the 9 patients underwent some form of treatment for their stenosis. Five of these 9 patients underwent bronchoscopy with tracheal balloon dilation, while 4 of the 9 patients underwent tracheal resection. In our study, the risk of tracheal stenosis requiring invasive intervention following tracheostomy was 1.5%.

There was a significant difference in patient disposition between the 2 groups (Table 2). The open group was discharged home more often and to a rehabilitation center less often when compared with the percutaneous group ($P = .007$); however, mortality rate was higher in the open group (15.5% vs 9.7%, $P = .030$).

We also conducted analyses comparing those patients with tracheal stenosis with those with no tracheal stenosis, independent of which method of tracheostomy was performed (Table 3). Patients who developed tracheal stenosis were younger (29.8 vs 45.2 years of age, $P = .021$), had a longer ICU LOS (28.3 vs 18.9 days, $P = .036$), and tended to require mechanical ventilation for a longer interval (26.7 vs 16.1 days, $P = .055$) compared with those who did not develop tracheal stenosis. There were, however, no differences between the groups in regard to sex, ISS, GCS score, mechanism of injury, interval between admission and tracheostomy formation, hospital LOS, disposition, or mortality.

Comments

While there is support in the literature of equivalent early complication rates between open and percutaneous techniques,[8,9] there is less evidence about their equivalency with regard to late complications such as tracheal stenosis. For this reason, there is still debate about which method provides superior patient outcomes. The incidence of symptomatic tracheal stenosis following OT or PT ranges in the literature from 0% to 10%.[4–6] The true incidence of tracheal stenosis is difficult to ascertain because it is often subclinical in nature. In our study, tracheal stenosis was identified based on clinical symptoms. Our study was similar to these published results, demonstrating equivalent symptomatic tracheal stenosis rates for OT and PT (1.9% vs 1.1%, respectively).

As stated earlier, several studies demonstrate complication rates that are equivalent for PT and OT. Our study supports the literature in this regard with an overall complication rate of 2.3% and 3.3%, respectively. The types of complications encountered during tracheostomy creation have been described in the literature and include peristomal bleeding, peristomal infection, loss of airway during procedure, surgical scar contracture, and tracheo-innominate artery fistula.[2,10] The complications reported in our study are in line with those previously described. Major complications in our study were defined as need for surgical intervention or death. Both of the reoperations in the

Table 3 Comparison of demographics, injury severity, mechanism of injury, hospitalization details, disposition, and death for patients with tracheal stenosis versus patients without tracheal stenosis

	Tracheal stenosis	No tracheal stenosis	
Parameter	Value	Value	*P* value
No. of subjects	9 (1.5%)	607 (98.5%)	
Age (years)*	29.8 ± 11.8	45.2 ± 20.6	.021
Sex (male)	6 (66.7%)	452 (74.5%)	.701
Injury severity score†	30.0 (19.5, 37.0)	25.0 (17.0, 33.0)	.175
Glasgow coma scale score†	3.0 (3.0, 13.0)	5.0 (3.0, 14.0)	.659
Mechanism of injury			1.000
Blunt	9 (100.0%)	572 (94.2%)	
Penetrating	0 (.0%)	31 (5.1%)	
Drowning	0 (.0%)	2 (.3%)	
Burn	0 (.0%)	2 (.3%)	
Admission to tracheostomy interval (days)	9.3 ± 7.9	7.0 ± 4.9	.175
Intensive care unit days*	28.3 ± 18.8	18.9 ± 13.3	.036
Mechanical ventilation days*	26.7 ± 21.7	16.1 ± 11.7	.055
Hospital length of stay (days)*	40.0 ± 21.7	26.9 ± 25.6	.127
Disposition			.604
Home/home with home health care/jail/mental health facility	2 (22.2%)	77 (12.7%)	
Rehabilitation center/select specialty hospital acute care/ other acute hospitals	7 (77.8%)	42 (69.4%)	
Skilled nursing unit/nursing home	0 (.0%)	28 (4.6%)	
Hospice/death	0 (.0%)	80 (13.2%)	
Death	0 (.0%)	75 (12.4%)	.610

*Mean ± standard deviation.
†Median (25th and 75th percentile).

open group were for surgical scar revision. The other major complications observed in our PT patients included loss of airway, bleeding, and tracheo-innominate artery fistula, all well-known risks of PT.

As we did not find the route of tracheostomy formation to influence the incidence of tracheal stenosis, we attempted to identify factors that may have contributed to stenosis formation. Numerous and variable risk factors for tracheal stenosis following intubation have been suggested in the literature and include trauma and inflammation at the endotracheal tube cuff site, excess granulation tissue around the tracheal stoma site or over a fractured cartilage, high tracheostomy site, prolonged intubation, traumatic intubation, or previous intubation or tracheostomy.[1] Both groups in this study received tracheostomy within 7 days of admission as per American College of Surgeons recommendations. Our study demonstrated that patients who developed tracheal stenosis tended to have longer mechanical ventilator requirements (26.7 vs 16.1 days, $P = .055$), with patients developing stenosis being on the ventilator on average 11 additional days. It could be hypothesized that additional ventilator days meant more time with an inflated tracheal cuff causing tracheal ischemia and stenosis. We did identify that younger age and longer length of ICU stay were associated with increased rate of tracheal stenosis; however, the reason for these findings is unclear and these findings were not observed in similar studies.

There are several limitations to our study. It is retrospective in nature with a relatively small sample size. Also, the study time frame included the widespread introduction of the percutaneous technique at our institution. The associated learning curve for a new procedure may have influenced the results or influenced which technique was used to create the tracheostomy in specific patients. Also, the percentage of patients seen after dismissal from the hospital was unknown.

Common shortcomings of other investigations into this topic are length and reliability of follow-up, thereby calling into question the accuracy of the reported incidence of tracheal stenosis. We have a unique practice environment in which there are 2 level-1 trauma centers serving the entire population center with extremely rural surroundings. This leads to an isolated trauma population for study. These centers inform each other of any complications or admissions from each other's population. Additionally, the next closest trauma center for follow-up for tracheal symptoms is 200 miles away. Because of this, we do not believe that any patients were lost to follow-up or transferred to the other trauma center with a complication of tracheal stenosis. Also, in our city, if a patient presented to one of the other large multispecialty groups, then that patient would be redirected to our clinic. Evidence for this rests in the fact that 5 of the 9 tracheal stenosis patients presented to our clinic in a delayed fashion well after hospital discharge. These factors help distinguish our follow-up results as compared with other studies. This said, there still exists the possibility that late occurring and/or subclinical tracheal stenoses may have

been missed as we did not contact study subjects for late-term follow-up or measure tracheal circumference in this retrospective study.

PT has been proven to be a safe and cost-effective method for elective tracheostomy. Our study demonstrates that the complication rate, particularly that of tracheal stenosis, after PT is equivalent to that observed after the traditional OT procedure.

References

1. Freeman BD, Isabella K, Lin N, et al. A meta-analysis of prospective trials comparing percutaneous and surgical tracheostomy in critically ill patients. Chest 2000;118:1412–8.
2. Friedman Y, Fildes J, Mizock B, et al. Comparison of percutaneous and surgical tracheostomies. Chest 1996;110:480–5.
3. Holdgaard HO, Pedersen J, Jensen RH, et al. Percutaneous dilational tracheostomy versus conventional surgical tracheostomy. A clinical randomized study. Acta Anaesthesiol Scand 1998;42:545–50.
4. Norwood S, Vallina VL, Short K, et al. Incidence of tracheal stenosis and other late complications after percutaneous tracheostomy. Ann Surg 2000;232:233–41.
5. Kornblith LZ, Burlew CC, Moore EE, et al. One thousand bedside percutaneous tracheostomies in the surgical intensive care unit: time to change the gold standard. J Am Coll Surg 2011;212:163–70.
6. Jackson LS, Davis JW, Kaups KL, et al. Percutaneous tracheostomy: to bronch or not to bronch–that is the question. J Trauma 2011;71:1553–6.
7. Raghuraman G, Rajan S, Marzouk JK, et al. Is tracheal stenosis caused by percutaneous tracheostomy different from that by surgical tracheostomy? Chest 2005;127:879–85.
8. Ciaglia P, Graniero KD. Percutaneous dilatational tracheostomy. Results and long-term follow-up. Chest 1992;101:464–7.
9. Cobean R, Beals M, Moss C, et al. Percutaneous dilatational tracheostomy. A safe, cost-effective bedside procedure. Arch Surg 1996;131:265–71.
10. Wood DE. Tracheostomy. Chest Surg Clin N Am 1996;6:749–64.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2015; 124(2):137-142.

Article

Using Pulmonary Function Data to Assess Outcomes in the Endoscopic Management of Subglottic Stenosis

Annals of Otology, Rhinology & Laryngology
2015, Vol. 124(2) 137–142
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489414548915
aor.sagepub.com
SAGE

Shannon M. Kraft, MD[1], Kevin Sykes, PhD, MPH[1], Andrew Palmer, CCC-SLP[2], and Joshua Schindler, MD[2]

Abstract

Objective: This study aimed to examine the authors' experience with endoscopic management of idiopathic subglottic stenosis (iSGS), and to identify pulmonary function test (PFT) values that can be used to quantify outcomes.

Methods: Retrospective review.

Results: Twenty-five patients with a new diagnosis of iSGS were seen between 2006 and 2012. Median age at surgery was 45.3 years (interquartile range [IQR], 38.5-67.0), and median body mass index was 28.7 kg/m^2 (IQR, 23.5-32.1). Forty-five procedures were performed. Median preoperative stenosis was 56.8% (Cotton-Myer grade 2). The typical stenosis began 15 mm below the true vocal folds and was 12 mm long. Median follow-up was 21.4 months (IQR, 5.1-43.1). For patients receiving multiple dilations, median time between procedures was 23.7 months. Four PFT parameters demonstrated significant improvement after intervention: (1) PEF (absolute change = 2.54 L/s), (2) PIF (absolute change = 1.57 L/s), (3) FEV1/PEF (absolute change = 0.44), and (4) FIF50% (absolute change = 1.71 L/s). PIF was the only parameter affected by using a larger balloon (P = .047).

Conclusion: PEF, PIF, FEV1/PEF, and FIF50% improved significantly after endoscopic incision and dilation of iSGS, and this could potentially be used as a metric by which to evaluate outcomes in the endoscopic management of subglottic stenosis.

Keywords

endoscopic, idiopathic subglottic stenosis, pulmonary function test

Introduction

Etiology/Pathophysiology

Idiopathic subglottic stenosis (iSGS) is a nonspecific fibroinflammatory process that results in progressive narrowing of the airway at the level of the cricoid and proximal trachea. Investigators have proposed a multitude of theories as to the etiology of the process since its first description in 1972.[1] The disorder occurs almost exclusively in women, typically presenting in the fourth or fifth decade.[2] This female preponderance led to the hypothesis that estrogen might play a role in the pathogenesis of iSGS, although the evidence for this has not borne out.[3,4]

Current theories regarding the etiology of iSGS have focused on non-gender-specific causes. Subglottic injury from laryngopharyngeal reflux (LPR) has been implicated as a possible cause.[5] Autoimmune disease, specifically a limited form of seronegative polyangiitis with granulomatosis (GPA), has also been proposed as a mechanism for iSGS.[6] Other theories on the etiology of the disorder include repetitive microtrauma from cough[7] and hereditary factors.[8] Ultimately a disease of the subglottic lamina propria, no single etiologic factor has been identified.

Endoscopic Management of iSGS

Although segmental resection of the diseased portion of the airway is considered the gold standard for obtaining long-term improvement in the airway in iSGS, it is a potentially morbid procedure.[9] Endoscopic surgery provides an attractive alternative that is less invasive, albeit less definitive. As surgeon experience has grown, several variations in technique have been explored. Dedo and

[1]Department of Otolaryngology, University of Kansas, Kansas City, Kansas, USA
[2]Department of Otolaryngology, Oregon Health and Science University, Portland, Oregon, USA

Corresponding Author:
Shannon M. Kraft, MD, Department of Otolaryngology–Head and Neck Surgery, University of Kansas Medical Center, 3901 Rainbow Boulevard, MS 3010, Kansas City, KS 66160, USA.
Email: skraft3@kumc.edu

Catten[3] described a microflap technique in which a CO_2 laser was used to make a mucosal flap followed by resection or ablation of the aberrant soft tissue between the flap and the cricoid. Shapshay et al[10] reported on the use of radial incisions with a CO_2 laser, followed by dilation with a rigid bronchoscope. Some surgeons favor a cold technique over the laser for making radial incisions.[11] Balloon dilation was introduced in the management of subglottic stenosis as a theoretically less traumatic alternative to passing rigid dilators.[12]

In addition to the different surgical techniques, there have been several investigations evaluating wound-healing modulators as adjunctive therapies. Most notable, inhaled, systemic, and locally injected steroids have been used extensively in an attempt to slow restenosis after dilation.[13] Mitomycin C, an alkylating agent, has been proven to prevent fibroblast proliferation[14] and has been used with varying degrees of success in endoscopic airway surgery.[15] Halofuginone, an inhibitor of collagen 1-α synthesis,[16] and 5-fluorouracil, an antimetabolite that inhibits fibroblast activity,[17] are also being investigated in animal models as potential adjunctive therapies.

Given the variety of options in the endoscopic management of iSGS, it has become obvious that an objective means by which to quantify operative outcomes and to compare the efficacy of different techniques or adjunctive therapies is sorely needed.

Assessing Outcomes in the Endoscopic Management of Subglottic Stenosis

In the late 1960s and early 1970s, there was a great deal of interest in using pulmonary function tests (PFTs) to aid in the diagnosis of upper airway obstruction (UAO). Initial investigations focused on identifying values that could differentiate UAO from lower airway disease.[18,19] It was ultimately Hyatt's[20] description of the flow-volume loop and the different patterns of obstruction (variable intrathoracic, variable extrathoracic, and fixed) that provided clinicians with a powerful tool for diagnosing and classifying UAO. There has been a renewed interest in the past 2 decades in using PFT data not only as a diagnostic tool but also as a means of quantifying the results of interventions in UAO.[21,22]

This retrospective study is designed to review our experience with using pulmonary function data in the management of patients with iSGS. Specifically, we aim to (1) describe our experience with iSGS, (2) identify which PFT parameters change following endoscopic intervention, (3) quantify the degree of improvement in airflow postoperatively using PFT data, and (4) determine if PFTs change in a predictable manner postoperatively as restenosis invariably occurs.

Methods

Inclusion/Exclusion Criteria

This study was approved by the institutional review board at Oregon Health and Science University. A retrospective chart review was performed, examining a single surgeon's experience with iSGS. Records for all newly diagnosed adult patients with iSGS referred to the senior author between January 1, 2006, and December 31, 2012, were reviewed. At least 1 standard endoscopic intervention (described below) and 1 pulmonary function test was required for inclusion in the study. Patients with a history of previous airway surgery, airway trauma, tracheotomy, positive serology (c-ANCA or ACE level), or intubation within the previous 12 months were excluded from the analysis.

Management of Stenosis

All patients included in the study were managed endoscopically in a standardized fashion. After confirming that the patient could be mask ventilated, anesthesia was induced and maintained using a total intravenous technique. A plastic tooth guard was used to protect the maxillary dentition. The airway was exposed using an Osshoff-Pilling laryngoscope, and the patient was suspended from the Mustard table. Ventilation was conducted via jet technique. Standard laser precautions were used throughout the procedure, to include placing moist gauze pads over the eyes and wet towels over any exposed skin.

A 4-mm rod lens telescope was used to evaluate and measure the stenosis. If the initial values were obtained by the resident or fellow, the senior author (J.S.) repeated and confirmed the measurements. The stenosis was described by its distance below the true vocal folds and by its total length. Once measurements were complete, 2-mm cup forceps were used to take a biopsy. The stenosis subsequently was infiltrated with 1 mL of triamcinolone suspension (40 mg/mL). A CO_2 laser was then used to make 3 to 5 radial incisions in the stenosis. Dilation of the stenosis was performed using a constant radial expansion (CRE) balloon. After confirming hemostasis and obtaining postoperative photo documentation, a small, cuffed endotracheal tube (5.0) was placed through the laryngoscope. The patient was taken out of suspension and the laryngoscope removed. The patient was then remanded to the care of the anesthesia team for recovery. Patients were discharged home from the postanesthesia care unit that same day with a fluticasone inhaler (220 mcg) and instructed to use 2 puffs twice daily until the inhaler was empty.

Data Collection and Statistical Analysis

After obtaining approval from the institutional review board, the senior surgeon's operative record was reviewed

Table 1. Clinical Presentation.

	Median	Interquartile Range
Race		
Caucasian	23	
Latina	2	
Age at surgery, mo	45.3	38.5-67.0
Weight, kg	70	61.5-85.5
Height, m	1.6	1.55-1.65
Body mass index, kg/m^2	28.7	23.5-32.1
Stenosis, %	58.60	38.9-78.4
Cotton-Myer grade	2	1-3
Distance below cords, mm	15	11.0-16.5
Length of stenosis, mm	12	9.0-17.0
Follow-up, mo	21.4	5.1-43.1
No. of procedures	2	1-3
Time between first & second surgeries, mo	23.7	15.6-31.8

to identify patients with a preoperative diagnosis of subglottic stenosis. Clinical records were then further screened to select patients who met the inclusion criteria. All data points, to include demographic, clinical, operative, laboratory, radiographic, and pulmonary functional data, were entered in a Microsoft Excel spreadsheet. Data were imported into and analyzed via SPSS version 20.0 (SPSS Inc, Chicago, Illinois, USA). A P value < .05 was considered to be significant.

As we could not assume a normal distribution for our small sample size, a Wilcoxon signed rank test was used to look for differences in the pre-dilation and post-dilation differences in pulmonary function data. Subsequently, a Kruskal-Wallis test was performed on those PFT parameters that improved significantly to determine the contribution of dilation size on the change. For those measures that changed significantly, postoperative PFT values were plotted as a function of time to assess the rate of change after intervention.

Results

Clinical Presentation and Evaluation of Stenosis

A total of 25 new patients with iSGS were seen between 2006 and 2011 at our clinic. As anticipated, all patients were female. Ninety-two percent (N = 23) were white and 8% (n = 2) Hispanic. The median age at the time of the first surgical intervention was 45.3 years (interquartile range [IQR], 38.5-67.0), with a median body mass index of 28.7 kg/m^2 (IQR, 23.5-32.1) (Table 1).

A total of 45 procedures were performed. Twelve patients had 1 procedure, 9 patients had 2 procedures, 2 patients had 3 procedures, 1 patient had 4 procedures, and 1 patient had a total of 7 procedures (median = 2). At the time of initial presentation, the typical stenosis was described as beginning 15 mm below the true vocal folds and measuring 12 mm long. Fifteen patients had preoperative computed tomography (CT) scans. The median degree of stenosis as determined by CT was 56.8% (Cotton-Myer grade 2). Four patients had grade 1 stenosis at presentation and 4 patients had grade 3 stenosis.

Median follow-up after surgery was 21.4 months (IQR, 5.1-43.1). For the 11 patients receiving at least 2 dilations, the median time between the first and second procedures was 23.7 months. Three patients ultimately proceeded to definitive cricotracheal resection.

Preoperative Versus Postoperative PFT Assessment

Seventeen of the 25 patients had a preoperative PFT in addition to at least 1 postoperative PFT performed within 8 weeks of surgery. Four parameters demonstrated a statistically significant improvement after intervention: (1) PEF (absolute change = 2.54 L/s), (2) PIF (absolute change = 1.57 L/s), (3) FEV1/PEF (absolute change = 0.44), and (4) FIF50% (absolute change = 1.71 L/s). FEV1, FVC, FEF25%-75%, and PEF/PIF did not change significantly (Table 2). Preoperative and postoperative PFT values were then examined in the context of balloon size. Improvement in the PIF was the only parameter that was affected by the size of dilation, with rank-order testing indicating a greater degree of improvement with use of a larger balloon (P = .047) (Table 3).

Changes in PFT Over Time

Seven patients of the 25 in this sample had at least 3 PFTs taken following the initial dilation. PEF, PIF, FEV1/PEF, and FIF50% from these samples were plotted as a function of time. There is a linear relationship between time and both PEF and FEV1/PEF (P = .0307 and P < .001, respectively). The slope of the line was unique to each patient (Figures 1 and 2). PIF and FIF50% generally decrease as the time from surgery increases, but a linear relationship could not be established (data not shown). Five of the 7 patients did have a second procedure but had not accumulated a sufficient number of subsequent PFTs during the study period for analysis. (The patients represented by a star and hexagon had only 1 procedure.)

Discussion

Not all patients are ideal candidates for endoscopic management of subglottic stenosis. Historically, previous failed dilations, stenosis length greater than 1 cm, circumferential stenosis, evidence of cartilage loss/damage, a history of severe bacterial infection with tracheotomy, posterior glottic stenosis with arytenoid fixation, and involvement of the inferior margin of the vocal folds were considered poor

Table 2. Comparison of Pre-dilation and Post-dilation Pulmonary Function Data.

PFT Parameter	Pre-dilation	Post-dilation	Median Change, Absolute	Median Change, %	*P* Value
FEV1, L/s	3.23	3.22	0.14	3.9	.177
FVC, L	2.4	2.57	0.25	10.0	.720
FEV1/FVC	0.74	0.77	0.5	5.8	.155
FEF25%-75%	2.05	2.88	0.41	21.2	.156
PEF, L/s	3.89	6.67	2.54	56.7	< .001
PIF, L/s	2.36	4.21	1.57	66.8	.001
PEF/PIF	1.66	1.67	0.13	8.4	.906
FEV1/PEF	0.82	0.39	0.443	56.0	.001
FIF50%, L/s	1.83	3.97	1.71	92.0	.001

Abbreviation: PFT, pulmonary function test.

Table 3. Influence of Dilation Size on Pulmonary Function Test (PFT) Parameters.

PFT Parameter	Dilation Size, mm	N	Median Change, Absolute	*P* Value
PEF, L/s	15.0	5	4.46	.497
	16.5	3	2.4	
	18.0	9	2.45	
PIF, L/s	15.0	5	2.47	.047
	16.5	3	1.35	
	18.0	9	1.51	
FEV1/PEF	15.0	5	–0.63	.441
	16.5	3	–0.62	
	18.0	9	–0.36	
FIF50%, L/s	15.0	4	2.13	.329
	16.5	3	1.38	
	18.0	9	1.58	

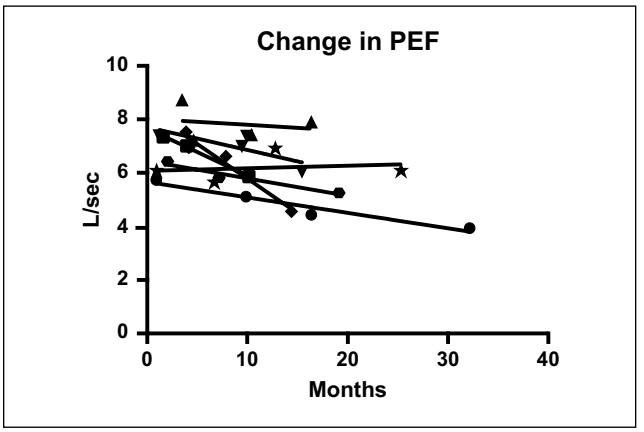

Figure 1. Change in PEF postoperatively. PEF declines in a linear fashion after intervention. The slope of the line is unique to each patient (*P* = .0307).

prognostic indicators.[23] Patients with isolated subglottic disease, however, can successfully be managed endoscopically in approximately 87% of cases.[24] The expectation is that more than 85% of these patients will have recurrence of their stenosis within 5 years, requiring a return to the operating suite.[25]

Clinically, the goal is to identify the techniques and adjuvant therapies that provide the greatest improvement in airflow and result in the slowest rate of restenosis. Not surprising, time between surgeries is a commonly reported metric used to demonstrate the efficacy of an intervention.[26,27] Although clinically relevant, confounding variables related to the surgical interval raise questions about its reliability for use in research. Surgeon availability certainly affects the time between procedures. Physically active patients are more likely to notice impairment of airflow than sedentary patients and may seek intervention sooner. Financial factors may affect a patient's decision to seek surgery, as well. It is clear that an objective measure is needed.

Using PFT data to evaluate airway stenosis is not a new concept. In the 1970s, a number of studies tried to identify which values or ratio of values could be used to diagnose UAO. Empey[18] reported on a series of 10 patients (most with bilateral vocal cord paralysis), noting that the FEV1/PEF ratio was greater than 10 in all cases and that the larger the ratio, the greater the degree of obstruction. In comparing UAO with chronic obstructive pulmonary disease, Rotman et al[19] identified 4 measurements that were found to differentiate upper from lower airway disease: (1) FEF50%/FIF50% > 1, (2) FEV1/PEFR > 10, (3) FIF50% < 100 L/min, and (4) FEV1/FEV0.5 > 1.5, with the latter 2 measures being less sensitive.

Pulmonary function tests have been used in a limited fashion to assess postoperative outcomes following tracheal resection[28] and endoscopic dilation,[29] but reports have been generally nonspecific about the degree of improvement and which measures are of greatest utility in quantifying results. To examine this issue, Wasserman et al[21] created a model of fixed obstruction using mouthpieces of decreasing inner

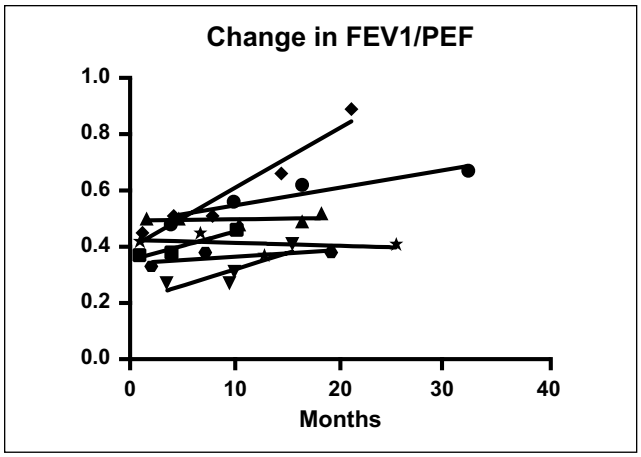

Figure 2. Change in FEV1/PEF postoperatively. FEV1/PEF similarly changes in a unique linear fashion after intervention ($P < .0001$).

diameters (15, 10, 8, and 6 mm) in line with a spirometer. PEF and PIF were found to be the most sensitive measures by which to assess changes in airway resistance. Nouraei et al[22] performed a similar experiment in 2007 in which the resistors were designed to generate a greater resistance to inspiratory flow than expiratory flow. The MEF_{50}/MIF_{50} and the ratio of the areas under the expiratory/inspiratory curves were the most sensitive and specific parameters.

Both the location of a stenosis and its characteristics (length, radius, boundary conditions) affect airflow, making every stenosis unique. In this study, we elected to specifically look at iSGS as a model of fixed obstruction. We chose to exclude glottic and tracheal stenosis from our analysis, both of which have some component of variability due to the contribution of Bernoulli forces and compression of membranous tracheal wall, respectively. Not surprising, we found that the PFT values that changed after intervention were the same as those reported by Wasserman et al. We suspect that an examination of PFT values affected by intervention for glottic or tracheal stenosis would yield results more closely resembling those recently reported by Nouraei and colleagues for a variable obstructive model.

In our evaluation, we examined the effect of radial incision and dilation on PFT values. Pulmonary function tests are an attractive means of evaluating stenosis patients because the study is inexpensive; widely available; and, unlike current techniques using CT for computational fluid dynamic studies, avoids the need for ionizing radiation. Whereas CT examines the structural component of stenosis, PFTs are a physiologic measure of the effect of stenosis on airflow and the patient's actual respiratory function. There are limitations to the PFT, the primary issue being that the study is dependent on patient effort. In addition, not all studies are conducted uniformly. Spirometry software is programmed to select the loop with the best expiratory effort for evaluation. Previous studies have demonstrated that the number reported by the computer for PIF and FIF50% did not represent the largest value in 50% and 69% of cases, respectively.[30,31] It is necessary to look at all loops or to establish a protocol for maximal inspiratory effort to mitigate this. Finally, some data suggest that the test may not be adequately sensitive for mild stenosis. Miller and Hyatt[32] reported that the trachea would have to be narrowed to 8 mm (~ 80%) prior to creating detectable changes in flow.

The retrospective nature of this study is an inherent limitation. As this is a single surgeon experience, the surgical technique was fairly consistent between patients. However, there was some variability in selecting the initial dilation size. The radius of the fluid conduit being one of the greatest influences on airflow has the potential to affect postoperative PFT values. Our data suggest that PIF may be influenced by dilation size. In addition, pulmonary function data were not collected at consistent time points during follow-up. Our data demonstrate that restenosis occurs in a linear manner that appears to be unique for each patient. It will be important in future studies to standardize the time and manner in which PFTs are administered in order to mitigate these confounding variables.

Conclusion

In the isolated subglottic stenosis model, PEF, PIF, FEV1/PEF, and FIF50% demonstrate a significant change after endoscopic intervention. The postoperative rate of change in the PEF and FEV1/PEF is linear but appears to be unique to each patient. PIF and FIF50% also change as restenosis occurs, but the lack of standardized protocols makes the rate of change more difficult to predict. These pilot data suggest that the change in PEF, FEV1/PEF, PIF, and FIF50% can be used to assess outcomes in the endoscopic management and treatment of iSGS. Furthermore, the rate of change in these values (PEF and FEV1/PEF, in particular) could potentially be used as an objective means of comparing the efficacy of different techniques or adjuvant therapies. Moving forward, prospective studies should focus on establishing uniform time points and protocols for collecting PFT data. Given the low incidence of the entity and the unique characteristics of each stenosis, garnering adequate numbers and long-term follow-up will likely require multi-institutional collaboration.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Brandenburg JH. Idiopathic subglottic stenosis. *Trans Am Acad Ophthalmol Otolaryngol*. 1972;76(5):1402-1406.
2. Mark EF, Meng F, Kradin RL. Idiopathic tracheal stenosis: a clinicopathologic study of 63 cases and comparison of the pathology with chondromalacia. *Am J Surg Path*. 2008;32(8):1138-1143.
3. Dedo HH, Catten MD. Idiopathic progressive subglottic stenosis: findings and treatment in 52 patients. *Ann Otol Rhinol Laryngol*. 2001;110(4):305-311.
4. Valdez TA, Shapsay SM. Idiopathic subglottic stenosis revisited. *Ann Otol Rhinol Laryngol*. 2008;111(8):690-695.
5. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope*. 1991;101(4, pt 2)(suppl 53):1-78.
6. Maldonado F, Loiselle A, DePew ZS, et al. Idiopathic subglottic stenosis: an evolving therapeutic algorithm. *Laryngoscope*. 2014;124(2):498-503.
7. Damrose EJ. On the development of idiopathic subglottic stenosis. *Med Hypoth*. 2008;71:122-125.
8. Dumoulin E, Stather DR, Gelfand G, Maranda B, Maceachern P, Tremblay A. Idiopathic subglottic stenosis: a familial predisposition. *Ann Thorac Surg*. 2013;95(3):1084-1086.
9. Ashiku SK, Kuzucu A, Grillo HC, et al. Idiopathic laryngotracheal stenosis: effective definitive treatment with laryngotracheal resection. *J Thorac Cardioavasc Surg*. 2004;127:99-107.
10. Shapshay SM, Beamis JF, Hybels RL, et al. Endoscopic treatment of subglottic and tracheal stenosis by radial laser incision and dilation. *Ann Otol Rhinol Laryngol*. 1987;96:661-664.
11. Parker NP, Bandyopadhyay D, Misono S, Goding GS Jr. Endoscopic cold incision, balloon dilation, mitomycin C application, and steroid injection for adult laryngotracheal stenosis. *Laryngoscope*. 2013;123(1):220-225.
12. Rutter MJ, Cohen AP, de Alarcon A. Endoscopic airway management in children. *Curr Opin Otolaryngol Head Neck Surg*. 2008;16(6):525-529.
13. Wicke C, Halliday B, Allen D, et al. Effects of steroids and retinoids on wound healing. *Arch Surg*. 2000;135(11):1265-1270.
14. Perepelitsyn I, Shapshay SM. Endoscopic treatment of laryngeal and tracheal stenosis—has mitomycin C improved the outcome? *Otolaryngol Head Neck Surg*. 2004;131(1):16-20.
15. Hata T, Hoshi T, Kanamori K, et al. Mitomycin, a new antibiotic from Streptomyces. I. *J Antibiot*. 1956;9(4):141-146.
16. McGaha TL, Phelps RG, Spiera H, Bona C. Halofuginone, an inhibitor of type-I collagen synthesis and skin sclerosis, blocks transforming-growth-factor-beta-mediated Smad3 activation in fibroblasts. *J Invest Dermatol*. 2002;118(3):461-470.
17. Cincik H, Gungor A, Cakmak A, et al. The effects of mitomycin C and 5-fluorouracil/triamcinolone on fibrosis/scar tissue formation secondary to subglottic trauma. *Am J Otolaryngol*. 2005;26(1):45-50.
18. Empey DW. Assessment of upper airways obstruction. *Br Med J*. 1972;3:503-505.
19. Rotman HH, Liss HP, Weg JG. Diagnosis of upper airway obstruction by pulmonary function testing. *Chest*. 1975;68(6):796-799.
20. Hyatt RE. Evaluation of major airway lesions using the flow volume loop. *Ann Otol Rhinol Laryngol*. 1975;84:635-642.
21. Wassermann K, Gitt A, Weyde J, Eckel HE. Lung function changes and exercise-induced ventilatory responses to external resistive loads in normal subjects. *Respiration*. 1995;62(4):177-184.
22. Nouraei SAR, Winterborne C, Nouraei M, et al. Quantifying the physiology of laryngotracheal stenosis: changes in pulmonary dynamics in response to graded extrathoracic resistive loading. *Laryngoscope*. 2007;117:581-589.
23. Simpson GT, Strong MS, Healy GB, et al. Predictive factors of success or failure in the endoscopic management of laryngeal and tracheal stenosis. *Ann Otol Rhinol Laryngol*. 1982;91:384-388.
24. Nouraei SAR, Sandhu GS. Outcome of a multimodality approach to the management of idiopathic subglottic stenosis. *Laryngoscope*. 2013;123:2474-2484.
25. Perotin JM, Jeanfaivre T, Thibout Y, et al. *Ann Thoracic Surg*. 2011;92:297-301.
26. Hseu AF, Benninger MS, Haffey TM, Lorenz R. Subglottic stenosis: a ten-year review of treatment outcomes. *Laryngoscope*. 2014;124(3):736-741.
27. Hirshoren N, Eliashar R. Wound-healing modulation in upper airway stenosis—myths and facts. *Head Neck*. 2009;31(1):111-126.
28. Farmer W, Littner MR, Gee BL. Assessment of therapy of upper airway obstruction. *Arch Intern Med*. 1977;137:309-312.
29. Roediger FC, Orloff LA, Courey MS. Adult subglottic stenosis: management with laser incisions and mitomycin-C. *Laryngoscope*. 2008;118(9):1542-1546.
30. Sterner JB, Morris MJ, Sill JM, Hayes JA. Inspiratory flow-volume curve evaluation for detecting upper airway disease. *Resp Care*. 2009;54(4):461-466.
31. Ewald FW, Tenholder MF, Waller R. Analysis of the inspiratory flow-volume curve: should it always precede the forced expiratory maneuver? *Chest*. 1994;106(3):814-818.
32. Miller RD, Hyatt RE. Obstructing lesions of the larynx and trachea: clinical and physiologic characteristics. *Mayo Clin Proc*. 1969;44(3):145-161.

Reprinted by permission of JAMA Otolaryngol Head Neck Surg. 2013; 139(1):76-81.

ORIGINAL ARTICLE

Clinical Manifestations and Treatment of Idiopathic and Wegener Granulomatosis–Associated Subglottic Stenosis

Stanford C. Taylor, BS; Daniel R. Clayburgh, MD, PhD; James T. Rosenbaum, MD; Joshua S. Schindler, MD

Objective: To compare and contrast the manifestations and surgical management of subglottic stenosis in patients with airway obstruction attributed to granulomatosis with polyangiitis (GPA), previously known as Wegener granulomatosis, and those with idiopathic subglottic stenosis (iSGS).

Design: Retrospective medical chart review. Review of subglottic stenosis cases seen in the otolaryngology department of an academic medical center from 2005 through 2010. Data were obtained on disease presentation, operative management. and findings.

Setting: Tertiary referral center.

Participants: A total of 24 patients with iSGS and 15 patients with GPA-associated subglottic stenosis (GPA-SGS).

Results: All individuals with iSGS were female, and 40% of patients with GPA-SGS were male ($P<.01$). Patients with iSGS tended to have a higher Myer-Cotton stenosis grade at the time of dilation than those with GPA-SGS ($P=.02$). Individuals with GPA-SGS were more likely to undergo tracheotomy as a result of disease-related complications than individuals with iSGS ($P<.01$). No patients with an open airway reconstruction in the iSGS group required follow-up mechanical dilation. In contrast, all patients with open airway reconstructions in the GPA-SGS group underwent more than 1 subsequent airway dilation ($P<.01$).

Conclusions: While surgical utilization is the mainstay of treatment in iSGS and GPA-SGS, iSGS occurs almost exclusively in females and presents with a greater degree of stenosis at the time of endoscopic dilation. In contrast, GPA-SGS is associated with greater rates of tracheotomy. Open airway reconstruction may be used in the treatment of iSGS and GPA-SGS but is much more effective in iSGS.

JAMA Otolaryngol Head Neck Surg. 2013;139(1):76-81

Author Affiliations: School of Medicine (Mr Taylor), Department of Otolaryngology–Head and Neck Surgery (Drs Clayburgh and Schindler), and Division of Arthritis and Rheumatic Diseases, Department of Medicine (Dr Rosenbaum), Oregon Health and Science University, Portland.

Acquired subglottic stenosis (SGS) describes airway narrowing in the area of the cricoid and is associated with prolonged intubation or external trauma. Other patients acquire SGS from a systemic autoimmune etiology, such as granulomatosis with polyangiitis (GPA), previously known as Wegener granulomatosis. However, in about 20% of cases there is no identifiable precipitant, and patients are therefore classified as having idiopathic subglottic stenosis (iSGS).[1] While posttraumatic SGS is often evident based on patient history, discriminating between iSGS and GPA-SGS can be diagnostically challenging, especially in cases in which SGS is the presenting symptom of GPA. Presenting symptoms of dyspnea, noisy breathing, and voice changes can occur in both groups.[2,3] Anatomically, the cohorts appear different, as scar formation in patients with GPA-SGS and iSGS typically involves soft tissue (membranous stenosis), but intubation and trauma-related SGS typically leads to firm, cartilaginous scar tissue.[4] GPA-SGS is diagnosed based on the presence of autoantibodies or characteristic findings on biopsy, but negative results cannot reliably rule out autoimmune-mediated SGS, especially when the disease has limited organ involvement.[5,6] Indeed, given the many similarities between patients with iSGS and those with GPA-SGS, some have suggested that iSGS may represent a spectrum of autoimmune-mediated SGS that is not yet fully understood and that GPA-SGS is but a part of this autoimmune SGS spectrum.

Despite the similarities noted between iSGS and GPA-SGS, there is a paucity of data in the literature directly comparing these entities, specifically with regard to therapeutic and surgical outcomes. Given

the systemic nature of GPA, it seems reasonable to assume that these patients would have more airway difficulty than those with idiopathic SGS. Furthermore, one would predict that patients with GPA would have a worse clinical response to surgical treatment, including a need for more procedures and more frequent failure of open procedure, although data to support this are lacking. To further clarify this, we undertook this study to compare the manifestations and surgical management for SGS in patients with airway obstruction attributed to GPA and those with iSGS.

METHODS

Following approval by the Oregon Health and Science University institutional review board, a retrospective review was performed of medical records of patients with GPA-SGS or iSGS seen in the otolaryngology department at our institution from 2005 through 2010. To identify patients, the department billing records and operative records were queried for all encounters associated with a diagnosis of GPA, laryngotracheal stenosis, or patients who had undergone endoscopic airway dilation or open airway reconstruction. Patients were classified as having GPA if (1) they had at least 1 clinical feature, such as SGS, consistent with the disease, and they were antineutrophil cytoplasmic antibody (ANCA) positive; (2) they had a biopsy finding consistent with GPA (granulomatous inflammation, vasculitis, and/or rapidly progressive glomerular nephritis); or (3) they manifested at least 2 signs of disease (laryngotracheal involvement, septal perforation, sinonasal involvement, nasolacrimal involvement, recurrent otitis media, or characteristic renal or pulmonary involvement). Patients were classified as having iSGS if they did not have a history of laryngotracheal trauma or tracheotomy and the airway narrowing could not be attributed to another cause, such as malignant disease or a systemic autoimmune condition.

Once patients were identified, all documentation, including pre-2005 encounters, was reviewed. Data were obtained on age at diagnosis, diagnostic procedures and laboratory tests, and therapeutic management, including immunosuppressive therapy and surgical procedures. SGS was diagnosed and evaluated by flexible fiber-optic examination or by intraoperative direct laryngoscopy. Extent of laryngotracheal involvement and gross characteristics of lesions were assessed. The Myer-Cotton staging system (MCS), which was originally developed as a pediatric SGS scale but has since been implemented in monitoring adult SGS, was used to describe the stenosis based on the percentage relative reduction in cross-sectional area of the subglottis. Four grades of stenosis are described: grade 1 lesions have less than 50% obstruction, grade 2 lesions have 51% to 70% obstruction, grade 3 lesions have 71% to 99% obstruction, and grade 4 lesions have no detectable lumen or complete stenosis.[7] Grade of stenosis was not documented in the operative record in 17 dilations. Airway dilations were performed with direct microlaryngoscopy using continuous radial expansion balloons (Boston Scientific), Jackson laryngeal dilators, or rigid bronchoscopic dilation. When comparing rates of surgical outcome, a minimum follow-up time of 6 months after the operative date was required for inclusion. Gastroesophageal reflux disease was diagnosed via esophagoscopy demonstrating esophagitis in 3 patients, pH probe testing in 5 patients, and clinical improvement of reflux symptoms with proton pump inhibitor therapy in 4 patients.

Descriptive statistics, *t* test, Fisher exact test, Mann-Whitney U test, and χ^2 analysis for categorical data were performed ($P < .05$ denoted significance).

Table 1. Background Information

Characteristic	GPA (n = 15)	Idiopathic (n = 24)	*P* Value
Sex			
Female	9	24	<.01
Male	6	0	
Ethnicity			
Hispanic	0	1	>.99
White, non-Hispanic	15	23	
Comorbidities			
GERD	4	12	.19
Asthma	1	2	>.99
Lifetime tracheotomy history	6	0	<.01
Age, median, y			
At GPA diagnosis	31.7	NA	
At SGS diagnosis	36.3	45.2	.24

Abbreviations: GERD, gastroesophageal reflux disease; GPA, granulomatosis with polyangiitis; NA, not applicable; SGS, subglottic stenosis.

RESULTS

A total of 39 patients were identified for the study, 24 with iSGS and 15 with GPA-associated subglottic airway obstruction (**Table 1**). Aside from the absence of males in the iSGS group, no significant differences in patient demographics were noted. Of note, 4 patients with GPA-SGS in our cohort (27%) were diagnosed as having GPA when younger than 20 years. At the date of last follow-up, 7 patients with GPA-SGS (47%) exhibited disease involvement restricted to the head and neck while 8 (53%) had systemic involvement, including renal and/or pulmonary manifestations. The cohort was followed for 177 patient-years. The mean and median periods of follow-up for the GPA-SGS group were 8.2 and 9.9 years, respectively. In comparison, the mean and median lengths of follow-up for the iSGS group were 2.8 and 1.8 years, respectively ($P < .01$).

Diagnosis of the 15 patients with GPA is illustrated as follows:

Patients who are ANCA positive, %	93
Patients with biopsy proven diagnosis, %	47
Patients who are ANCA positive and with diagnostic biopsy, No.	7
Patients with ANCA and nondiagnostic biopsy, No.	7
Patients diagnosed by clinical features alone, No.	1

The patient diagnosed by clinical features alone was male and, as is typical with GPA,[8] had additional disease involving the nose and sinuses. Furthermore, although his autoantibody ANCA titers were not positive, they were interpreted as having an atypical pattern. Given these considerations, his SGS was attributed to GPA rather than an idiopathic etiology.

The severity and location of stenosis observed during dilation was assessed. Patients with iSGS were found to have significantly worse stenosis based on MCS grading than patients with GPA-SGS (**Table 2**). There were no significant differences in the location of stenosis seen at initial dilation, although there was a trend toward more circumferential stenoses in the patients with GPA (**Table 3**).

©2013 American Medical Association. All rights reserved.

Table 2. Myer-Cotton Staging (MCS) at Time of Endoscopic Dilation

Characteristic	No. (%)		
	GPA	Idiopathic	Combined
Patients undergoing endoscopic dilation. No.	14	23	38
Dilations with known MCS, No.	36	50	86
MCS[a]			
1	8 (22.2)	2 (4)	10 (11.6)
2	16 (44.4)	14 (28)	30 (34.9)
3	12 (33.3)	34 (68)	46 (53.5)
4	0	0	0

Abbreviation: GPA, granulomatosis with polyangiitis.
[a] $P = .02$.

Table 3. Stenosis Location at Time of First Endoscopic Dilation

Characteristic	No. (%)		
	GPA	Idiopathic	Combined
Patients undergoing endoscopic dilation, No.	14	23	38
Dilations with known stenosis location at time of first endoscopic dilation, No.	11	19	30
Stenosis location[a]			
Anterior	1 (9.1)	8 (42.1)	9 (30)
Posterior	1 (9.1)	2 (10.5)	3 (10)
Circumferential	9 (81.2)	9 (47.4)	18 (60)

Abbreviation: GPA, granulomatosis with polyangiitis.
[a] $P = .42$.

Surgical treatment of SGS consisted of both endoscopic dilation and cricotracheal resection. Endoscopic dilation technique did not vary substantially between groups, aside from less frequent use of the carbon dioxide laser in the GPA group (**Table 4**). Patients with GPA underwent a mean of 3.53 surgical dilations per patient compared with 2.54 in those with iSGS ($P = .44$). Seven patients with GPA-SGS (47%) required fewer than 2 airway dilations compared with 11 of those with iSGS (46%) ($P < .99$).

Definitive operative resection or reconstruction was attempted in both groups; 5 of those with GPA-SGS (33%), and 6 of those with iSGS (25%). While no patients with an open airway reconstruction in the iSGS group required follow-up mechanical dilation, all patients with open airway reconstructions in the GPA-SGS group underwent more than 1 subsequent airway dilation ($P < .01$). Following open airway reconstruction, 1 patient with GPA-SGS underwent subsequent tracheotomy. Open airway reconstruction led to permanent decannulation of 2 previously tracheotomy-dependent patients with GPA-SGS. Six patients with GPA-SGS (40%) underwent tracheotomy as a result of disease-related complications and 2 (13%) remained tracheotomy dependent at the date of last follow-up. No patients with iSGS required tracheotomy as a result of a disease-related complications ($P < .01$).

We examined the impact of various factors on the success of airway procedures within each group (**Table 5** and **Table 6**). The presence of gastroesophageal reflux disease (GERD) and the operative use of carbon dioxide laser were not found to have an impact on the rate of surgical utilization. Male patients with GPA-SGS had a shorter time until additional procedures were needed than female patients with GPA-SGS, while the presence of a previous tracheostomy showed a nonsignificant trend toward worse outcomes. When patients from both cohorts (GPA-SGS and iSGS) were pooled into a single group, none of these factors (sex and history of tracheostomy or GERD) had any significant impact on time until additional procedures were needed.

We analyzed the utilization of systemic immunotherapy within both groups. All patients with GPA-SGS and 50% of iSGS individuals received systemic immunotherapy at some point throughout follow-up as part of disease management; all patients with iSGS received corticosteroids, while patients with GPA received a mix of corticosteroids (n = 13), methotrexate sodium (n = 11), and cyclophosphamide (n = 9). Use of immunosuppressive medication was not associated with longer procedure-free intervals.

COMMENT

When traumatic causes are not readily identifiable by patient history, determining the etiology of SGS can be diagnostically challenging. While many patients with nontraumatic SGS may have a systemic autoimmune condition such as GPA, many others will have an unrevealing autoimmune workup. In the absence of any identifiable cause, these patients are considered to have iSGS, although there is some speculation that this may be due to some unknown autoimmune mechanism. This study was conducted to better define the similarities and differences in presentation and therapeutic management of iSGS and GPA-SGS.

Previously, it has been demonstrated that GPA-SGS affects men and women equally.[2] In contrast, iSGS almost exclusively affects women[3,9,10]; it is thought to predominantly affect women owing to estrogen-mediated alterations to wound-healing responses in the subglottic airway.[11] Our study is consistent with these observations, further confirming the tendency for iSGS to disproportionately affect men and for GPA-SGS to affect both men and women. Interestingly, we observed that male GPA-SGS patients underwent more frequent subglottic airway surgical procedures than female patients with GPA-SGS. This may be due, in part, to previous trends noting that male patients with GPA tend to develop a more severe form of the disease than female patients with GPA.[12]

The median age of initial presentation in patients with GPA-SGS (36.3 years) was almost 9 years younger than that of the iSGS group (45.2 years). Although not statistically significant ($P = .24$), this finding is broadly in line with that of previous research, and we speculate that had our cohort been larger, the observations would have been significant. A study of 52 patients with iSGS found the average age of initial presentation to be 43.5 years,[9]

©2013 American Medical Association. All rights reserved.

Table 4. Operative Technique at Time of Endoscopic Dilation

Characteristic, No.	GPA	Idiopathic	Combined	*P* Value
Patients undergoing endoscopic dilation	14	23	38	NA
Total dilations	48	55	103	NA
Operative use of laser	15	35	50	<.01
Intraoperative injections				
No injection	10	5	15	.10
Triamcinolone	31	39	70	.53
Mitomycin C	1	1	2	>.99
Triamcinolone and mitomycin C	6	10	16	.59

Abbreviations: GPA, granulomatosis with polyangiitis; NA, not applicable.

Table 5. Therapeutic Airway Procedure Frequency[a]

	Granulomatosis With Polyangiitis, No.					Idiopathic, No.				
Characteristic	Patients	Procedures	*P* Value	DBP, Mean	*P* Value	Patients	Procedures	*P* Value	DBP, Mean	*P* Value
All patients	15	48		557		24	46		495	NA
GERD history										
Yes	4	9	.74	565	.96	12	32	.35	506	.60
No	11	39	.74	555	.96	12	14		470	
Lifetime tracheotomy history										
Yes	6	33	.24	367	.11	NA	NA	NA	NA	NA
No	9	15	.24	975	.11	NA	NA	NA	NA	NA
Sex										
Male	6	31	.24	318	.04	NA	NA	NA	NA	NA
Female	9	17	.24	994	.04	NA	NA	NA	NA	NA

Abbreviations: DBP, days between procedures; GERD, gastroesophageal reflux disease; NA, not applicable.
[a] Excludes surgical procedures with less than 6 months' postoperative follow-up.

whereas the median age at which SGS was diagnosed in a cohort of patients with GPA was 26 years.[2] Furthermore, patients with GPA and SGS are frequently diagnosed as having GPA at a very young age; in fact, up to 44% are diagnosed before the age of 20 years.[13] 27% of patients with GPA-SGS in our cohort were diagnosed as having GPA when younger than 20 years.

GERD has been implicated in the development of SGS and has been identified as a probable precipitant of iSGS.[14-17] However, some question the existence of a direct association.[9] GERD has also been explored as a possible cause of GPA-SGS, but evidence of a definitive link has yet to be identified.[18] The most compelling data to date come from a study by Blumin and Johnston[19] demonstrating pepsin in the larynx and trachea in 59% of patients with iSGS, but none in matched control patients. Half of our iSGS group either had a history of, or was empirically treated for, GERD, which was statistically no different from the comparison GPA-SGS group. Furthermore, the rate of surgical utilization between those with a diagnosis of GERD and those without was no different in both groups. While our results fail to demonstrate a difference in the rate of GERD and SGS in the iSGS and GPA-SGS groups, understanding the impact of GERD on the development of SGS will be best accomplished through continued prospective studies.

Operative management strategies for subglottic stenosis are focused on improving the airway, either via endoscopic dilation of the stenosis, excision of the stenosis with laryngotracheal reconstruction, or bypassing the stenosis with tracheostomy. Carbon dioxide laser resection and/or intralesional corticosteroid injection are common adjuvant treatments to endoscopic dilation. Interestingly, in our series we found that laser resection was utilized more frequently in patients with iSGS than in those with GPA-SGS. This may in part be explained by practices of the operating surgeon or a reluctance to use the carbon dioxide laser if there is a possibility of active

Table 6. Endoscopic Dilation Frequency[a] Based on Myer-Cotton Staging (MCS) at Time of Endoscopic Dilation

Characteristic	GPA	Idiopathic	*P* Value
Patients undergoing endoscopic dilation, No.	14	23	NA
Dilations with known MCS	36	50	NA
Days between procedures based on MCS, mean, No.			
1	829	358	NA[b]
2	562	602	.03
3	462	477	.23
4	NA	NA	NA

Abbreviations: GPA, granulomatosis with polyangiitis; NA, not applicable.
[a] Excludes surgical procedures with less than 6 months' postoperative follow-up.
[b] Only 1 dilation in the idiopathic MCS 1 group with more than 6 months' postoperative follow-up.

GPA within the stenosis. While previous studies have shown the successful use of the carbon dioxide laser for GPA-SGS,[20] a general principle in the treatment of GPA-SGS is to avoid as much airway manipulation as possible when active disease is present.

Intralesional corticosteroid injections at the time of manual dilation were documented in 85% of cases. Previous research of an intratracheal dilation-injection technique using glucocorticoids in GPA-SGS has shown this to be effective[21] and possibly a preferred method of immunosuppressive therapy in GPA isolated strictly to the subglottis.[2] While the role of intralesional corticosteroids in iSGS is less clear, it is often considered an adjunct to dilation to prolong the time between procedures. Definitive treatment of iSGS is thought to be most likely achieved with open airway reconstruction.[3,10] More recently, mitomycin C has also been used as an inhibitor of fibroblastic-mediated scar formation in laryngotracheal stenosis.[11,22,23] In our cohort, only 2 patients received mitomycin C; thus, it is impossible to derive any conclusions about this therapy. Future research will be needed to better define its role in the management of GPA-SGS and iSGS.

Our data demonstrated more severe stenosis as measured by MCS at the time of dilation in patients with iSGS than those with GPA-SGS, with 33% of dilations in patients with GPA-SGS and 68% of dilations in patients with iSGS classified as MCS 3. Although there is a perceived reluctance to operate on patients with GPA, our experience indicates they often undergo dilation for smaller degrees of stenosis. This may indicate that patients with GPA-SGS experience more clinically significant symptoms than those with iSGS for a given grade of stenosis. Alternatively, underlying sinonasal or pulmonary involvement with consequent increased work of breathing in individuals with GPA-SGS could explain the larger diameter airway at the time of dilation. Patients with GPA-SGS may also have longer or more irregular sections of stenosis that result in more turbulence and poorer airflow than those with iSGS with comparatively discrete and symmetric stenoses. Further research will be needed to explore the differences between the dyspnea in these 2 groups.

Although the MCS at the time of dilation was different between the 2 groups, we found identical percentages of patients with GPA-SGS and iSGS undergoing more than 1 endoscopic dilation and open airway reconstruction. Our data further indicate that iSGS cases classified as having MCS 2 have more days between mechanical airway dilations than those with GPA-SGS. A similar trend, albeit not significant, is noted in average days between dilations in patients classified as having MCS 3.

The percentage of patients with GPA-SGS undergoing open airway reconstruction in our group is similar to those of other studies,[13,20] and we recently reported on the efficacy of airway reconstruction in GPA-related laryngotracheal stenosis.[24] It is important to note that open airway reconstruction was much more effective for iSGS than for GPA-SGS. In patients with iSGS, open airway reconstruction could be considered definitive management, with no need for tracheostomy afterward and rare need for further airway interventions. However, in GPA-SGS, further dilation is the norm; as we have shown previously, the major benefit of open airway reconstruction for GPA-SGS is to effect decannulation.[24] Forty percent of patients with GPA-SGS in our cohort required tracheotomy as part of disease treatment. This is consistent with other research demonstrating that between 41% and 52% of patients with GPA-SGS require tracheotomy.[2,13,20] No patients with iSGS in our cohort required tracheotomy owing to disease-related complications, which is less than a previous study showing a 20% tracheotomy rate.[25]

Owing to the nature of retrospective medical chart reviews, our study has several inherent limitations. While specialists at tertiary referral centers follow patients with chronic medical conditions longitudinally for many years, the same is not true of some conditions, such as SGS, that may resolve after 1 or more treatments. The mean length of follow-up for patients with iSGS at our institution was 2.8 years compared with 8.2 years in those with GPA-SGS. This discrepancy, although informational for comparing disease chronicity between groups, does not allow for an accurate comparison of rate of surgical dilations over time. Thus, we viewed the observation that patients with GPA-SGS undergo less frequent surgical utilization as being due to loss of follow-up and attributable to the nature of retrospective reviews involving tertiary referral centers. It should be noted we also explored the possibility of systemic immunosuppressive therapy, which was used by all patients with GPA-SGS and half of those with iSGS, as an additional factor contributing to the decreased rate of surgical utilization in the GPA-SGS group. However, owing to a lack of numbers and prescribing variability between patients and procedures, the effect of immunosuppressive therapy on time between airway procedures could not be accurately assessed. The true rate of surgical utilization and systemic immunosuppressive therapy efficacy in these patients will be best determined by future prospective studies. In further considering the operative demand of both diseases, it is important to note that the median number of airway procedures in both the GPA-SGS and iSGS groups was 1 surgical intervention per patient. Thus, while both groups had individuals requiring chronic follow-up and multiple airway procedures, 1 operation provided definitive treatment for many of the patients.

When individually considering the duration of follow-up of the GPA-SGS group, the finding of an average of 0.47 surgical dilations per patient-year of follow-up could be viewed as representative for patients with GPA-SGS requiring referral to a head and neck surgeon. A similarly constructed previous study that followed patients for an average of 6.4 years demonstrated a comparable rate of surgical utilization in those with GPA-SGS: 0.36 surgical procedures per patient-year of follow-up.[20]

In conclusion, although several similarities exist between GPA-SGS and iSGS, iSGS occurs more often in women and presents with a greater degree of stenosis. GPA-SGS requires more long-term management and is associated with a higher rate of tracheotomy. While open airway reconstruction may be used in both iSGS and GPA-SGS, it is much more effective in iSGS. Within the GPA-SGS group, the rate of surgical utilization in individuals with GPA-SGS requiring tracheotomy was not significantly different from those whose disease did not re-

quire tracheotomy and male patients with GPA-SGS required more frequent subglottic airway procedures than female patients with GPA-SGS.

Submitted for Publication: July 13, 2012; final revision received September 27, 2012; accepted October 15, 2012.
Correspondence: Joshua S. Schindler, MD, Department of Otolaryngology–Head and Neck Surgery, Oregon Health and Science University, 3181 SW Sam Jackson Park Rd, Mail Code PV-01, Portland, OR 97239 (schindlj @ohsu.edu).
Author Contributions: All authors had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Taylor, Clayburgh, Rosenbaum, and Schindler. *Acquisition of data:* Taylor, Clayburgh, Rosenbaum, and Schindler. *Analysis and interpretation of data:* Taylor, Clayburgh, Rosenbaum, and Schindler. *Drafting of the manuscript:* Taylor and Clayburgh. *Critical revision of the manuscript for important intellectual content:* Clayburgh, Rosenbaum, and Schindler. *Statistical analysis:* Taylor and Clayburgh. *Study supervision:* Rosenbaum and Schindler.
Conflict of Interest Disclosures: Dr Rosenbaum is a paid consultant for Genentech.
Funding/Support: Research by Dr Rosenbaum is, in part, funded by Research to Prevent Blindness (New York), Genentech, the Stan and Madelle Rosenfeld Family Trust, the William and Mary Bauman Foundation, and the William C. Kuzell Foundation.
Previous Presentation: This study was presented at the Annual American Broncho-Esophagological Association Meeting; April 19, 2012; San Diego, California.
Additional Contributions: Jess Mace, MPH, provided assistance with the statistical analysis.

REFERENCES

1. Herrington HC, Weber SM, Andersen PE. Modern management of laryngotracheal stenosis. *Laryngoscope*. 2006;116(9):1553-1557.
2. Langford CA, Sneller MC, Hallahan CW, et al. Clinical features and therapeutic management of subglottic stenosis in patients with Wegener's granulomatosis. *Arthritis Rheum*. 1996;39(10):1754-1760.
3. Grillo HC, Mark EJ, Mathisen DJ, Wain JC. Idiopathic laryngotracheal stenosis and its management. *Ann Thorac Surg*. 1993;56(1):80-87.
4. Roediger FC, Orloff LA, Courey MS. Adult subglottic stenosis: management with laser incisions and mitomycin-C. *Laryngoscope*. 2008;118(9):1542-1546.
5. Finkielman JD, Lee AS, Hummel AM, et al; WGET Research Group. ANCA are detectable in nearly all patients with active severe Wegener's granulomatosis. *Am J Med*. 2007;120(7):643-, e9-e14.
6. Devaney KO, Travis WD, Hoffman G, Leavitt R, Lebovics R, Fauci AS. Interpretation of head and neck biopsies in Wegener's granulomatosis: a pathologic study of 126 biopsies in 70 patients. *Am J Surg Pathol*. 1990;14(6):555-564.
7. Myer CM III, O'Connor DM, Cotton RT. Proposed grading system for subglottic stenosis based on endotracheal tube sizes. *Ann Otol Rhinol Laryngol*. 1994; 103(4, pt 1):319-323.
8. Hoffman GS, Kerr GS, Leavitt RY, et al. Wegener granulomatosis: an analysis of 158 patients. *Ann Intern Med*. 1992;116(6):488-498.
9. Dedo HH, Catten MD. Idiopathic progressive subglottic stenosis: findings and treatment in 52 patients. *Ann Otol Rhinol Laryngol*. 2001;110(4):305-311.
10. Ashiku SK, Kuzucu A, Grillo HC, et al. Idiopathic laryngotracheal stenosis: effective definitive treatment with laryngotracheal resection. *J Thorac Cardiovasc Surg*. 2004;127(1):99-107.
11. Valdez TA, Shapshay SM. Idiopathic subglottic stenosis revisited. *Ann Otol Rhinol Laryngol*. 2002;111(8):690-695.
12. Stone JH; Wegener's Granulomatosis Etanercept Trial Research Group. Limited versus severe Wegener's granulomatosis: baseline data on patients in the Wegener's granulomatosis etanercept trial. *Arthritis Rheum*. 2003;48(8):2299-2309.
13. Lebovics RS, Hoffman GS, Leavitt RY, et al. The management of subglottic stenosis in patients with Wegener's granulomatosis. *Laryngoscope*. 1992;102 (12, pt 1):1341-1345.
14. Jindal JR, Milbrath MM, Shaker R, Hogan WJ, Toohill RJ. Gastroesophageal reflux disease as a likely cause of "idiopathic" subglottic stenosis. *Ann Otol Rhinol Laryngol*. 1994;103(3):186-191.
15. Maronian NC, Azadeh H, Waugh P, Hillel A. Association of laryngopharyngeal reflux disease and subglottic stenosis. *Ann Otol Rhinol Laryngol*. 2001;110 (7, pt 1):606-612.
16. Poetker DM, Ettema SL, Blumin JH, Toohill RJ, Merati AL. Association of airway abnormalities and risk factors in 37 subglottic stenosis patients. *Otolaryngol Head Neck Surg*. 2006;135(3):434-437.
17. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope*. 1991;101(4, pt 2)(suppl 53):1-78.
18. Church AC, Goldsmith K, Sivasothy P. Aspiration and development of subglottic stenosis in patients with Wegener's granulomatosis. *J Laryngol Otol*. 2010; 124(4):393-396.
19. Blumin JH, Johnston N. Evidence of extraesophageal reflux in idiopathic subglottic stenosis. *Laryngoscope*. 2011;121(6):1266-1273.
20. Gluth MB, Shinners PA, Kasperbauer JL. Subglottic stenosis associated with Wegener's granulomatosis. *Laryngoscope*. 2003;113(8):1304-1307.
21. Hoffman GS, Thomas-Golbanov CK, Chan J, Akst LM, Eliachar I. Treatment of subglottic stenosis, due to Wegener's granulomatosis, with intralesional corticosteroids and dilation. *J Rheumatol*. 2003;30(5):1017-1021.
22. Smith ME, Elstad M. Mitomycin C and the endoscopic treatment of laryngotracheal stenosis: are two applications better than one? *Laryngoscope*. 2009; 119(2):272-283.
23. Perepelitsyn I, Shapshay SM. Endoscopic treatment of laryngeal and tracheal stenosis: has mitomycin C improved the outcome? *Otolaryngol Head Neck Surg*. 2004;131(1):16-20.
24. Wester JL, Clayburgh DR, Stott WJ, Schindler JS, Andersen PE, Gross ND. Airway reconstruction in Wegener's granulomatosis-associated laryngotracheal stenosis. *Laryngoscope*. 2011;121(12):2566-2571.
25. Giudice M, Piazza C, Foccoli P, Toninelli C, Cavaliere S, Peretti G. Idiopathic subglottic stenosis: management by endoscopic and open-neck surgery in a series of 30 patients. *Eur Arch Otorhinolaryngol*. 2003;260(5):235-238.

Hsc

Home Study Course

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Empowering otolaryngologist—head and neck surgeons to deliver the best patient care

1650 Diagonal Road, Alexandria, Virginia 22314-2857 U.S.A.
www.entnet.org

AcademyU® Your Otolaryngology Education Source | www.entnet.org/hsc

Laryngology, Voice Disorders, and Bronchoesophagology

Hsc
Home Study Course

Home Study Course

Section 6
November 2016

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

© 2016 American Academy of Otolaryngology—Head and Neck Surgery Foundation
Empowering otolaryngologist-head and neck surgeons to deliver the best patient care

THE
HOME STUDY COURSE
IN
OTOLARYNGOLOGY — HEAD AND NECK SURGERY

SECTION 6

Laryngology, Voice Disorders and Bronchoesophagology

November 2016

SECTION FACULTY:

Thomas L. Carroll, MD**
Brad W. deSilva, MD**
Priya Krishna, MD, MS**
C. Kwang Sung, MD, MS**
Lee M. Akst, MD
Maggie A. Kuhn, MD
Ashli K. O'Rourke, MD, MS
David E. Rosow, MD

American Academy of Otolaryngology - Head and Neck Surgery Foundation
1650 Diagonal Road, Alexandria, VA 22314

Section 6 suggested exam deadline: January 3, 2017
Expiration Date: August 4, 2017; CME credit not available after that date

Introduction
The Home Study Course is designed to provide relevant and timely clinical information for physicians in training and current practitioners in otolaryngology - head and neck surgery. The course, spanning four sections, allows participants the opportunity to explore current and cutting edge perspectives within each of the core specialty areas of otolaryngology.

The **Selected Recent Material** represents primary fundamentals, evidence-based research, and state of the art technologies in Laryngology, Voice Disorders and Bronchoesophagology. The scientific literature included in this activity forms the basis of the assessment examination.

The number and length of articles selected are limited by editorial production schedules and copyright permission issues, and should not be considered an exhaustive compilation of knowledge Laryngology, Voice Disorders and Bronchoesophagology.

The **Additional Reference Material** is provided as an educational supplement to guide individual learning. This material is not included in the course examination and reprints are not provided.

Needs Assessment
AAO-HNSF's education activities are designed to improve healthcare provider competence through lifelong learning. The Foundation focuses its education activities on the needs of providers within the specialized scope of practice of otolaryngologists. Emphasis is placed on practice gaps and education needs identified within eight subspecialties. The *Home Study Course* selects content that addresses these gaps and needs within all subspecialties.

Target Audience
The primary audience for this activity is physicians and physicians-in-training who specialize in otolaryngology-head and neck surgery.

Outcomes Objectives

1. Evaluate the utility of computed tomography in the evaluation of patients with idiopathic vocal fold paresis.
2. Communicate the current practice patterns for otolaryngologists in diagnosing unilateral vocal fold paresis.
3. Discuss the various etiologies of unilateral vocal fold paralysis and concurrent dysphagia findings that these patients exhibit.
4. Evaluate the utility of stroboscopy in evaluating patients with laryngeal dysplasia, vocal fold paresis, and vocal fold paralysis.
5. Articulate the effectiveness of photoangiolytic laser treatment for Reinke's edema and expected voice outcomes following treatment.
6. Determine the benefits of office-based biopsy of laryngeal lesions versus surgical intervention for these pathologies.
7. Describe the anatomic changes that occur with vocal fold atrophy in the aging larynx and expected voice outcomes following voice therapy for this disorder.
8. Evaluate the utility of impedance testing in patients with extraesophageal reflux symptoms.
9. Measure body mass index and communicate how it predicts tracheal size.
10. Recognize the outcomes of tracheal stenosis and other airway complications in trauma patients that undergo percutaneous versus open tracheostomy.
11. Measure pulmonary function testing in patients that have undergone endoscopic treatment for subglottic stenosis.

Medium Used
The Home Study Course is available in electronic or print format. The activity includes a review of outcomes objectives, selected scientific literature, and a self-assessment examination.

Method of Physician Participation in the Learning Process
The physician learner will read the selected scientific literature, reflect on what they have read, and complete the self-assessment exam. After completing this section, participants should have a greater understanding of Laryngology, Voice Disorders and Bronchoesophagology as they affect the head and neck area, as well as useful information for clinical application.

Estimated time to complete this activity: 40.0 hours

Accreditation Statement
The American Academy of Otolaryngology—Head and Neck Surgery Foundation (AAO-HNSF) is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

Credit Designation
The AAO-HNSF designates this enduring material for a maximum of 40.0 *AMA PRA Category 1 Credit(s)*™. Physicians should claim credit commensurate with the extent of their participation in the activity.

ALL PARTICIPANTS must achieve a post-test score of 70% or higher for a passing completions to be recorded and a transcript to be produced. Residents' results will be provided to the Training Program Director.

PHYSICIANS ONLY: In order to receive *Credit* for this activity **a post-test score of 70% or higher is required**. Two retest opportunities will automatically be available if a minimum of 70% is not achieved.

Disclosure
The American Academy of Otolaryngology Head and Neck Surgery/Foundation (AAO-HNS/F) supports fair and unbiased participation of our volunteers in Academy/Foundation activities. All individuals who may be in a position to control an activity's content must disclose all relevant financial relationships or disclose that no relevant financial relationships exist. All relevant financial relationships with commercial interests[1] that directly impact and/or might conflict with Academy/Foundation activities must be disclosed. Any real or potential conflicts of interest[2] must be identified, managed, and disclosed to the learners. In addition, disclosure must be made of presentations on drugs or devices, or uses of drugs or devices that have not been approved by the Food and Drug Administration. This policy is intended to openly identify any potential conflict so that participants in an activity are able to form their own judgments about the presentation.

[1] A "Commercial interest" is any entity producing, marketing, re-selling, or distributing health care goods or services consumed by, or used on, patients.

[2] "Conflict of interest" is defined as any real or potential situation that has competing professional or personal interests that would make it difficult to be unbiased. Conflicts of interest occur when an individual has an opportunity to affect education content about products or services of a commercial interest with which they have a financial relationship. **A conflict of interest depends on the situation and not on the character of the individual.**

2016 SECTION 6 LARYNGOLOGY, VOICE DISORDERS AND BRONCHOESOPHAGOLOGY FACULTY

Co-Chairs:

Thomas L. Carroll, MD, Director, The Voice Program at Brigham and Women's Hospital, Instructor, Department of Otolaryngology, Harvard Medical School, Boston, Massachusetts.
Disclosure: Consultant: Pentax Medical.

Brad W. deSilva, MD, Assistant Professor, Residency Program Director, Voice and Swallow Disorders Clinic. Ohio State University Wexner Medical Center, Columbus, Ohio.
Disclosure: No relationships to disclose.

Priya Krishna, MD, MS, Co-Director, Loma Linda University Center for Voice and Swallowing, Assistant Professor, Department of Otolaryngology, Loma Linda University, Loma Linda, California.
Disclosure: Consultant: Merz North America, Olympus Medical.

C. Kwang Sung, MD, MS, Assistant Professor**,** Department of Otolaryngology-Head and Neck Surgery, Stanford University School of Medicine, Stanford, California.
Disclosure: Consultant: Olympus Medical; Expert Witness: Girardi – Keese.

Faculty:

Lee M. Akst, MD, Director, Division of Laryngology and Assistant Professor, Department of Otolaryngology, Johns Hopkins Department of Otolaryngology-Head and Neck Surgery, Baltimore, Maryland.
Disclosure: Consultant: Merz Pharmaceuticals and Olympus.

Maggie A. Kuhn, MD, Assistant Professor, University of California, Davis, Department of Otolaryngology – Head & Neck Surgery, Sacramento, California
Disclosure: No relationships to disclose.

Ashli K. O'Rourke, MD, MS, Assistant Professor, Evelyn Trammell Institute for Voice and Swallowing, Department Otolaryngology – Head and Neck Surgery, Medical University of South Carolina, Charleston, South Carolina.
Disclosure: Other: Carolina Speech Services; Consultant: Medtronic.

David E. Rosow, MD, Director, Division of Laryngology and Voice, Assistant Professor, Department of Otolaryngology, University of Miami Miller School of Medicine, Miami, Florida. Assistant Professor, Department of Vocal Performance, University of Miami Frost School of Music, Coral Gables, Florida.
Disclosure: No relationships to disclose.

Planner(s):

Linda Lee, AAO─HNSF Education Program Manager	No relationships to disclose
Stephanie Wilson, Stephanie Wilson Consulting, LLC; Production Manager	No relationships to disclose
Richard V. Smith, MD, chair, AAO-HNSF Education Steering Committee	Expert Witness: various legal firms
Catherine R. Lintzenich, MD, chair, AAO-HNSF Laryngology & Bronchoesophagology Education Committee	No relationships to disclose

This 2016 Section 6 Home Study Course does not include discussion of off-label uses of drugs or devices which have not been approved by the United States Food and Drug Administration:

Disclaimer

The information contained in this activity represents the views of those who created it and does not necessarily represent the official view or recommendations of the American Academy of Otolaryngology – Head and Neck Surgery Foundation.

January 3, 2017: Suggested Section 6 Exam submission deadline; course closed August 4, 2017.

EVIDENCE BASED MEDICINE

The AAO-HNSF Education Advisory Committee approved the assignment of the appropriate level of evidence to support each clinical and/or scientific journal reference used to authenticate a continuing medical education activity. Noted at the end of each reference, the level of evidence is displayed in this format: **[EBM Level 3]**.

Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001)	
Level 1	Randomized[1] controlled trials[2] or a systematic review[3] (meta-analysis[4]) of randomized controlled trials[5].
Level 2	Prospective (cohort[6] or outcomes) study[7] with an internal control group or a systematic review of prospective, controlled trials.
Level 3	Retrospective (case-control[8]) study[9] with an internal control group or a systematic review of retrospective, controlled trials.
Level 4	Case series[10] without an internal control group (retrospective reviews; uncontrolled cohort or outcome studies).
Level 5	Expert opinion without explicit critical appraisal, or recommendation based on physiology/bench research.

Two *additional ratings* to be used for articles that do not fall into the above scale. Articles that are informational only can be rated N/A , and articles that are a review of an article can be rated as Review. All definitions adapted from Glossary of Terms, Evidence Based Emergency Medicine at New York Academy of Medicine at www.ebem.org.

[1] A technique which gives every patient an equal chance of being assigned to any particular arm of a controlled clinical trial.
[2] Any study which compares two groups by virtue of different therapies or exposures fulfills this definition.
[3] A formal review of a focused clinical question based on a comprehensive search strategy and structure critical appraisal.
[4] A review of a focused clinical question following rigorous methodological criteria and employing statistical techniques to combine data from independently performed studies on that question.
[5] A controlled clinical trial in which the study groups are created through randomizations.
[6] This design follows a group of patients, called a "cohort", over time to determine general outcomes as well as outcomes of different subgroups.
[7] Any study done forward in time. This is particularly important in studies on therapy, prognosis or harm, where retrospective studies make hidden biases very likely.
[8] This might be considered a randomized controlled trial played backwards. People who get sick or have a bad outcome are identified and "matched" with people who did better. Then, the effects of the therapy or harmful exposure which might have been administered at the start of the trial are evaluated.
[9] Any study in which the outcomes have already occurred before the study has begun.
[10] This includes single case reports and published case series.

OUTLINE

NOVEMBER 2016 SECTION 6

LARYNGOLOGY, VOICE DISORDERS AND BRONCHOESOPHAGOLOGY

I. LARYNGOLOGY
   - A. Neurolaryngology: Injection and Diagnostics
   - B. Stroboscopy
   - C. Office-Based Procedures
   - D. Aging Larynx

II. BRONCHOESOPHAGOLOGY
   - A. Esophageal Dysphagia
   - B. Esophagopharyngeal Reflux
   - C. Tracheobronchial Disorders

TABLE OF CONTENTS
Selected Recent Materials - Reproduced in this Study Guide

NOVEMBER 2016 SECTION 6:
LARYNGOLOGY, VOICE DISORDERS, AND BRONCHOESOPHAGOLOGY

ADDITIONAL REFERENCE MATERIAL…….....i-ii

I. LARYNGOLOGY

A. Neurolaryngology: Injection and Diagnostics

Domer AS, Leonard R, Belafsky PC. Pharyngeal weakness and upper esophageal sphincter opening in patients with unilateral vocal fold immobility. *Laryngoscope*. 2014; 124(10):2371-2374. EBM level 4.....1-4

Summary: Individuals with unilateral vocal fold immobility of iatrogenic and idiopathic etiologies with subjective dysphagia demonstrate objective evidence of pharyngeal weakness. The increased prevalence of aspiration in this population may not solely be the result of impaired airway protection.

Li M, Chen S, Wang W, et al. Effect of duration of denervation on outcomes of ansa-recurrent laryngeal nerve reinnervation. *Laryngoscope*. 2014; 124(8):1900-1905. EBM level 4.....5-10

Summary: Although delayed laryngeal reinnervation is proved valid for unilateral vocal fold paralysis, surgical outcome is better if the procedure is performed within 2 years after nerve injury than after 2 years.

Paddle PM, Mansor MB, Song PC, Franco RA Jr. Diagnostic yield of computed tomography in the evaluation of idiopathic vocal fold paresis. *Otolaryngol Head Neck Surg*. 2015; 153(3):414-419. EBM level 4.....11-16

Summary: This study reviewed charts of 174 patients with a diagnosis of idiopathic unilateral vocal fold paresis (IUVFP) who underwent CT scan from skull base to mediastinum in a tertiary laryngology practice over a 10-year period. Of the 174 patients, 5 patients had a cause for their paresis identified on CT. This equated to a diagnostic yield of 2.9% (95% confidence interval, 0.94% to 6.6%). In addition, 48 patients had other incidental lesions identified that required further follow-up, investigation, or treatment. This equated to an incidental yield of 27.6% (95% confidence interval, 21.1% to 34.9%). This review demonstrates a low diagnostic yield and a high incidental yield. These findings suggest that the routine use of CT in the evaluation of idiopathic vocal fold paresis should be given careful consideration and that a tailored approach to investigation with good otolaryngologic follow-up is warranted.

Spataro EA, Grindler DJ, Paniello RC. Etiology and time to presentation of unilateral vocal fold paralysis. *Otolaryngol Head Neck Surg*. 2014; 151(2):286-293. EBM level 4.....17-24

Summary: Records were reviewed for 10 years at a single institution for unilateral vocal fold paralysis (UVFP); 938 patients met inclusion criteria. Of this group, 522 patients (55.6%) had UVFP due to surgery; 158 patients (16.8%) had UVFP associated with thyroid/parathyroid surgery, while in 364 patients (38.8%), UVFP was due to non-thyroid surgery. Of the total group, 416 patients (44.4%) had nonsurgical etiologies, 124 patients (13.2%) had idiopathic UVFP, and 621 patients (66.2%) had left-sided UVFP. Thyroidectomy remains the leading cause of surgery-related UVFP. Patients typically are seen within 3 to 4 months of onset.

Wu AP, Sulica L. Diagnosis of vocal fold paresis: current opinion and practice. *Laryngoscope*. 2015; 125(4):904-908. EBM level 5.....25-29

Summary: Wu and Sulica surveyed expert laryngologists who diagnosed vocal fold paresis predominantly on stroboscopic examination. Gross motion abnormalities had the highest positive predictive value. Laryngeal electromyography was infrequently used to assess for vocal fold paresis.

B. Stroboscopy

Djukic V, Milovanovic J, Jotic AD, Vukasinovic M. Stroboscopy in detection of laryngeal dysplasia: effectiveness and limitations. *J Voice*. 2014; 28(2):262.e13-262.e21. EBM level 3a.....30-38

Summary: This is a large study examining pre- and posttreatment stroboscopic findings in a prospectively collected group of 112 patients with dysplasia. There were fairly stringent exclusion criteria. The main finding was that abnormal amplitude of vocal fold vibration was significantly associated with recurrence. The type of cordectomy performed for the dysplasia and involvement of the vibratory segment was also associated with recurrence of dysplasia. Most recurrence occurred in a moderate dysplasia group. One limitation was that there were smaller numbers of patients in each dysplasia category; however, the study did highlight that caution should be exercised in the posttreatment follow-up period, and that stroboscopy has to be used in combination with other methods for an accurate diagnosis but can be helpful in predicting recurrence.

Reiter R, Pickhard A, Sander S, Brosch S. Prognostic relevance of mucosal waves in patients with unilateral vocal fold paralysis. *Ann Otol Rhinol Laryngol*. 2016; 125(4):331-335. EBM level 4.....39-43

Summary: This is a retrospective review of 100 patients with unilateral vocal fold paralysis looking at stroboscopic characteristics and relationship to eventual outcome. The majority of patients (75%) recovered function, and in all patients who recovered, there was an existing mucosal wave. In patients who did not recover function, only 10% had a mucosal wave.

Rosow DE, Sulica L. Laryngoscopy of vocal fold paralysis: evaluation of consistency of clinical findings. *Laryngoscope*. 2010; 120(7):1376-1382. EBM level 2.....44-50

Summary: This study sent videostroboscopy examination results from patients with unilateral vocal fold paralysis (VFP) to 22 blinded laryngologists and asked them to rate the results on twelve different criteria. The interrater reliability for each criterion was then calculated. The criteria with the best interrater agreement were glottic insufficiency, vocal fold bowing, and salivary pooling, which showed moderate agreement. All other criteria showed fair or poor agreement. The authors concluded that while it would be ideal to have a standardized rating scale for evaluation of VFP, the lack of interrater agreement across a wide range of laryngologists with different training and different backgrounds suggests that this may be very difficult to achieve.

Simpson CB, May LS, Green JK, et al. Vibratory asymmetry in mobile vocal folds: is it predictive of vocal fold paresis? *Ann Otol Rhinol Laryngol*. 2011; 120(4):239-242. EBM level 4.....51-54

Summary: This study is a retrospective review of 23 patients with symptoms suggestive of glottic insufficiency and stroboscopy examinations showing normal vocal fold mobility and vibratory asymmetry. All patients underwent laryngeal electromyography (LEMG) to determine presence of paresis. A total of 19 patients (83%) had evidence of paresis on LEMG. Blinded reviewers evaluated stroboscopy examinations for presence of paresis, but their ability to predict the distribution (sidedness) of the paresis was 37% or worse. The authors note that their findings suggest that all clinical and stroboscopic diagnoses of vocal fold paresis should be followed up with LEMG as the gold standard for diagnosis.

C. Office-Based Procedures

Croake DJ, Stemple JC, Uhl T, et al. Reliability of clinical office-based laryngeal electromyography in vocally healthy adults. *Ann Otol Rhinol Laryngol*. 2014; 123(4):271-278. EBM level 3.....55-62

Summary: Using a quantitative analysis protocol to inform an essentially qualitative technique, the study results indicated that there was generally poor to fair reliability in the laryngeal electromyography (LEMG) signal over testing sessions. Vocal intensity was an important variable that affected LEMG signal reliability. Standardization of LEMG protocols using vocal control parameters and quantitative analyses may help improve LEMG reliability in clinical settings.

Koszewski IJ, Hoffman MR, Young WG, et al. Office-based photoangiolytic laser treatment of Reinke's edema: safety and voice outcomes. *Otolaryngol Head Neck Surg*. 2015; 152(6):1075-1081. EBM level 4.....63-69

Summary: This study provides a retrospective analysis of patients undergoing office-based laser treatment of endoscopically proven Reinke's edema. Nineteen patients met criteria for the study inclusion. Five procedures were truncated due to patient intolerance. Phonatory frequency range increased (N = 12, $p = 0.003$), while percent jitter decreased (N = 12, $p = 0.004$). Phonation threshold pressure decreased after treatment (N = 4, $p = 0.049$). The Voice Handicap Index also decreased (N = 14, $p = 0.001$).

Richards AL, Sugumaran M, Aviv JE, et al. The utility of office-based biopsy for laryngopharyngeal lesions: comparison with surgical evaluation. *Laryngoscope*. 2015; 125(4):909-912. EBM level 4.....70-73

Summary: Office biopsy for laryngopharyngeal lesions may offer early detection and avoid operative intervention in some cases; however, for suspected dysplastic or malignant lesions, direct microlaryngoscopy should be the standard of care to ensure adequate full-thickness sampling and staging. For benign pathology, office biopsy is a safe and viable alternative to direct microlaryngoscopy and biopsy/excision.

Verma SP, Dailey SH. Office-based injection laryngoplasty for the management of unilateral vocal fold paralysis. *J Voice*. 2014; 28(3):382-386. EBM level 4.....74-78

Summary: This study is a retrospective chart review of 82 consecutive office-based injection laryngoplasty (OBIL) attempts on 57 patients. The most common route of access was transoral (85.6%). All OBILs were able to be completed. Injectates used were hyaluronic acid derivatives (57.3%), calcium hydroxyapatite (16%), and Cymmetra (16.5%). Three complications (3.7%) occurred. Thirty percent of patients ultimately elected for thyroplasty or ansa reinnervation, 22% found their condition to self-resolve, 14% died, and 25% were lost to follow-up.

D. Aging Larynx

Branco A, Todorovic Fabro A, Gonçalves TM, Garcia Martins RH. Alterations in extracellular matrix composition in the aging larynx. *Otolaryngol Head Neck Surg*. 2015; 152(2):302-307. EBM level 4.....79-84

Summary: This cadaver study proposes to further characterize extracellular matrix composition (ECM) changes in the aged vocal fold. Through immunohistochemistry, an overall increase in ECM mediated by increased collagen as well as decreased elastin were demonstrated. This work further highlights the histologic changes responsible for age-related voice changes.

Lee YC, Lee JS, Kim SW, et al. Influence of age on treatment with proton pump inhibitors in patients with laryngopharyngeal reflux disease: a prospective multicenter study. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(12):1291-1295. EBM level 4.....85-89

Summary: The authors aim to evaluate differences in laryngopharyngeal reflux (LPR) symptom severity among different age cohorts as well as response to treatment. The authors demonstrate that patients over 60 years of age experience greater symptoms and impact on quality of life from LPR; however, they seem to achieve less benefit from proton pump inhibitor therapy.

Yamauchi A, Yokonishi H, Imagawa H, et al. Vocal fold vibration in vocal fold atrophy: quantitative analysis with high-speed digital imaging. *J Voice*. 2015; 29(6):755-762. EBM level 2.....90-97

Summary: This study shows that high-speed digital imaging gives more insight into characterizing atrophic vocal folds.

Zeigler A, Verdolini Abbott K, Johns M, et al. Preliminary data on two voice therapy interventions in the treatment of presbyphonia. *Laryngoscope*. 2014; 124(8):1869-1876. EBM level 2b.....98-105

Summary: Zeigler et al demonstrate that vocal function exercise and PhoRTE voice therapy techniques appear to be effective in atrophy patients.

II. BRONCHOESOPHAGOLOGY

A. Esophageal Dysphagia

Kocdor P, Siegel ER, Tulunay-Ugur OE. Cricopharyngeal dysfunction: a systematic review comparing outcomes of dilatation, botulinum toxin injection, and myotomy. *Laryngoscope*. 2016; 126(1):135-141. EBM level 2a.....106-112

Summary: This systematic review of cohort studies evaluated the outcomes between different interventions for cricopharyngeal (CP) dysfunction, including CP dilation, botulinum toxin injections, and myotomy. The authors found that there was a significant increase in the odds of success and decreased complication rates with endoscopic myotomy versus open myotomy. They also found that myotomy was more effective than botulinum toxin injections.

Miles A, McMillan J, Ward K, Allen J. Esophageal visualization as an adjunct to the videofluoroscopic study of swallowing. *Otolaryngol Head Neck Surg*. 2015; 152(3):488-493. EBM level 4.....113-118

Summary: Miles et al hypothesize that esophageal disorders are the cause for dysphagia in many patients and propose that studying the esophagus as part of videofluoroscopic study of swallowing will yield greater diagnosis of abnormalities. Their findings suggest that esophageal disease is common and sometimes is the only abnormality in patients with cervical dysphagia. Furthermore, esophageal abnormalities frequently coexist with oral and pharyngeal disorders.

Moawad FJ, Veerappan GR, Dias JA, et al. Randomized controlled trial comparing aerosolized swallowed fluticasone to esomeprazole for esophageal eosinophilia. *Am J Gastroenterol*. 2013; 108(3):366-372. EBM level 1b.....119-125

Summary: This single-blinded, randomized controlled trial compared the efficacy of swallowed aerosolized steroids (fluticasone) to a proton pump inhibitor (omeprazole) for the treatment of eosinophilic esophagitis (EOE)–an important cause of esophageal dysphagia. Gastroesophageal reflux disease (GERD) patients were stratified equally into each arm. GERD patients with eosinophilia had improvement in their dysphagia symptoms and eosinophilia in biopsy specimens with PPI treatment alone. Current guidelines recommend failed trial of PPI prior to formal diagnosis of EOE due to this "newer" entity of PPI-responsive EOE.

Peng KA, Feinstein AJ, Salinas JB, Chhetri DK. Utility of the transnasal esophagoscope in the management of chemoradiation-induced esophageal stenosis. *Ann Otol Rhinol Laryngol*. 2015; 124(3):221-226. EBM level 4.....126-131

Summary: This article establishes the applications and safety of the transnasal esophagoscope (TNE) for chemoradiation-induced pharyngoesophageal swallowing dysfunction. Through use of a modified dysphagia score, the Functional Outcome Swallowing Scale (FOSS), the authors also suggest efficacy of TNE-based procedures in the population.

B. Esophagopharyngeal Reflux

Jetté ME, Gaumnitz EA, Birchall MA, et al. Correlation between reflux and multichannel intraluminal impedance pH monitoring in untreated volunteers. *Laryngoscope*. 2014; 124(10):2345-2351. EBM level 1b.....132-138

Summary: This study shows that the Reflux Finding Score (RFS) is not specific to detect laryngopharyngeal reflux in healthy volunteers, suggesting that other things can cause laryngeal inflammation.

Kavitt RT, Yuksel ES, Slaughter JC, et al. The role of impedance monitoring in patients with extraesophageal symptoms. *Laryngoscope*. 2013; 123(10):2463-2468. EBM level 2b.....139-144

Summary: Kavitt et al demonstrate that typical gastroesophageal reflux disease (GERD) testing parameters cannot predict the same conclusions when impedance is performed while the patient is on proton pump inhibitor therapy and recommend caution on over-interpreting impedance data in laryngopharyngeal reflux.

Reichel O, Dressel H, Wiederänders K, Issing WJ. Double-blind, placebo-controlled trial with esomeprazole for symptoms and signs associated with laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2008; 139(3):414-420. EBM level 1.....145-151

Summary: This study is a prospective, double-blinded randomized controlled trial examining esomeprazole versus placebo in managing the symptoms and signs of laryngopharyngeal reflux (LPR). A total of 62 patients with LPR were enrolled, and ultimately 30 patients were in the esomeprazole group, with 28 in the control group. Study subjects were given either esomeprazole 20 mg twice daily or identical placebo. They underwent laryngoscopic examinations and completed a Reflux Symptom Index (RSI) at 6 weeks and 12 weeks. There was minimal difference between the two groups at 6 weeks, but at 12 weeks, there was a significant difference in RSI as well as the Reflux Finding Score on laryngoscopy. The authors also found a high percentage of placebo patients (42%) experienced complete relief of symptoms at the end of the 3-month trial, though the percentage of the esomeprazole group was significantly higher (over 78%). One possible limitation is that pH monitoring was not used to diagnose LPR.

Youssef TF, Ahmed MR. Treatment of clinically diagnosed laryngopharyngeal reflux disease. *Arch Otolaryngol Head Neck Surg*. 2010; 136(11):1089-1092. EBM level 1.....152-155

Summary: This randomized controlled trial looked at the difference between monotherapy with esomeprazole versus triple therapy with esomeprazole, amoxicillin, and clarithromycin in treating patients with laryngopharyngeal reflux disease (LPRD) that have been found to be positive for *H. pylori* stool antigen (HPSA). The authors first determined that there is no statistical difference in symptoms between HPSA-positive and HPSA-negative patients with LPRD. Next, patients with clinically diagnosis of LPRD confirmed on pH testing were divided into two treatment groups based on HPSA status. HPSA-negative patients received esomeprazole 40 mg once daily for 4 weeks, while HPSA-positive patients were divided into monotherapy and triple therapy groups. The monotherapy treatment was identical to the HPSA-negative treatment, while the triple therapy group received esomeprazole 40 mg daily, 1 g amoxicillin daily, and 500 mg clarithromycin daily for 4 weeks. The HPSA-negative group showed improvement in 97% of patients, while in the HPSA-positive group, patients on monotherapy showed improvement in 40%, and 90% of those on triple therapy showed improvement. A major weakness of the study is that "improvement" is based on self-reporting by patients and by blinded laryngoscopic evaluation by the senior author. While these are no doubt important factors to take into consideration, they do not make use of objective testing, such as repeat 24-hour pH probe or validated questionnaires.

C. Tracheobronchial Disorders

D'Anza B, Knight J, Greene JS. Does body mass index predict tracheal airway size? *Laryngoscope*. 2015; 125(5):1093-1097. EBM level 4.....156-160

Summary: This study reviewed information on 123 patients who underwent tracheotomy over a 4-year period who also had CT imaging of the trachea in the 3 months preceding tracheostomy. The size of the endotracheal tube at time of the tracheotomy was also noted. Measurements were taken at the level of the first tracheal ring, as this was the most likely area for cuff-related injury of the airway. Important findings from the study were that airway area was correlated with height, and body mass index was inversely related to tracheal width after controlling for gender and age.

Gelbard A, Francis DO, Sandulache VC, et al. Causes and consequences of adult laryngotracheal stenosis. *Laryngoscope*. 2015; 125(5):1137-1143. EBM level 4.....161-167

Summary: This study looked at 340 patients with tracheal or laryngeal stenosis at two different sites. The etiology categories were idiopathic, iatrogenic, autoimmune, and trauma. The trauma group had significantly younger patients, whereas the idiopathic group had significantly more females. Comorbidities such as cardiovascular disease, peripheral vascular disease, and diabetes were more prevalent in the iatrogenic group. The idiopathic group also had the least-severe degree of laryngotracheal stenosis, with significantly fewer patients (none in this study) having had tracheostomy. As expected, patients with higher-grade stenosis (Cotton-Myer grades III or IV) had higher odds of being tracheostomy-dependent. The presence of tracheomalacia increased the odds of requiring a tracheostomy in the iatrogenic group.

Kettunen WW, Helmer SD, Haan JM. Incidence of overall complications and symptomatic tracheal stenosis is equivalent following open and percutaneous tracheostomy in the trauma patient. *Am J Surg*. 2014; 208(5):770-774. EBM level 3.....168-172

Summary: This is a large (N = 616) retrospective comparative study of the rate of tracheal stenosis in trauma patients who underwent either percutaneous (N = 351) versus open (N = 265) tracheostomy. The authors found no significant difference in the rate of tracheal stenosis in the open (1.9%) versus percutaneous (1.1%) groups. They did find that patients who developed tracheal stenosis were younger ($p = 0.02$) and had longer mechanical ventilation periods ($p = 0.055$). In addition, mortality was significantly higher in open tracheostomy patients, but this may be secondary to selection bias since patients with higher acuity of illness may be more likely to undergo open procedures.

Kraft SM, Sykes K, Palmer A, Schindler J. Using pulmonary function data to assess outcomes in the endoscopic management of subglottic stenosis. *Ann Otol Rhinol Laryngol*. 2015; 124(2):137-142. EBM level 4.....173-178

Summary: This retrospective case series described the utility of using pulmonary function tests to evaluate the efficacy of interventions for idiopathic subglottic stenosis. The pulmonary function parameters of PEF, PIF, FEV1/PEF, and FIF50% appeared to be the most valuable in judging response to endoscopic management and were significantly improved after airway dilation. PIF was the only parameter that was significantly associated with balloon size used for dilation. This study suggests that changes in PFTs are individualistic and need to be compared pre- and postprocedure for each patient (ie, there was no proposed "cut off" for intervention).

Taylor SC, Clayburgh DR, Rosenbaum JT, Schindler JS. Clinical manifestations and treatment of idiopathic and Wegener granulomatosis-associated subglottic stenosis. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(1):76-81. EBM level 4.....179-184

Summary: This article compares the clinical presentation, treatments, and outcomes in patients with subglottic stenosis (SGS) due to granulomatosis with polyangiitis (GPA, previously Wegener granulomatosis) versus idiopathic SGS. Although retrospective, this article has some interesting comparisons between the groups after similar treatment. For example, following open airway reconstruction, no idiopathic SGS patients required subsequent endoscopic dilations, while all GPA-SGS patients required subsequent interventions and had a higher rate of tracheotomy.

2016 SECTION 6 ADDITIONAL REFERENCES

Badia PI, Hillel AT, Shah MD, et al. Computed tomography has low yield in the evaluation of idiopathic unilateral true vocal fold paresis. *Laryngoscope*. 2013; 123(1):204-207.

Belafsky PC, Rees CJ, Allen J, Leonard RJ. Pharyngeal dilation in cricopharyngeus muscle dysfunction and Zenker diverticulum. *Laryngoscope*. 2010; 120(5):889-894.

Calloway HE, Kimbell JS, Davis SD, et al. Comparison of endoscopic versus 3D CT derived airway measurements. *Laryngoscope*. 2013; 123:2136-2141.

Carroll TL, Fedore LW, Aldahlawi MM. pH impedance and high-resolution manometry in laryngopharyngeal reflux disease high-dose proton pump inhibitor failures. *Laryngoscope*. 2012; 122(11):2473-2481.

Cheng J, Elden L. Outcomes in children under 12 months of age undergoing adenotonsillectomy for sleep-disordered breathing. *Laryngoscope*. 2013; 123(9):2281-2284.

Chin CJ, George C, Lannigan R, Rotenberg BW. Association of CPAP bacterial colonization with chronic rhinosinusitis. *J Clin Sleep Med*. 2013; 9(8):747-750.

Cohen SM, Misono S. Use of specific neuromodulators in the treatment of chronic, idiopathic cough: a systematic review. *Otolaryngol Head Neck Surg*. 2013; 148(3)374-382.

Deckard N, Yeh J, Soares DJ, et al. Utility of two-stage laryngotracheal reconstruction in the management of subglottic stenosis in adults. *Ann Otol Rhinol Laryngol*. 2013; 122(5):322-329.

Del Pero MM, Chaudhry A, Rasmussen N, et al. A disease activity score for ENT involvement in granulomatosis with polyangiitis (Wegener's). *Laryngoscope*. 2013; 123(3):622-628.

Eun YG, Shin SY, Kim SW. Effects of uvulopalatopharyngoplasty with or without radiofrequency tongue base reduction on voice in patients with obstructive sleep apnea. *Laryngoscope*. 2013; 123(7):1806-1810.

Friedman AD, Kobler JB, Landau-Zemer T, et al. High-force simulated intubation fails to dislocate cricoarytenoid joint in ex vivo human larynges. *Ann Otol Rhinol Laryngol*. 2012; 121(11):746-753.

Hartzell LD, Richter GT, Glade RS, Bower CM. Accuracy and safety of tracheoscopy for infants in a tertiary care clinic. *Arch Otolaryngol Head Neck Surg*. 2010; 136(1):66-69.

Huang WJ, Shu CH, Chou KT, et al. Evaluating the autonomic nervous system in patients with laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2013; 148(6):997-1002.

Johnston N, Yan JC, Hoekzema CR, et al. Pepsin promotes proliferation of laryngeal and pharyngeal epithelial cells. *Laryngoscope*. 2012; 122(6):1317-1325.

Kelchner LN, Brehm SB, de Alarcon A, Weinrich B. Update on pediatric voice and airway disorders: assessment and care. *Curr Opin Otolaryngol Head Neck Surg*. 2012; 20(3):160-164.

Kezirian EJ, Hohenhorst W, de Vries N. Drug-induced sleep endoscopy: the VOTE classification. *Eur Arch Otorhinolaryngol*. 2011; 268(8):1233-1236.

Krausert CR, Ying D, Choi SH, et al. Effect of vocal fold injury location on vibratory parameters in excised canine larynges. *Otolaryngol Head Neck Surg*. 2013; 148(1):89-95.

Marklund M, Verbraecken J, Randerath W. Non-CPAP therapies in obstructive sleep apnoea: mandibular advancement device therapy. *Eur Respir J*. 2012; 39(5):1241-1247.

Maturo SC, Braun N, Brown DJ, et al. Intraoperative laryngeal electromyography in children with vocal fold immobility: results of a multicenter longitudinal study. *Arch Otolaryngol Head Neck Surg*. 2011; 137(12):1251-1257.

Murry T, Branski RC, Yu K, et al. Laryngeal sensory deficits in patients with chronic cough and paradoxical vocal fold movement disorder. *Laryngoscope*. 2010; 120(8):1576-1581.

Nouraei SAR, Sandhu GS. Outcome of a multimodality approach to the management of idiopathic subglottic stenosis. *Laryngoscope*. 2013; 123(10):2474-2484.

Rao DA, Dellaripa PF. Extrapulmonary manifestations of sarcoidosis. *Rheum Dis Clin N Am*. 2013; 39(2):277-297.

Shah RN, Deal AM, Buckmire RA. Multidimensional voice outcomes after type I Gore-Tex thyroplasty in patients with nonparalytic glottic incompetence: a subgroup analysis. *Laryngoscope*. 2013; 123(7):1742-1745.

Smith DF, Vikani AR, Benke JR, et al. Weight gain after adenotonsillectomy is more common in young children. *Otolaryngol Head Neck Surg*. 2013; 148(3):488-493.

Sufyan AS, Kincaid JC, Wannemuehler TJ, Halum SL. The interarytenoid spatial relationship: accuracy and interrater reliability for determining sidedness in cases of unilateral adductor paresis. *J Voice*. 2013; 27(1):90-94.

Ummarino D, Vandermeulen L, Roosens B, et al. Gastroesophageal reflux evaluation in patients affected by chronic cough: Restech versus multichannel intraluminal impedance/pH metry. *Laryngoscope*. 2012; 123(4):980-984.

Vaca M, Mora E, Cobeta I. The aging voice: influence of respiratory and laryngeal changes. *Otolaryngol Head Neck Surg*. 2015; 153(3):409-413.

Van de Heyning PH, Badr MS, Baskin JZ, et al. Implanted upper airway stimulation device for obstructive sleep apnea. *Laryngoscope*. 2012; 122(7):1626-1633.

Vertigan AE, Gibson PG. The role of speech pathology in the management of patients with chronic refractory cough. *Lung*. 2012; 190(1):35-40.

Wang X, Ouyang Y, Wang Z, et al. Obstructive sleep apnea and risk of cardiovascular disease and all-cause mortality: a meta-analysis of prospective cohort studies. *Int J Cardiol*. 2013; 169(3):207-214.

Zerbib F, Roman S, Bruley Des Varannes S, et al. Normal values of pharyngeal and esophageal 24-hour pH impedance in individuals on and off therapy and interobserver reproducibility. *Clin Gastroenterol Hepatol*. 2013; 11(4):366-372.

Reprinted by permission of Laryngoscope. 2014; 124(10):2371-2374.

The Laryngoscope
© 2014 The American Laryngological,
Rhinological and Otological Society, Inc.

Pharyngeal Weakness and Upper Esophageal Sphincter Opening in Patients With Unilateral Vocal Fold Immobility

Amanda S. Domer, MS, CCC-SLP; Rebecca Leonard, PhD, CCC-SLP; Peter C. Belafsky, MD, PhD, MPH

Objectives/Hypothesis: To evaluate pharyngeal strength and upper esophageal sphincter opening in patients with unilateral vocal fold immobility (UVFI).

Study Design: Case control study.

Methods: Charts of individuals with UVFI who underwent a videofluoroscopic swallow study were reviewed. To exclude confounding variables associated with pharyngeal weakness, inclusion was limited to patients with iatrogenic and idiopathic UVFI. Data abstracted included patient demographics, etiology of UVFI, pharyngeal constriction ratio (PCR), and upper esophageal sphincter (UES) opening (UESmax). Data were compared to age/gender-matched controls with no history of dysphagia or UVFI. Discrete variables were analyzed using a chi-square test of independence, and an independent samples *t* test was used to compare the UVFI and control groups ($P = 0.05$). A one-way analysis of variance (ANOVA) was used to compare iatrogenic and idiopathic UVFI groups.

Results: The mean age of the cohort (n = 25) was 61 (±14 SD) years and 52% was female. The etiologies of UVFI were iatrogenic (n = 17) and idiopathic (n = 8). Thirty-eight percent of UVFI patients (n = 25) aspirated compared to 0% of controls ($P < 0.05$). The mean PCR for the UVFI group was 0.14 (±0.02) compared to 0.06 (±.01) for controls ($P < 0.05$). The mean UESmax for the UVFI group was 0.82 cm (±0.04) compared to 1.0 cm (±0.05) for controls ($P > 0.05$).

Conclusion: Individuals with UVFI of iatrogenic and idiopathic etiologies with subjective dysphagia demonstrate objective evidence of pharyngeal weakness. The increased prevalence of aspiration in this population may not be solely the result of impaired airway protection.

Key Words: Dysphagia, aspiration, vocal fold immobility, swallowing disorder, unilateral vocal fold immobility, UVFI.

Level of Evidence: 3b.

***Laryngoscope*, 124:2371–2374, 2014**

INTRODUCTION

Glottal competence is essential in airway protection during deglutition. If glottal closure is ineffective as a result of unilateral vocal fold immobility (UVFI), airway protection during the swallow may be compromised. UVFI may result from damage to the 1) brainstem nuclei, 2) vagus nerve, or 3) recurrent laryngeal nerve. Etiologies of UVFI include surgical trauma/iatrogenic (40%), tumor/neoplasm (30%), unknown/idiopathic (11%), trauma (8%), central nervous system dysfunction (4%), radiation (3%), inflammatory conditions (2%), and cardiovascular disease (2%).[1] Individuals with UVFI may present with aphonia (i.e., absence of voice), dysphonia (i.e., voice impairment), and/or dysphagia (i.e., swallowing impairment).

The precise etiology of dysphagia in patients with UVFI is uncertain. It is generally accepted that UVFI results in diminished airway protection. If airway protection is ineffective, an individual may aspirate material into the lungs, which may result in respiratory infection and/or death due to aspiration pneumonia. Approximately 33% to 42% of individuals with UVFI have been identified to aspirate.[2–5] Diminished airway protection is presumed to be the primary cause of swallowing dysfunction in patients with UVFI.[3,5] The integrity of other important biomechanics of the swallow, such as upper esophageal sphincter opening and pharyngeal contractility, however, has not been adequately evaluated in this patient population. Due to the highly intricate nature of the nerves and muscles in the pharynx and larynx, as well as the complex kinematics of the swallow, we hypothesize that features aside from impaired airway protection alone may contribute to increased occurrence of aspiration in this population.

This has been hypothesized in previous studies, which have demonstrated subjective findings in addition to impaired glottic closure that the authors stated contributed to a patient's increased risk of aspiration. One study that included patients with UVFI of both central and peripheral origins identified poor pharyngeal movement in patients with peripheral (i.e., recurrent laryngeal nerve injury, vagus nerve injury, or idiopathic etiologies) UVFI.[4] Another study identified decreased

From the Department of Communication Sciences & Disorders, University of South Florida (A.S.D.), Tampa, Florida; and the Department of Otolaryngology–Head & Neck Surgery, University of California, Davis (R.L., P.C.B.), Sacramento, California, U.S.A.

Editor's Note: This Manuscript was accepted for publication May 20, 2014.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Peter C. Belafsky, MD, PhD, MPH, Department of Otolaryngology–Head & Neck Surgery, University of California, Davis, 2521 Stockton Blvd, Suite 7200, Sacramento, CA 95817.
E-mail: peter.belafsky@ucdmc.ucdavis.edu

DOI: 10.1002/lary.24779

TABLE I.
Objective Measures and Definitions for a Dynamic Videofluoroscopic Swallow Study.

Objective Measures on Dynamic Videofluoroscopic Swallow Study (DSS)	Definition
Total pharyngeal transit time (TPT)	The time between the head of the bolus passing the posterior nasal spine to the time the tail of the bolus passes through the UES.
Upper esophageal sphincter opening (UESmax)	UES opening. The narrowest point of opening between C3 and C6 during maximal distention for bolus passage.
Pharyngeal constriction ratio (PCR)	A surrogate measure of pharyngeal strength. Specifically, a ratio of pharyngeal area measured in lateral fluoroscopic view at the point of maximal pharyngeal constriction during the swallow to the pharyngeal area measured with the bolus held in the oral cavity.
Hyoid to larynx approximation (HLx)	The difference in distance between the anterior margin of the hyoid bone with a 1cc bolus held in the oral cavity to maximal approximation of the hyoid and larynx during swallow and maximal hyoid to larynx approximation. A clear and consistently visible landmark on the anterior thyroid cartilage, such as calcification, was used as an alternative to the subglottic air column if it could not easily be visualized. Maximal approximation usually occurred just after maximal hyoid excursion.
Hyoid excursion (Hmax)	The distance traveled by the hyoid to the point of maximal elevation during a swallow from its position during hold.

sensation as a contributing factor to aspiration.[5] Other studies have provided evidence to support the notion that aspiration may not be solely related to vocal fold immobility, as demonstrated by continued dysphagia and/or aspiration after patients underwent surgical intervention to improve glottic closure.[6,7] There have yet to be objective kinematic and temporal measurements obtained from patients with subjective dysphagia as a result of UVFI related solely to vagus nerve injury. The purpose of this investigation was to evaluate pharyngeal strength and upper esophageal sphincter (UES) opening in individuals with UVFI caused by idiopathic or iatrogenic injuries to the vagus nerve.

MATERIALS AND METHODS

A clinical swallowing database consisting of individuals with dysphagia who underwent a dynamic videofluoroscopic swallow study (DSS) was reviewed to identify individuals with UVFI between January 1, 1999, and June 1, 2012. The Institutional Review Board of the University of California, Davis, approved use of this database for clinical research. All patients with UVFI were confirmed by videolaryngoscopy and/or strobovideolaryngoscopy. Individuals were excluded if they were under 18 years of age, had suspected vagus nerve injury of central origin (e.g., cerebrovascular accident, neuromuscular disease, brain tumor, etc.), head or neck cancer (i.e., except that isolated to the thyroid, which resulted in only surgical removal of all or part of the thyroid), and/or anterior approach cervical spine surgery. The purpose for excluding these populations was the possibility of a more complex swallowing disorder not necessarily limited to injury of the vagus nerve. This resulted in patients with iatrogenic and idiopathic UVFI. None of the patients had undergone a vocal fold medialization procedure prior to the videofluoroscopic swallow study. The timing between onset of UVFI and time of evaluation was not recorded.

All swallow studies were preformed using a properly collimated OEC Medical Systems 9800 Radiographic/Fluoroscopic unit that provided a 63 kV, 1.2 mA-type output for the full field of view mode (12-inch input phosphor diameter). In accordance with our standard protocol, a metal ring of known diameter was taped to the chin or neck of the patient for measurement calibration purposes. Lateral views were obtained while the patient, seated in an examination chair, was administered liquid barium (EZpaque barium sulfate suspension, 60% w/v; EZ-EM, Inc., Westbury, NY) boluses of 1 cc, 3 cc, and 20 cc and a 3-cc paste bolus (EZ-paste, EZ-EM, Inc.) measured with a syringe or graduated medicine cup. The patient was then turned to obtain anteroposterior views and administered liquid barium boluses of 3 cc and 20 cc. Studies were recorded on a Sony Md-1000 DVD recorder (Sony Corp. America, New York, NY) and were played back using Quick Time (7.7.1; Apple, Cupertino, CA). Measures were obtained from digitized images using ImageJ software (National Institutes of Health, Bethesda, MD) and software tools from Iconico, Inc (New York, NY). Specific measurement techniques have been previously described in detail.[8,9] An experienced unblinded clinician (i.e., the same clinician who conducted the videofluoroscopic swallow studies) analyzed all studies; however, because this study was retrospectively completed, there was no information available related to this study at the time of evaluation. All measures for the current study were obtained from the lateral view.

The primary outcome measures were upper esophageal sphincter opening (UESmax) and the pharyngeal constriction ratio (PCR). The PCR is a validated surrogate measure of pharyngeal strength on fluoroscopy; and an elevated PCR suggests pharyngeal weakness.[9] The secondary outcome measures were larynx to hyoid approximation (HLx), hyoid displacement (Hmax), and total pharyngeal transit time (TPT) (see Table I for definitions). The data from each variable were compared to age and gender-matched controls with no history of dysphagia. Discrete variables were analyzed with a chi-square test of independence and an independent samples t test was used to compare the control and combined UVFI groups with alpha set at 0.05. A one-way analysis of variance (ANOVA) was used to compare the iatrogenic and idiopathic UVFI groups with the control group. A Bonferroni correction was applied to adjust for multiple comparisons with alpha set at 0.01. There is a probability of 0.05 that a type I error has been made in the set of tests.

RESULTS

A total of 137 individuals with UVFI were identified from the clinical database. There were 25 subjects who met strict inclusion and exclusion criteria (i.e., did not

TABLE II.
Mean, Standard Deviation, and *P* Value for Iatrogenic and Idiopathic Groups.

Iatrogenic UVFI vs. Idiopathic UVFI	Iatrogenic (n = 17) Mean (SD)	Idiopathic (n = 8) mean (SD)	*P* Value
UESmax (cm)	0.85 (0.05)	0.76 (0.07)	1.00
PCR	0.12 (0.02)	0.18 (0.04)	0.34
HLx (cm)	1.25 (0.14)	1.58 (0.15)	0.55
Hmax (cm)	1.91 (0.16)	1.84 (0.16)	1.00
TPT (seconds)	1.36 (0.09)	2.66 (0.93)	0.04

$\alpha = 0.01$.
*Statistical significance.
HLx = hyoid to larynx approximation; Hmax = hyoid excursion; PCR = pharyngeal constriction ratio; SD = standard error; TPT = total pharyngeal transit time; UESmax = upper esophageal sphincter opening; UVFI = unilateral vocal fold immobility.

have dysphagia complaints that could be explained by any other etiology in their medical history). The etiology of UVFI was identified as iatrogenic in 17 individuals and idiopathic in eight individuals. The cohort was 52% female with a mean age of 61 (±14) years.

Comparison of the Idiopathic Group, Iatrogenic Group, and Control Group

The idiopathic group (n = 8) was 50% female with a mean age of 64 (±14) years (Table II). The iatrogenic group (n = 17) was 53% female with a mean age of 60 years (±14) years. There was no difference in age or gender between groups ($P > 0.05$).

The mean UESmax was 0.76 (±0.07) cm for the idiopathic group, 0.85 (±0.05) cm for the iatrogenic group, and 1.0 (±0.05) cm for the control group. There were no significant differences between any groups for UESmax ($P > 0.01$).

The PCR was 0.18 (±0.04) for the idiopathic group, 0.12 (±0.02) for the iatrogenic group, and 0.06 (±0.01) for the control group. PCR was significantly greater for the idiopathic group compared with the control group ($P < 0.01$). PCR for the iatrogenic group was not significant, but less than the control group ($P > 0.01$). There was not a significant difference between the idiopathic or iatrogenic groups ($P > 0.01$).

The HLx was 1.58 (±0.15) cm for the idiopathic group, 1.25 (±0.14) cm for the iatrogenic group, and 1.42 (±0.12) cm for the control group ($P > 0.01$). There were no significant differences between any groups for HLx ($P > 0.01$).

The Hmax was 1.84 (±0.16) cm for the idiopathic group, 1.91 (±0.16) cm for the iatrogenic group, and 2.22 (±0.18) cm for the control group ($P > 0.01$). There were no significant differences between any groups for Hmax ($P > 0.01$).

The mean TPT was 2.66 (±0.93) seconds for the idiopathic group, 1.36 (±0.09) seconds for the iatrogenic group, and 1.01 (±0.06) seconds for the control group ($P > 0.01$). There was a significant difference between the idiopathic UVFI and iatrogenic UVFI groups ($P < 0.01$). There was a significant difference between the idiopathic group and the control group. There was not a significant difference between the iatrogenic and the control group.

Thirty-eight percent of individuals with idiopathic UVFI and 35% of individuals with iatrogenic UVFI aspirated at least once during the videofluoroscopic swallow study. There was not a significant difference between either of the UVFI groups ($P > 0.05$); however, there was a significant difference between the idiopathic group and control group ($P < 0.05$), as well as the iatrogenic group and control group ($P < 0.05$).

UVFI Group Compared With the Control Group

The mean UESmax opening was 0.82 (±0.04) cm for the UVFI group compared to 1.00 (±0.05) cm for controls ($P > 0.05$ see Table III). The pharyngeal constriction ratio was 0.14 (±0.02) for the UVFI group compared to 0.06 (±0.01) for controls ($P < 0.05$). Larynx to hyoid approximation was 1.35 (±0.11) cm for the UVFI group and 1.42 (±0.12) cm for the control group ($P > 0.05$). The mean for hyoid displacement was 1.89 (±0.12) cm for the UVFI group and 2.22 (±0.17) cm for the control group ($P > 0.05$). The mean TPT was 1.78 (±0.32) seconds for the UVFI group and 1.01 (±0.06) seconds for the control group ($P < 0.05$). Thirty-six percent of individuals with UVFI aspirated at least once during the videofluoroscopic swallow study compared to 0% of controls ($P < 0.05$).

DISCUSSION

The data in the current investigation provided evidence to suggest that individuals with UVFI of iatrogenic and idiopathic etiologies may present with additional biomechanical findings that may increase the prevalence of aspiration. The group of individuals with UVFI of idiopathic and iatrogenic etiologies demonstrated significantly prolonged TPT and elevated PCRs, suggesting delayed bolus transit and pharyngeal weakness. Additionally, individuals with UVFI of idiopathic etiology demonstrated significantly prolonged TPT, increased PCR (i.e., pharyngeal weakness), and decreased UESmax compared to controls. These findings support the notion that factors other than glottal

TABLE III.
Mean, Standard Deviation, and *P* Value for UVFI Group and Controls.

UVFI vs. Controls	UVFI (n = 25) Mean (SD)	Controls (n = 25) Mean (SD)	*P* Value
UESmax (cm)	0.82 (0.05)	1.00 (0.05)	0.94
PCR	0.14 (0.02)	0.06 (0.01)	0.03*
HLx (cm)	1.35 (0.11)	1.42 (0.12)	0.94
Hmax (cm)	1.89 (0.12)	2.21 (0.17)	0.16
TPT (seconds)	1.78 (0.32)	1.01 (0.06)	0.02*

$\alpha = 0.05$
*Statistical significance.
HLx = hyoid to larynx approximation; Hmax = hyoid excursion; PCR = pharyngeal constriction ratio; SD = standard error; TPT = total pharyngeal transit time; UESmax = upper esophageal sphincter opening; UVFI = unilateral vocal fold immobility.

competence influence the increased prevalence of aspiration observed in patients with UVFI.

The results of the present study confirm the qualitative findings of Jang et al. with objective data. In that study, individuals with UVFI of peripheral origin (i.e., recurrent laryngeal nerve injury, vagus nerve injury, or idiopathic etiologies) presented with subjective suggestion of abnormal laryngeal elevation and epiglottic inversion, residue in the valleculae, residue in the pyriform sinuses, and aspiration.[4]

The percentage of aspiration in this study was also comparable to previous studies, which identified aspiration in approximately 33% to 42% of individuals with UVFI.[2–5] It is important to note that the present study was performed in an outpatient tertiary care center and most of the previous work evaluating aspiration in patients with UVFI was performed in acute care settings. Although information about the length of time from the onset of vocal fold immobility was not available in these studies, we suspect that the individuals included in the present investigation may have had a more prolonged duration of UVFI in comparison to previous work. Nonetheless, the percentage of aspiration in the current investigation was similar to previous findings.

When evaluating UVFI between groups, the only significant finding was total pharyngeal transit time. This may be the result of a higher vagal injury in the idiopathic group, compared to the iatrogenic group. However, this confirmed that, although the iatrogenic group was more likely isolated to recurrent laryngeal nerve and/or superior laryngeal nerve injury than the idiopathic group, the finding of increased PCR did not vary significantly between groups. Additionally, the iatrogenic group was approaching significance compared to the control group, and significance may be achieved with a larger sample size. Therefore, pharyngeal weakness may exist in individuals with UVFI of both idiopathic and iatrogenic etiologies.

This study was not without limitations. Electromyography was not utilized to determine the site of lesion causing UVFI. Also, in an effort to keep the groups as homogenous as possible, the sample size was small. However, the groups that were chosen were intended to represent individuals with UVFI limited to vagus nerve injury. None of the individuals underwent a vocal fold medialization procedure prior to study, and all individuals presented with a dysphagia complaint. In addition, this study was retrospective, so a future prospective investigation with a larger sample size is required to confirm these results.

CONCLUSION

Individuals with UVFI of iatrogenic and idiopathic etiologies with subjective dysphagia demonstrate objective evidence of pharyngeal weakness. The increased prevalence of aspiration in this population may not be solely the result of impaired airway protection.

BIBLIOGRAPHY

1. Chen HC, Jen YM, Wang CH, Lee JC, Lin YS. Etiology of vocal cord paralysis. *ORL J Otorhinolaryngol Relat Spec* 2007;69:167–171.
2. Leder SB, Suiter DM, Duffey D, Judson BL. Vocal fold immobility and aspiration status: a direct replication study. *Dysphagia* 2012;27:265–270.
3. Heitmiller RF, Tseng E, Jones B. Prevalence of aspiration and laryngeal penetration in patients with unilateral vocal fold motion impairment. *Dysphagia* 2000;15:184–187.
4. Jang YY, Lee SJ, Jeon JY, Lee SJ. Analysis of video fluoroscopic swallowing study in patients with vocal cord paralysis. *Dysphagia* 2012;27:185–190.
5. Tabaee A, Murry T, Zschommler A, Desloge RB. Flexible endoscopic evaluation of swallowing with sensory testing in patients with unilateral vocal fold immobility: incidence and pathophysiology of aspiration. *Laryngoscope* 2005;115:565–569.
6. Bhattacharyya N, Kotz T, Shapiro J. Dysphagia and aspiration with unilateral vocal cord immobility: incidence, characterization, and response to surgical treatment. *Ann Otol Rhinol Laryngol* 2002;111:672–679.
7. Yip HT, Leonard R, Kendall, KA. Impact of vocal fold medialization on pharyngeal transit time and incidence of aspiration. *Otolaryngol Head Neck Surg* 2004; 131
8. Kendall K, McKenzie S, Leonard R. Dynamic swallow study: objective measures and normative data. In: Leonard R, Kendall K, eds. *Dysphagia Assessment and Treatment Planning: A Team Approach*. San Diego, CA: Singular Publishing; 1998.
9. Leonard R, Rees CJ, Belafsky P, Allen J. Fluoroscopic surrogate for pharyngeal strength: the pharyngeal constriction ratio (PCR). *Dysphagia* 2011;26:13–17.

Reprinted by permission of Laryngoscope. 2014; 124(8):1900-1905.

The Laryngoscope
© 2014 The American Laryngological,
Rhinological and Otological Society, Inc.

Effect of Duration of Denervation on Outcomes of Ansa-Recurrent Laryngeal Nerve Reinnervation

Meng Li, MD; Shicai Chen, MD; Wei Wang, MD; Donghui Chen, MD; Minhui Zhu, MD; Fei Liu, MD; Caiyun Zhang, MD; Yan Li, MD; Hongliang Zheng, MD

Objectives/Hypothesis: To investigate the efficacy of laryngeal reinnervation with ansa cervicalis among unilateral vocal fold paralysis (UVFP) patients with different denervation durations.

Study Design: We retrospectively reviewed 349 consecutive UVFP cases of delayed ansa cervicalis to the recurrent laryngeal nerve (RLN) anastomosis. Potential influencing factors were analyzed in multivariable logistic regression analysis. Stratification analysis performed was aimed at one of the identified significant variables: denervation duration.

Methods: Videostroboscopy, perceptual evaluation, acoustic analysis, maximum phonation time (MPT), and laryngeal electromyography (EMG) were performed preoperatively and postoperatively. Gender, age, preoperative EMG status and denervation duration were analyzed in multivariable logistic regression analysis. Stratification analysis was performed on denervation duration, which was divided into three groups according to the interval between RLN injury and reinnervation: group A, 6 to 12 months; group B, 12 to 24 months; and group C, $>$ 24 months.

Results: Age, preoperative EMG, and denervation duration were identified as significant variables in multivariable logistic regression analysis. Stratification analysis on denervation duration showed significant differences between group A and C and between group B and C ($P < 0.05$)—but showed no significant difference between group A and B ($P > 0.05$) with regard to parameters overall grade, jitter, shimmer, noise-to-harmonics ratio, MPT, and postoperative EMG. In addition, videostroboscopic and laryngeal EMG data, perceptual and acoustic parameters, and MPT values were significantly improved postoperatively in each denervation duration group ($P < 0.01$).

Conclusions: Although delayed laryngeal reinnervation is proved valid for UVFP, surgical outcome is better if the procedure is performed within 2 years after nerve injury than that over 2 years.

Key Words: Vocal cord paralysis, laryngeal reinnervation, ansa cervicalis, recurrent laryngeal nerve, denervation duration.

Level of Evidence: 4.

***Laryngoscope*, 124:1900–1905, 2014**

INTRODUCTION

Unilateral vocal fold paralysis (UVFP) is a condition commonly seen in otolaryngology clinics. The most frequent cause of UVFP is injury to the recurrent laryngeal nerve (RLN).[1] UVFP can present as various degrees of dysphonia and dysphagia, and has a significant impact on a patient's quality of life. There are various surgical methods for treating UVFP. These include vocal fold injection, thyroplasty, arytenoid adduction, and laryngeal reinnervation, which is an effective surgical procedure with better long-term outcomes because it restores neural connections to laryngeal muscles and thus maintains the bulk, tension, and position of the paralyzed vocal fold.[2,3] Our previous large-scale study, as well as reports from other researchers, demonstrated satisfactory or good voice outcomes in patients who underwent ansa cervicalis–RLN anastomosis.[4–6]

Experimental studies have shown that reinnervation surgery can restore laryngeal function with excellent results when performed immediately following nerve resection.[7,8] However, in clinical practice, there is usually a considerable delay between RLN injury and presentation for reinnervation surgery. In addition, 6 to 12 months are usually allowed for possible spontaneous recovery of the paralyzed vocal fold or compensation from the contralateral vocal fold, even when UVFP is diagnosed early. Surgical intervention is considered only in cases with unsatisfactory spontaneous recovery. At present, there are no definitive clinical data regarding the longest allowable period between the onset of RLN injury and nerve reconstruction to achieve functional recovery of the adductor muscle.

In the present study, we analyzed several potential influencing factors of laryngeal reinnervation, including gender, age of patients, preoperative maximal voluntary motor-unit recruitment (VMUR) of laryngeal muscles,

From the Department of Otolaryngology–Head and Neck Surgery, Changhai Hospital, Second Military Medical University, Shanghai, 200433, People's Republic of China

Editor's Note: This Manuscript was accepted for publication January 27, 2014.

This work was supported by Grants 81100724, 81170899, 81070775 for science research from National Natural Science Foundation of China, and Grant 10XD1405500 from the Science and Technology Commission of Shanghai Municipality. The authors have no other financial relationships, or conflicts of interest to disclose.

Send correspondence to Hongliang Zheng, MD, Department of Otolaryngology–Head and Neck Surgery, Changhai Hospital, Second Military Medical University, Shanghai 200433, People's Republic of China. E-mail: zheng_hl2004@163.com

DOI: 10.1002/lary.24623

and denervation duration using multivariate logistic regression analysis method in a large series of 349 UVFP patients who underwent delayed laryngeal reinnervation. We also performed further stratification analysis aimed at one of the identified significant variables—denervation duration—in order to investigate the effect of denervation duration on the surgical outcome of laryngeal reinnervation.

MATERIALS AND METHODS

Patient Characteristics

Our study was approved by the institutional review board of Second Military Medical University, Shanghai, People's Republic of China. The medical records of 349 UVFP patients (94 males and 255 females; mean age 44.0 years, ranging 17–69 years) who underwent anastomosis of the main branch of the ansa cervicalis to the RLN between January 1996 and January 2011, and who were followed for at least 2 years, were reviewed. The etiology of UVFP in this series of patients was RLN injury during thyroid surgery. Informed consent was obtained from all patients involved in this study. Patients who were lost to follow-up were excluded. There was a minimum waiting period of 6 months following onset of RLN injury to allow for possible spontaneous recovery or compensation. The median denervation course was 16.1 months (range, 6–45 months). When stratified by denervation duration, the number of patients in each denervation duration group was: 172 (49.3%) patients with a denervation duration 6 to 12 months (group A); 108 (30.9%) patients with a denervation duration 12 to 24 months (group B), and 69 (19.8%) patients with a denervation duration >24 months (group C). The median follow-up period after laryngeal reinnervation was 70.8 months (range, 24–156 months).

Surgical Procedure

The surgical procedure has been elaborated in our previous report.[4] Briefly, under general anesthesia, the ipsilateral ansa cervicalis was explored, and the main branch was transected at the bifurcation and freely mobilized for preparation of anastomosis. The RLN was dissected at a point sufficiently far from the injury site to provide a tension-free anastomosis and then transected. Under an operating microscope, the distal RLN stump was anastomosed to the main branch of the ansa cervicalis using nylon 11-0 thread in three to five epineural sutures.

Videostroboscopy

All patients were observed via a videostroboscope (RICHARD WOLF GmbH, model 5570, Knittlingen, Germany) during "eee" phonation at a comfortable loudness and pitch for as long as possible, and dynamic videos were recorded preoperatively and postoperatively. Three experienced laryngologists who had not performed any of the surgeries reviewed all of the videos. The videos were randomized, and the reviewers were blinded to whether the videos were preoperative or postoperative. Visual laryngeal analysis included glottal closure (0, complete; 1, slightly incomplete; 2, moderately incomplete; 3, severely incomplete), vocal fold position, vocal fold edge of paralyzed side, phrase symmetry, and regularity. Consensus of the reviewers was reached on the visual appearance of the larynx. Our previous studies demonstrated that the above parameters were consistent in presenting reinnervation outcome of vocal fold paralysis, among which the parameter glottal closure was the most representative one[4]; therefore, only the parameter "glottal closure" was included when performing statistical analysis using a univariable analysis and multivariable logistic regression analysis.

Vocal Function Assessment

Vocal function assessment included perceptual evaluation, acoustic analysis, and maximum phonation time (MPT) measurement. Preoperative and postoperative voice samples containing sustained vowels /a/ and connected speech samples were used for perceptual evaluation and acoustic analysis. The recording equipment consisted of a digital audiotape recorder and a dynamic microphone (Tiger Electronics Inc., North Reading, MA). Five laryngologists who had been trained in grade, roughness, breathiness, asthenia, and strain (GRBAS) rating performed voice perceptual evaluation using a perceptual rating scale (GRBAS) for voice quality and characteristics. The ratings were accomplished in a blinded fashion, with patient voice samples arranged in a random manner. Each listener was asked to score connected speech samples for overall grade, roughness, breathiness, asthenia, and strain using a voice-quality scale for each parameter (0, normal; 1, mild; 2, moderate; 3, severe). The values were averaged among the five listeners. Our previous studies demonstrated that the interrater and intrarater reliability was acceptable (interrater reliability >0.76; intrarater reliability >0.81).[4,9] In addition, the above five parameters of perceptual evaluation were consistent in presenting vocal outcome of vocal fold paralysis, among which the parameter overall grade was the most representative one.[4] Therefore, only the parameter "overall grade" was included when performing statistical analysis using a univariable analysis and multivariable logistic regression analysis.

The acoustic parameters of sustained vowel /a/ were evaluated using Praat software (Boersma, Paul & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.1.12, retrieved from http://www.praat.org/). The acoustic parameters were mean noise-to-harmonics ratio (NHR) and measures of phonatory stability—jitter (local) and shimmer (local). MPT was defined as the duration of sustained phonation of the vowel /a/ after maximum inspiration and was measured preoperatively and postoperatively.[4]

Laryngeal Electromyography

A four-channel electromyograph and concentric needle electrodes (Dantec Counterpoint, Copenhagen, Denmark) were used for the laryngeal electromyography (EMG) recordings. To test for proper needle position, the unaffected vocal fold was examined first. The electromyographic activity of the bilateral thyroarytenoid (TA) muscles was recorded during the following two stages: while breathing quietly when relaxed, and while pronouncing the vowel /eee/ with the greatest exertion, then sniff. One board-certified otolaryngologist performed the EMG, and a neurologist operated the EMG machine and interpreted the EMG results. The neurologist rated the VMUR using the following scale: 0, full interference; 1, mixed interference; 2, simple interference; and 3, without motor unit potential.[4]

Statistical Analysis

The perceptual evaluation, acoustic analysis, and MPT data did not follow normal distribution and were presented as median (low quartile, upper quartile). We sought to evaluate influencing factors for the surgical outcome of laryngeal reinnervation using multivariable logistic regression methods. Potential influencing factors were examined in univariable

TABLE I.
Univariable Analysis of Influencing Factors on the Surgical Outcome of Laryngeal Reinnervation.

	Overall Grade	Jitter	Shimmer	NHR	MPT	Post-VMUR	Glottal Closure
Sex	0.7131	0.1844	0.1394	0.3456	< ***0.001***	0.1804	0.5141
Age	< ***0.001***	0.3441	< ***0.001***	***0.0010***	< ***0.001***	< ***0.001***	< ***0.001***
Pre-VMUR	< ***0.001***	0.3807	0.0630	0.1554	< ***0.001***	< ***0.001***	< ***0.001***
Denervation duration	***0.0022***	< ***0.001***	< ***0.001***	< ***0.001***	***0.0007***	***0.0010***	0.5410

Data in Table I represent *P* value calculated by univariable analysis.
$P < 0.05$ is deemed as statistically significant.
MPT = maximum phonation time; NHR = noise-to-harmonics ratio; VMUR = voluntary motor-unit recruitment.

analyses before building the multivariable logistic regression model. Univariable logistic regression models containing each covariate were fit. Covariates of known clinical significance or with P value < 0.05 in the univariable logistic models were selected as candidates for the multivariable model. Statistical significance was considered α-level 0.05.

RESULTS

Univariable Results

Table I displays the results of demographic or preoperative medical variables in univariable analysis. Briefly, gender was only significantly associated with MPT. Age was significantly associated with overall grade, shimmer, NHR, MPT, postoperative VMUR, and glottal closure. Preoperative VMUR was significantly associated with overall grade, MPT, postoperative VMUR, and glottal closure, whereas denervation duration was significantly associated with overall grade, jitter, shimmer, NHR, MPT, and postoperative VMUR.

Multivariable Logistic Regression Model Results

Based on univariable logistic regression results, significant candidate variables were identified ($P < 0.05$) for assessment in the multivariable logistic regression model. Table II shows the results of multivariable logistic regression. These indicate that age of patients; severity of nerve injury, which was presented as preoperative VMUR; and denervation duration had impact on the surgical effect of laryngeal reinnervation. Stratification analysis on other significant variables such as age of patients and preoperative VMUR will be presented in another report (unpublished). In the present study, we performed further stratification analysis aimed at one of the identified significant variables—denervation duration. The results are as follows (see Table III).

Videostroboscopic Findings

On preoperative videostroboscopy, the majority of cases in each group had severely incomplete glottal closure, which did not differ significantly among the three groups. On videostroboscopy performed 2 years after the reinnervation operation, most patients showed complete glottal closure. The postoperative stroboscopic findings were significantly improved in total sample and within each group compared with the corresponding preoperative findings ($P < 0.01$). However, denervation duration was not a significant variable with regard to multivariable logistic regression analysis in glottal closure.

Vocal Function Assessment

The preoperative values of perceptual evaluation parameter (overall grade) and acoustic parameters—jitter (local), shimmer (local), and the NHR—showed no significant differences among the three groups. Two years after the reinnervation operation, the postoperative value of overall grade was significantly improved within each of the three groups compared with the corresponding preoperative values ($P < 0.001$). Postoperative values of jitter (local), shimmer (local), and the NHR were significantly lower within each group compared with the corresponding preoperative values ($P < 0.05$). The postoperative MPT was significantly longer than the preoperative MPT in each group ($P < 0.05$).

Denervation duration was identified as a significant variable with regard to parameters (overall grade), jitter, shimmer, NHR, and MPT in multivariable logistic

TABLE II.
Multivariable Analysis of Influencing Factors on the Surgical Outcome of Laryngeal Reinnervation.

	Overall Grade	Jitter	Shimmer	NHR	MPT	Post-VMUR	Glottal Closure
Sex	0.6692	0.1124	**0.0133**	0.2320	**< 0.001**	0.4301	0.7663
Age	**< 0.001**	0.2588	**< 0.001**	**< 0.001**	**< 0.001**	**0.0080**	**0.0004**
Pre-VMUR	**< 0.001**	0.3374	0.4935	**0.0091**	**< 0.001**	**< 0.001**	**< 0.001**
Duration of denervation	< **0.001**	< **0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**0.0004**	0.1809

Data represent *P* value calculated by multivariable analysis.
$P < 0.05$ is deemed as statistically significant.
MPT = maximum phonation time; NHR = noise-to-harmonics ratio; VMUR = voluntary motor-unit recruitment.

TABLE III.
Stratification Analysis on Denervation Duration.

Parameters	Groups	Changes Between Preoperative and Postoperative Values: Preoperative Median (QL, QU)	Changes Between Preoperative and Postoperative Values: Postoperative Median (QL, QU)	*P* Values of Preoperative and Postoperative Comparison	*P* Values of Intergroups Comparison
Overall Grade	A	2.2 (2.0–2.4)	0.0 (0.0–0.4)	< 0.001	*P* total < 0.01
	B	2.2 (2.0–2.4)	0.0 (0.0–0.6)	< 0.001	*P* ab > 0.05; *P* bc < 0.01
	C	2.2 (2.0–2.4)	0.4 (0.2–0.7)	< 0.001	*P* ac < 0.01
	Overall sample	2.2 (2.0–2.4)	0.2 (0.0–0.6)	< 0.001	
Jitter	A	1.87 (1.44–2.66)	0.26 (0.21–0.35)	< 0.001	*P* total < 0.001
	B	1.75 (1.33–2.34)	0.24 (0.15–0.30)	< 0.001	*P* ab > 0.05
	C	1.85 (1.48–2.05)	0.49 (0.39–0.83)	< 0.001	*P* ac < 0.05; *P* bc < 0.05
	Overall sample	1.82 (1.40–2.47)	0.27 (0.22–0.42)	< 0.001	
Shimmer	A	9.47 (8.49–10.90)	2.86 (2.44–3.61)	< 0.001	*P* total < 0.001
	B	9.55 (8.55–11.12)	3.12 (2.34–3.72)	< 0.001	*P* ab > 0.05; *P* bc < 0.05
	C	9.37 (8.37–10.87)	4.82 (4.04–5.54)	< 0.001	*P* ac < 0.05
	Overall sample	9.49 (8.49–10.90)	3.22 (2.52–4.29)	< 0.001	
NHR	A	0.18 (0.14–0.29)	0.02 (0.01–0.02)	< 0.001	*P* total < 0.001
	B	0.17 (0.14–0.25)	0.02 (0.01–0.02)	< 0.001	*P* ab > 0.05; *P* ac < 0.05
	C	0.20 (0.13–0.23)	0.04 (0.03–0.07)	< 0.001	*P* bc < 0.05
	Overall sample	0.18 (0.14–0.25)	0.02 (0.01–0.03)	< 0.001	
MPT	A	6.02 (4.68–6.79)	17.17 (14.97–21.46)	< 0.001	*P* total < 0.001
	B	5.59 (4.46–6.89)	17.37 (14.93–20.75)	< 0.001	*P* ab > 0.05; *P* ac < 0.05
	C	5.45 (4.54–6.69)	14.75 (11.10–17.22)	< 0.001	*P* bc < 0.05
	Overall sample	5.73 (4.60–6.79)	16.56 (14.42–20.33)	< 0.001	
Post-VMUR	A	2.0 (1.0–2.0)	0.0 (0.0–0.0)	< 0.001	*P* total < 0.01
	B	2.0 (1.0–2.0)	0.0 (0.0–0.0)	< 0.001	*P* ab > 0.05; *P* bc < 0.01
	C	2.0 (1.0–2.0)	0.0 (0.0–2.0)	< 0.001	*P* ac < 0.001
	Overall sample	2.0 (1.0–2.0)	0.0 (0.0–1.0)	< 0.001	
Glottal closure	A	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	*P* total > 0.05
	B	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	*P* ab > 0.05
	C	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	*P* bc > 0.05; *P* ac > 0.05
	Overall sample	3.0 (3.0–3.0)	0.0 (0.0–0.0)	< 0.001	

MPT = maximum phonation time; NHR = noise-to-harmonics ratio; QL= low quartile; QU= upper quartile; VMUR = voluntary motor-unit recruitment.
$P < 0.05$ is deemed as statistically significant.

regression analysis. In further stratification analysis, there was a significant difference between group A and C and group B and C ($P < 0.05$), but not between group A and B ($P > 0.05$).

Electromyographic Findings

All patients received laryngeal electromyographic examinations preoperatively. However, postoperative EMG results were available in only 148, 97, and 41 patients in group A, B, and C, respectively. The electrical activity in the TA muscles of the affected vocal folds during EMG in the present study was divided into two types: spontaneous activity and VMUR. Spontaneous activity, which included positive waves, fibrillations, and complex repetitive discharges, could not be recorded; and VMUR were improved in each group postoperatively ($P < 0.05$). Denervation duration was identified as a significant variable with regard to the parameter postoperative VMUR in multivariable logistic regression analysis. In further stratification analysis, there was a significant difference between group A and C and group B and C ($P < 0.05$), but not between group A and B ($P > 0.05$).

DISCUSSION

Laryngeal reinnervation with ansa cervicalis to RLN anastomosis was reintroduced by Crumley[2] and followed by many other investigators,[4–6] and it has been presented as a good procedure to achieve a normal voice quality. Previous studies in animal models[7,10,11] and in humans[12] have shown that reinnervation can potentially restore laryngeal function when it is performed immediately after RLN injury. In the clinical setting, most UVFP patients suffer from hoarseness and aspiration for several months or even years before they seek medical help. Even in UVFP patients who are diagnosed early,

surgery is usually postponed for at least 6 months to allow for possible spontaneous recovery or compensation from the contralateral vocal fold. However, the studies on regeneration of other peripheral nerves showed that the degree of functional reinnervation lessens as the period of denervation increases; and there appears to be a time range beyond which effective reinnervation declines dramatically.[13] But a significant body of evidence indicates that this does not necessarily apply to the larynx.[8,14,15] Clinical and experimental evidences have demonstrated that spontaneous regeneration commonly takes place after RLN injuries.[16–18] Although this type of reinnervation is usually nonfunctional and seldom occurs with laryngeal mobility (termed subclinical reinnervation), it can help to halt or even reverse muscle atrophy and/or fibrosis caused by denervation.[16] Therefore, the researchers think that denervation duration does not affect the surgical outcome of laryngeal reinnervation in a linear fashion. However, so far we have not seen any report regarding the stratification analysis of denervation duration on the laryngeal reinnervation effect. Thus, it is of great clinical importance to explore whether the same situation in the regeneration of other peripheral nerves also happens to the recurrent laryngeal nerve—that regeneration capacity declines progressively as the denervation duration increases.

Delayed reinnervation procedures have proven effective after peripheral nerve injury in animal experiments. For example, selective reinnervation of the posterior cricoarytenoid muscle with a phrenic nerve transfer has been feasible after a 9-month delay in cat models; however, functional recovery was less successful than with immediate reinnervation.[14] We previously reported that laryngeal reinnervation is still possible to some degree, even after an 18-month denervation period in dogs; however, the degree of RLN regeneration is less than those with an 8-month denervation period.[8] In the clinical studies, Maronian et al. reported on nine patients, eight of whom had an interval between RLN injury and surgery that exceeded 12 months. These patients had a normal or improved voice after laryngeal reinnervation. The longest denervation interval in that series was 9 years, and the postoperative voice in that case was improved.[15] Olson et al. reported excellent acoustic and perceptual results in patients with the maximal interval of 6 years between injury and surgery.[19] Our study of a large sample of UVFP patients, in which the longest denervation course was more than 3 years, confirmed that delayed reinnervation can be effective.[4] Nevertheless, the relationship between denervation duration and degree of functional recovery of the laryngeal muscle in UVFP patients remains to be elucidated.

Results of the present study showed that there was no significant difference with regard to glottal closure among the three groups. This was probably due to a lack of standardization of inspiratory effort while the patients were undergoing videostroboscopy examination. In addition, a difference in the vertical plane of the vocal folds can result in a significant glottic gap, even when the apparent closure as viewed from above seems adequate.[17] However, all of the parameters of vocal function assessment, including perceptual evaluation, objective acoustic analysis, and aerodynamics parameter MPT, showed that denervation duration was an influential factor to the surgical outcome of laryngeal reinnervation. Data of postoperative motor-unit recruitment also support the vocal function results. The perceptual and acoustic parameters showed no significant difference postoperatively among patients with denervation intervals of less than 24 months, and the parameter values in these patients were better than those in patients with longer denervation intervals. These results indicate that delayed reinnervation is still effective. There are several reasons that may support delayed laryngeal reinnervation. There may be an inherent cellular mechanism for preserving the structure of denervated laryngeal muscles.[20] Johns et al. found that 6 months after RLN resection there was no significant difference in maximal isometric force of the TA muscle between the experimental and control cats,[21] possibly due to spontaneous regeneration of the RLN. Our previous study indicated a strong tendency for regeneration in the RLN following injury, which may at least partially reinnervate the laryngeal muscle, helping to maintain its structural integrity and function and to alleviate excessive muscle atrophy and fibrosis.[16] In addition, laryngeal muscle stem cells provide persistent regenerative potential for delayed laryngeal reinnervation for up to 2 years after denervation, as revealed by our previous study.[22] The population of activated muscle stem cells in the laryngeal muscles may be more resistant to apoptosis than those in limb muscles, which may contribute to regenerative myogenesis in denervated laryngeal muscles through compensatory mechanisms. [22]

Nevertheless, after 2 years of denervation, the surgical outcomes were less favorable in the present study, although most postoperative parameters in these patients showed improvement compared with the corresponding preoperative values. As fixation of the cricoarytenoid joint was precluded preoperatively in these cases, the compromise of the recovery of voice quality might have been due to insufficient laryngeal reinnervation, which was confirmed by postoperative EMG.

One cause of poor functional recovery after excessive long-term muscle denervation is the failure of many regenerating axons to elongate and/or make synaptic connections with denervated muscle fibers. The ability of nerve sheaths to support axon regeneration to long-term denervated muscle fibers may progressively deteriorate because of: a decrease in the number of Schwann cells to a level that cannot provide adequate support for regenerating axons[23]; degeneration and collagenization of endoneurial tubes, which may obstruct axonal regeneration[24]; and an inability of the basal lamina to be renewed without Schwann cell–axon contact.[25] These factors contribute to a profound reduction in the number of axons that eventually reach denervated muscles.[26] Another possible explanation is occupation of the denervated muscle end plates by axons coming from adjacent nerves or by fibers of autonomous origin, precluding delayed reinnervation.[27] In addition, muscle fiber

atrophy and irreversible muscle fibrosis and degeneration of muscle endplates during long-term denervation may hinder successful reinnervation.[28] Even after reinnervation, muscle fibers may fail to resume their normal size, possibly because of the progressive exhaustion of satellite cells whose activity and status may be the key determinant of skeletal muscle regeneration potential. Our previous work showed that the levels of myoD and myogenin, which are markers for activated satellite cells, were upregulated 6 to 12 months after denervation and then downregulated over time, ultimately becoming undetectable by 2 years after denervation. This indicates a decreased myogenic ability after a 2-year denervation duration, [22] which might explain the better surgical outcome of laryngeal reinnervation on patients with a denervation duration less than 2 years than on patients with a denervation duration longer than 2 years. Therefore, for UVFP patients with a denervation course of more than 2 years, it may be better to combine reinnervation surgery with arytenoid adduction.[29,30]

In addition, there are some other factors that may also affect the surgical outcome of laryngeal reinnervation, such as the age of patients and the severity of nerve injury.[31–33] Crumley recommended that patients' age should be less than 70 in order to ensure the effectiveness of laryngeal reinnervation,[2] while Paniello et al. revealed that patients under age 52 had significantly better voice recovery than those over age 52.[34] Stratification analysis on age in the present series of patients revealed that laryngeal reinnervation is less effective when patients are older than 60 years old. Details of further stratification analysis on these two identified influential factors will be presented in other reports (unpublished).

CONCLUSION

The data from this study indicate that surgical outcome of laryngeal reinnervation is affected by denervation duration, the age of patients, and the severity of nerve injury. Although delayed reinnervation is effective, surgical outcome is better when the procedure is performed within 2 years after nerve injury than when the procedure is performed over 2 years.

BIBLIOGRAPHY

1. Crumley RL. Unilateral recurrent laryngeal nerve paralysis. *J Voice* 1994; 8:79–83.
2. Crumley RL. Update: ansa cervicalis to recurrent laryngeal nerve anastomosis for unilateral laryngeal paralysis. *Laryngoscope* 1991;101:384–387; discussion 388.
3. Zheng H, Li Z, Zhou S, Cuan Y, Wen W. Update: laryngeal reinnervation for unilateral vocal cord paralysis with the ansa cervicalis. *Laryngoscope* 1996;106:1522–1527.
4. Wang W, Chen D, Chen S, et al. Laryngeal reinnervation using ansa cervicalis for thyroid surgery-related unilateral vocal fold paralysis: a long-term outcome analysis of 237 cases. *PLoS One* 2011;6:e19128.
5. Lee WT, Milstein C, Hicks D, Akst LM, Esclamado RM. Results of ansa to recurrent laryngeal nerve reinnervation. *Otolaryngol Head Neck Surg* 2007;136:450–454.
6. Lorenz RR, Esclamado RM, Teker AM, et al. Ansa cervicalis-to-recurrent laryngeal nerve anastomosis for unilateral vocal fold paralysis: experience of a single institution. *Ann Otol Rhinol Laryngol* 2008;117:40–45.
7. Zheng H, Li Z, Zhou S, Cuan Y, Wen W, Lan J. Experimental study on reinnervation of vocal cord adductors with the ansa cervicalis. *Laryngoscope* 1996;106:1516–1521.
8. Zheng H, Zhou S, Chen S, Li Z, Cuan Y. An experimental comparison of different kinds of laryngeal muscle reinnervation. *Otolaryngol Head Neck Surg* 1998;119:540–547.
9. Wang W, Chen S, Chen D, et al. Contralateral ansa cervicalis-to-recurrent laryngeal nerve anastomosis for unilateral vocal fold paralysis: a long-term outcome analysis of 56 cases. *Laryngoscope* 2011;121:1027–1034.
10. Mahieu HF, van Lith-Bijl JT, Groenhout C, Tonnaer JA, de Wilde P. Selective laryngeal abductor reinnervation in cats using a phrenic nerve transfer and ORG 2766. *Arch Otolaryngol Head Neck Surg* 1993;119: 772–776.
11. van Lith-Bijl JT, Stolk RJ, Tonnaer JA, Groenhout C, Konings PN, Mahieu HF. Selective laryngeal reinnervation with separate phrenic and ansa cervicalis nerve transfers. *Arch Otolaryngol Head Neck Surg* 1997; 123:406–411.
12. Yumoto E, Sanuki T, Kumai Y. Immediate recurrent laryngeal nerve reconstruction and vocal outcome. *Laryngoscope* 2006;116:1657–1661.
13. Keilhoff G, Fansa H. Successful intramuscular neurotization is dependent on the denervation period. A histomorphological study of the gracilis muscle in rats. *Muscle Nerve* 2005;31:221–228.
14. van Lith-Bijl JT, Stolk RJ, Tonnaer JA, Groenhout C, Konings PN, Mahieu HF. Laryngeal abductor reinnervation with a phrenic nerve transfer after a 9-month delay. *Arch Otolaryngol Head Neck Surg* 1998; 124:393–398.
15. Maronian N, Waugh P, Robinson L, Hillel A. Electromyographic findings in recurrent laryngeal nerve reinnervation. *Ann Otol Rhinol Laryngol* 2003;112:314–323.
16. Chen D, Chen S, Wang W, Zhang C, Zheng H. Spontaneous regeneration of recurrent laryngeal nerve following long-term vocal fold paralysis in humans: histologic evidence. *Laryngoscope* 2011;121:1035–1039.
17. Woodson GE. Spontaneous laryngeal reinnervation after recurrent laryngeal or vagus nerve injury. *Ann Otol Rhinol Laryngol* 2007;116:57–65.
18. Kupfer RA, Old MO, Oh SS, Feldman EL, Hogikyan ND. Spontaneous laryngeal reinnervation following chronic recurrent laryngeal nerve injury. *Laryngoscope* 2013;123:2216–2227.
19. Olson DE, Goding GS, Michael DD. Acoustic and perceptual evaluation of laryngeal reinnervation by ansa cervicalis transfer. *Laryngoscope* 1998; 108:1767–1772.
20. Miyamaru S, Kumai Y, Ito T, Yumoto E. Effects of long-term denervation on the rat thyroarytenoid muscle. *Laryngoscope* 2008;118:1318–1323.
21. Johns MM, Urbanchek M, Chepeha DB, Kuzon WM Jr, Hogikyan ND. Thyroarytenoid muscle maintains normal contractile force in chronic vocal fold immobility. *Laryngoscope* 2001;111:2152–2156.
22. Donghui C, Shicai C, Wei W, et al. Functional modulation of satellite cells in long-term denervated human laryngeal muscle. *Laryngoscope* 2010; 120:353–358.
23. Salonen V, Lehto M, Kalimo M, Penttinen R, Aro H. Changes in intramuscular collagen and fibronectin in denervation atrophy. *Muscle Nerve* 1985;8:125–131.
24. Sunderland S, Bradley KC. Endoneurial tube shrinkage in the distal segment of a severed nerve. *J Comp Neurol* 1950;93:411–420.
25. Bunge MB, Williams AK, Wood PM. Neuron-Schwann cell interaction in basal lamina formation. *Dev Biol* 1982;92:449–460.
26. Fu SY, Gordon T. Contributing factors to poor functional recovery after delayed nerve repair: prolonged denervation. *J Neurosci* 1995;15:3886–3895.
27. Hydman J, Remahl S, Bjorck G, Svensson M, Mattsson P. Nimodipine improves reinnervation and neuromuscular function after injury to the recurrent laryngeal nerve in the rat. *Ann Otol Rhinol Laryngol* 2007; 116:623–630.
28. Anzil AP, Wernig A. Muscle fibre loss and reinnervation after long-term denervation. *J Neurocytol* 1989;18:833–845.
29. Chhetri DK, Gerratt BR, Kreiman J, Berke GS. Combined arytenoid adduction and laryngeal reinnervation in the treatment of vocal fold paralysis. *Laryngoscope* 1999;109:1928–1936.
30. Hassan MM, Yumoto E, Kumai Y, Sanuki T, Kodama N. Vocal outcome after arytenoid adduction and ansa cervicalis transfer. *Arch Otolaryngol Head Neck Surg* 2012;138:60–65.
31. Korte N, Schenk HC, Grothe C, Tipold A, Haastert-Talini K. Evaluation of periodic electrodiagnostic measurements to monitor motor recovery after different peripheral nerve lesions in the rat. *Muscle Nerve* 2011;44:63–73.
32. Kovacic U, Sketelj J, Bajrovic FF. Chapter 26: Age-related differences in the reinnervation after peripheral nerve injury. *Int Rev Neurobiol* 2009; 87:465–482.
33. van Lith-Bijl JT, Mahieu HF, Stolk RJ, Tonnaer JA, Groenhout C, Konings PN. Laryngeal abductor function after recurrent laryngeal nerve injury in cats. *Arch Otolaryngol Head Neck Surg* 1996;122:393–396.
34. Paniello RC, Edgar JD, Kallogjeri D, Piccirillo JF. Medialization versus reinnervation for unilateral vocal fold paralysis: a multicenter randomized clinical trial. *Laryngoscope* 2011;121:2172–2179.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 153(3):414-419.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–
Head and Neck Surgery
2015, Vol. 153(3) 414–419
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599815593268
http://otojournal.org
SAGE

Diagnostic Yield of Computed Tomography in the Evaluation of Idiopathic Vocal Fold Paresis

Paul M. Paddle, MD, MPH[1], Masaany B. Mansor, MD[1], Phillip C. Song, MD[1], and Ramon A. Franco Jr, MD[1]

No sponsorships or competing interests have been disclosed for this article.

Received October 26, 2014; revised May 6, 2015; accepted June 5, 2015.

Abstract

Objective. To determine the diagnostic yield of computed tomography (CT) in establishing an etiology in patients with idiopathic unilateral vocal fold paresis (IUVFP). To determine the proportion of CT scans yielding incidental findings requiring further patient management.

Study Design. Case series with chart review.

Setting. Tertiary laryngology practice.

Subjects. Laryngology clinic patients under the care of the 2 senior authors.

Methods. All clinic patients were identified who had a diagnosis of IUVFP and underwent CT of the skull base to the upper mediastinum from 2004 to 2014. Demographic, historical, examination, and investigation data were extracted. CT reports and endoscopic recordings were reviewed. Patients were excluded if there were insufficient clinical findings recorded or if there was a known neurologic disorder, complete vocal fold immobility, or bilateral involvement.

Results. A total of 174 patients with IUVFP who had also undergone contrast-enhanced CT were identified. Of the 174 patients, 5 had a cause for their paresis identified on CT. This equated to a diagnostic yield of 2.9% (95% confidence interval, 0.94% to 6.6%). Of the 174 patients, 48 had other incidental lesions identified that required further follow-up, investigation, or treatment. This equated to an incidental yield of 27.6% (95% confidence interval, 21.1% to 34.9%).

Conclusion. This is the second and largest study to evaluate the diagnostic yield of CT in the evaluation of IUVFP. It demonstrates a low diagnostic yield and a high incidental yield. These findings suggest that the routine use of CT in the evaluation of idiopathic vocal fold paresis should be given careful consideration and that a tailored approach to investigation with good otolaryngologic follow-up is warranted.

Keywords

idiopathic unilateral vocal fold paresis, computed tomography, diagnostic yield, incidental yield

Vocal fold paresis implies vocal fold hypomobility due to neurologic injury, with a peripheral etiology in 90% of cases. It may result from weakness of the vagus nerve or its superior or recurrent laryngeal branches. This may occur anywhere in its course—from the lower motor neurons in the nucleus ambiguus of the medulla through the jugular foramen, neck, and mediastinum. Vocal fold paresis is unilateral in 90% of cases.[1]

Paresis is the most common cause of vocal fold hypomobility, being present in 90% of cases.[1] Other causes of vocal fold hypomobility include myopathies (4%) and cricoarytenoid joint dysfunction (6%). Paresis is an increasingly recognized phenomenon in patients with laryngologic complaints. Previous studies reported mild vocal fold hypomobility in 46% of patients with vocal complaints, 71% of singing teachers with complaints of technical difficulties, and 23% of singing teachers with no vocal complaints.[1,2]

Diagnosis of vocal fold paresis requires a high index of suspicion. The symptoms of vocal fold paresis are more varied and subtle than paralysis. Classic symptoms of glottic insufficiency—such as breathy dysphonia, diplophonia, aspiration, and dysphagia—may be absent or muted. Instead, the patient may complain of a loss of quality volume and range, vocal instability, and increased phonatory effort.[3] Atypical symptoms, such as globus, chronic cough, and laryngospasm, are also described.

Unlike the findings of an established vocal fold paralysis, such as vocal fold atrophy, bowing, and arytenoid prolapse, the examination findings of unilateral vocal fold paresis are subtle and difficult to discern from nonpathologic asymmetries. Findings may include asymmetric vocal fold range and velocity of movement, decreased ipsilateral false vocal

[1]Harvard Medical School, Division of Laryngology, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, USA

Corresponding Author:
Paul M. Paddle, MD, Harvard Medical School, Division of Laryngology, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA.
Email: paulpaddle@me.com

fold and supraglottic contraction, pushing rotation of the petiole toward the weakened side, and asymmetric fatigability on repetitive movement. Findings associated with compensatory hyperfunction may also be present, including contralateral supraglottic contraction, and benign vocal fold lesions such as nodules, polyps, or pseudocysts.[4] On stroboscopy, asymmetry of mucosal wave motion may be the only manifestation. The diagnostic accuracy of these signs is controversial, but when they are identified on nasoendoscopy, along with a suggestive history, a diagnosis of paresis is made.

Idiopathic paresis is diagnosed when no cause is found on thorough history and examination. In our clinic, paresis with a history of preceding upper respiratory infection is defined as idiopathic, as there is no definitive test to confirm causality.

Laryngeal electromyography (LEMG) is used as a diagnostic and prognostic tool in cases of vocal fold paralysis. There is no consensus on the use of LEMG in the context of vocal fold hypomobility. When performed correctly, LEMG can confirm the presence and laterality of a neuropathy and identify neuromuscular junction abnormalities and myopathies, as well as ongoing degeneration or regeneration. Some advocates propose that it be employed systematically in paresis,[3] while others use it in situations where the results would alter patient management.[5] Certainly, it does not obviate the role of imaging studies in the evaluation of vocal fold paresis.

The rationale of imaging in paresis is twofold: First, paresis may be an early sentinel of an underlying pathology that, where identified, would require further investigation and management in its own right, particularly neoplasia. Second, finding an underlying pathologic process may guide management of the paresis itself. The role of computed tomography (CT) in the evaluation of vocal fold paralysis is well established, given a high overall diagnostic yield (35% to 62%)[6,7] and a high proportion of neoplastic causes (13% to 33%).[8-11] However, its role in the evaluation of paresis is not clearly established, and current practice seems to be extrapolated from the paralysis literature. A single previous study assessed the diagnostic yield of CT in the investigation of paresis.[12] In our institution, CT is performed when there is a clinical diagnosis of paresis but the cause remains ''idiopathic'' after thorough history and examination—that is, no clear history of preceding nerve injury or other compressive or infiltrative lesion and no evidence of a cause on otolaryngologic, neurologic, and chest examination and video endoscopy. Patients may also refuse or strongly desire a CT study.

Our study objectives are twofold: first, to establish a diagnostic yield in performing CT in patients with idiopathic vocal fold paresis; second, to establish a percentage yield of incidental lesions requiring further management in this cohort of patients. This has important clinical, cost, and medicolegal implications.

Method

This study was approved by the Massachusetts Eye and Ear Infirmary Institutional Review Board. With a precision-based sample size calculation based on an expected diagnostic yield of approximately 2.0%,[12] an acceptable precision of 1.99%, and a confidence level of 95%, an estimated 191 patients were required. The practice records from January 2004 to January 2014 of 2 senior laryngologists from a single tertiary practice were reviewed. All adult patients were identified who had a clinical diagnosis of idiopathic unilateral vocal fold paresis (IUVFP) and underwent contrast-enhanced CT from skull base to mediastinum. Patients were excluded if they had bilateral vocal fold hypomobility due to the decreased reliability of clinical assessment and the higher likelihood of a central etiology.[13] Patients were also excluded if there was a history of a neurologic diagnosis, such as laryngeal dystonia or tremor, myoclonus, parkinsonism, stroke, or other central neurologic process.

In each case, a diagnosis of paresis was made by a senior laryngologist, using the above-described symptoms and signs. CT images and reports were reviewed. In each case, any etiology for paresis and any incidental finding were recorded. An etiology for paresis was defined as any lesion along the expected extracranial course of the ipsilateral superior or recurrent laryngeal nerve or vagus, which could be causing pathologic compression, invasion, stretch, or inflammation. An incidental lesion was defined as any clinically silent lesion, not associated with the diagnosis of paresis, but that could lead to further diagnostic or therapeutic intervention.

Longitudinal review of files was also undertaken to identify evolution of findings or interval evidence of an etiology for the diagnosed paresis. Due to the evolution in endoscopic diagnostic criteria for paresis over the 10 years of the study, a sensitivity analysis was performed comparing the mean diagnostic yield of the first 5 years with that of the second 5 years. The null hypothesis of no difference between the means was tested with an unpaired 2-sample *t* test. Excel 2010 and Stata 10.0 were used for data storage and statistical analysis.

Results

Patients (n = 237) with unilateral paresis were identified over the period January 2004 to January 2014. Of these, 174 (73%) underwent contrast-enhanced CT scans of skull base to mediastinum and were included in the study. The other 63 (27%) did not undergo CT due to either a clear etiology of their paresis or patient refusal. There was no systematic difference in the demographic characteristics of the CT and non-CT workup patient populations. In the CT workup group of patients, the mean age at diagnosis was 54.5 years (range, 21 to 82). There were a greater proportion of women (56%), while laterality of paresis was evenly distributed, with 51% of lesions being left sided (**Table 1**).

The most common patient symptom was hoarseness. Symptoms of glottic inefficiency were also common, such as vocal fatigue, increased phonatory effort, and decreased projection (**Table 2**). Patients less commonly complained of loss of range, cough, laryngospasm, globus/dysphagia, and pain.

Table 1. Demographic Data of IUVFP by Workup Groups: CT and Non-CT.[a]

	CT	Non-CT
Total unilateral paresis	174	63
Age, y[b]	54.5 (21-82)	53.9 (18-75)
Women	98 (56)	33 (52)
Left laterality	88 (51)	33 (52)

Abbreviations: CT, computed tomography; IUVFP, idiopathic unilateral vocal fold paresis.
[a]Results presented as n (%), except where noted otherwise.
[b]Mean (range).

Table 2. Symptoms of Paresis Subjects Included in this Study.

Symptom	n (%)
Hoarseness	144 (83)
Vocal fatigue	113 (65)
Increased phonatory effort	67 (39)
Decreased vocal projection	63 (36)
Loss of range	45 (26)
Cough	44 (25)
Dysphagia	37 (21)
Pain: odynophagia / odynophonia / laryngeal strain	20 (11)
Breathlessness during voicing	13 (7)
Laryngospasm	11 (6)

Table 3. Videostroboscopic Findings of Paresis Subjects Included in this Study.

Examination Feature	n (%)
Asymmetry of velocity of movement	138 (79)
Increased glottic show	129 (74)
Asymmetry of range of movement	117 (67)
Phase asymmetry	91 (66)
Supraglottic hyperfunction	72 (41)
Incomplete closure	65 (37)
Bowing/atrophy of vocal fold	45 (26)
Deviation	26 (15)
Increased vibratory amplitude	17 (12)

On video endoscopy, common findings were subtle vocal fold range-of-motion asymmetries, asymmetrically increased glottic show, and asymmetric velocity of motion. These were seen in 79%, 74%, and 67% of patients, respectively. Asymmetries at rest were less prevalent as were signs of bowing and incomplete closure (**Table 3**).

Of the 174 patients, 5 had CT that revealed an etiologic lesion for their vocal fold paresis, a diagnostic yield of 2.9% (95% confidence interval, 0.94 to 6.6; **Table 4**). Of these 5 CT-positive cases, 1 was positive for malignancy. The patient had an exophytic thyroid nodule with possible compression of the recurrent laryngeal nerve in the ipsilateral tracheoesophageal groove. This nodule was positive for papillary thyroid carcinoma on fine-needle aspiration. The patient underwent total thyroidectomy and adjuvant radioactive iodine. His paresis did not improve on serial follow-up.

Four CT-positive cases were benign: 1 was due to previous thoracic aortic aneurysm repair with dense scarring on CT in the aortopulmonary window. Two cases were due to tracheoesophageal groove masses. One mass was an exophytic thyroid nodule and associated tracheoesophageal groove lymph node. The patient underwent a right hemithyroidectomy and prelaryngeal lymph node dissection. The final histopathology was a benign follicular adenoma. The other tracheoesophageal groove case was due to a large parathyroid adenoma that, on removal, was seen to be stretching the recurrent laryngeal nerve. A final case was due to an undiagnosed Arnold Chiari II malformation with tentorium crowding and tonsillar herniation. This patient was referred to neurosurgery and underwent urgent posterior fossa decompression. It is interesting to note that there were no other neurologic symptoms or signs nor evidence of bilateral paresis. In all of the above 3 benign cases undergoing surgery, there was no recovery of function of the nerve after surgical intervention. The diagnostic yield equates to a number needed to treat of 34. In other words, to find 1 patient with a vocal fold paresis-associated lesion, 34 patients had to undergo CT.

In contrast, 48 of 174 patients had a new incidental finding on CT that required further management. Further management was defined as serial clinical examination, repeat imaging, a diagnostic procedure, or operation. This equates to an incidental yield of 27.6% (95% confidence interval, 23.7% to 37.8%).

Of these 48 patients, 40 underwent clinical and or serial imaging follow-up alone; 5 underwent fine-needle aspiration alone; and 3 underwent surgery. The range of incidental lesions included pulmonary nodules, thyroid nodules, and other mediastinal and cervical lesions, predominantly lymphadenopathy (**Table 5**). Over the mean 2.95 years of follow-up (SD, 1.52), none of these patients developed a symptomatic or clinically significant pathology. Of the 3 patients who underwent surgery, 1 underwent hemithyroidectomy for a follicular adenoma that had no extracapsular extension and was not compressing on the tracheoesophageal groove. A second patient underwent total thyroidectomy for a dominant intrathyroid nodule that was positive for papillary carcinoma on fine-needle aspiration, and the third patient underwent resection of a benign, submucosal false fold lipoma. The number needed to ''harm'' was 4.

A sensitivity analysis of diagnostic yield revealed a yield of 2.2% for the first 5 years of the study, compared with a yield of 5.1% for the second 5 years of the study. An unpaired 2-sample *t* test of the difference between these 2 means (2.9%) resulted in a *P* value of .34.

Table 4. Yield of CT in IUVFP.

Yield	Positive CT, n	Yield, % (95% CI)	Needed to Treat/Harm, n
Diagnostic	5[a]	2.9 (0.94-6.6)	34
Incidental	48[b]	27.6 (21.1-34.9)	4

Abbreviations: CI, confidence interval; CT, computed tomography; IUVFP, idiopathic unilateral vocal fold paresis.
[a]Diagnostic of benign lesion, n = 4; diagnostic of a malignancy, n = 1.
[b]Underwent clinical or imaging follow-up alone, n = 40; underwent fine-needle aspiration alone, n = 5; required operation, n = 3.

Table 5. Proportion of Incidental Lesions by Anatomic Group.

Incidental Lesion Type	Proportion of All Incidental Lesions
Thyroid abnormalities: nodules, cysts, enlargement	29
Pulmonary lesions: nodules, granulomas, pleural plaques, hilar lymphadenopathy	40
Mediastinal lesions: thoracic aortic aneurysms, mediastinal lymphadenopathy	15
Cervical abnormalities: laryngocele, thyroglossal duct cyst, cervical lymphadenopathy	13
Miscellaneous: vertebral lesions	4

Table 6. Etiology of Vocal Fold Paralysis vs Paresis (in Percentages).

	Paralysis	Paresis			
Etiology	MacGregor[10] (n = 1308)	Koufman[4] (n = 50)	Heman-Ackah[1] (n = 46)	Badia[12] (n = 176)	Present Study (n = 237)
Total iatrogenic	22	20	4.3	39.2	3.4
Total neoplastic lesions	21.7	6	13	1.1	2.9
Total nonneoplastic benign disease	39.9	6	54.3	13.6	5.7
Idiopathic[a]	16.4	68	28.3	46	88

[a]Includes viral neuritis.

Discussion

The role of CT in the evaluation of vocal fold paralysis is well established. Its near routine use is justified by a high diagnostic yield (35% to 62%)[6,7] and a high proportion of cases due to underlying neoplasia (13% to 33%).[8,10] Its role in paresis, however, is unclear, partly because the prevalence of a neoplastic etiology in published studies is a comparatively low (1.1% to 6%; **Table 6**)[1,4,12] and partly because there is a paucity of studies examining the diagnostic yield of CT paresis. A previous study found the diagnostic yield of CT in IUVFP to be 1.7%.[12]

The overall diagnostic yield of CT in the evaluation of paresis in the current study was 2.9% (95% confidence interval, 0.94% to 6.6%). Such values may justify the use of routine CT evaluation. A similar percentage yield (1% to 4%) is seen by many authors as justification for the routine use of magnetic resonance imaging in the evaluation of asymmetric sensorineural hearing loss.[14] This argument for routine CT in paresis might be strengthened from a qualitative perspective, when one considers that 4 of 5 (80%) of our CT-positive cases had a pathology that required operation, which untreated may have led to serious morbidity. Conversely, the study by Badia et al[12] described a final yield of 0%, as the single CT-positive case was benign and required no intervention.

The argument against the routine use of CT in the evaluation of IUVFP is strengthened when one considers the discovery of incidental lesions and the potential morbidities due to the performance of additional diagnostic tests and interventions. This study identified an incidental yield of 27.6%. The majority of these patients with incidental imaging findings (83.3%) underwent serial examination and imaging alone (including serial thyroid ultrasound, CT of the neck and chest, and magnetic resonance imaging of the brain), thus exposing them to additional perhaps unnecessary radiation and expense. In all of these imaging-alone cases, the incidental lesions did not evolve into clinically significant pathologies over the mean laryngologic follow-up period of 2.95 years (SD, 1.52). Five patients underwent

fine-needle aspiration for thyroid nodules. In all but 1 case, these lesions were benign, and no direct complications from the aspiration procedure were recorded. Two patients underwent surgery for benign, otherwise asymptomatic lesions. In both cases, there were no complications from surgery. Only 1 of the 48 incidental cases was an incidental finding clearly beneficial to the patient. In this case, the patient had a serendipitous incidental finding of a thyroid malignancy, which was successfully treated with total thyroidectomy. Although it is not possible to definitively state how many of these lesions may have become clinically significant at a later date, none of the lesions became symptomatic or required further action over the period of follow-up. And while it can be argued that the incidental discovery of an early thyroid malignancy or a small descending thoracic aortic aneurysm is beneficial for the patient and may lead to improved survival, it cannot be argued that this justifies routine CT in the evaluation of IUVFP. One must be clear that, in this context, CT investigation is a diagnostic test, trying to answer the question of etiology of paresis, and is not a screening investigation.

While not a focus of this study, the underlying etiologies of vocal fold paresis further inform the finding of a low diagnostic yield. Compared with paresis, a larger proportion of paralysis cases is due to neoplastic causes, and a smaller proportion is idiopathic.[8,15] In the 4 published studies (including this one), there is an increasing proportion of idiopathic cases and a much smaller proportion of cases due to neoplasia. In our study, the proportion of idiopathic cases that remained idiopathic was 88%, and the proportion of cases due to neoplasia was 2.9%. Thus, there is already a low likelihood that CT will uncover an underlying malignancy.

This is the largest published study to evaluate the role of CT imaging in the investigation of unilateral vocal fold paresis. It has an adequate sample size for the desired precision and has a rigorous study method. Only 1 other study addressed the diagnostic yield of CT in the evaluation of IUVFP. Other strengths of this study include the long-term follow-up of patients with incidental findings and the use of 2 laryngologists from a single institution. Given the controversy in the literature in making a clinical diagnosis of paresis, having only 2 assessors may limit generalizability, but this singular clinical definition of paresis maximizes internal validity.

A possible weakness of this study is the lack of routine LEMG in the confirmation of a diagnosis of paresis, possibly leading to misclassification bias. Certain authors suggested that LEMG is an essential diagnostic tool in the evaluation of paresis and espoused its use in every case,[3,4] emphasizing that there is a marked discrepancy between clinical observations of paresis and LEMG findings, with discordance in 25% to 40% of cases.[2,4] This discrepancy, however, is not in the absence or presence of neuropathy in the larynx but on the paretic side and the nerve involved. In fact, there is excellent concordance between a clinical diagnosis of paresis and LEMG findings of the presence of a neuropathy, with 1 study demonstrating LEMG confirmation of a clinically diagnosed mild paresis in 86.4% of cases[1,5] and another demonstrating an LEMG-confirmed neuropathy in 100% of clinically diagnosed pareses. Dursun et al demonstrated that a thorough neurolaryngeal examination in the hands of an experienced laryngologist can diagnose superior laryngeal nerve paresis in 98% of cases with characteristic examination findings.[16] Merati et al demonstrated that 92% of patients with clinical vocal fold motion impairment had a neuropathy.[17] Furthermore, the ability to obtain reliable and accurate LEMG results in the larynx requires a significant degree of experience and expertise, specifically in the interpretation of laryngeal data, and it is heavily dependent on accurate and consistent needle placement, which limits its usefulness and availability in many centers. Last, it is an invasive test, not without attendant morbidity, and it does not often address the etiology of the paresis. Thus, it would not obviate the need for further imaging and laboratory testing in confirmed cases.

A possible source of selection bias lies in the fact that 27% of paresis cases did not undergo CT. These cases likely had a clear etiology on history and examination, such as preceding surgical injury or known cervicothoracic malignancy. While this may raise the diagnostic yield of CT by excluding cases with a higher likelihood of a negative finding, they are not true idiopathic cases, and CT in this context would not likely provide additional diagnostic information, which is the key clinical question.

Last, a possible source of misclassification bias lies in the fact that the clinical diagnosis of vocal fold paresis has evolved over the 10-year sample period. That is, we have had a higher index of suspicion and perhaps a lower threshold for the diagnosis of paresis in recent years, and patients with more subtle findings of paresis may have been excluded in earlier years of the study. However, a sensitivity analysis was performed comparing the diagnostic yields of the first 5 years of the study (2.2%) and the second 5 years (5.1%), and it revealed no statistically significant difference between these 2 periods (2.9%, $P = .34$).

In the light of the above findings, our current investigation of IUVFP involves a tailored approach. Patients who have additional localizing symptoms and examination findings are investigated in a targeted manner with imaging and appropriate blood tests. In true clinically idiopathic cases, we now offer the patient an informed choice between repeat videostroboscopic evaluation and initial CT, having discussed the diagnostic and incidental yields of imaging. If there is a convincing history of a significant preceding upper respiratory infection and definitive coincident onset of laryngologic symptoms, we are more likely to advise repeat observation over up-front CT, as the underlying etiology is more likely a viral neuritis. Reliability of follow-up is also an important consideration.

On serial examination, if a paresis evolves into paralysis or bilateral findings, our suspicion of an underlying nefarious lesion is heightened, and we then perform CT and other investigations as appropriate. Future prospective multicenter studies validating clinical diagnosis of paresis against laryngeal EMG and CT are warranted.

Conclusion

This study demonstrates that contrast-enhanced CT of the base of the skull to the upper mediastinum has a low yield of 2.9% in the initial evaluation of idiopathic vocal fold paresis. CT has a high rate (27.6%) of incidental asymptomatic cervicothoracic findings. These findings suggest that routine use of CT in the evaluation of idiopathic vocal fold paresis should be given careful consideration and that a tailored approach to investigation with good otolaryngologic follow-up is warranted.

Author Contributions

Paul M. Paddle, study design, data acquisition, analysis and interpretation, drafting and revision, final approval and accountability for all aspects of the work; **Masaany B. Mansor**, study design, data acquisition, manuscript drafting, final approval and accountability for all aspects of the work; **Phillip C. Song**, study design, manuscript revision, final approval and accountability for all aspects of the work; **Ramon A. Franco Jr**, study design, data interpretation, manuscript revision, final approval and accountability for all aspects of the work.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Heman-Ackah YD, Batory M. Determining the etiology of mild vocal fold hypomobility. *J Voice*. 2003;17:579-588.
2. Heman-Ackah YD, Barr A. Mild vocal fold paresis: understanding clinical presentation and electromyographic findings. *J Voice*. 2006;20:269-281.
3. Sulica L, Blitzer A. Vocal fold paresis: evidence and controversies. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:159-162.
4. Koufman JA, Postma GN, Cummins MM, Blalock PD. Vocal fold paresis. *Otolaryngol Head Neck Surg*. 2000;122:537-541.
5. Heman-Ackah YD, Barr A. The value of laryngeal electromyography in the evaluation of laryngeal motion abnormalities. *J Voice*. 2006;20:452-460.
6. Terris DJ, Arnstein DP, Nguyen HH. Contemporary evaluation of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1992;107:84-90.
7. Ramadan HH, Wax MK, Avery S. Outcome and changing cause of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1998;118:199-202.
8. Yumoto E, Minoda R, Hyodo M, Yamagata T. Causes of recurrent laryngeal nerve paralysis. *Auris Nasus Larynx*. 2002; 29:41-45.
9. Rosenthal LH, Benninger MS, Deeb RH. Vocal fold immobility: a longitudinal analysis of etiology over 20 years. *Laryngoscope*. 2007;117:1864-1870.
10. MacGregor FB, Roberts DN, Howard DJ, Phelps PD. Vocal fold palsy: a re-evaluation of investigations. *Journal Laryngol Otol*. 1994;108:193-196.
11. Richardson BE, Bastian RW. Clinical evaluation of vocal fold paralysis. *Otolaryngol Clin North Am*. 2004;37:45-58.
12. Badia PI, Hillel AT, Shah MD, Johns MM 3rd, Klein AM. Computed tomography has low yield in the evaluation of idiopathic unilateral true vocal fold paresis. *Laryngoscope*. 2013; 123:204-207.
13. Hillel AD, Benninger M, Blitzer A, et al. Evaluation and management of bilateral vocal cord immobility. *Otolaryngol Head Neck Surg*. 1999;121:760-765.
14. Jiang ZY, Mhoon E, Saadia-Redleaf M. Medicolegal concerns among neurotologists in ordering MRIs for idiopathic sensorineural hearing loss and asymmetric sensorineural hearing loss. *Otol Neurotol*. 2011;32:403-405.
15. Altman JS, Benninger MS. The evaluation of unilateral vocal fold immobility: is chest X-ray enough? *J Voice*. 1997;11:364-367.
16. Dursun G, Sataloff RT, Spiegel JR, Mandel S, Heuer RJ, Rosen DC. Superior laryngeal nerve paresis and paralysis. *J Voice*. 1996;10:206-211.
17. Merati AL, Shemirani N, Smith TL, Toohill RJ. Changing trends in the nature of vocal fold motion impairment. *Am Journal Otolaryngol*. 2006;27:106-108.

Reprinted by permission of Otolaryngol Head Neck Surg. 2014; 151(2):286-293.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–Head and Neck Surgery
2014, Vol. 151(2) 286–293
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2014
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599814531733
http://otojournal.org
SAGE

Etiology and Time to Presentation of Unilateral Vocal Fold Paralysis

Emily A. Spataro, MD[1], David J. Grindler, MD[1], and Randal C. Paniello, MD[1]

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract

Objective. To determine the etiology, laterality, and time to presentation of unilateral vocal fold paralysis (UVFP) at a tertiary care institution over 10 years.

Study Design. Case series with chart review.

Setting. Academic medical center.

Subjects and Methods. All patients seen between 2002 and 2012 by the Department of Otolaryngology at the Washington University School of Medicine (WUSM), with a diagnosis of unilateral vocal fold paralysis, were included. Medical records were reviewed for symptom onset date, presentation date(s), and etiology of UVFP.

Results. Of the patients, 938 met inclusion criteria and were included. In total, 522 patients (55.6%) had UVFP due to surgery; 158 (16.8%) were associated with thyroid/parathyroid surgery, while 364 (38.8%) were due to nonthyroid surgery. Of the patients, 416 (44.4%) had nonsurgical etiologies, 124 (13.2%) had idiopathic UVFP, and 621 (66.2%) had left-sided UVFP. The diagnosis was more common on the left side in cases of intrathoracic surgeries and malignancies, as expected, but also in idiopathic, carotid endarterectomy, intubation, and skull base tumors. In total, 9.8% of patients presented first to an outside otolaryngologist at a median time of 2.1 months after onset, but these patients presented to WUSM at a median time of 9.5 months. Overall, 70.6% of patients presented to a WUSM otolaryngologist within 3 months of onset.

Conclusion. Iatrogenic injury remains the most common cause of UVFP. Thyroidectomy remains the leading cause of surgery-related UVFP. Patients are typically seen within 3-4 months of onset; however, a significant delay exists for those referred to WUSM.

Keywords

vocal cord, paralysis, etiology

Received October 31, 2013; revised February 3, 2014; accepted March 25, 2014.

The etiology of unilateral vocal fold paralysis (UVFP) is of great interest to the otolaryngologist and has been reported in many studies over the past 40 years.[1-19] Etiologies include thyroid surgery, nonthyroid surgery, trauma, neurologic disease, malignancy, intubation, infection, inflammatory diseases, and idiopathic causes. Among past studies, there is great discrepancy between the most common causes of vocal fold paralysis. The most common have included lung malignancies,[1-5] idiopathic causes,[2,6-10] thyroid surgery,[11-16] and nonthyroid surgeries.[15-18] In 2 recent large retrospective chart review studies, thyroid surgery was the single most common cause of UVFP, but nonthyroid surgeries as a group more commonly cause UVFP.[15,16]

The etiology of UVFP is important because it affects the natural course, treatment, and outcome of the condition. Both the mechanism and degree of injury are important, ranging from neuropraxia, where complete recovery is expected, to complete transection, which may require surgical intervention.[20] Outcomes are affected by contralateral vocal fold compensation, as well as the degree of reinnervation and synkinesis established.[20] In a recent review of idiopathic UVFP, most improvement of vocal fold function and voice occurred within the first year of injury.[21]

Treatment of UVFP includes voice therapy, permanent and nonpermanent medialization procedures, and reinnervation. If the etiology suggests the nerve was not transected, then some degree of recovery of laryngeal nerve function is expected, and nonpermanent treatments are generally recommended until 6 to 12 months after onset of paralysis, whereas if complete transection has occurred, permanent medialization or reinnervation procedures may be undertaken sooner.[20] In addition, voice and airway are affected by the degree of synkinesis present. Synkinesis is caused by

[1]Department of Otolaryngology–Head and Neck Surgery, Washington, University School of Medicine, St Louis, Missouri, USA

This article was presented at the 2013 AAO-HNSF Annual Meeting & OTO EXPO; September 29–October 3, 2013; Vancouver, British Columbia, Canada.

Corresponding Author:
Randal C. Paniello, MD, Department of Otolaryngology–Head and Neck Surgery, Washington University School of Medicine, 660 S. Euclid Avenue, Campus Box 8115, St Louis, MO 63110, USA.
Email: paniellor@ent.wustl.edu

reinnervation of opposing muscle groups by the same nerve, leading the muscles to contract simultaneously. In canine models, reinnervating axons begin reaching the vocal fold muscles within 3 months of injury. Therefore, treatments to prevent unfavorable synkinesis would need to be administered within this time frame.[22,23]

The objective of this study was to determine how often patients with UVFP present to the Washington University School of Medicine (WUSM) within 3 to 4 months of symptom onset. These patients might be eligible for a clinical trial of early intervention for the prevention of synkinesis. A review of literature revealed no previous studies focusing on time of presentation of patients with UVFP. Etiology and laterality data were also collected and reported.

Methods

Approval for the study was obtained from the WUSM Institutional Review Board. Adult patients (>18 years of age) seen between January 1, 2002, and January 1, 2012, with a diagnosis of unilateral vocal fold paresis or paralysis (based on *International Classification of Diseases, Ninth Revision* and *Current Procedural Terminology* codes) were identified via a query of electronic medical records. In most cases, the diagnosis of UVFP was made by the attending physician based only on examination of the larynx; electromyography or cricoarytenoid joint palpation was performed in only a small percentage of patients. In all cases, the diagnosis was confirmed by flexible fiberoptic examination or videostroboscopy.

Each patient's chart was reviewed and data were collected and stored in an electronic database. Data included age at initial visit, sex, date of visit to a WUSM otolaryngologist, date of presentation to a non-WUSM otolaryngologist (in any), primary symptom, date of symptom onset, side of paralysis, etiology of paralysis, initial and subsequent treatment received, date of initial treatment, voice improvement, and fiberoptic examination vocal fold movement outcomes. From this information, the time intervals from symptom onset to initial presentation to an otolaryngologist and the time interval from initial presentation to referral to WUSM (if made), were calculated. The specific reason for referral to the university was not usually recorded in the chart, but very few had been previously treated. Diagnosis of idiopathic UVFP was confirmed by negative imaging along the course of the vagus and recurrent laryngeal nerves. Patients with incomplete records were excluded.

Data were analyzed by first determining percentages of men and women in the study population, laterality of UVFP, and etiology. To evaluate laterality, a calculation for left-right skew was devised as follows:

$$\text{Skew} = (\#\ \text{cases on right}/\text{total}\ \#\ \text{cases}) - 0.5.$$

With this formula, it can be seen that if there is a perfect 50-50 split, the skew is zero; as the proportion of left-sided cases increases, the value becomes more negative (moves to the left), and so on, with a maximum/minimum value of ± 0.5. This value was determined and plotted for each etiology. Statistical significance was determined based on a null hypothesis that the frequency on each side was the same, using a χ^2 test.

Median presentation time to an outside otolaryngologist was compared with median referral presentation time to a WUSM otolaryngologist. For this analysis, patients who initially presented to the WUSM were excluded. Due to wide variance in presentation times and nonnormally distributed data, the mean presentation times were not presented. To compare median presentation times between the 2 groups, a paired Wilcoxon rank sum test was used. Histograms were also generated incorporating all data, displaying number of patients and cumulative percentage of patients by time of presentation. Percentages of patients per etiology who presented at 2, 3, and 4 months after symptom onset were also calculated.

Results

Of the charts reviewed, 938 patients met inclusion criteria for this study; 497 (53%) were women and 441 (47%) were men. The average age of patients was 56.9 years (range, 18-93 years). Overall, 621 (66.2%) patients had left-sided UVFP, while 317 (33.8%) patients had right-sided UVFP.

Table 1 displays the etiologies of unilateral vocal fold paralysis. In total, 522 (55.6%) patients had UVFP due to iatrogenic effects related to surgery. The most frequently observed surgery related to UVFP was thyroid/parathyroid surgery, noted in 158 (16.8%) patients. Lung surgery (n = 73 [7.8%]), cardiac surgery (n = 58 [6.2%]), and cervical spine surgery (n = 48 [5.1%]) were the next most common surgical causes of UVFP.

In total, 358 patients (38.2%) had UVFP due to causes not directly related to surgical intervention. Malignancy was the cause of UVFP in 167 (17.8%) of patients. Lung malignancy (n = 73 [7.8%]), metastatic malignancy (n = 24 [2.6%]), skull base malignancy (n = 18 [1.9%]), and direct invasion by thyroid malignancy (n = 14 [1.5%]) were most common. Idiopathic UVFP was noted in 124 patients (13.2%). Other less common causes of UVFP included intubation (n = 58 [6.2%]), trauma (n = 30 [3.2%]), cerebral vascular accident (CVA; n = 18 [1.9%]), and neck radiation (n = 8 [0.9%]).

Table 2 shows the laterality of UVFP based on etiology. In total, 622 (66.2%) patients had left-sided UVFP. This table shows the difference between right- and left-sided UVFP. Left-skewed etiologies of UVFP, represented by negative values, and right-skewed etiologies of UVFP, represented by positive values, are plotted in **Figure 1**. In addition to the expected left-sided predominance of intrathoracic etiologies (lung surgery, cardiac surgery, esophageal surgery, and lung malignancy), other significantly left-sided causes included idiopathic, intubation, carotid surgery, and skull base malignancy. There were no etiologies that were significantly skewed to the right.

Table 3 shows the median time of presentation for the 92 patients (9.8% of study population) who initially

Table 1. Etiology of Unilateral Vocal Fold Paralysis.

Etiology	No. (% of Total)
Surgery	
Cardiac surgery	58 (6.2)
Carotid surgery	22 (2.3)
Cervical spine surgery	48 (5.1)
Tracheostomy	2 (0.2)
Esophageal surgery	37 (3.9)
Lung surgery	73 (7.8)
Mediastinal surgery	17 (1.8)
Laryngeal surgery	4 (0.4)
Lateral neck surgery	61 (6.5)
Parathyroid surgery	18 (1.9)
Thyroid surgery	140 (14.9)
Skull base surgery	18 (1.9)
Intracranial surgery	24 (2.6)
Total	522 (55.6)
Malignancy	
Laryngeal cancer	20 (2.1)
Esophageal cancer	11 (1.2)
Lung cancer	73 (7.8)
Skull base tumor	18 (1.9)
Lymphoma	1 (0.1)
Mediastinal mass	5 (0.5)
Metastatic cancer	24 (2.6)
Parotid cancer	1 (0.1)
Thyroid cancer—direct invasion	14 (1.5)
Total	167 (17.8)
Idiopathic	124 (13.2)
Intubation	58 (6.2)
Trauma	30 (3.2)
CVA	18 (1.9)
Transesophageal echocardiogram	1 (0.1)
IJ catheter placement	1 (0.1)
Infected vagal nerve stimulator	1 (0.1)
Neck infection	1 (0.1)
Right skull base osteomyelitis	1 (0.1)
Neck radiation	8 (0.9)
Lung radiation	2 (0.2)
Thoracic deformity	1 (0.1)
Ankylosing spondylitis	1 (0.1)
Sarcoidosis	2 (0.2)
Total	938 (100.0)

Abbreviations: CVA, cerebral vascular accident; IJ, internal jugular.

presented to an outside otolaryngologist and were later referred. The median time of presentation to an outside otolaryngologist was 2.1 months, while the median time of presentation to a WUSM otolaryngologist was 9.5 months ($P < .001$). Given the very low rate of patients previously treated, we conclude that this 7.4-month difference in median is the period during which patients were observed by the outside otolaryngologist for possible recovery prior to referral to WUSM. The etiology with the largest delay between symptom onset and treatment was trauma at 563 months, while several etiologies had delays of only 0.5 months. Etiologies with the greatest percentage of patients presenting to an outside otolaryngologist before a WUSM otolaryngologist included carotid surgery (18.2%), laryngeal surgery (25.0%), parathyroid surgery (16.7%), thyroid surgery (12.1%), laryngeal cancer (20.0%), thyroid cancer (28.6%), idiopathic causes (16.9%), CVA (22.2%), and neck radiation (25.0%).

Figure 2 displays histograms of time of presentation to any otolaryngologist and to a WUSM otolaryngologist within 3 years of symptom onset, as well as cumulative percentages of patients who presented within this time window. It can be seen that 81% of patients present within 6 months, 89% within 1 year, and 93% within 2 years. In **Figure 3**, the same data focus on the first 4 months after onset, during which 44% present within the first month (many during the same hospital stay during which the paralysis began), 63% within 2 months, 71% within 3 months, and 75% within 4 months. These are the patients for whom an early intervention strategy might be an option. The cumulative plots for all patients and for WUSM-only have similar contours because the WUSM referral group comprises 90.2% of the patients.

Table 4 shows the first 4-month presentation data by etiology, excluding those groups with less than 10 patients. Etiologies with the greatest percentage of patients presenting to the WUSM within a 4-month period included esophageal cancer (90.9% present within 4 months), skull base surgery (88.9%), esophageal surgery (86.5%), intubation (86.2%), lung surgery (84.9%), and lung cancer (82.2%). Etiologies with the lowest percentage of patients presenting to the WUSM within a 4-month period included idiopathic causes (54.8%), CVA (55.6%), thyroid cancer (57.1%), and carotid surgery (59.1%). In patients who had UVFP caused by thyroid surgery, 66.4% presented to the WUSM within 4 months, and in patients with parathyroid surgery, 72.2% presented to the WUSM within 4 months.

Discussion

Etiology of UVFP

In this large retrospective study of UVFP, most of the etiologic findings were similar to 2 other large series, by Rosenthal et al[15] and Takano et al,[16] as shown in **Table 5**. Surgical/iatrogenic causes of UVFP are more common than nonsurgical causes, and thyroid/parathyroid surgeries are implicated more often than other types of surgery but do not comprise most surgical etiologies overall. Intubation injuries and idiopathic UVFP frequencies are similar in all 3 series, and the condition occurs on the left side in nearly two-thirds of cases. Among nonsurgical cases, malignancy was the most common category, most often lung cancer. Malignancy of the lung was the most common cause in 3 previous studies.[1-3]

The risk of iatrogenic injury to the recurrent laryngeal nerve in different surgical procedures has been widely

Table 2. Unilateral Vocal Fold Paralysis Laterality.[a]

Etiology	n	Right	Left	Skew	*P*-value[b]
Thyroid surgery	140	64	76	–0.04	.31
Lung surgery	73	15	58	–0.29	**<.01**
Lateral neck surgery	61	25	36	–0.09	.16
Cardiac surgery	58	9	49	–0.34	**<.01**
Cervical spine surgery	48	27	21	0.06	.39
Esophageal surgery	37	6	31	–0.34	**<.01**
Carotid surgery	22	6	16	–0.23	**.03**
Skull base surgery	18	12	6	0.17	.16
Parathyroid surgery	18	9	9	0	1.00
Mediastinal surgery	17	10	7	0.09	.47
Idiopathic	124	39	85	–0.19	**<.01**
Lung cancer	73	6	67	–0.42	**<.01**
Intubation	58	17	41	–0.21	**<.01**
Trauma	30	14	16	–0.03	.72
Metastatic cancer	24	10	14	–0.08	.41
Laryngeal cancer	20	7	13	–0.15	.18
Skull base tumor	18	4	14	–0.28	**.02**
CVA	18	7	11	–0.11	.35
Thyroid cancer—direct invasion	14	6	8	–0.07	.59
Esophageal cancer	11	4	7	–0.14	.37

Abbreviation: CVA, cerebral vascular accident.
[a]Etiologies with less than 10 patients not shown.
[b]*P* values based on χ^2 comparison with 50-50 L-R split with same N. Bold, $P < .05$.

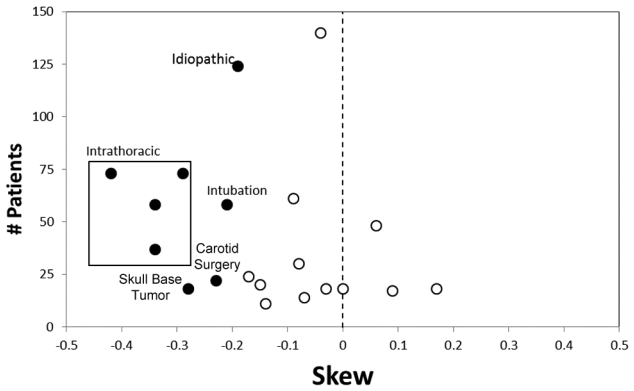

Figure 1. Left-right skew by etiology. Solid circles, statistically significantly skewed; open circles, not significant. Skew = (# cases on right / total cases for etiology) – 0.5. If left = right, skew = 0 (as seen for 18 parathyroid cases); if all cases occurred on left, skew = –0.5. Grouped are the intrathoracic causes of unilateral vocal fold paralysis (lung malignancy, thoracic, cardiac, and esophageal surgery).

reported. In a recent review by Misono and Merati,[24] the risk of vocal fold paralysis in thyroidectomy was between 0.8% and 2.3%; anterior cervical spine injury, less than 1%; cardiac/aortic surgery, 2%; mediastinoscopy, 0.2% to 6%; esophagectomy, 11%; and carotid endarterectomy, 4%.

Idiopathic UVFP was the etiology in 124 (13.2%) patients. Some older studies reported idiopathic etiologies to be the most common cause of UVFP,[2,6-10] but a recent review of the literature reported a rate of idiopathic UVFP of 24% ± 10%.[21] The decrease in idiopathic UVFP is likely due to better imaging capabilities to find small lesions along the nerve, as well as the shift toward surgery-related UVFP, which increases the proportion of cases with a clear etiology.

The data in this study by comparison to earlier studies of UVFP etiologies show the trend toward increasing surgical rather than malignant or idiopathic causes. This trend may be due to several factors. Since this study was conducted at a large tertiary referral center, as were the other recent studies by Rosenthal et al[15] and Takano et al,[16] a greater amount of surgery, as well as more complicated surgery, was likely being performed. An increased number of complicated surgeries both increase the risk of injury to the recurrent laryngeal nerve, as well as the relative numbers of UVFP due to surgical causes. In addition, as diagnostic imaging capabilities continue to improve, UVFP due to malignant or idiopathic causes will continue to decrease, as tumors are identified before causing UVFP and the course of the vagus and recurrent nerve can be imaged to determine the etiology of cases formerly diagnosed as idiopathic.

In analysis of laterality, 621 (66.2%) patients had left-sided UVFP, consistent with previous studies showing left-sided UVFP ranging from 59% to 81%.[1,2,7,8,11,15,16] Intrathoracic etiologies had the expected left-sided predominance, due to the anatomic course of the left recurrent laryngeal nerve. The statistically significant left-sidedness of some other etiologies may be a little harder to explain:

Table 3. Presentation Time for 92 Patients Referred to the WUSM.[a]

Etiology	n	Fraction of Original Total (%)	Outside ENT Presentation, Median, mo	WUSM ENT Presentation, Median, mo	Difference	Range in Difference
Carotid surgery	4	4/22 (18.2)	1.8	8.5	6.7	2-178
Cervical spine surgery	5	5/48 (10.4)	3.0	7.0	4.0	0.5-43
Lateral neck surgery	5	5/61 (8.2)	4.0	15.0	11.0	2-116
Parathyroid surgery	3	3/18 (16.7)	3.0	8.0	5.0	5-49
Thyroid surgery	17	17/140 (12.1)	1.5	12.0	10.5	0.5-111
Lung surgery	5	5/73 (6.8)	2.0	92.0	90.0	1-104
Cardiac surgery	4	4/58 (6.9)	2.5	6.5	4.0	2-5
Laryngeal cancer	4	4/20 (20.0)	1.5	4.5	3.0	3-7
Thyroid cancer	4	4/14 (28.6)	2.1	6.0	3.9	1.75-19
Idiopathic	21	21/124 (16.9)	3.0	12.0	9.0	0.5-51
Intubation	4	4/58 (6.9)	2.0	3.8	1.8	1-3
CVA	4	4/18 (22.2)	2.5	14.0	11.5	3-150
Total[b]	92	92/938 (9.8)	2.1	9.5	7.4	0.5-563

Abbreviations: ENT, ear, nose, and throat; WUSM, Washington University School of Medicine.
[a]Excludes etiologies with less than 3 patients.
[b]Includes 12 outside referral patients not listed above.

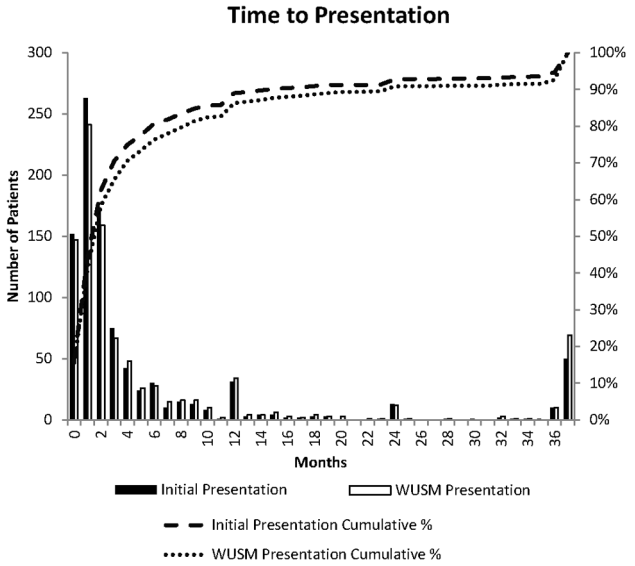

Figure 2. Histogram and cumulative plot of time from unilateral vocal fold paralysis symptom onset to presentation to the Washington University School of Medicine (WUSM; open bars) or to outside otolaryngologist prior to referral to the WUSM (solid bars; 9.8% of cases). Data shown for first 3 years following onset.

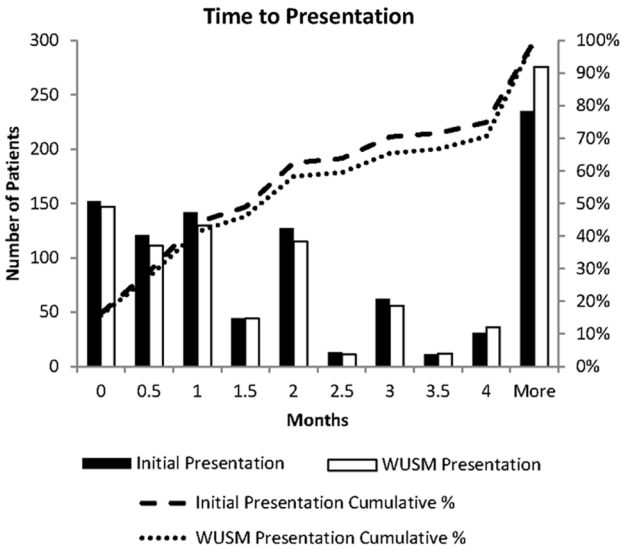

Figure 3. Histogram and cumulative plot from **Figure 2**, focusing on the first 4 months after symptom onset. WUSM, Washington University School of Medicine.

Carotid endarterectomy. Stroke and death rates in carotid endarterectomy have been shown to be higher for left-sided vs right-sided surgeries. Explanations for this observation include that due to the dominance of the left cerebral hemisphere, left-sided events are more symptomatic than right-sided events, and therefore more surgeries occur on the left side. There also may be a role in surgeon handedness, making left-sided carotid endarterectomies more technically difficult for right-handed surgeons.[25]

Intubation. More people in general are right-handed; there may be some greater tendency for right-handed anesthetists to traumatize the left hemilarynx more than the right during intubation. The mechanism of UVFP with intubation is not known but may relate to an acute event (at the time of intubation) or a longer-term event (while the tube is in place) that may be related to the cuff pressure or tube positioning. Some postintubation vocal fold immobility may

Table 4. Percentage of Patients Presenting to Washington University School of Medicine within 2, 3, and 4 Months from Symptom Onset.[a]

Etiology	n	No. (%)		
		Within 2 Months	Within 3 Months	Within 4 Months
Intracranial surgery	24	17 (70.8)	17 (70.8)	17 (70.8)
Skull base surgery	18	15 (83.3)	15 (83.3)	16 (88.9)
Carotid surgery	22	10 (45.5)	13 (59.1)	13 (59.1)
Cervical spine surgery	48	23 (47.9)	26 (54.2)	31 (64.6)
Lateral neck surgery	61	33 (54.1)	36 (59.0)	40 (65.6)
Parathyroid surgery	18	13 (72.2)	13 (72.2)	13 (72.2)
Thyroid surgery	140	81 (57.9)	86 (61.4)	93 (66.4)
Lung surgery	73	51 (69.9)	57 (78.1)	62 (84.9)
Mediastinal surgery	17	12 (70.6)	12 (70.6)	13 (76.5)
Cardiac surgery	58	19 (32.8)	41 (70.7)	44 (75.9)
Esophageal surgery	37	29 (78.4)	31 (83.8)	32 (86.5)
Skull base tumor	18	10 (55.6)	10 (55.6)	12 (66.7)
Laryngeal cancer	20	11 (55.0)	13 (65.0)	13 (65.0)
Thyroid cancer—direct invasion	14	6 (42.9)	6 (42.9)	8 (57.1)
Lung cancer	73	46 (63.0)	54 (74.0)	60 (82.2)
Esophageal cancer	11	9 (81.8)	10 (90.9)	10 (90.9)
Metastatic cancer	24	17 (70.8)	19 (79.2)	19 (79.2)
Idiopathic	124	48 (38.7)	64 (51.6)	68 (54.8)
Intubation	58	40 (69.0)	46 (79.3)	50 (86.2)
Trauma	30	14 (46.7)	18 (60.0)	19 (63.3)
CVA	18	7 (38.9)	9 (50.0)	10 (55.6)

Abbreviation: CVA, cerebral vascular accident.
[a]Etiologies with less than 10 patients not shown.

Table 5. Some Comparisons between the 3 Largest Series of Unilateral Vocal Fold Paralysis Etiologies.[a]

Etiology	Rosenthal et al[15] (n = 643)	Takano et al[16] (n = 797)	Present Study (n = 938)
Surgery, total	36.5	51.1	55.6
Thyroid/parathyroid	12.4	14.1	16.8
Nonthyroid	24.1	37.1	38.8
Intubation	5.8	7.3	6.2
Malignancy	18.4	9.9	17.8
Idiopathic	18.5	16.8	13.2
Left side	60.9[b]	64.0	66.2

[a]All values are percentages.
[b]Data reported for only 56.4% of patients in study.

have actually been cricoarytenoid joint ankylosis or dislocation that was misdiagnosed as UVFP.

Idiopathic. This is commonly presumed to be a post-viral neuropathy, but this would not explain a left-sided preference unless the distal left recurrent laryngeal nerve (RLN) is somehow more susceptible to viral attack than the right. Alternately, there may be some mechanism in which sudden neck extension or other postural change causes the left RLN to be stretched around the aorta. There may also be differences in intrathoracic connective tissue near the aorta that predispose to this sort of injury.

Skull base tumor. This finding may simply be related to the relatively low number of patients in this series.[18] We can find no references indicating a biological explanation for a left-sided predilection for skull base tumors.

Time to Presentation

To our knowledge, this study is the first large series to investigate time to presentation of these patients. It was found that three-fourths of patients with UVFP seen at the WUSM present within 3 or 4 months of the onset of symptoms. While this is encouraging, it also indicates that there

is potential for earlier presentation in the remaining 25% of patients. In patients who initially presented to an outside otolaryngologist, there was a significant referral delay of 7.4 months in median presentation time to the WUSM. This also highlights an opportunity for education of the referring physicians of the value of early intervention. For example, a recent study by Young et al[14] found that patients with UVFP had better vocal function after undergoing temporary vocal fold injection (VFI) even after direct benefit of VFI had dissipated, regardless of whether vocal fold mobility had recovered. Yung et al,[26] Arviso et al,[27] and Friedman et al[28] all reported lower rates of thyroplasty in patients who underwent temporary VFI. Bhattacharyya et al[29] found that early medialization within 1 to 4 days after onset of UVFP after thoracic surgery decreased the rate of pneumonia and led to a shorter length of stay compared with late medialization. Early intervention for UVFP clearly improves patient outcomes.

Early intervention could also involve treatment to prevent synkinesis. In animal studies, it has been found that following UVFP, some degree of reinnervation is evident within 3 months of injury.[22] One strategy that has been proposed is to perform a chemical blockade of reinnervation of the posterior cricoarytenoid muscle using a neurotoxic drug such as vincristine.[30,31] In an animal model, this was found to improve adductor recovery if given at 3 months postinjury but not at 5 months,[23] indicating there is a window of opportunity for treatment, after which it becomes too late for effective early intervention strategies. The present study shows that 71% of patients would be eligible for such intervention with current referral patterns. A clinical trial is the next step to determine whether this approach can help these patients.

A limitation of this study is that the surgical care at a tertiary care referral center skews data due to the greater number of difficult cases with a greater likelihood of nerve injury during surgery. In addition to missing data from those with incomplete charts, there are also an unknown number of patients who may have had a vocal fold paralysis but, due to quick recovery of voice, never sought treatment at a tertiary care facility.

Future directions of this study include analysis of the initial treatment and outcomes for each etiology of UVFP. Outcomes include voice improvement and return of vocal fold motion by fiberoptic examination. This analysis would allow further correlation with specific UVFP etiologies with the natural history of the disease, effectiveness of treatment, and type of treatment received. Outcomes specific to the length of time from symptom onset to treatment can also be assessed. This assessment would determine if delay in treatment adversely affects outcomes.

Conclusion

This retrospective medical record review of 938 patients with UVFP over the past 10 years is the largest series to date. It expands on the previous reports of UVFP etiology, with surgery and specifically thyroid surgery being the most common causes of UVFP. This study also reflects the growing contribution of nonthyroid surgeries accounting for a significant amount of injury to the recurrent laryngeal nerve, especially on the left side. Presently, 71% of patients with UVFP are seen within 3 months of RLN injury and would be eligible for early intervention procedures. Patients referred from outside otolaryngologists present, on average, after a significant delay.

Author Contributions

Emily A. Spataro, data analysis, manuscript preparation; **David J. Grindler**, data collection and analysis, manuscript preparation; **Randal C. Paniello**, original idea, final manuscript approval and editing.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: NIH (R01 DC010884)—salary support for senior author (R.C.P.).

References

1. Titche L. Causes of recurrent laryngeal nerve paralysis. *Arch Otolaryngol*. 1976;102:249-261.
2. Benninger MS, Gillen JB, Altman JS. Changing etiology of vocal fold immobility. *Laryngoscope*. 1998;108:1346-1350.
3. Ramadan HH, Wax MK, Avery S. Outcome and changing cause of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1998;118:199-202.
4. Kearsley JH. Vocal Cod paralysis—an aetiologic review of 100 cases over 20 years. *Aust NZ J Med*. 1981;11:663-666.
5. Barondess JA, Pompel P, Schley W. A study of vocal cord paralysis. *Trans Am Clin Climatol Assoc*. 1985;97:141-148.
6. Maisel RH, Ogura JH. Evaluation and treatment of vocal cord paralysis. *Laryngoscope*. 1974;84:302-316.
7. Hirose H, Sawashima M, Yoshioka H. Clinical observations on 750 cases of laryngeal palsy. *Tokyo J Med Sci*. 1981;88:33-37.
8. Hirose H. Clinical observations on 600 cases of recurrent laryngeal nerve paralysis. *Auris Nasus Larynx*. 1978;5:39-48.
9. Leon X, Venegas MP, Orus C, Quer M, Maranillo E, Sanudo JR. Glottic immobility: retrospective study of 229 cases. *Acta Otolaryngol Esp*. 2001;52:486-492.
10. Yamada M, Hirano M, Ohkubo H. Recurrent laryngeal nerve paralysis: a 10-year review of 564 patients. *Auris Nasus Larynx*. 1983;10(suppl):S1-S15.
11. Parnell FW, Bandenbug JH. Vocal cord paralysis: a review of 100 cases. *Laryngoscope*. 1970;80:1036-1045.
12. Chen HC, Jen YM, Wang CH, et al. Etiology of vocal cord paralysis. *ORL J Otorhinolaryngol Relat Spec*. 2007;69:167-171.
13. Sielska-Badurek E, Domeracka-Kolodziej A, Zawadzka R, Debowska-Jarzebska E. Vocal fold paralysis in the medical university of Warsaw's ambulatory of phoniatry in years 2000-2011. *Otolaryngol Pol*. 2012;66:313-317.
14. Young VN, Smith LJ, Rosen C. Voice outcome following acute unilateral vocal cord paralysis. *Ann Otol Rhinol Laryngol*. 2013;2122:197-204.

15. Rosenthal LHS, Benninger MS, Deeb RH. Vocal fold immobility: a longitudinal analysis of etiology over 20 years. *Laryngoscope*. 2007;117:1864-1870.
16. Takano S, Nito T, Tamaruya N, Kimura M, Tayama N. Single institutional analysis of trends over 45 years in etiology of vocal fold paralysis. *Auris Nasus Larynx*. 2012;39:597-600.
17. Terris D, Anstein D. Nguyen H. Contemporary evaluation of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1992;107:84-90.
18. Tucker H. Vocal cord paralysis-1979: etiology and management. *Laryngoscope*. 1980;90:585-590.
19. Myssiorek D. Recurrent laryngeal nerve paralysis: anatomy and etiology. *Otolaryngol Clin North Am*. 2004;37:25-44.
20. Rubin AD, Sataloff RT. Vocal fold paresis and paralysis. *Otolaryngol Clin North Am*. 2007;40:1109-1131.
21. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118: 1303-1307.
22. Shindo M, Herzon G, Hanson D, et al. Effects of denervation on laryngeal muscles: a canine model. *Laryngoscope*. 1992; 102:663-669.
23. Paniello RC. Effect on laryngeal adductor function of vincristine block of posterior cricoarytenoid muscle 3-5 months after recurrent laryngeal nerve injury. *Annals Otol Rhinol Laryngol*. In press.
24. Misono S, Merati A. Evaluation and management of unilateral vocal cord paralysis. *Otolaryngol Clin North Am*. 2012;45: 1083-1108.
25. Girard LP, Feasby TE, Eliasziw M, et al. Complication rates after left- versus right-sided carotid endarterectomy. *Circ Cardiovasc Qual Outcomes*. 2009;2:642-647.
26. Yung KC, Likhterov I, Courey MS. Effect of temporary vocal fold injection medialization on the rate of permanent medialization laryngoplasty in unilateral vocal fold paralysis patients. *Laryngoscope*. 2011;121:2191-2194.
27. Arviso LC, Johns MM III, Mathison CC, Klein AM. Longer-term outcomes of injection laryngoplasty in patients with potentially recoverable vocal fold paralysis. *Laryngoscope*. 2010;120:2237-2240.
28. Friedman AD, Burns JA, Henton JT, Zeitels SM. Early versus late injection medialization for unilateral vocal cord paralysis. *Laryngoscope*. 2010;120:2042-2046.
29. Bhattacharyya N, Batirel H, Swanson SJ. Improved outcomes with early vocal fold medialization for vocal fold paralysis after thoracic surgery. *Auris Nasus Larynx*. 2003;30:71-75.
30. Paydarfar JA, Paniello RC. Functional study of four neurotoxins as inhibitors of post-traumatic nerve regeneration. *Laryngoscope*. 2001;111:844-850.
31. McRae BR, Kincaid JC, Illing EA, et al. Local neurotoxins for prevention of laryngeal synkinesis after recurrent laryngeal nerve injury. *Ann Otol Rhinol Laryngol*. 2009;118:887-893.

Reprinted by permission of Laryngoscope. 2015; 125(4):904-908.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Diagnosis of Vocal Fold Paresis: Current Opinion and Practice

Amy P. Wu, MD; Lucian Sulica, MD

Objectives/Hypothesis: No accepted standard exists for the diagnosis of vocal fold paresis (VFP). Laryngeal specialists are surveyed to establish expert opinion on diagnostic methodology and criteria.
Study Design: Cross-sectional survey.
Methods: Questionnaires were distributed at laryngology conferences in fall 2013. Responses were collated anonymously and subjected to cross-tabulated data analysis.
Results: Fifty-eight responses completed by posttraining physicians whose practice focused in laryngology $\geq 75\%$ were analyzed. One (1.7%) relied principally on laryngeal electromyography, one (1.7%) on history, 10 (17%) on laryngoscopy, and 42 (72%) on strobovideolaryngoscopy for diagnosis. Only 12 (21%) performed laryngeal electromyography on $> 50\%$ of vocal fold paresis patients. Laryngeal electromyography sensitivity was considered moderate ($61 \pm 3.7\%$, $\sigma = 28$). Laryngoscopic/stroboscopic findings considered to have the strongest positive predictive value for VFP were slow/sluggish vocal fold motion ($75 \pm 3.0\%$, $\sigma = 23$), decreased adduction ($67 \pm 3.5\%$, $\sigma = 27$), decreased abduction ($65 \pm 3.4\%$, $\sigma = 26$), and decreased vocal fold tone ($61 \pm 3.5\%$, $\sigma = 26$). Asymmetric mucosal wave amplitude ($52 \pm 4.2\%$, $\sigma = 32$), asymmetric mucosal wave phase ($60 \pm 4.1\%$, $\sigma = 31$), hemilaryngeal atrophy ($60 \pm 4.0\%$, $\sigma = 31$), and asymmetric mucosal wave frequency ($49 \pm 4.0\%$, $\sigma = 30$) generated greatest disagreement.
Conclusions: Surveyed expert laryngologists diagnose vocal fold paresis predominantly on stroboscopic examination. Gross motion abnormalities had the highest positive predictive value. Laryngeal electromyography was infrequently used to assess for vocal fold paresis.
Key Words: Vocal fold paresis, laryngeal electromyography, laryngoscopy, stroboscopy.
Level of Evidence: 5

***Laryngoscope*, 125:904–908, 2015**

INTRODUCTION

Vocal fold paresis (VFP) is a partial motor denervation of the vocal fold causing variable degrees of compromised glottal function.[1] Appreciation of its clinical significance has grown hand in hand with an increasingly sophisticated understanding of laryngeal neuropathy, and it continues to evolve. Although not different in kind but only in degree from vocal fold paralysis, VFP is often considered separately; the spectrum of difficulties it causes is different, and perhaps most important, its diagnosis is more challenging and controversial. Paresis is usually diagnosed based on qualitative findings on laryngoscopy, stroboscopy, and/or laryngeal electromyography (LEMG). Endoscopic diagnosis typically rests on the observation of asymmetries of laryngeal motion. However, some asymmetry in laryngeal motion may be without clinical significance. Electromyographic findings may be indistinct as well.[1–4] In the absence of a standard for diagnosis, investigation of important aspects of paresis including causes, incidence, natural history, and effectiveness of treatment is challenging.

In the context of these limitations, expert clinical consensus may provide a useful basis to initiate discourse regarding VFP. The purpose of this investigation is to describe expert opinion regarding the diagnosis of VFP by means of a survey of practicing laryngeal specialists.

MATERIALS AND METHODS

A 29-item, 4-part questionnaire (see Appendix 1) was designed to characterize responders' experience, training and practice setting (part 1), assess diagnostic strategy (part 2), evaluate opinion regarding the positive predictive value of various laryngoscopic signs (part 3), and evaluate option regarding the sensitivity of LEMG (part 4). The roster of laryngoscopic signs was compiled from clinical experience, consultation with colleagues, and a review of the literature regarding VFP diagnosis and VFP-associated lesions. It included vocal fold hypomobility,[1,2,5] glottic insufficiency,[1,5] unilateral atrophy,[1] supraglottic hyperfunction,[2,6] mucosal wave asymmetries,[5,7] glottic axis deviation,[2,5] vocal fold height mismatch, arytenoid rotation,[8] and presence of a contact lesion[2,9,10] or pseudocyst.[2,11] Examples are illustrated in Figures 1 to 3. Opinion regarding each sign and LEMG sensitivity was assessed using a visual analog scale.

Additional Supporting Information may be found in the online version of this article.

From the New York Center for Voice and Swallowing Disorders, Department of Otolaryngology–Head and Neck Surgery, St Luke's–Roosevelt Hospital (A.P.W.); and the Sean Parker Institute for the Voice, Department of Otolaryngology–Head and Neck Surgery, Weill Cornell Medical College (L.S.), New York, New York, U.S.A.

Editor's Note: This Manuscript was accepted for publication October 7, 2014.

Scheduled to be presented at the 136th Annual Meeting of the American Laryngological Association, Las Vegas, NV, U.S.A., May 14–15, 2014.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Lucian Sulica, MD, 1305 York Avenue, 5th Floor, New York, NY 10021. E-mail: lus2005@med.cornell.edu

DOI: 10.1002/lary.25004

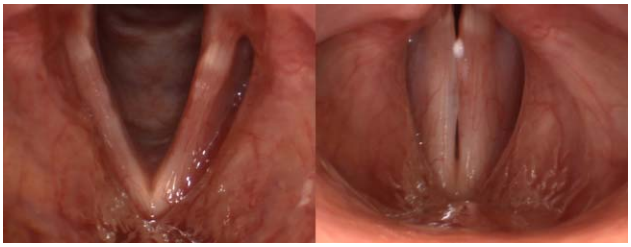

Fig. 1. A 44-year-old man with atrophy of the left hemilarynx, manifested as an enlarged laryngeal ventricle, and the beginning of a right vocal fold pseudocyst.

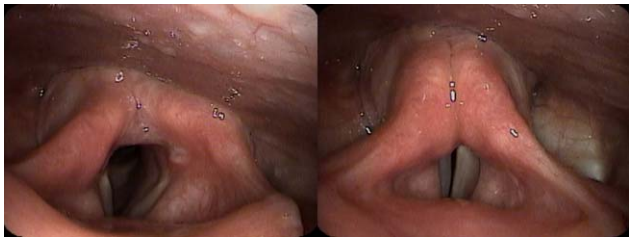

Fig. 3. A 30-year-old woman with a left-sided contact lesion and decreased left vocal fold adduction.

This questionnaire was distributed to attendees at three national laryngology conferences that took place in 2013: 1) the Neurolaryngology Study Group at the American Academy of Otolaryngology Annual Meeting, Vancouver, BC; 2) the Fall Voice Conference, Atlanta, GA; and 3) Advances and Controversies in Laryngology, Elk Grove, IL. Duplicate responses were avoided, and confidentiality was preserved by separately tracking responder identity. Responses were collated anonymously and subjected to cross-tabulated data analysis. Only otolaryngologists were surveyed.

Physicians whose practices did not predominantly focus on laryngology (< 75% laryngology), who did not see patients with dysphonia as a chief complaint, or who never diagnosed patients with VFP were excluded. Responders still in training, whether in residency or fellowship, were also excluded. Incomplete surveys were removed from analysis. One response with internally conflicting responses (a greater number of monthly paresis diagnoses than new dysphonic patients) was excluded. Statistical analysis using descriptive frequencies and cross-tabulations were performed with 2011 Microsoft Excel.

RESULTS

A total of 88 responses were received. After applying the exclusion criteria, a total of 58 entries were included for data analysis. Respondent demographics are summarized in Table I. The vast majority of respondents are laryngology fellowship-trained otolaryngologists with predominantly adult practices in the United States. Practice specifics related to VFP are summarized in Table II. Most respondents (72%) base diagnosis principally on stroboscopy. Only 12 (21%) respondents performed LEMG on greater than 50% of their VFP patients. Figure 4 and Table III summarize the respondents' opinion regarding positive predictive value of laryngoscopic findings in VFP patients. In general, abnormalities of gross vocal fold motion were considered to be most highly predictive, followed by stroboscopic signs. Associated lesions, namely contact lesions and pseudocysts, were not considered highly diagnostic of paresis. Respondents rated the sensitivity of LEMG at $61 \pm 3.7\%$, $\sigma = 28$.

Fig. 2. A 38-year-old woman with atrophy of the left vocal fold, incomplete glottic closure, unilateral (right) supraglottic hyperfunction, and impairment of arytenoid rotation on the left.

DISCUSSION

The prevalence and clinical importance of VFP is not yet established, which should not be surprising when diagnostic criteria remain under discussion, but it has the potential to be substantial. Among our respondents, VFP was diagnosed 8 times in a typical month, or roughly 100 times per year, and was found in about one-sixth of new patients presenting with voice complaints. By comparison, the largest reported series of VFP only consist of under 50 patients per year reviewed.[2,3,5] The prevalence of paresis has been proposed to be similar to paralysis on the basis of similar pathophysiology.[1] The prevalence of paralysis varies widely from report to report and is dependent on practice environment, geographic location, patient selection, and a host of other factors. Studies from one center have reported 46% prevalence of mild vocal hypomobility among patients with vocal complaints, 15% to 23% among singing teachers without vocal complaints, and 71% among singing teachers with technical difficulty complaints.[12–14] Simpson et al.[5] drew cases from a series of 739 patients presenting to their tertiary laryngology practice over a 4-year period with a chief complaint of dysphonia. Of 195 (26.4%) patients initially diagnosed with VFP or paralysis by videostroboscopy, only 13 or 1.8% of the overall dysphonic patients had LEMG-confirmed unilateral or bilateral VFP. Koufman et al.[15] reviewed 415 patients who underwent LEMG over a 5-year period. This group found "abnormal diagnostic LEMG" results (presumed to

TABLE I.
Respondent Demographics (N = 58).

Years in practice posttraining	11 ± 1.1 ($\sigma = 8.6$, range 0.1–31)
Laryngology fellowship trained	54 (93%)
Practicing in the United States	56 (97%)
Percent of practice which is laryngology	$93 \pm 1.0\%$
Percent adult patients	$91 \pm 1.1\%$

TABLE II.
Practice Related to VFP

New patients with voice-related complaint/month	49 ± 3.2, $\sigma = 25$
VFP diagnosis/month	8.5 ± 1.6, $\sigma = 12$
Diagnosis of VFP Rests Principally On:	
History	1 (1.7%)
Laryngoscopy (continuous light)	10 (17%)
Strobovideolaryngoscopy	42 (72%)
LEMG	1 (1.7%)
% Patients diagnosed with VFP who had videostroboscopy	$96 \pm 1.6\%$, $\sigma = 12$
% patients diagnosed with VFP who had LEMG	$26 \pm 4.0\%$, $\sigma = 31$

LEMG = laryngeal electromyography; VFP = vocal fold paresis.

represent paresis or paralysis in large part) in 346 (83%) patients. Sataloff et al.[14] reviewed 751 patients who underwent LEMG for all causes over a 4-year period. This series contained 689 suspected cases of paresis/paralysis by videostrobscopy, with LEMG confirming the diagnosis in 661 patients (95.9%). The variation among these three series reveals substantial differences among practitioners regarding diagnosis and testing.

Respondents indicated that they principally relied on laryngoscopy, usually under stroboscopic light, to make the diagnosis of VFP. Although LEMG is the only way to definitively diagnose laryngeal neuropathy objectively in vivo, the vast majority of respondents evidently felt that laryngoscopic criteria were sufficiently reliable to support the diagnosis alone. Only one respondent routinely relied on LEMG for diagnosis, and only the minority of patients ever had LEMG at all. Many reasons may prevent the use of LEMG, lack of availability and expertise prominent among them, but respondents felt that the sensitivity of LEMG was not high. There is little doubt that LEMG is highly specific for neuropathy. Findings of fibrillations, positive sharp waves, or polyphasic motor unit action potentials are unambiguous signs of neurologic impairment. Unfortunately in paresis, such clearly abnormal findings may be absent or obscured. Decreased recruitment of otherwise normal-appearing motor unit action potentials may be the only abnormality present. Because this relative change may be small and mimicked by incomplete muscle activation or suboptimal needle placement, there remains a role for physician judgment and inevitably error. Moreover, the maximal interference pattern in striated muscle is typically present at only 30% of maximum isometric contraction, which creates the possibility that even substantial paresis may be obscured during testing. Thus, although LEMG can provide important information that laryngoscopy cannot, it is not clear that it is a more accurate diagnostic tool than laryngoscopy in the diagnosis of VFP.

Reliance on laryngoscopy begs the question of which findings are considered important. To say that one may find signs of paresis in virtually every larynx is only a mild exaggeration. Unlike systems such as the extraocular muscles, mild discoordination in the larynx probably carries little functional and evolutionary disadvantage as long as glottic closure for airway protection is brisk and effective. Thus, much asymmetry in vocal fold

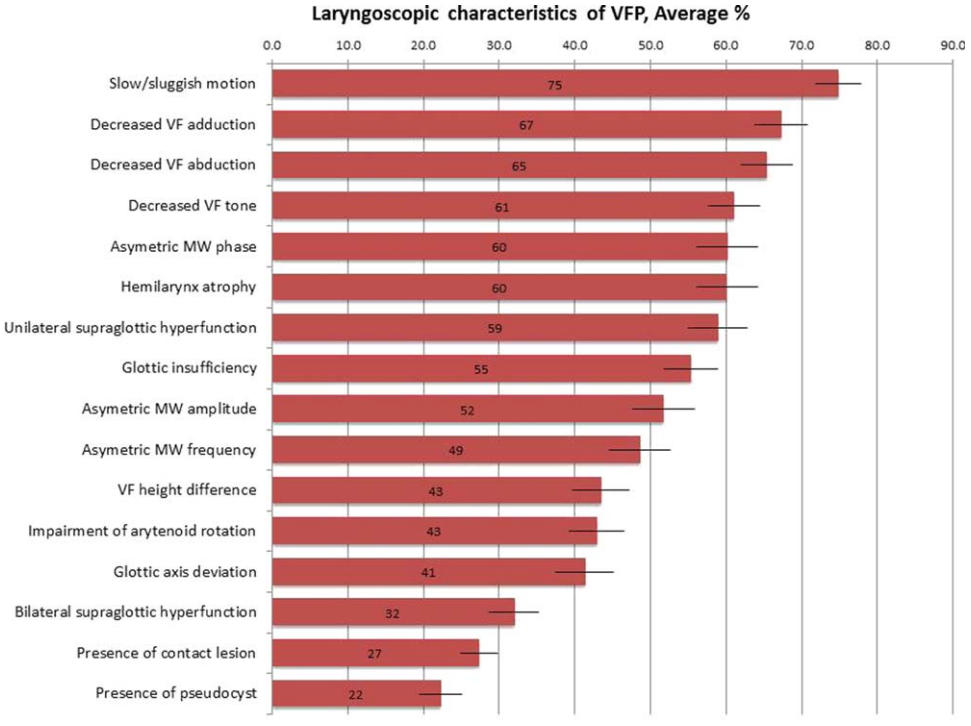

Fig. 4. Positive predictive value of laryngoscopic signs for vocal fold paresis. MW = mucosal wave; VF = vocal fold.

TABLE III.
Respondent Opinion Regarding Positive Predictive Value of Laryngoscopic Findings in VFP.

Examination finding	Average %	Error	σ
Slow/sluggish motion	74.9	3.0	22.8
Decreased VF adduction	67.3	3.5	26.7
Decreased VF abduction	65.4	3.4	26.1
Decreased VF tone	61.1	3.5	26.3
Asymetric MW phase	60.2	4.1	31.3
Hemilarynx atrophy	60.1	4.0	30.8
Unilateral supraglottic hyperfunction	58.9	3.9	29.4
Glottic insufficiency	55.4	3.5	26.8
Asymetric MW amplitude	51.7	4.2	31.7
Asymetric MW frequency	48.6	4.0	30.6
VF height difference	43.5	3.8	28.6
Impairment of arytenoid rotation	42.9	3.6	27.8
Glottic axis deviation	41.3	3.8	29.3
Bilateral supraglottic hyperfunction	32.0	3.3	25
Presence of contact lesion	27.3	2.4	18.1
Presence of pseudocyst	22.3	2.8	21.4

MW = mucosal wave; VF = vocal fold; VFP = vocal fold paresis.

motion may be without clinical significance. Further, Roy et al. showed that laryngoscopic findings are not consistent from case to case, even in experimentally induced isolated unilateral superior laryngeal nerve palsy, a condition probably more homogeneous than that which presents clinically.[17] Respondents identified deficits of gross motion as having the highest positive predictive value for VFP, despite reporting heavy reliance on stroboscopic examination in practice. This may reflect merely that stroboscopy is the standard clinical examination for patients with a voice complaint in the specialized practices of these physicians rather than the use of examination under stroboscopic light to identify VFP. Mucosal wave phase asymmetry was deemed the most useful stroboscopic sign, ranking only fifth in order of preference despite a report that identified it as correlating very well with LEMG abnormalities.[7] Stroboscopic signs (phase, amplitude, and frequency) were also marked by the greatest divergence of opinion regarding significance, as reflected by the standard error. The few signs that have been the subject of systematic analysis in the literature, namely arytenoid rotation[8] and unilateral ventricular fold hyperfunction,[6] were not regarded as among the most useful. Reports have proposed a relationship between contact lesions[9,10] and vocal fold pseudocysts[11] and VFP. Despite this, respondents thought that the potential for VFP to be present when such lesions were identified was very low.

Overall, this investigation reveals that paresis is frequently diagnosed and appears to be a significant clinical entity in laryngology practices. Diagnosis appears to be made on the basis of qualitative findings on laryngoscopy, principally deficits of gross vocal fold motion. Although stroboscopy is widely used, stroboscopic signs are not considered the most reliable signs to identify VFP. Electrophysiologic testing is not used often. Plainly, there exists no clear consensus on how the diagnosis of VFP should be made in a given patient, and establishing one will be a challenge in the absence of tests or findings that are both reasonably specific and sensitive. Under these circumstances, and given the frequency of asymmetric motion in the larynx, VFP is at risk of being diagnosed uncritically when no other obvious reason for a patient's complaint is evident to the examiner.

The survey format is subject to substantial recall bias and may give a false impression—likely falsely elevated—of the prevalence of paresis. This survey explicitly did not distinguish between superior laryngeal nerve paresis and recurrent laryngeal nerve paresis, frequently separated in the literature, which may have caused surveyed physicians to assign less positive predictive value to the signs under consideration than a more specific diagnosis. Reasons for the relatively rare use of LEMG were not investigated; these may have little to do with reservations regarding LEMG utility. Most importantly, the format of the survey necessarily does not well reflect the method of diagnosis of VFP in clinical practice. Such a diagnosis is rarely made on the basis of a single element of the evaluation or a single sign considered by itself, but depends on an educated critical synthesis of the clinical evidence. Physicians may form an impression of the likelihood of a given diagnosis based on the history, which then informs the physical examination. In fact, the perceived likelihood of VFP based on symptoms and clinical evolution of the complaint may significantly affect the perceived positive predictive value of a given laryngoscopic sign. Despite these limitations, this data may form a useful basis for further consideration of this challenging topic.

CONCLUSION

Surveyed laryngologists diagnose VFP frequently, relying principally on laryngeal strobovideolaryngoscopy to make the diagnosis. Among laryngoscopic signs, gross motion abnormalities were judged to have the highest positive predictive value for VFP, followed by abnormalities in the mucosal wave. Opinion varied most about the importance of these. LEMG was infrequently used to assess for VFP and was considered to have only moderate sensitivity for the diagnosis. Given the perceived clinical importance of VFP, directed investigation is necessary to refine diagnostic accuracy.

BIBLIOGRAPHY

1. Sulica L., Vocal Fold Paresis: An Evolving Clinical Concept. *Curr Otorhinolaryngol Rep* 2013;1:158–162.
2. Koufman JA, Postma GN, Cummins MM, Blalock PD. Vocal fold paresis. *Otolaryngol Head Neck Surg* 2000;122:537–541.
3. Heman-Ackah YD, Barr A. Mild vocal fold paresis: understanding clinical presentation and electromyographic findings. *J Voice* 2006;20:269–281.
4. Rubin AD, Praneetvatakul V, Heman-Ackah Y, Moyer CA, Mandel S, Sataloff RT. Repetitive phonatory tasks for identifying vocal fold paresis. *J Voice* 2005;19:679–686.
5. Simpson CB, Cheung EJ, Jackson CJ. Vocal fold paresis: clinical and electrophysiologic features in a tertiary laryngology practice. *J Voice* 2009; 23:396–398. doi: 10.1016/j.jvoice.2007.10.011.
6. Belafsky PC, Postma GN, Reulbach TR, Holland BW, Koufman JA. Muscle tension dysphonia as a sign of underlying glottal insufficiency. *Otolaryngol Head Neck Surg* 2002;127:448–451.

7. Simpson CB, May LS, Green JK, Eller RL, Jackson CE. Vibratory asymmetry in mobile vocal folds: is it predictive of vocal fold paresis? *Ann Otol Rhinol Laryngol* 2011;120:239–242.
8. Sufyan AS, Kincaid JC, Wannemuehler TJ, Halum SL. The interarytenoid spatial relationship: accuracy and interrater reliability for determining sidedness in cases of unilateral adductor paresis. *J Voice* 2013;27:90–94. doi: 10.1016/j.jvoice.2012.08.003.
9. Carroll TL, Gartner-Schmidt J, Statham MM, Rosen CA. Vocal process granuloma and glottal insufficiency: an overlooked etiology? *Laryngoscope* 2010;120:114–120. doi: 10.1002/lary.20656.
10. Halum SL, Miller P, Early K. Laryngeal granulomas associated with superior laryngeal nerve paresis. *J Voice* 2010;24:490–493. doi: 10.1016/j.jvoice.2008.12.010.
11. Koufman JA, Belafsky PC. Unilateral or localized Reinke's edema (pseudocyst) as a manifestation of vocal fold paresis: the paresis podule. *Laryngoscope* 2001;111:576–580.
12. Heman-Ackah YD, Batory M. Determining the etiology of mild vocal fold hypomobility. *J Voice* 2003;17:579–588.
13. Heman-Ackah YD, Dean CM, Sataloff RT. Strobovideolaryngoscopic findings in singing teachers. *J Voice* 2002;16:81–86.
14. Sataloff RT, Hawkshaw MJ, Johnson JL, Ruel B, Wilhelm A, Lurie D. Prevalence of abnormal laryngeal findings in healthy singing teachers. *J Voice* 2012;26:577–583. doi: 10.1016/j.jvoice.2012.01.002.
15. Koufman JA, Postma GN, Whang CS, et al., Diagnostic laryngeal electromyography: The Wake Forest experience 1995–1999. *Otolaryngol Head Neck Surg* 2001;124:603–606.
16. Brewer DW, et al., Unilateral recurrent laryngeal nerve paralysis: A reexamination. *J Voice* 1991;5:178–185.
17. Roy N, Barton ME, Smith ME, Dromey C, Merrill RM, Sauder C. An in vivo model of external superior laryngeal nerve paralysis: laryngoscopic findings. *Laryngoscope* 2009;119:1017–1032. doi: 10.1002/lary.20193.

Reprinted by permission of J Voice. 2014; 28(2):262.e13-262.e21.

Stroboscopy in Detection of Laryngeal Dysplasia Effectiveness and Limitations

***,†Vojko Djukic, *,†Jovica Milovanovic, *,†Ana D. Jotic, and †Milan Vukasinovic,** **†Belgrade, Serbia*

Summary: Vocal fold pathology changes the appearance and vibratory patterns observed during stroboscopic examination, but a strict correlation between the vibratory pattern and the dysplasia type does not exist. The aims of this study were to determine the role of stroboscopy in vocal fold dysplasia assessment and to determine whether stroboscopy is the deciding factor when performing laryngomicroscopy with biopsy in suspicious lesions. This prospective controlled study involved 112 patients with laryngeal dysplasia treated over a 2-year period at a tertiary medical center. Patient data and clinical, stroboscopy, laryngomicroscopy, and histopathologic reports were reviewed. During the stroboscopy, glottic occlusion, phase symmetry, periodicity, amplitude, mucosal wave, and nonvibratory segments were followed. Laryngomicroscopy with different types of endoscopic cordectomies (types I–III) was performed as a therapeutic measure, with a 12-month follow-up period. Nonvibrating segments were present in 15.1% of the patients with mild dysplasia and in 38.5% of the patients with moderate dysplasia. In 45.5% of the patients with severe dysplasia (carcinoma *in situ*), nonvibrating segments were absent. The amplitude of vocal fold vibrations in patients with mild dysplasia ($P = 0.03$) was a significant factor indicative of recurrent disease, but none of the stroboscopic signs was significant for the disease progression. Severe dysplasia can be related to both nonvibrating and vibrating vocal fold segments. Stroboscopy cannot be used reliably for classifying laryngeal dysplasia and may indicate the need to perform laryngomicroscopy with biopsy in suspicious vocal fold lesions. The warning factors for recurrence and progression of dysplasia are treatment modality, abnormal amplitude of vibration, and nonvibrating segment.

Key Words: Laryngeal dysplasia–Stroboscopy–Nonvibrating segment.

INTRODUCTION

Despite all the efforts made in discovering and classifying vocal fold lesions, uncertainty exists when determining which lesions are malignant or premalignant. These lesions are usually described as chronic laryngitis, parakeratosis, leukoplakia, erythroplakia, or dyskeratosis. A number of histologic results can be found under the same clinical appearance; therefore, the histologic nature of these lesions is completely unpredictable until a biopsy is performed. Malignant transformation rates range from 6% to 22%, and the rates increase with the severity of the precancerous lesion.[1–3] Therefore, the early detection of these lesions is of paramount importance.

Another difficulty in diagnosing these lesions is that there is no universally accepted histopathologic classification system. In the current literature and clinical practice, there are several widely accepted classification systems: the 2005 World Health Organization (WHO), Squamous Intraepithelial Neoplasia, Laryngeal Intraepithelial Neoplasia, and the Ljubljana Classification of Squamous Intraepithelial Lesions systems. This disparity makes it difficult to compare the diagnostic and follow-up studies. The WHO system uses three tiers of dysplasia: mild, moderate, and severe. Severe dysplasia includes what has been previously reported as noninvasive carcinoma (carcinoma *in situ* [CIS]) and severe dysplasia.[4] The progression and transformation to invasive carcinoma is one of the important outcome measures for intraepithelial lesions. Correlating molecular parameters with clinical outcome was recently suggested as a gold standard for classifying dysplasia. Some authors have stated that any histopathologic classification of this millennium should also depend on additional evidence, such as the genetic and molecular structural changes of the cells that contribute to the malignant transformation.[5]

Stroboscopy is considered to be an important part of diagnosing patients with laryngeal dysplasia. Nevertheless, we must note that a strict correlation between a vocal fold vibratory pattern and a certain type of lesion does not exist. Vocal fold pathology may produce changes in the appearance and vibratory patterns observed during stroboscopic examination. Interpreting the stroboscopic examination involves systematic judgment and describing the different vibratory pattern signs. These signs, which were first identified by Hirano and Bless,[6] included the fundamental frequency and periodicity, amplitude of horizontal excursion, glottal closure, symmetry of bilateral movement, mucosal wave, and nonvibrating portions of the vocal fold. Recently, Kelley et al have attempted to improve or refine the basic stroboscopic rating form and develop criteria to improve the reliability of selected stroboscopic signs.[7] Few studies have indicated which stroboscopic signs are more significant than others in evaluating the vibratory pattern of vocal folds with premalignant lesions. The aim of this study was to determine the importance of stroboscopy in diagnosing vocal fold dysplasia and ascertain if it can reliably estimate a level of dysplasia and be the deciding factor when performing laryngomicroscopy with biopsy. We also wanted to determine whether other factors, such as treatment modality and stroboscopic

Accepted for publication July 16, 2013.

From the *Medical Faculty Belgrade, University of Belgrade, Belgrade, Serbia; and the †Clinic for Otorhinolaryngology and Maxillofacial Surgery, Clinical Centre of Serbia, Belgrade, Serbia

Address correspondence and reprint requests to Ana D. Jotic, Clinic for Otorhinolaryngology and Maxillofacial Surgery, Clinical Centre of Serbia, Pasterova 2, 11000 Belgrade, Serbia. E-mail: anajotic@yahoo.com

Journal of Voice, Vol. 28, No. 2, pp. 262.e13-262.e21
0892-1997/$36.00
© 2014 The Voice Foundation
http://dx.doi.org/10.1016/j.jvoice.2013.07.006

signs, could be used to anticipate if disease recurrence or progression will occur.

MATERIALS AND METHODS

This prospective study included 112 patients who were treated over a 2-year period (between January 1, 2010 and December 31, 2011, with a 12-month follow-up period) in the Clinic for Otorhinolaryngology and Maxillofacial Surgery at the Clinical Centre of Serbia in Belgrade. This study was approved by the Institutional Ethical Committee, and all patients provided written informed consent before their inclusion in the study.

The following inclusion criteria were applied: the presence of a vocal fold lesion of any grade of dysplasia according to the WHO classification (mild, moderate, and severe dysplasia), a vocal fold lesion on the superior surface and free edge of the membranous part of the vocal fold, lesions ranging in size from 2 to 10 mm and up to 2 mm in thickness, normal motility of the vocal folds and arytenoid, no previous or simultaneous vocal fold lesions (inflammatory, dysplastic, carcinoma, or otherwise), and no previous laryngeal surgery, radiotherapy, or endotracheal intubation. All patient data, including clinical, stroboscopy, and laryngomicroscopy examinations and histopathologic reports were evaluated.

Stroboscopy was performed with the ATMOS Strobo 21 LED, ATMOS Cam 31 DV Data, and Laryngoscope 70° resp. 90° (ATMOS MedizinTechnik GmbH & Co., Lenzkirch, Germany) during modal pitch at comfortable intensity on sustained vowel /i/. The following parameters were rated:

1. glottic occlusion (1, sufficient or 2, insufficient),
2. phase symmetry (1, symmetrical or 2, asymmetrical opening and closing of the other vocal fold mirrors),
3. periodicity (1, regular or 2, irregular successive vibrations),
4. amplitude (1, normal; 2, decreased; or 3, increased),
5. mucosal wave (1, normal with 30–50% lateral travel; 2, increased with lateral travel greater than 50%; or 3, decreased with lateral travel less than 30%),
6. nonvibratory segment (1, presence or 2, absence of nonvibratory segment in the vocal fold or a portion thereof).

Laryngomicroscopy and different types of endoscopic cordectomy with cold instruments (types I–III according to recommended European Laryngological Society (ELS) classification for endoscopic cordectomies)[8] were performed using a Carl Zeiss Surgical OPMI Sensera optical microscope (Carl Zeiss Meditec Inc, Dublin, CA) under general endotracheal anesthesia.

The follow-up period for every patient was 12 months. During this period, a control examination with stroboscopy was performed monthly, and all patients with established recurrent vocal fold lesions on their control examinations underwent a laryngomicroscopy with complete lesion removal and histopathologic analysis. Any histologic progression of the lesions was noted.

PASW Statistics 18 program (IBM Corporation, New York, NY) was used for the data analysis. To determine the statistical significance of change in dynamics between the stroboscopic signs before the treatment and after the follow-up period, the McNemar and the Wilcoxon signed-rank tests were used. To determine a correlation between the chosen predicting factors and dysplasia, a multivariate regression analysis was performed. To assess which of the stroboscopic signs was most useful in predicting the histopathologic outcome and the degree of dysplasia, logistical regression was used. P values <0.05 were considered statistically significant.

RESULTS

The study included 98 males (87.5%) and 14 females (12.5%), with an average age of 55.65 years. There were 105 (93.7%) smokers, 95 (90.5%) of whom were males and 10 (9.5%) were females. Considering histopathologic results according to the WHO classification, 53 (47.3%) patients were classified as mild, 26 (23.2%) as moderate, and 33 (29.5%) as severe dysplasia.

Stroboscopic signs for patients with mild dysplasia before any treatment and after 12 months of follow-up because of recurrent disease are shown in Table 1. Considering phase symmetry, periodicity, amplitude of the vocal fold vibrations, and mucosal wave appearance, there were significant changes in the number of patients before the treatment and after the follow-up (McNemar or Wilcoxon signed-rank test, $P < 0.00$). Nonvibrating segments were present in eight (15.1%) patients before the treatment and in nine (17.0%) patients after the treatment ($P = 1.000$, McNemar test).

Considering the number of patients in the group with moderate dysplasia (Table 2), the changes in glottic occlusion and the presence of nonvibrating segment were not statistically significant, but the changes in the number of patients considering phase symmetry, periodicity, amplitude of vocal fold vibrations, and the mucosal wave appearance were statistically significant (McNemar or Wilcoxon signed-rank test, $P < 0.00$). In the group with moderate dysplasia, nonvibrating segments were present in 38.5% of the patients before the treatment and in 23.1% of the patients after the 12-month follow-up.

The results were similar in a group with severe dysplasia (Table 3). There were significant changes in the number of patients considering periodicity, amplitude of vocal fold vibrations, mucosal wave appearance, and the existence of nonvibrating segments (McNemar or Wilcoxon signed-rank test, $P < 0.00$). In this group, McNemar test could not be performed for the phase symmetry because all patients had asymmetric vibrations of the vocal fold vibrations before the treatment. Nonvibrating segments were present in 54.5% patients before the treatment and in 24.2% of patients after the 12-month follow-up. Most stroboscopic parameters were statistically significantly improved in all three patient groups.

Considering the treatment options, our patients underwent cordectomy types I–III, according to ELS classification for endoscopic cordectomies, the microscopic appearance of the change, and the assessment of the vertical expansion of the lesion (Table 4). Type I cordectomy was performed in 64.1% of the patients with mild dysplasia, 25.4% of the patients with

TABLE 1.
Stroboscopic Signs for Patients With Mild Dysplasia Before Treatment and After 12 Months of Follow-Up or Before Retreatment

	Mild Dysplasia		
Stroboscopic Signs	Before Treatment (%)	After 12 Months (%)	Sig.
Glottic occlusion			0.791
Sufficient	38/53 (71.7)	40/53 (75.5)	
Insufficient	15/53 (28.3)	13/53 (24.5)	
Phase symmetry			0.000*
Symmetrical	6/53 (11.3)	36/53 (67.9)	
Asymmetrical	47/53 (88.7)	17/53 (32.1)	
Periodicity			0.000*
Regular	6/53 (11.3)	36/53 (67.9)	
Irregular	47/53 (88.7)	17/53 (32.1)	
Amplitude			0.000*
Normal	10/53 (18.9)	35/53 (66)	
Decreased	25/53 (47.1)	14/53 (26.4)	
Increased	18/53 (34)	4/53 (7.6)	
Mucosal wave			0.000*
Normal with 30–50% lateral travel	13/53 (24.5)	35/53 (66)	
Increased with lateral travel greater than 50%	32/53 (60.4)	17/53 (32.1)	
Decreased with lateral travel less than 30%	8/53 (15.1)	1/53 (1.9)	
Nonvibratory segment			1.000
Presence	8/53 (15.1)	9/53 (17.0)	
Absence	45/53 (84.9)	44/53 (83.0)	

Abbreviation: Sig., statistical significance.
*$P < 0.05$.

TABLE 2.
Stroboscopic Signs for Patients With Moderate Dysplasia Before Treatment and After 12 Months of Follow-Up or Before Retreatment

	Moderate Dysplasia		
Stroboscopic Signs	Before Treatment (%)	After 12 Months (%)	Sig.
Glottic occlusion			1.000
Sufficient	16/26 (61.5)	17/26 (65.4)	
Insufficient	10/26 (38.5)	9/26 (34.6)	
Phase symmetry			0.007*
Symmetrical	3/26 (11.6)	14/26 (53.8)	
Asymmetrical	23/26 (88.4)	12/26 (46.2)	
Periodicity			0.021*
Regular	4/26 (15.4)	14/26 (53.8)	
Irregular	22/26 (84.6)	12/26 (46.2)	
Amplitude			0.001*
Normal	3/26 (11.6)	14/33 (42.4)	
Decreased	18/26 (69.2)	12/33 (57.6)	
Increased	5/26 (19.2)	0/33 (0)	
Mucosal wave			0.029*
Normal with 30–50% lateral travel	6/26 (23.1)	14/33 (42.4)	
Increased with lateral travel greater than 50%	18/26 (69.2)	11/33 (54.6)	
Decreased with lateral travel less than 30%	2/26 (7.7)	1/33 (3)	
Nonvibratory segment			0.344
Presence	10/26 (38.5)	6/26 (23.1)	
Absence	16/26 (61.5)	20/26 (76.9)	

Abbreviation: Sig., statistical significance.
*$P < 0.05$.

TABLE 3.
Stroboscopic Signs for Patients With Severe Dysplasia Before Treatment and After 12 Months of Follow-Up or Before Retreatment

	Severe Dysplasia		
Stroboscopic Signs	Before Treatment (%)	After 12 Months (%)	Sig.
Glottic occlusion			1.000
Sufficient	25/33 (75.8)	24/33 (72.7)	
Insufficient	8/33 (24.2)	9/33 (27.3)	
Phase symmetry			—
Symmetrical	0/33 (0)	19/33 (57.6)	
Asymmetrical	33/33 (100.0)	14/33 (42.4)	
Periodicity			0.000*
Regular	1/33 (3.0)	19/33 (57.6)	
Irregular	32/33 (97.0)	14/33 (42.4)	
Amplitude			0.000*
Normal	1/33 (3)	19/33 (57.6)	
Decreased	29/33 (87.9)	14/33 (42.4)	
Increased	3/33 (9.1)	0/33 (0)	
Mucosal wave			0.000*
Normal with 30–50% lateral travel	1/33 (3)	19/33 (57.6)	
Increased with lateral travel greater than 50%	0/33 (0)	0/33 (0)	
Decreased with lateral travel less than 30%	32/33 (97)	14/33 (42.4)	
Nonvibratory segment			0.013*
Presence	18/33 (54.5)	8/33 (24.2)	
Absence	15/33 (45.5)	25/33 (75.8)	

Abbreviation: Sig., statistical significance.
*$P < 0.05$.

moderate dysplasia, and 36.4% of the patients with severe dysplasia. Type II cordectomy was performed in 35.9% of the patients with mild dysplasia, 34.6% of the patients with moderate dysplasia, and 57.6% of the patients with severe dysplasia. Type III cordectomy was performed in only two (6%) patients with severe dysplasia.

Recurrence of the disease occurred after 8 months in six patients, after 9 months in five patients, after 10 months in four patients, after 11 months in eight patients, and after 12 months in five patients. Most patients with recurrence were in a group with moderate dysplasia (Table 5). Disease progression was noted in 10 patients with recurrence. Invasive carcinoma developed in four patients: one from group with moderate dysplasia and three from group with severe dysplasia.

Multivariate regression analysis was performed to determine how factors such as cordectomy type and stroboscopic signs (glottic occlusion, phase symmetry, periodicity, amplitude of vibrations, mucosal wave, and nonvibrating segment) correlated with the histopathologic verification of different types of dysplasia (Table 6). Some factors, such as the type of cordectomy and the existence of nonvibrating segment, were set apart from others and were proven to be significantly different in various levels of dysplasia ($P < 0.05$).

Logistic regression was used to determine whether the recurrence and progression of the disease could be anticipated by cordectomy type and stroboscopic signs (glottic occlusion, phase symmetry, periodicity, amplitude of vibrations, mucosal wave, and nonvibrating segment) (Table 7). Of all the analyzed factors, only the amplitude of vocal fold vibrations in group with mild dysplasia ($P = 0.03$) was statistically significant for recurrent disease. In this group, some factors indicated a higher risk of recurrence but not statistically significant enough. Patients with asymmetry in vocal fold vibrations and irregular vocal fold vibrations, with abnormal amplitude of vocal fold vibrations, and the existing nonvibrating segment were at higher risk of recurrence. None of the factors was statistically

TABLE 4.
Treatment Modalities for the Different Grades of Dysplasia

Cordectomy Types	Mild Dysplasia (%)	Moderate Dysplasia (%)	Severe Dysplasia (%)	All Patients (%)
I	34/53 (64.1)	17/26 (65.4)	12/33 (36.4)	63/112 (56.2)
II	19/53 (35.9)	9/26 (34.6)	19/33 (57.6)	47/112 (42.0)
III	0/53 (0)	0/26 (0)	2/33 (6.0)	2/112 (1.8)

TABLE 5.
Patients With Recurrent Disease, Progression of the Disease, and Developed Invasive Carcinoma After 12 Months of Follow-Up

	Recurrence (%)	Progression of the Disease (%)	Malignant Transformation (%)
Mild dysplasia	12/53 (22.6)	4/53 (7.5)	0/53 (0)
Moderate dysplasia	9/26 (34.6)	3/26 (11.5)	1/26 (3.8)
Severe dysplasia	7/33 (21.2)	3/33 (9.1)	3/33 (9.1)
All patients	28/112 (25)	10/112 (8.9)	4/112 (3.6)

significant in patients with moderate dysplasia, but factors, such as cordectomy type, abnormal amplitude of the vocal fold vibrations, and the existence of nonvibrating segment, placed the patients at greater risk of recurrent disease. In patients with severe dysplasia type of cordectomy, insufficient glottic occlusion and abnormal amplitude of the vocal fold vibrations carried higher risk of recurrent disease than others, but this result was not statistically significant. In this group, logistic regression could not be performed for phase symmetry because all patients had asymmetric vibrations of the vocal fold vibrations. For all patients, regardless of the degree of dysplasia, abnormal amplitude of vocal fold vibrations ($P = 0.01$) was a significant factor connected with recurrence. Considering the other factors, the type of cordectomy and the existence of vibratory segment placed the patients at greater risk of recurrence, but this result was not statistically significant.

Regarding the disease progression, none of the considered factors proved to be statistically significant. Some factors placed the patients at greater risk of progression (Table 8). The existence of nonvibrating segment carried higher risk of progression in group with mild and moderate dysplasia and in all patients regardless of the degree of dysplasia. In a group of patients with severe dysplasia, logistic regression could not be performed for the phase symmetry because all patients had asymmetric vibrations of the vocal fold vibrations.

DISCUSSION

In our study, there is male predominance in laryngeal dysplasia (87.5% males). The average age of our patients was 55.63 years, and most patients were in their sixth and seventh decades of life. There were 105 (93.7%) smokers. These facts aligned with other studies, with slight variations; therefore, we can generally expect this demography.[2,9–11] Malignant transformation occurred in 3.6% of the patients after 12 months of follow-up. Most patients were from the group with severe dysplasia. Ricci et al[2] observed a recurrence rate of 14.1% for all dysplasia patients. Dispenza et al observed a recurrence rate of 13.2% for patients with LIN1 and 29% for patients with LIN2 after a 1-year follow-up.[9] Malignant transformation in those studies ranged from 6.48%[2] to 16%.[9,10,12] Weller et al observed a progression rate (according to severity) of 21% with severe dysplasia compared with 14% for mild and moderate dysplasia, which differed from our results.

Gamboa et al[13] conducted a study on stroboscopic assessment of chronic laryngitis in 27 patients (eight of whom had different degrees of laryngeal dysplasia). Among the 15 cases with absence of mucosal wave in the stroboscopic exploration, 60% of the cases had severe dysplasia with squamous cell carcinoma. The authors concluded that the stroboscopic results were related to the pathologic results.

Atypical mucosal waves, as viewed through stroboscopy, should travel one-half of the width of the superior surface of the vocal fold during modal phonation. A reduced mucosal wave and decreased amplitude during modal phonation signifies stiffness, which may result from a lesion, edema, or scar. The vocal fold epithelium normally shows five to 10 cell layers and a thickness of 100–200 μm. Arens et al[14] determined that vocal fold mucosa shows progressive thickening from normal epithelium (147 μm) over the different epithelial dysplasia grades (grade I epithelial dysplasia, 258 μm; grade II epithelial dysplasia, 301 μm; and CIS, 445 μm) up to early invasive carcinoma (974 μm). This result can explain the increasing number of patients with decreased mucosal wave in the three dysplasia groups. Colden et al[15] also conducted a study to determine whether stroboscopy is a reliable method for differentiating invasive glottic carcinoma from intraepithelial atypia and determining the depth of cancer invasion. The authors examined 62 keratotic lesions (45 intraepithelial and 17 carcinomas). The reduced amplitude of vocal fold vibration and/or mucosal wave propagation associated with keratosis did not reliably predict the presence of cancer or the depth of cancer invasion. Reductions in the amplitude of vocal fold vibration and in mucosal wave magnitude were noticed in intraepithelial atypia, despite the fact that there was no invasion into the superficial lamina propria. The authors concluded that the reduced epithelial flexibility could be caused by voluminous keratosis without dysplasia and that abnormalities of the superficial lamina propria could be provoked by inflammation or fibrovascular scarring; for this reason, the absence of mucosal wave was not synonymous with malignancy.

The existence of atypical vocal fold vibration patterns was also reported in normophonic speakers.[16] Nonvibrating segments were associated with the existence of malignant infiltration of the vocal fold epithelium and basal membrane. In our patients, nonvibrating segments were present in 15.1% of the patients with mild hyperplasia, 38.5% of the patients with moderate hyperplasia, and 54.5% of the patients with severe hyperplasia. Shaw and Deliyski[17] determined the presence of atypical magnitude and symmetry of the mucosal waves in the vocal fold vibration of normophonic speakers. In their study, mucosal wave absence was noted in at least 21% of vocal fold vibration

TABLE 6.
Multivariate Regression Analysis of Correlation of Dysplasia With Stroboscopic Signs and Type of Treatment

	Cordectomy Type	Glottic Occlusion	Phase Symmetry	Periodicity	Amplitude	Mucosal Wave	Nonvibratory Segment
t	5.31	7.65	1.39	2.36	6.69	5.288	10.39
OR (95% CI)	1.22 (0.76 to 1.68)	1.88 (1.39 to 2.37)	0.80 (−0.34 to 1.94)	1.24 (0.19 to 2.28)	2.03 (1.43 to 2.63)	1.68 (1.05 to 2.30)	2.950 (2.39 to 3.51)
Sig.	0.006*	0.791	0.076	0.265	0.474	0.636	0.000*

Abbreviations: CI, confidence interval; OR, odds ratio.
* $P < 0.05$.

samples from normophonic speakers. The authors strongly advised that caution should be used when determining the abnormality of mucosal wave variations during clinical visualization procedures. One concern in our study is that in 45.5% of the patients with histopathologically determined CIS, nonvibrating segments were absent, which is a significant number.

Treatment involves removing the lesion with epithelium, basal membrane, and lamina propria and, depending on the type of cordectomy, deeper underlying structures. Vocal fold scaring was examined on animal models. Rousseau et al[18,19] described the development of a vocal fold scar 6 months after surgical injury in canine and rabbit models. As early as 2 months after the surgical removal of the epithelium and lamina propria, no significant difference in collagen density was noted, but at 6 months after injury, collagen density was significantly increased in the surgically injured animals compared with those with normal vocal folds. By 6 months, the procollagen and elastin levels had achieved the densities observed in normal vocal folds, although the elastin fibers remain fragmented and disorganized. The basal layer of the mucosal epithelium continues to experience remodeling in the later stages of wound healing, whereas the intercellular epithelial space undergoes remodeling earlier during the acute stage of wound healing.[20]

Kishimoto et al[21] investigated the maturation process of vocal fold scarring after cordectomy in 10 patients (eight with early laryngeal carcinoma and two with laryngeal dysplasia) using videostroboscopy. The patients were treated with cordectomy types I–III. Improvements in amplitude of mucosal wave were visible 6 months after the procedure and continued to improve up to 14 months after the procedure. Twelve months after the initial treatment was a reasonable time to assess the treatment results in our study. Indeed, there were improvements in phase symmetry, periodicity, amplitude of vocal fold vibrations, and the regularity of mucosal wave. The number of patients with nonvibratory segment decreased. At the end of the follow-up period, there were 23 (20.53%) patients with detected nonvibrating segment. Four patients who developed invasive carcinoma were among these patients. In other patients, this result could be explained by the vocal fold scarring process, particularly because in these patients, type II and type III cordectomies were performed as a treatment of choice. This is yet another limiting factor for stroboscopy use because it cannot reliably distinguish the vocal fold process resulting from the existence of a nonvibrating segment.

Many voice disorders are marked by either aperiodicity or fluctuating frequency and, therefore, cannot be visualized with stroboscopy.[22] There are a growing number of articles that emphasize the importance of different and more effective methods in evaluating irregular vocal fold vibrations and the propagation and existence of the mucosal wave, such as electroglottography, high-speed digital imaging, videokymography, or digital kymography. Mucosal wave propagates in both vertical and horizontal directions, and quantifying the vertical displacement is crucial for understanding the effect of pathologies on the mucosal wave. Stroboscopy, videokymography, and high-speed digital imaging only provide a two-dimensional image

TABLE 7.
Logistic Regression Analysis of Different Stroboscopic Signs and Type of Treatment and Recurrent Disease

	Pretreatment			
	Mild Dysplasia	Moderate Dysplasia	Severe Dysplasia	All Patients
Cordectomy type				
Sig.	0.84	0.34	0.17	0.13
OR (95% CI)	1.15 (0.29 to 4.49)	2.45 (0.39 to 15.49)	3.03 (0.61 to 15.02)	1.96 (0.82 to 4.67)
Glottic occlusion				
Sig.	0.66	0.2	0.49	0.40
OR (95% CI)	0.73 (0.18 to 2.92)	0.33 (0.06 to 1.78)	2.21 (0.22 to 21.78)	0.68 (0.27 to 1.68)
Phase symmetry				
Sig.	0.71	0.96	—	0.84
OR (95% CI)	0.65 (0.07 to 6.21)	0.93 (0.07 to 11.99)		0.85 (0.16 to 4.33)
Periodicity				
Sig.	0.71	0.66	1	0.58
OR (95% CI)	0.65 (0.07 to 6.21)	0.58 (0.05 to 6.58)	0	0.64 (0.13 to 3.16)
Amplitude				
Sig.	0.03*	0.22	0.6	0.01*
OR (95% CI)	2.93 (1.08 to 7.95)	2.80 (0.53 to 14.7)	2.03 (0.14 to 28.86)	2.79 (1.25 to 6.22)
Mucosal wave				
Sig.	0.94	0.77	1	0.83
OR (95% CI)	0.96 (0.34 to 2.71)	0.79 (0.17 to 3.66)	0	0.91 (0.39 to 2.11)
Nonvibratory segment				
Sig.	0.06	0.65	0.49	0.35
OR (95% CI)	4.62 (0.95 to 22.51)	1.47 (0.28 to 7.63)	0.55 (0.10 to 2.97)	1.53 (0.63 to 3.72)

Abbreviations: CI, confidence interval; OR, odds ratio.
*$P < 0.05$.

TABLE 8.
Logistic Regression Analysis of the Different Stroboscopic Signs and Type of Treatment and Disease Progression

	Pretreatment			
	Mild Dysplasia	Moderate Dysplasia	Severe Dysplasia	All patients
Cordectomy type				
Sig.	0.13	0.99	0.92	0.78
OR (95% CI)	0.16 (0.02 to 1.68)	0	1.10 (0.14 to 8.73)	0.84 (0.25 to 2.79)
Glottic occlusion				
Sig.	0.88	0.85	0.7	0.97
OR (95% CI)	1.2 (0.11 to 12.54)	1.28 (0.10 to 16.34)	0.61 (0.048 to 7.76)	0.97 (0.23 to 4.01)
Phase symmetry				
Sig.	0.99	0.99	—	0.99
OR (95% CI)	0	0		0
Periodicity				
Sig.	0.99	0.99	0.99	0.99
OR (95% CI)	0	0	0	0
Amplitude				
Sig.	0.66	0.79	0.75	0.55
OR (95% CI)	1.37 (0.33 to 5.7)	1.33 (0.15 to 12.09)	1.84 (0.04 to 73.47)	1.41 (0.46 to 4.26)
Mucosal wave				
Sig.	0.75	0.59	1	0.56
OR (95% CI)	0.77 (0.15 to 3.95)	0.53 (0.05 to 5.48)	0	0.68 (0.19 to 2.48)
Nonvibratory segment				
Sig.	0.07	0.31	0.45	0.21
OR (95% CI)	7.17 (0.84 to 60.79)	3.75 (0.29 to 47.99)	0.38 (0.03 to 4.69)	2.29 (0.62 to 8.48)

Abbreviations: CI, confidence interval; OR, odds ratio.

of the mucosal wave, whereas digital kymography can be used to provide a complete three-dimensional profile of vocal fold vibration dynamics.[23,24] Although these new techniques are superior to stroboscopy, there are some limitations to their use. High-speed imaging systems are still too expensive to be widely used in clinical practice, gathering larger data sets is problematic because of that fact, and there are no general accepted clinical protocols in laryngology for these techniques.[25]

After performing the multivariate regression analysis, some factors, such as the type of cordectomy and the existence of a nonvibrating segment, were set apart from others and were proven to significantly correlate with various levels of dysplasia ($P < 0.05$). However, after logistic regression of all chosen factors (ie, cordectomy type and stroboscopic signs—glottic occlusion, phase symmetry, periodicity, amplitude of vibrations, mucosal wave, and nonvibrating segment), only the abnormal amplitude of vocal fold vibrations was observed to occur significantly more frequently for recurrent disease in the group with mild dysplasia and in all patients. None of the factors was significant for disease progression. The presence of some factors placed the patients at higher risk of recurrence and progression of the disease. The patient group with mild dysplasia and abnormal vocal fold vibration amplitudes had a 2.93 times greater risk of recurrence, and the group with nonvibrating segments was at 4.62 times greater risk compared with patients without those stroboscopic signs. Nonvibrating segment placed those patients at a 7.17 times greater risk of disease progression than those patients without nonvibrating segment during stroboscopic examination. In the group with moderate dysplasia, patients with insufficient glottic occlusion and abnormal amplitude of vocal fold vibrations were at a greater risk of recurrence. In that group, the patients with insufficient glottic occlusion, with abnormal amplitude of vocal fold vibrations, and the existence of nonvibrating segment were at a greater risk of disease progression. In the patient group with severe dysplasia, the greater risk of recurrence and disease progression aligned with the type of cordectomy and abnormal amplitude of vocal fold vibrations. These findings could also be the result of a relatively small number of patients in the different dysplasia groups, which is one of the limitations of this study. With a larger number of patients, some of the stroboscopic signs could be more prominent. Chang et al[26] conducted a study on a small (18 patients) and nonhomogenous group of patients with laryngeal dysplasia and carcinoma to determine whether the clinical features and clinical appearance of the lesions at presentation correlated with the outcomes of treatment in terms of cure rate and voice outcome. They noted that the clinical appearance of the lesion at presentation, as judged by either still light endoscopy or stroboscopy, did not correlate with disease recurrence. The lesion appearance on still light endoscopy and vibratory characteristics on stroboscopy also did not correlate with the disease-free interval or voice outcome after endoscopic resection.

Stroboscopy is a subjective method in terms of a stroboscopic parameter rating system, and the person conducting the procedure should be well trained to reduce variation and bias. Because of the increasing popularity of stroboscopy equipment in the general otolaryngology office, it is useful to point out some limitations of stroboscopy that can benefit less experienced examiner. In this article, we showed that a large and clinically significant number of cases with CIS with absence of nonvibrating segments can be overlooked when relying solely on stroboscopy. Caution must be exercised when assessing stroboscopic findings, particularly during the posttreatment follow-up period, or if other more sophisticated means of diagnostics are unavailable.

CONCLUSION

Stroboscopy cannot be used reliably for classifying laryngeal dysplasia. Some stroboscopic signs cannot be used as an indication for performing or not performing laryngomicroscopy with biopsy in cases of any suspicious vocal fold lesions. In the absence of more expensive and advanced diagnostic methods, vocal fold dysplasia could be precisely classified only by histopathology analysis. The patient age, treatment modality, and stroboscopic signs, such as abnormal amplitude of vocal fold vibration and the existence of nonvibrating segment, can be considered as warning factors for recurrence and disease progression.

REFERENCES

1. Gallo A, De Vincentiis M, Della Rocca C, Moi R, Simonelli M, Minni A, Shaha AR. Evolution of precancerous laryngeal lesions: a clinicopathological study with long-term follow-up on 259 patients. *Head Neck*. 2001;23:42–47.
2. Ricci G, Molini E, Faralli M, Simoncelli C. Retrospective study on precancerous laryngeal lesions: long-term follow-up. *Acta Otorhinolaryngol Ital*. 2003;23:362–367.
3. Isenberg JS, Crozier DL, Dailey SH. Institutional and comprehensive review of laryngeal leukoplakia. *Ann Otol Rhinol Laryngol*. 2008;117:74–79.
4. Shanmugaratnam K, Sobin LH. The World Health Organization histological classification of tumours of the upper respiratory tract and ear. A commentary on the second edition. *Cancer*. 1993;71:2689–2697.
5. Bosman FT. Dysplasia classification: pathology in disgrace? *J Pathol*. 2001;194:143–144.
6. Hirano M, Bless DM. *Videostroboscopic Examination of the Larynx*. San Diego, CA: Singular Publishing Group Inc; 1993.
7. Kelley RT, Colton RH, Casper J, Paseman A, Brewer D. Evaluation of stroboscopic signs. *J Voice*. 2011;25:490–495.
8. Remacle M, Van Haverbeke C, Eckel H, Bradley P, Chevalier D, Djukic V, de Vicentiis M, Friedrich G, Olofsson J, Peretti G, Quer M, Werner J. Proposal for revision of the European Laryngological Society classification of endoscopic cordectomies. *Eur Arch Otorhinolaryngol*. 2007;264:499–504.
9. Dispenza F, De Stefano A, Marchese D, Martines F, Dispenza C. Management of laryngeal precancerous lesions. *Auris Nasus Larynx*. 2012;39:280–283.
10. Montgomery J, White A. A decade of laryngeal dysplasia in Paisley, Scotland. *Eur Arch Otorhinolaryngol*. 2012;269:947–951.
11. Spielmann PM, Palmer T, McClymont L. 15-Year review of laryngeal and oral dysplasias and progression to invasive carcinoma. *Eur Arch Otorhinolaryngol*. 2010;267:423–427.
12. Weller M, Nankivell P, McConckey C, Paleri V, Mehanna M. The risk and interval to malignancy of patients with laryngeal dysplasia: a systematic review of case series and meta-analysis. *Clin Otolaryngol*. 2010;35:364–372.
13. Gamboa J, Echeverría L, Molina B, Cobeta I. Stroboscopic assessment of chronic laryngitis. *Acta Otorrinolaringol Esp*. 2006;57:266–269.
14. Arens C, Glanz H, Wönckhaus J, Hersemeyer K, Kraft M. Histologic assessment of epithelial thickness in early laryngeal cancer or precursor lesions and its impact on endoscopic imaging. *Eur Arch Otorhinolaryngol*. 2007;264:645–649.
15. Colden D, Jarboe J, Zeitels SM, Bunting G, Hillman RE, Spanou K. Stroboscopic assessment of vocal fold keratosis and glottic cancer. *Ann Otol Rhinol Laryngol*. 2001;110:293–298.

16. Haben CM, Kost K, Papagiannis G. Mucosal wave asymmetries in the clinical voice laboratory. *J Otolaryngol*. 2002;31:275–280.
17. Shaw HS, Deliyski DD. Mucosal wave: a normophonic study across visualization techniques. *J Voice*. 2006;22:23–33.
18. Rousseau B, Hirano S, Scheidt TD, Welham NV, Thibeault SL, Chan RW, Bless DM. Characterization of vocal fold scarring in a canine model. *Laryngoscope*. 2003;113:620–627.
19. Rousseau B, Hirano S, Chan RW, Welham NV, Thibeault SL, Ford C, Bless DM. Characterization of chronic vocal fold scarring in a rabbit model. *J Voice*. 2004;18:116–124.
20. Hansen JK, Thibeault SL. Current understanding and review of the literature: vocal fold scarring. *J Voice*. 2006;20:110–120.
21. Kishimoto Y, Hirano S, Tateya I, Kanemaru S, Ito J. Temporal changes in vocal functions of human scarred vocal folds after cordectomy. *Laryngoscope*. 2010;120:1597–1601.
22. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008;117:413–424.
23. George NA, de Mul FF, Qiu Q, Rakhorst G, Schutte HK. Depth kymography: high-speed calibrated 3D imaging of human vocal fold vibration dynamics. *Phys Med Biol*. 2008;53:2667–2675.
24. Krausert CR, Olszewski AE, Taylor LN, McMurray JS, Dailey SH, Jiang JJ. Mucosal wave measurement and visualization techniques. *J Voice*. 2011; 25:395–405.
25. Verikas A, Uloza V, Bacauskiene M, Gelzinis A, Kelertas E. Advances in laryngeal imaging. *Eur Arch Otorhinolaryngol*. 2009;266: 1509–1520.
26. Chang J, Fang TJ, Yung K, et al. Clinical and histologic predictors of voice and disease outcome in patients with early glottic cancer. *Laryngoscope*. 2012;122:2240–2247.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2016; 125(4):331-335.

Article

Annals of Otology, Rhinology & Laryngology
2016, Vol. 125(4) 331–335
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489415613801
aor.sagepub.com
SAGE

Prognostic Relevance of Mucosal Waves in Patients With Unilateral Vocal Fold Paralysis

Rudolf Reiter, MD, PhD[1], Anja Pickhard, MD, PhD[2], Sylvia Sander[3], and Sibylle Brosch, MD, PhD[1]

Abstract

Objectives: To analyze the prognostic relevance of mucosal waves (MWs) for recovery of unilateral vocal fold paralysis (UVP).

Methods: The charts and stroboscopic examinations of 100 consecutive patients with a complete UVP were reviewed retrospectively. All had a minimal (estimated <3 mm) mucosal gap on stroboscopy. A positive or negative MW on the paralyzed vocal fold was associated with complete recovery to full adduction and abduction. All patients were followed for at least 12 months.

Results: Causes of the paralysis were iatrogenic/traumatic (n = 82), malignancy associated (n = 10), and idiopathic (n = 8). In patients with positive MW at diagnosis (n = 80), the chance of recovery of unilateral vocal fold paralysis was 91.25%, whereas the chance of recovery with a negative mucosal wave (n = 20) was only 10%.

Conclusion: Positive MWs in stroboscopy are a predictor for recovery of (iatrogenic/traumatic) unilateral vocal fold paralysis and should be used in routine diagnostic assessment.

Keywords

unilateral vocal fold paralysis, mucosal wave, microstroboscopy, recovery

Introduction

Unilateral vocal fold paralysis (UVP) is a challenge for otolaryngologists and phoniatricians. Insufficient glottic closure during phonation can lead to severe vocal impairment with dysphonia and reduced vocal intensity. The etiology of UVP is often traumatic (especially in thyroid surgery) followed by malignancy-associated and idiopathic paralysis.[1] Recovery of vocal fold paralysis is mainly observed within 12 months after onset, as shown in a literature review of 717 cases with an idiopathic UVP.[2,3] Recovery of postoperative vocal fold paralysis in patients with thyroidectomy usually occurs within the first 6 months,[4] but according to a literature review, up to 11% do not recover.[5] Knowing the prognosis of UVP is helpful in planning therapy such as voice therapy and augmentation, either early temporary or permanent.[6-8] Stroboscopy is a tool for imaging the vibration of the vocal folds during phonation, especially the mucosal waves.[9,10] By analyzing the vibrations of the vocal folds, an assessment can be made as to the state of the mucosa (cover) and the underlying laryngeal muscle tone (body).[9,11] In patients with UVP, reduced vocal fold movement (adduction or abduction), vocal fold bowing, incomplete glottis closure, or a vibratory asymmetry, and absent mucosal waves (MW) can be observed in videostroboscopy. Recurrence of MW suggests reinnervation of the paretic vocal fold.[12-15] The aim of this study was to analyze the prognostic relevance of MW for recovery of mobility to full adduction and abduction in unilateral vocal fold paralysis in 100 patients.

Material and Methods

A retrospective review was made of a select group of 100 consecutive patients who had a UVP with complete vocal

[1]Section of Phoniatrics and Pedaudiology, Department of Otolaryngology, Head and Neck Surgery, University of Ulm, Ulm, Germany
[2]Department of Otolaryngology Head and Neck Surgery, Technical University, Munich, Germany
[3]Institute of Epidemiology and Medical Biometry, University of Ulm, Germany

Corresponding Author:
Rudolf Reiter, Section of Phoniatrics and Pedaudiology, Department of Otolaryngology, Head and Neck Surgery, University of Ulm, Frauensteige 12, 89070 Ulm, Germany.
Email: rudolf.reiter@uniklinik-ulm.de

fold immobility and a minimal mucosal gap (estimated <3 mm by 2 experienced phoniatricians) in videolaryngostroboscopy between January 2008 and January 2013. They were identified in our "Electronic Patient Record of the University Hospital Ulm"—a specially designed electronic health record. As we have a close connection to a large over-regional center of (thyroid) surgery, all their patients with voice problems are routinely seen in our department, usually the day after surgery, and could be included in this study. Taken together, all stroboscopic examinations were done 1 day to a maximum of 3 days after surgery (iatrogenic paralysis) and about up to 3 weeks after onset of the symptoms (other paralyses).

The videolaryngostroboscopic vocal fold examination was done (90° endostroboscope 5052, Wolf, Hamburg, Germany) and documented (rpSzene, Rehder, Hamburg, Germany). For measurement of the glottal gap and MW presence on the paralyzed vocal fold by a phoniatrician, phonation trials were performed with a sound pressure level of 65 dB and 100 Hz for men and 200 Hz for women. Only patients diagnosed with a UVP who had a vocal fold positioned with a minimal mucosal gap (estimated <3 mm) by videolaryngostroboscopy were included in this study group. It was only in these patients that evaluation by videolaryngostroboscopy was technically feasible because the vocal fold could get into contact (couple). Analysis of the presence of a MW in the paralyzed vocal fold was done by 2 experienced phoniatricians.

Patients with chronic laryngitis/leucoplakia, carcinoma, and scar formation of the vocal folds were not included as these pathologies influence mucosal waves. Similarly, patients with a medialization or augmentation of the paralyzed vocal fold were excluded. Positive MW (pMW) and negative MW (nMW) on the paralyzed vocal fold in stroboscopy were associated with complete recovery of the paralysis to full adduction and abduction. All patients were followed for at least 12 months at an interval of 3 to 6 months.

Statistical data analysis was performed with Microsoft Excel 2003 and SAS 9.3 (SAS Institute, Cary, North Carolina, USA). In the descriptive statistical analysis median, minimum and maximum were calculated for quantitative variables. For qualitative variables, absolute and relative frequencies with corresponding exact 95% confidence interval were calculated. The study was approved by the local Ethics Committees at the University of Ulm.

Results

Patients, Etiology of UVP, and Therapy

A hundred patients (median age 49.6 years; range, 16-81 years; 28 males/72 females) with a UVP were included. The group was divided iatrogenic/traumatic group (n = 82; median age 46.5 years; range, 16-76 years; 18 males/64 females), a malignancy-associated UVP (n = 10; median age 51.3 years; range, 16-70 years; 4 males/6 females), and an idiopathic UVP (n = 8; median age 43.5 years; range, 20.5-65.7 years; 6 males/2 females). The etiology of the 82 patients with iatrogenic/traumatic was thyroidectomy in 61 patients (56 thyroid hypertrophy and 5 patients with thyroid carcinoma), cardiac/carotid surgery in 14 patients, and spine surgery in 7 patients. In 10 patients, a UVP was observed as a primary symptom of a malignancy (8 carcinoma of the thyroid gland, 2 bronchial carcinoma), and in 8 cases, idiopathic paralysis was diagnosed. These patients received voice therapy (n = 88) or observation (n = 12). Patients with a medialization or augmentation of the paralyzed vocal fold were not included.

Time of Recovery

Out of the 100 patients, a complete recovery was seen in 75 patients with none in the remaining 25 patients. This recovery was observed at a median of 6.6 months (range, 1-14 months) after diagnosis. In the group of the 67 patients with iatrogenic/traumatic paresis, complete recovery occurred at a median of 6.5 months (range, 2-14 months). In thyroid hypertrophy surgery associated paralysis (n = 43), recovery was observed at a median of 4.6 months (range, 2-12 months) and in surgery because of a thyroid carcinoma (n = 5) in 10.4 months (range, 4-14 months). The 1 malignancy-associated UVP recovered after 10 months and the idiopathic paralysis (n = 7) after a median of 5.7 months (range, 3-9).

Relationship of Recovery With MW

As mentioned before, recovery was observed in 75% (75/100; 95% CI, 65.3-83.1). In total, 80 patients had pMW and 20 nMW.

In all patients with pMW at diagnosis, the chance of recovery of UVP was 91.25% (73/80; 95% CI, 82.8%-96.4%) (Figure 1), whereas the chance of recovery of the UVP was only 10% (2/20; 95% CI, 1.2%-31.7%) in patients with nMW (Figure 2, Table 1).

In the subgroup of the 82 patients with iatrogenic/traumatic UVP, a recovery was observed in 81.7% (67/82; 95% CI, 71.6%-89.4%), especially in 90.4% (66/73; 95% CI, 81.2%-96.1%) in patients with pMW. In 3 cases with nMW that did not resolve, the nerve was cut.

The best chance of recovery was in the group of the patients with idiopathic UVP in 87.5% (7/8; 95% CI, 47.4%-99.7%). All patients who recovered had pMW. Only 1 patient (10%, 1/10; 95% CI, 0.3%-44.5%) recovered in malignancy-associated UVP. This patient had nMW. Unilateral vocal fold paralyses with pMW were not observed in this group (Tables 2-4).

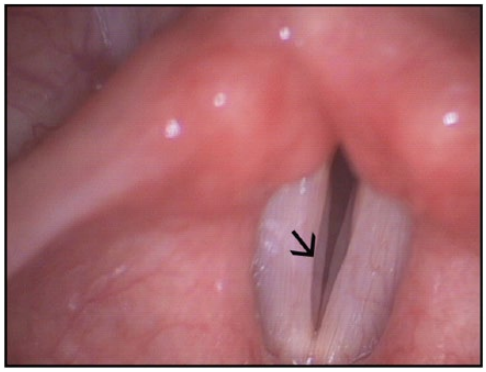

Figure 1. A patient with a unilateral vocal fold paralysis (UVP) on the right side. Positive mucosal waves (MWs) are seen in microstroboscopy on the paralyzed right (→) and the non-paralyzed vocal left fold.

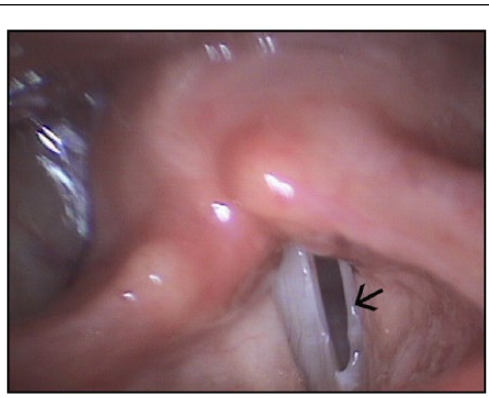

Figure 2. A patient with a unilateral vocal fold paralysis (UVP) on the left side. Positive mucosal waves (MWs) are seen in microstroboscopy on the non-paralyzed right side but are absent in the paralyzed left vocal fold (→).

Table 1. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in All Patients.

UVP	Positive MW	Negative MW	Total
Recovery	73	2	75
No recovery	7	18	25
Total	80	20	100

Discussion

The main reasons for UVP are iatrogenic/traumatic paralysis (about 4/5, especially in thyroid surgery) followed by malignancy-associated and idiopathic paralysis, as shown recently in a cohort of 400 patients[1] and confirmed in our study. Iatrogenic paralysis after (thyroid) surgery represents the majority of patients because we are associated with a large supra-regional center of thyroid surgery where difficult cases including revision operations are performed. All patients with voice problems after surgery were sent to our department and could be included in this study.

Unilateral vocal fold paralysis often causes severe impairment of the voice with dysphonia and reduced intensity because of insufficient glottic closure during phonation. However, UVPs have a potential of resolution that usually occurs within 12 months and in most cases within the first 6 months[2-4] (also confirmed in our study).

The overall rate of recovery was 75 of 100 (75%) in all patients and 67 of 82 (82%) in the iatrogenic/traumatic group, respectively. Studies have shown that permanent paralysis remains in 15% after thyroid surgery and in up to 20% in 717 patients with an idiopathic paralysis—similar to our study.[2,4]

In order to plan therapy and keep the patient informed, it is advantageous to know if any resolution of paralysis is likely. Laryngeal electromyography (LEMG) is an established method of assessing the neuromuscular status of the paralyzed vocal fold. Early evidence of reinnervation or innervation in paresis can be shown by this method. Laryngeal electromyography criteria for poor prognosis were the presence of spontaneous activity and absence or reduced recruitment of motor unit potentials.[16] In a meta-analysis by Rickert et al[16] and an analysis by Sittel et al,[17] LEMG predicted defective recovery defined as absence of completely free vocal fold mobility in up to 94%. However, in a study with a small cohort, it was shown that LEMG findings predict only 44.4% of the resolved cases,[18,19] which is less satisfactory. A drawback of the LEMG is the fact that it takes time and is an invasive procedure not well tolerated by patients. Furthermore, it requires expensive equipment and an experienced investigator.[9]

By analyzing the vibrations of the vocal folds, the state of the mucosa (cover), and the underlying laryngeal, muscle tone (body) can be assessed.[9,11] In microlarygostroboscopy, the mucosal wave represents the clinical correlate for muscular tonicity of the vocal cord. When present, recurrent nerve function is at least in a sense present, that the epithelium of the vocal cord can exactly follow the musculus vocalis movements in pitch as well as in intensity. However, mucosal waves can only be judged when there is no phonation gap ≤3 mm in UVP, enabling the mucosa to get in contact with each other. Furthermore, MWs are reduced or absent if mucosa (cover) adheres/sticks to the musculus vocalis and/or ligamentum vocale (body). This phenomenon is observed in patients with chronic laryngitis/leucoplakia, carcinoma, and scar formation of the vocal folds. The warning factors for progression of dysplasia to invasive carcinomas is a nonvibrating segment/absence or reduction of the MW in videostroboscopy.[20,21] A scarred vocal fold has an absent or limited MW[22] as well, and for this reason, such patients were excluded from our study.

To our knowledge, this is the first clinical study to report the prognostic relevance of MW in laryngostroboscopy for

Table 2. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in Iatrogenic/Traumatic UVP.

UVP	Positive MW	Negative MW	Total
Recovery	66 (42 thyroid hypertrophy, 5 thyroid carcinoma, 12 cardiac/carotid surgery, 7 spine surgery)	1 (1 thyroid hypertrophy)	67
No recovery	7 (6 struma, 1 cardiac/carotds surgery)	8 (7 thyroid hypertrophy, 1 cardiac/carotid surgery)	15
Total	73	9	82

Table 3. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in Malignancy-Associated UVP.

UVP	Positive MW	Negative MW	Total
Recovery	0	1 (1 thyroid carcinoma)	1
No recovery	0	9 (7 thyroid carcinoma, 2 bronchial carcinoma)	9
Total	0	10	10

Table 4. Recovery of Unilateral Vocal Fold Paralysis (UVP) Dependent on Mucosal Wave (MW) in Idiopathic UVP.

UVP	Positive MW	Negative MW	Total
Recovery	7	0	7
No recovery	0	1	1
Total	7	1	8

recovery of a (iatrogenic/traumatic) UVP. It was shown that the chance of recovery of UVP was over 90% when pMW were present at diagnosis in patients with a UVP, whereas only in 10% if not. Laryngostroboscopy is always done routinely in patients with UVP by an experienced phoniatrician or laryngologist. It is a noninvasive procedure that takes only a few minutes.

A disadvantage of laryngostrobsocopy is the fact that in our experience, reliable stroboscopic signals are only obtained in patients with the paralyzed vocal fold close to midline during phonation (glottis gap ≤3 mm during phonation). This is the reason why only 61 out of 100 patients could be analysed in a study by Harries and Morrison.[9] In a recent study with 400 patients with a vocal fold paralysis, it was shown that microstroboscopy was technically feasible in 76% of the patients because the vocal folds could get into contact (couple).[1] These data show that a routinely performed microlarygostroboscopy is a very good, noninvasive alternative to EMG in patents with UVP and—in addition—much better tolerated and less expensive.

Conclusion

The authors conclude that positive mucosal waves in microstroboscopy are a simple predictor for recovery of (iatrogenic/traumatic) UVP and should be used in routinely for diagnosis and prognosis

Acknowledgment

We thank John Bennett, FRCS, for critically reading this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Reiter R, Pickhard A, Smith E, et al. Vocal cord paralysis—analysis of a cohort of 400 patients. *Laryngorhinootologie*. 2015;94(2):91-96.
2. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118(7):13030-13037.
3. Misono S, Merati AL. Evidence-based practice: evaluation and management of unilateral vocal fold paralysis. *Otolaryngol Clin North Am*. 2012;45(5):1083-1108.
4. Steurer M, Passler C, Denk DM, Schneider B, Niederle B, Bigenzahn W. Advantages of recurrent laryngeal nerve identification in thyroidectomy and parathyroidectomy and the importance of preoperative and postoperative e laryngoscopic examination in more than 1000 nerves at risk. *Laryngoscope*. 2002;112(1):124-133.
5. Dralle H, Sekulla C, Lorenz K, Brauckhoff M, Machens A; German IONM Study Group. Intraoperative monitoring of the recurrent laryngeal nerve in thyroid surgery. *World J Surg*. 2008;32(7):1358-1366.
6. Reiter R, Brosch S. Laryngoplasty with hyaluronic acid in patients with unilateral vocal fold paralysis. *J Voice*. 2013;26:785-791.
7. Friedman AD, Burns JA, Heaton JT, Zeitels SM. Early versus late injection medialization for unilateral vocal cord paralysis. *Laryngoscope*. 2010;120(10):2042-2046.
8. Yung KC, Likhterov I, Courey MS. Effect of temporary vocal fold injection medialization on the rate of permanent medialization laryngoplasty in unilateral vocal fold paralysis patients. *Laryngoscope*. 2011;121(10):2191-2194.

9. Harries ML, Morrison M. The role of stroboscopy in the management of a patient with a unilateral vocal fold paralysis. *J Laryngol Otol*. 1996;110(2):141-143.
10. Sercarz JA, Berke GS, Ming Y, Gerratt BR, Natividad M. Videostroboscopy of human vocal fold paralysis. *Ann Otol Rhinol Laryngol*. 1992;101(7):567-577.
11. Hirano M, Yoshida Y, Yoshida T, Tateishi O. Strobofiberscopic video recording of vocal fold vibration. *Ann Otol Rhinol Laryngol*. 1985;94:588-590.
12. Simpson CB, Cheung EJ, Jackson CJ. Vocal fold paresis: clinical and electrophysiologic features in a tertiary laryngology practice. *J Voice*. 2009;23(3):396-398.
13. Simpson CB, May LS, Green JK, Eller RL, Jackson CE. Vibratory asymmetry in mobile vocal folds: is it predictive of vocal fold paresis? *Ann Otol Rhinol Laryngol*. 2011;120(4):239-242.
14. Altman KW. Laryngeal asymmetry on indirect laryngoscopy in a symptomatic patient should be evaluated with electromyography. *Arch Otolaryngol Head Neck Surg*. 2005;131(4):356-360.
15. Kosztyła-Hojna B, Rogowski M. Usefulness of videolaryngo-stroboscopy in the diagnosis of laryngeal pathology. *Pol Merkur Lekarski*. 2003;14(83):413-416.
16. Rickert SM, Childs LF, Carey BT, Murry T, Sulica L. Laryngeal electromyography for prognosis of vocal fold palsy: a meta-analysis. *Laryngoscope*. 2012;122(1):158-161.
17. Sittel C, Stennert E, Thumfart WF, Dapunt U, Eckel HE. Prognostic value of laryngeal electromyography in vocal fold paralysis. *Arch Otolaryngol Head Neck Surg*. 2001;127(2):155-160.
18. Blitzer A, Crumley RL, Dailey SH, et al. Recommendations of the Neurolaryngology Study Group on laryngeal electromyography. *Otolaryngol Head Neck Surg*. 2009;140(6):782-793.
19. Munin MC, Rosen CA, Zullo T. Utility of laryngeal electromyography in predicting recovery after vocal fold paralysis. *Arch Phys Med Rehabil*. 2003;84(8):1150-1153.
20. Djukic V, Milovanovic J, Jotic AD, Vukasinovic M. Stroboscopy in detection of laryngeal dysplasia effectiveness and limitations. *J Voice*. 2014;28(2):262.e13-262.
21. El-Demerdash A, Fawaz SA, Sabri SM, Sweed A, Rabie H. Sensitivity and specificity of stroboscopy in preoperative differentiation of dysplasia from early invasive glottic carcinoma. *Eur Arch Otorhinolaryngol*. 2015;272(5):1189-1193.
22. Chhetri DK, Berke GS. Injection of cultured autologous fibroblasts for human vocal fold scars. *Laryngoscope*. 2011;121(4):785-792.

Reprinted by permission of Laryngoscope. 2010; 120(7):1376-1382.

The Laryngoscope
© 2010 The American Laryngological, Rhinological and Otological Society, Inc.

Laryngoscopy of Vocal Fold Paralysis: Evaluation of Consistency of Clinical Findings

David E. Rosow, MD; Lucian Sulica, MD

Objectives/Hypothesis: Laryngoscopy is the principal tool for the clinical assessment of vocal fold paralysis (VFP). Yet no consistent, unified vocabulary to describe laryngoscopic findings exists, compromising the evaluation and comparison of cases, outcomes, and treatment. The goal of this investigation was to evaluate laryngoscopic findings in VFP for inter- and intra-rater consistency.

Study Design: Prospective survey-based study.

Methods: Half-minute excerpts from stroboscopic exams of 22 patients with VFP were mailed to 22 fellowship-trained laryngologists. Each reviewer was sent exams in randomized order, with three random repeats included to determine intra-rater reliability. Twelve laryngoscopic criteria were assessed and recorded on preprinted sheets. Eleven criteria were binary in nature (yes/no); glottic insufficiency was rated on a four-point scale (none/mild/moderate/severe). Raters were blinded to clinical history, each other's ratings, and to their own previous ratings. Inter-rater agreement was calculated by Fleiss' kappa.

Results: Twenty reviewers (91%) replied. Intra-rater reliability by reviewer ranged from 66% to 100% and by laryngoscopic criterion from 77% to 100%. Of the laryngoscopic criteria used, glottic insufficiency (κ = 0.55), vocal fold bowing (κ = 0.49), and salivary pooling (κ = 0.45) showed moderate agreement between reviewers. Arytenoid stability (κ = 0.1), arytenoid position (κ = 0.12), and vocal fold height mismatch (κ = 0.12) showed poor agreement. The remainder showed slight to fair agreement.

Conclusions: Inter-rater agreement on commonly used laryngoscopic criteria is generally fair to poor. Glottic insufficiency, vocal fold bowing, and salivary pooling demonstrated the most agreement among responding laryngologists. These findings suggest a need for a standardized descriptive scheme for laryngoscopic findings in VFP.

Key Words: Vocal fold paralysis, vocal cord paralysis, laryngoscopy, reliability, agreement.

Level of Evidence: 2b

***Laryngoscope*, 120:1376–1382, 2010**

INTRODUCTION

Laryngoscopy is the mainstay investigation in the diagnosis of vocal fold paralysis (VFP), and not infrequently the sole diagnostic evaluation on which direct treatment (as opposed to treatment of underlying cause) is based. For much of the history of laryngology, complex nosological schemes have been constructed around the laryngoscopic appearance of VFP. Semon's law, for instance, held that differences in vocal fold position were the product of differential vulnerability of adductor and abductor fibers of the recurrent laryngeal nerve.[1] Wagner and Grossman maintained that the position of the paralyzed vocal fold was indicative of the integrity of the superior laryngeal nerve.[2,3] Such constructs were abandoned as increasing anatomical knowledge and careful physiological investigations invalidated their assumptions. In the course of this progress, systematic analysis of the laryngoscopic appearance of VFP has apparently been abandoned too, as unrewarding in the face of the evident complexity of the neuropathology underlying VFP.

Yet, it is clear to any clinician that VFP manifests itself laryngoscopically in many different ways. Terms like height and length mismatch, arytenoid prolapse, flaccidity, posterior gap, and others that plainly refer to physical characteristics of the appearance of the paralyzed vocal fold make their appearance in the professional discourse with some frequency. Woodson, in a seminal study of the paralyzed vocal fold, described several such features: foreshortening, arytenoid displacement, decreased vocal process contact, bowing, and ventricular hyperfunction.[4] Recent literature has addressed vocal process height asymmetry.[5,6] Both the configuration and degree of glottic insufficiency related to VFP have been presented as important in the selection of treatment techniques.[7,8]

From the Department of Otorhinolaryngology, Weill Cornell Medical College, New York, New York, U.S.A.

Editor's Note: This Manuscript was accepted for publication March 17, 2010.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Lucian Sulica, MD, Director, Laryngology/Voice Disorders, Department of Otorhinolaryngology, Weill Cornell Medical College, 1305 York Avenue, New York, NY 10021. E-mail: lus2005@med.cornell.edu

DOI: 10.1002/lary.20945

Rating Sheet for Evaluators

Examiner #_____

Patient #_____

Volitional Adduction	Present	Not Present
Vocal Process Contact	Normal	Impaired/Decreased
Arytenoid Stability ("Jostle Sign")	Normal (no jostle)	Impaired/Decreased
Arytenoid Position	Normal	Displaced
Vocal Fold Tone	Normal	Decreased
Vocal Fold Atrophy	Not Present	Present
Vocal Fold Bowing	Not Present	Present
Vocal Fold Shortening	Not Present	Present
Salivary Pooling	Not Present	Present
Height Mismatch	Not Present	Present
Ventricular (False vocal fold) contraction	Not Present	Present
Glottic Insufficiency	None Mild	Moderate Severe

Fig. 1. Rating sheet for evaluators.

Despite the acknowledgment of variability in the laryngoscopic appearance of VFP implicit in these terms, no widely accepted rating system, or even a consistent, unified vocabulary to describe such variability exists. Consistency and reproducibility is fundamental in the evaluation and comparison of cases, their outcomes, and their treatment; even a brief reflection on the House-Brackmann scale for grading facial paralysis reveals the broad potential utility of such a standardized approach.

The goal of this investigation was to evaluate characteristics of the laryngoscopic appearance of VFP with respect to inter- and intra-rater consistency, and to identify features for which clinical consensus exists, which might lend themselves to a useful, standardized description system for VFP.

MATERIALS AND METHODS

Selection of Examinations

Strobovideolaryngoscopies of patients with a clinical diagnosis of VFP based on history, physical examination, and laryngoscopy were selected from a corpus of such examinations recorded during the course of routine evaluation. All recording was made under stroboscopic light using either a rigid glass rod peroral laryngoscope (Model 9106; KayPentax, Lincoln Park, NJ) or a distal chip flexible transnasal laryngoscope (VNL-1170K; Pentax Medical, Montvale, NJ). Only patients with VFP of known cause were included. Nineteen had pathology or injury limited to the recurrent laryngeal nerve, and three had paralysis from vagal neuropathy. For inclusion, the examination had to feature a sustained, unobstructed view of the glottis, arytenoids, aryepiglottic folds, and pyriform sinuses. A 20-second sample of each exam, containing at least one example each of phonatory adduction and postphonatory abduction as well as several cycles of phonatory vibration, was selected and saved. The pitch and intensity capabilities of patients were variable from exam to exam, as one would expect in cases of VFP. However, as evaluators' ratings of identical examinations were assessed in this study, no effort was made to standardize these parameters among examinations.

Patients who had been treated for their VFP in any way, including injection augmentation, framework surgery, and reinnervation were excluded. Cases of vocal fold paresis in which significant gross vocal fold mobility remained, even if it was clearly less than normal, were excluded. The authors recognize that the distinction between paralysis and paresis is not always sharply defined and does not necessarily reflect the underlying neurologic status.

Each exam was numbered, and randomly ordered lists of these exams, one for each potential reviewer, were generated. In each list, three exams were selected to be repeated by a random number generator, which created a new set of three numbers for each of the reviewers and brought the total number of exams to 25. Therefore, every reviewer had a randomly selected series of repeat examinations to test their intra-rater reliability. The purpose of de novo, random selection of the repeat exams for every reviewer was to eliminate any possibility that one exam might be more easily identified on repeat viewing than the others. Such a scenario would bias the entire sample and yield an artificially

TABLE I.
Examiners and Their Intra-Rater Reliabilities Determined Via the Three Repeat Examinations Given to All Participants.

Examiner	% Agreement	Pearson	Spearman
A	0.944	0.932	0.898
B	0.806	0.794	0.775
C	0.889	0.836	0.839
D	0.833	0.849	0.768
E	0.917	0.906	0.906
F	0.861	0.846	0.764
G	1.000	1.000	1.000
H	0.944	0.949	0.957
I	0.667	0.616	0.446
J	0.778	0.700	0.678
K			
L	0.694	0.605	0.526
M	0.778	0.711	0.661
N	0.778	0.752	0.652
O			
P	0.833	0.773	0.681
Q	0.861	0.893	0.775
R	0.750	0.700	0.623
S	0.889	0.890	0.803
T	0.750	0.675	0.650
U	1.000	1.000	1.000
V	0.833	0.631	0.622
Average	0.840	0.803	0.751

Examiners K and O did not participate in the study.

high intra-rater reliability. By randomizing the repeat exams across all reviewers, this potential source of bias was eliminated. Each list of exams was then burned to a DVD with all identifying information removed; the file names on the disc simply appeared in order as 01.avi, 02.avi, and so on.

Selection of Reviewers

Twenty-two fellowship-trained laryngologists were asked to participate in the study. A package containing the disc of videos, 25 prelabeled rating sheets, an institutional review board waiver, and a return mailer was sent to each. Results from returned rating sheets were entered into Excel spreadsheets (Microsoft Corp., Redmond, WA) for analysis.

Examination Rating

Each stroboscopic exam was rated with respect to 12 features (Fig. 1). A standard vocabulary to describe the paralyzed vocal fold does not exist. Therefore, criteria were selected from a literature review of the diagnosis and treatment of VFP, including those articles cited above[4–8] and others,[9,10] and informal consultation with colleagues. No formal definition of each term was presented; the study relied on common clinical usage to inform raters' perception of the meaning of each term rather than any formal training. We also acknowledge that some terms might be in part redundant or overlapping—for example, tone, atrophy and bowing—but sought to evaluate the utility of each of these concurrently. Responses to 11 of 12 categories were binary (normal/not normal or present/not present); the exception was glottic insufficiency, which was rated on a four-point scale of none/mild/moderate/severe. However, for the purposes of statistical analysis, answers in this category were grouped in binary fashion into none/mild or moderate/severe. Reviewers were instructed to view and rate examinations sequentially; individual exams could be reviewed an unlimited number of times, but once a new examination was started, examiners were instructed not to return to any prior examinations or alter ratings. Raters were therefore blinded to clinical history, each other's ratings, and their own previous ratings. Reviewers were also asked not to include written justification for their answers on the rating forms.

Statistical Analysis

Inter-rater reliability was determined using the kappa statistic as described by Fleiss[11,12]; values closer to 0 represent poor agreement, whereas those close to 1 represent near-perfect agreement. Although no uniformly agreed-upon scale exists for Fleiss' kappa, Fleiss described a scale where values >0.75 represent excellent agreement, 0.40 to 0.75 represent fair to good agreement, and values <0.40 represent poor agreement.[12] When an examiner omitted a rating for one of the 12 categories for an examination (accidental or otherwise), all of that examination's rankings in that category were excluded from kappa

TABLE II.
Laryngoscopic Criteria and Their Associated Inter- and Intra-Rater Reliability Scores.

	Inter-Rater	Intra-Rater		
Laryngoscopic Criterion	Fleiss	% Agreement	Pearson	Spearman
Volitional adduction	0.335	0.900	−0.053	−0.053
Vocal process contact	0.303	0.817	0.445	0.445
Arytenoid stability (jostle)	0.097	0.833	0.615	0.615
Arytenoid position	0.119	0.817	0.629	0.629
Vocal fold tone	0.310	0.900	0.744	0.744
Vocal fold atrophy	0.326	0.867	0.726	0.726
Vocal fold bowing	0.488	0.883	0.714	0.714
Vocal fold shortening	0.225	0.817	0.610	0.610
Salivary pooling	0.454	0.900	0.762	0.762
Height mismatch	0.123	0.733	0.457	0.457
Ventricular contraction	0.217	0.883	0.756	0.756
Glottic insufficiency	0.550	0.733	0.818	0.798

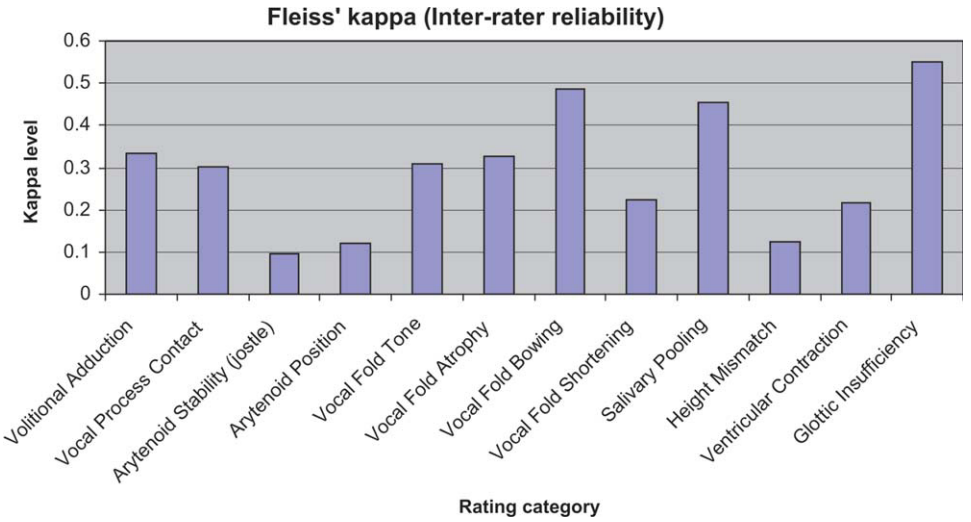

Fig. 2. Inter-rater reliability as determined by Fleiss' kappa. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

analysis. This exclusion assured a constant denominator in all the statistical calculations.

Intra-rater reliability for each examiner was determined by comparing the 12 laryngoscopic criteria in each of the three repeated patients, for a total of 36 comparison points. Intra-rater reliability for each criterion was determined in a similar fashion, with the denominator determined by adding up the 20 examiners' three repeated tests, by criterion, for a total of 60 comparison points. Three complementary methods were used to assess intra-rater reliability, both for each examiner and for each laryngoscopic criterion. The overall percent agreement was calculated, simply as the number of points of agreement divided by the total. This was compared to two known measures of correlation, Pearson product moment coefficient and Spearman corrected rank correlation coefficient.

This investigation was approved by the institutional review board of Weill Cornell Medical College.

RESULTS

Twenty of 22 examiners returned the survey, for a 91% response rate. The overall intra-rater reliability for each individual examiner varied between 44% and 100%, with an average internal consistency of 75% to 84%, depending on the statistical method used (Table I). As measured by all three statistics, 18 of 20 examiners (90%) showed >60% internal consistency (Table I).

The intra-rater reliability for each stroboscopic criterion had, for the most part, a very similar range of 44% to 100% (Table II), The single exception was volitional adduction, a category in which not a single examiner rated an exam as normal on both viewings. As a result, despite a 90% rate of intra-rater agreement, this category was found to have near-0 intra-rater correlation by both Spearman and Pearson correlation coefficients. Overall, height mismatch, vocal fold shortening, and vocal process contact had the lowest intra-rater reliability, whereas the ratings of salivary pooling, glottic insufficiency, ventricular contraction, and vocal fold tone were generally consistent.

Inter-rater reliability for each stroboscopic criterion was determined by kappa analysis. As represented in Figure 2, these kappa values ranged from 0.10 (poor

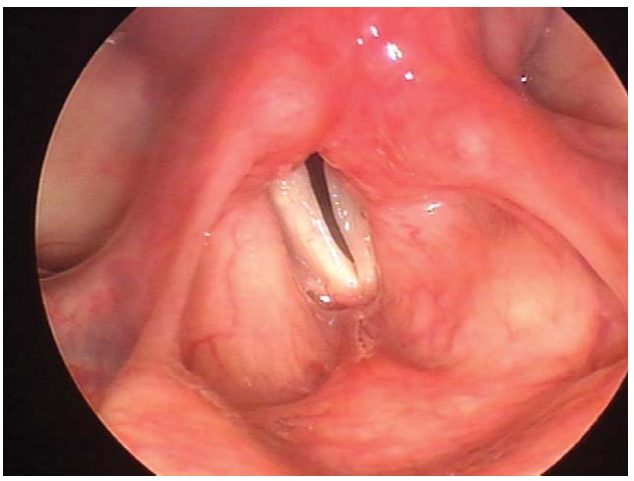

Fig. 3. Vocal process contact impaired. This case generated the most consistent rating for impaired vocal process contact. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

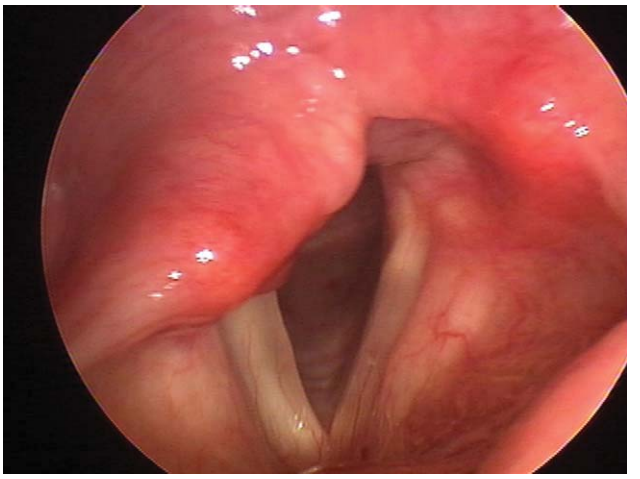

Fig. 4. Arytenoid position displaced. This case generated the most consistent rating for displaced arytenoid position. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

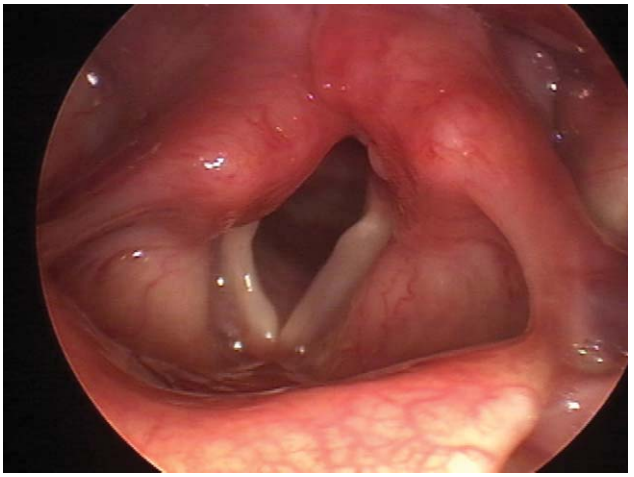

Fig. 5. Vocal fold tone decreased. This case generated the most consistent rating for decreased vocal fold tone. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

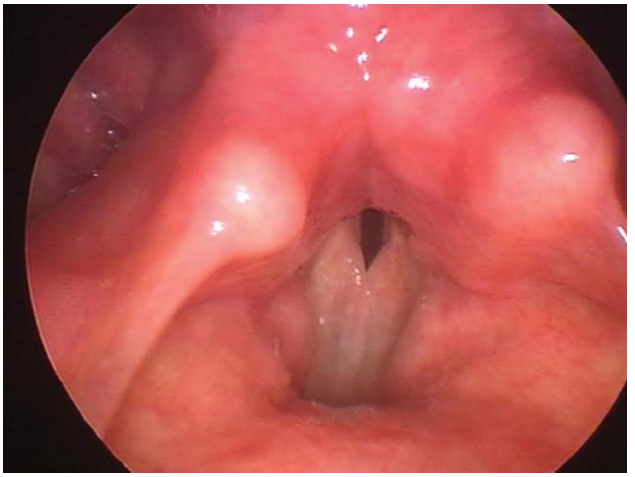

Fig. 7. Glottic insufficiency: none to mild. This case generated the most consistent rating for no or mild glottic insufficiency. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

agreement) to 0.55 (moderate agreement). Ratings for arytenoid stability, arytenoid position, and height mismatch generated the poorest values, whereas glottic insufficiency, vocal fold bowing, and salivary pooling resulted in the best (Fig. 3–Fig. 8].

DISCUSSION

Traditionally conceptualized as an all-or-none phenomenon, VFP has been shown by ample clinical and laboratory investigation to represent a continuum of neurogenic dysfunction encompassing partial denervation, complete denervation, and variable degrees and patterns of reinnervation. It should come as no surprise, then, that its clinical appearance too is highly variable. This is not synonymous with random, however; this variability no doubt reflects the considerable heterogeneity in the neurologic dysfunction that underlies the immobile vocal fold. Historical efforts to decipher the laryngoscopic appearance have fallen short, compromised by oversimplifications and an incomplete understanding of the relevant pathophysiology. Given these limitations, we find these failures neither particularly surprising nor discouraging.

Lest a re-examination of laryngoscopy in VFP be considered unnecessary or irrelevant, it is important to appreciate that existing neurodiagnostic techniques have also been defeated by the complex neurologic picture underlying the paralyzed vocal fold. Electromyography has proved to be as qualitative as laryngoscopy, and it is similarly susceptible to individual variation in interpretation. Although it has yielded crucial insight in the

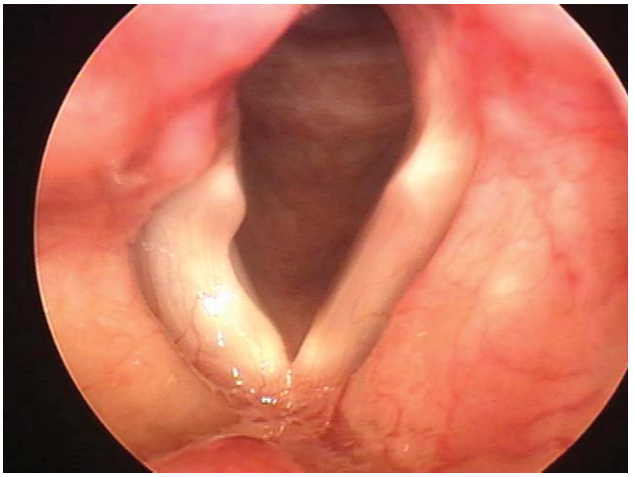

Fig. 6. Vocal fold shortened. This case generated the most consistent rating for shortened vocal fold. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

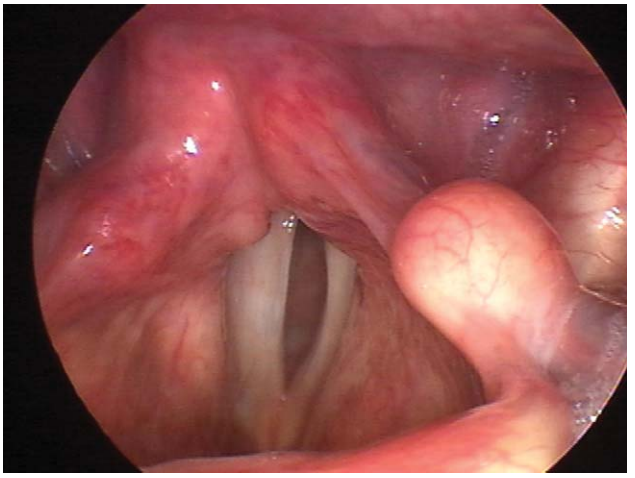

Fig. 8. Glottic insufficiency: moderate to severe. This case generated the most consistent rating for moderate to severe glottic insufficiency. All ratings were made from dynamic examinations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

pathophysiology of VFP, it has proved particularly disappointing as a prognostic tool, at best only partially helpful in informing treatment of individual patients.

This study was undertaken with the understanding that any re-examination of the laryngoscopy of VFP must be carried out using a commonly agreed-upon array of findings. The evolving discussion about vocal fold paresis, or partial paralysis of the vocal fold, reveals clearly that lack of consensus regarding clinical findings prevents conclusions regarding diagnosis, much less prognosis, treatment, and outcomes.[13]

No definitions of the terms used were provided to the reviewers for the simple reason that no such definitions exist. The terms used in this study have entered the literature informally, and the medical discourse has generally assumed that a broad and consistent understanding of these descriptors exists. This study is a formal testing of that assumption.

This investigation reveals that the evaluation of laryngoscopic appearance of VFP remains a personal and individual activity. As demonstrated by multiple correlation calculations, most evaluators were relatively consistent in their own evaluations across the entire range of features presented. These results were roughly equivalent to those of Rosen, who found that two thirds of voice professionals reviewing stroboscopic exams had intra-rater reliability scores less than 0.80.[14] Thus, it is possible that individual practitioners might use laryngoscopic features to analyze cases for diagnosis, as for example to identify degree of denervation, distribution of involvement across laryngeal muscles, and for selection and timing of treatment in a reasonably reliable manner. It remains to be proven, of course, that pathophysiologic aspects such as degree and distribution of neural compromise indeed have consistent laryngoscopic correlates. Incidentally, this study does not demonstrate whether individual observations across multiple examinations are reliable, or if changes over time in the same case can be consistently identified.

On the other hand, inter-rater variability revealed considerable lack of consensus regarding all aspects but salivary pooling, bowing, and a simplified rating of the degree of glottic insufficiency. Our study might even have been biased in favor of greater inter-rater agreement by the inclusion of the audio track in the video samples sent to reviewers. Such additional information might provide clues to blinded reviewers who are ultimately being studied for their video perceptual analysis alone. Future work in this area will require removal of all audio from samples sent to reviewers. Not only is this lack of agreement discouraging for the development of a unified rating system for this disorder, it also calls into question existing assumptions in the literature about consensus in the rating of features such as posterior gap (an important factor in the selection of patients for arytenoid adduction surgery), vocal fold height (hypothesized to be relevant to rehabilitation technique), and other features referred to in the discussion of the evaluation of unsatisfactory results of medialization.[7,15–17] Generalizations from study to study might be compromised by patient populations that are not comparable or equivalent. Also, treatment recommendations or descriptions of outcome based on laryngoscopic features are likely to be of limited utility. The prospects for agreement on vocal fold paresis, where clinical variability would be expected to be greater than in VFP at the same time that the degree of abnormality would be less, appear to be extremely poor.

Based on our results, degree of glottic insufficiency, vocal fold bowing, salivary pooling, and perhaps to a lesser extent volitional adduction, vocal fold tone, and vocal fold atrophy appear to be the best candidates for development into a standardized system of rating VFP. A rating or classification system for VFP based on only the three most consistently appreciated criteria might not be discriminating enough to be useful in diagnosis or treatment. We hypothesize that more formal development of rating categories, including explicit definitions and examples, would generate greater inter-rater agreement, for the terms and concepts evaluated in this investigation have received relatively little formal attention despite commonplace clinical use. We intend to explore this further before trialing an integrated rating system. At the same time, we recognize the possibility that individual variation in laryngeal anatomy and possibly in innervation, and the heterogeneity of neuropathic dysfunction might yet defeat such an effort.

CONCLUSION

However, although individuals are often consistent in their own evaluation of laryngoscopic features of VFP, little consensus appears to exist among physicians regarding these same findings. This raises the possibility that many assumptions about the significance of laryngoscopic features might not be reliable. This is an obvious challenge to arriving at a unified understanding of the laryngoscopic appearance of the disorder and will need to be addressed.

Results suggest that degree of glottic insufficiency, vocal fold bowing, and salivary pooling appear to be laryngoscopic features in cases of VFP with the highest inter-rater reliability. With further investigation and standardization, these might form a basis for the development of a clinically useful rating scheme.

Acknowledgments

We gratefully acknowledge the participation of the following as raters or consultants in this study: Drs. Linda Gerber, Gregory Postma, Michael Johns, Edward Damrose, Marshall Smith, Tanya Meyer, Gaelyn Garrett, Richard Kelly, Mark Courey, Seth Dailey, Ajay Chitkara, Sid Khosla, Pieter Noordzij, Libby Smith, Phillip Song, James Thomas, Albert Merati, Andrew McWhorter, Joel Blumin, Michiel Bove, and Milan Amin.

BIBLIOGRAPHY

1. Semon F. Clinical remarks on the proclivity of the abductor fibres of the the recurrent laryngeal nerve to become

affected sooner than the adductor fibres, or even exclusively, in cases of undoubted central or peripheral injury or disease of the roots or trunks of the pneumogastric, spinal accessory, or recurrent nerves. *Arch Laryngol New York* 1881;2:197–222.
2. Wagner R. Die medianstellung des stimmbandes bei Recurrenslahmung. *Archiv fur Pathologische Anatomie und Physiologie* 1890;120:437–459.
3. Grossman M. Experimentelle beitrage zur lehre von der posticuslahmung. *Archiv Laryngologie und Rhinologie* 1897;6:282–360.
4. Woodson GE. Configuration of the glottis in laryngeal paralysis. I: clinical study. *Laryngoscope* 1993;103(11 pt 1): 1227–1234.
5. Tsai V, Celmer A, Berke GS, Chhetri DK. Videostroboscopic findings in unilateral superior laryngeal nerve paralysis and paresis. *Otolaryngol Head Neck Surg* 2007;136: 660–662.
6. Rubin AD, Praneetvatakul V, Heman-Ackah Y, Moyer CA, Mandel S, Sataloff RT. Repetitive phonatory tasks for identifying vocal fold paresis. *J Voice* 2005;19: 679–686.
7. Netterville JL, Stone RE, Luken ES, Civantos FJ, Ossoff RH. Silastic medialization and arytenoid adduction: the Vanderbilt experience. A review of 116 phonosurgical procedures. *Ann Otol Rhinol Laryngol* 1993;102: 413–424.
8. Woodson GE, Picerno R, Yeung D, Hengesteg A. Arytenoid adduction: controlling vertical position. *Ann Otol Rhinol Laryngol* 2000;109:360–364.
9. Benninger MS, Crumley Rl, Ford CN, et al. Evaluation and treatment of the unilateral paralyzed vocal fold. *Otolaryngol Head Neck Surg* 1994;111:497–508.
10. Woo P. Arytenoid adduction and medialization laryngoplasty. *Otolaryngol Clin North Am* 2000;33:817–839
11. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–382.
12. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd ed. New York, NY: Wiley; 1981.
13. Sulica L, Blitzer A. Vocal fold paresis: evidence and controversies. *Curr Opin Otolaryngol Head Neck Surg* 2007;15: 159–162.
14. Rosen CA. Stroboscopy as a research instrument: development of a perceptual evaluation tool. *Laryngoscope* 2005; 115:423–428.
15. Woo P, Pearl AW, Hsiung MW, Som P. Failed medialization laryngoplasty: management by revision surgery. *Otolaryngol Head Neck Surg* 2001;124:615–621.
16. Anderson TD, Spiegel JR, Sataloff RT. Thyroplasty revision: frequency and predictive factors. *J Voice* 2003;17: 442–448.
17. Cohen JT, Bates DD, Postma GN. Revision Gore-Tex medialization laryngoplasty. *Otolaryngol Head Neck Surg* 2004;131:236–240.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2011; 120(4):239-242.

Annals of Otology, Rhinology & Laryngology 120(4):239-242.
© 2011 Annals Publishing Company. All rights reserved.

Vibratory Asymmetry in Mobile Vocal Folds: Is It Predictive of Vocal Fold Paresis?

C. Blake Simpson, MD; Linda Seitan May, MD; Jill K. Green, MS; Robert L. Eller, MD; Carlayne E. Jackson, MD

Objectives: The purpose of this study was to determine whether the videostroboscopic finding of vibratory asymmetry in mobile vocal folds is a reliable predictor of vocal fold paresis. In addition, the ability of experienced reviewers to predict the distribution (left/right/bilateral) of the paresis was investigated.

Methods: This is a retrospective chart review of all patients who presented to our clinic during a 3-year period with symptoms suggestive of glottal insufficiency (vocal fatigue or reduced vocal projection) accompanied by the videostroboscopic findings of bilateral normal vocal fold mobility and vibratory asymmetry. Twenty-three of these patients underwent diagnostic laryngeal electromyography of the thyroarytenoid and cricothyroid muscles to determine the presence of vocal fold paresis.

Results: Nineteen of the 23 patients (82.6%) were found to have electrophysiological evidence of vocal fold paresis, either unilaterally or bilaterally, when videostroboscopic asymmetry was present in mobile vocal folds. However, the three expert reviewers' ability to predict the distribution (left/right/bilateral) of the paresis was poor (26.3%, 36.8%, and 36.8%, respectively).

Conclusions: The videostroboscopic finding of vibratory asymmetry in mobile vocal folds is a reliable predictor of vocal fold paresis in most cases. However, the ability of expert reviewers to determine the distribution (left/right/bilateral) of the paresis using videostroboscopic findings is poor. This study highlights the value of laryngeal electromyography in arriving at a correct diagnosis in this clinical situation.

Key Words: electromyography, videostroboscopy, vocal fold paralysis, vocal fold paresis.

INTRODUCTION

Vocal fold paresis (VFP) is a well-established, albeit controversial, entity. Its incidence is not well established, but it is likely rare. The few reports that are available in the literature have shown a range of as many as 29 cases in a year to as few as 13 cases over 4 years in tertiary laryngology practices.[1-4] Although all of these studies used laryngeal electromyography (LEMG) to confirm the diagnosis, clinicians often use subtle asymmetries on videostroboscopy as indicators that paresis is likely present. During videostroboscopic examination, reduced vocal fold movement (adduction or abduction), vocal fold bowing, incomplete glottal closure, and vibratory asymmetry can all be associated with VFP.[4,5] Rubin et al[6] have also described the use of repetitive phonatory tasks to induce fatigue as a means of bringing out hypomobility in paretic vocal folds. As pointed out by Sulica and Blitzer, however, "Separating innocent asymmetries [on laryngoscopy] from significant findings may present the greatest challenge in defining vocal fold paresis."[7(p159)] The clinical setting of glottal insufficiency symptoms and grossly intact vocal fold mobility has previously been described. In these cases, vibratory asymmetry may be the only laryngoscopic clue to suggest VFP.[7] Identification of the asymmetry may help guide the clinician toward performing LEMG and eventually confirming a diagnosis of VFP.

The purpose of this study was to determine whether the videostroboscopic finding of vibratory asymmetry in mobile vocal folds was a reliable predictor of VFP. In addition, the ability of experienced reviewers to predict the distribution (left/right/bilateral) of the paresis was investigated.

METHODS

Institutional Review Board approval was obtained

From the Departments of Otolaryngology–Head and Neck Surgery (Simpson, May, Green) and Neurology (Jackson), University of Texas Health Science Center–San Antonio, and the Department of Otolaryngology–Head and Neck Surgery, Wilford Hall Medical Center, Lackland Air Force Base (Eller), San Antonio, Texas.

Presented at the meeting of the American Broncho-Esophagological Association, Las Vegas, Nevada, April 28-29, 2010.

Correspondence: C. Blake Simpson, MD, Dept of Otolaryngology–Head and Neck Surgery, University of Texas Health Science Center, Medical Arts and Research Center, 8300 Floyd Curl Dr, San Antonio, TX 78229.

from our institution before the study period. A retrospective chart review was carried out for all patients who presented to our clinic during a 3-year period and underwent LEMG for suspected vocal fold paresis.

Over the study period, 48 patients with suspected VFP underwent diagnostic LEMG. Of those, 23 patients met the study criteria with symptoms of VFP (vocal fatigue or reduced vocal projection) accompanied by the videostroboscopic findings of bilateral normal vocal fold mobility and vibratory asymmetry. The diagnostic LEMG examinations included an evaluation of the motor unit morphology and recruitment of motor unit potentials (MUPs) for the thyroarytenoid and cricothyroid muscles. Interpretation of the LEMG findings was done by a neurologist (C.E.J.) who was blinded to the findings of the laryngoscopic examination. In all cases, abnormal LEMG findings were considered to be present when there were large-amplitude polyphasic MUPs and incomplete recruitment of MUPs. All abnormal LEMG findings were then classified as left, right, or bilateral, depending on the side of involvement. We did not distinguish between recurrent laryngeal nerve (RLN) and superior laryngeal nerve (SLN) neuropathy for the purposes of this portion of the study. In other words, if the RLN, SLN, or both showed electrophysiological evidence of denervation, the findings were considered "abnormal" for that side.

Our endoscopic clinical examination protocol was as follows. All of the patients underwent videostroboscopy by means of a flexible laryngoscope with a distal chip (Olympus ENF-VQ, Olympus Surgical, Orangeburg, New York) rhinolaryngoscope, and most also had rigid laryngoscopy with a 70° rigid endoscope (KayPENTAX, Lincoln Park, New Jersey). The patients were instructed to phonate /i/ at low, modal, and high frequencies. When indicated, the technique of "unloading" as described by Koufman[8] was also used to help reveal more subtle vibratory asymmetry that may have been hidden under compensatory muscle tension patterns.

When retrospective evaluation of the endoscopic segments was carried out, the following protocol was used. The best-quality videostroboscopic examination (either flexible or rigid) was used for each case. Of the 48 cases in which LEMG was performed for suspected paresis, 23 examinations that were considered to show isolated vibratory asymmetry were selected for the study. The other 25 cases, which showed vocal fold immobility, partial immobility, videostroboscopic evidence of incomplete closure, or vocal fold lesions, were excluded.

TABLE 1. VOCAL FOLD PARESIS DEMOGRAPHICS AND LEMG FINDINGS

Age (y)	*Gender*	*Duration*	*LEMG Findings*	*Cause of Paresis*
62	F	1 y	B RLN + SLN	Idiopathic
67	F	1 y	B RLN	Idiopathic
30	M	9 y	L RLN + SLN	Idiopathic
36	M	36 y	B RLN	Congenital
28	M	4 mo	B RLN	Idiopathic
65	M	6 y	B RLN + SLN	Idiopathic
36	F	10 y	B RLN	Idiopathic
69	F	2 mo	B RLN	Idiopathic
35	F	1 y	B RLN	Idiopathic
36	M	7 y	B RLN	Idiopathic
44	F	9 y	R RLN	Idiopathic
29	F	1.5 y	L RLN	Idiopathic
58	F	9 mo	L RLN	Idiopathic
37	F	1 y	B RLN	Idiopathic
51	F	5 y	L RLN	Idiopathic
43	F	16 mo	R RLN	Idiopathic
76	M	6 mo	B RLN	Idiopathic
58	M	14 mo	B RLN	Idiopathic
54	F	4 mo	L SLN	Traumatic

LEMG — laryngeal electromyography; B — bilateral; RLN — recurrent laryngeal nerve paresis; SLN — superior laryngeal nerve paresis; L — left; R — right.

The videos were edited to include only segments in which the vocal folds were in a fully adducted position and were engaged in vibratory activity. We decided not to show footage of vocal fold mobility, in order to help exclude any possible bias that could occur from interpreting vocal fold movement. The video segments were then randomized and were interpreted by three reviewers with extensive experience in videostroboscopic interpretation. Each video segment was reviewed, and the following questions were addressed: 1) Is asymmetry of vibration (amplitude or mucosal wave) present? 2) If vibration is asymmetric, which side has the increased amplitude and/or mucosal wave? and 3) On which side would you predict the paresis to be present?

The LEMG results were used as the gold standard for the diagnosis of VFP. Interpretation of the videostroboscopic findings by our reviewers was then compared to this gold standard to determine the predictive value of subjective vibratory asymmetry on videostroboscopic examination.

RESULTS

Of the 19 patients with a diagnosis of LEMG-confirmed VFP (Table 1), the mean patient age was 48.5 years (range, 28 to 76 years). Twelve of the patients were female (63.2%) and had a mean age of 48.8 years, and 7 patients were male (36.8%) and had a mean age of 47 years. The mean time interval from

TABLE 2. LEMG RESULTS AND REVIEWERS' INTERPRETATION

Patient	*Reviewer 1*	*Reviewer 2*	*Reviewer 3*	*LEMG*
1	L	R	R	B
2	R	R	R	B
3	R	R	R	L
4	R	B	R	Normal
5	R	B	B	B
6	R	R	R	Normal
7	R	R	R	B
8	L	L	L	B
9	L	R	B	R
10	B	B	B	B
11	R	B	B	B
12	L	L	L	B
13	R	B	L	R
14	L	B	B	Normal
15	R	B	R	L
16	L	L	R	L
17	R	B	B	B
18	R	L	L	Normal
19	L	B	B	L
20	R	B	R	R
21	R	L	R	B
22	R	B	B	B
23	R	B	L	L

L — left-sided paresis; R — right-sided paresis; B — bilateral paresis.

the onset of symptoms to presentation to our clinic was 4.8 years (range, 2 months to 36 years). The cause of the paresis was idiopathic in the vast majority of cases (17 of 19 or 89.5%), and the remaining cases were congenital (1 of 19 or 5.2%) or traumatic (1 of 19 or 5.2%). In terms of neural involvement, the majority of cases involved the RLN only. Ten cases were bilateral RLN paresis, and 5 cases were unilateral RLN paresis. The remaining cases were 2 cases of bilateral combined RLN and SLN paresis, 1 case of unilateral combined RLN and SLN paresis, and 1 case of unilateral SLN paresis.

Of the 23 patients with symptoms of glottal insufficiency and isolated vibratory asymmetry on videostroboscopy, 19 (82.6%) were found to have electrophysiological evidence of denervation of one or both vocal folds (Table 2). However, the individual reviewers' ability to correctly predict the distribution of the paresis was quite poor. Given three options (bilateral, left, or right), each reviewer was unable to correctly predict the side in most cases (reviewer 1, 5 of 19 correct; reviewer 2, 7 of 19 correct; and reviewer 3, 7 of 19 correct). With all examination evaluations combined, the side of paresis was correctly predicted in only 33.3% of cases (19 of 57).

DISCUSSION

The idea behind this study was to answer a common question that is posed in our multidisciplinary clinics. As a general rule, the voice team (which includes the senior author, speech pathologist, and resident physician) reviews the videostroboscopic examination of the patient and discusses the subjective interpretation of the vibratory parameters. In most cases of suspected VFP, the clinicians can agree that vibratory asymmetry is present, and LEMG will later confirm the diagnosis. However, the reliability of using vibratory asymmetry to correctly predict the presence of VFP has not been examined. Although we can usually agree on the presence of vibratory asymmetry, there is often a debate about the sidedness of the suspected paresis. Conventional thinking suggests that the denervated side will have an increased amplitude and/or mucosal wave due to the laxity of the paretic vocal fold. Despite this consensus, we have noted that many times the clinicians do not agree as to which side(s) is involved.

Obviously, the clinical diagnosis of some cases of VFP is fairly straightforward when based on videostroboscopic findings and clinical history. In the setting of gross hypomobility and glottal insufficiency, the diagnosis is not often in question. However, when there are no readily apparent differences in vocal fold mobility, the diagnosis can be more difficult to make, or may not be suspected by the clinician at all. In these cases, vibratory asymmetry may be the only clue that VFP is present.[7] This finding may help guide the clinician toward performing LEMG and establishing a correct diagnosis.

Our clinical protocol for patients with symptoms suggestive of glottal insufficiency and an increased amplitude and/or mucosal wave or "chasing wave" (asymmetry of vibration) is to recommend LEMG. Obviously, not all patients with this combination of symptoms and findings agree to undergo or follow up for diagnostic LEMG, so we are not able to comment on the positive predictive value of vibratory asymmetry in these cases. Nonetheless, when vibratory asymmetry prompted LEMG testing in our series, the clinical "hunch" ended up being correct in 83% of cases. However, the ability of experienced clinicians to correctly predict which side was involved was quite poor (33.3%). This is exactly the percentage one would expect if the clinician's determination were randomly generated; ie, there is a 1-in-3 chance of predicting the outcome correctly. The difficulty partially arises from using the subjective observation that one side demonstrates increased vibratory amplitude (often thought to be a manifestation of reduced muscular tone in a denervated vocal fold). By necessity, that determination involves using the contralateral side as a control, ie, the side with the "normal tone." In many cas-

es, however, this side may also be affected, making the assumption of unilaterality erroneous. Despite this problem, there were many cases in which the reviewer correctly predicted that the paresis was unilateral, but the predicted side (ie, distribution of involvement) was incorrect.

Relying solely on laryngoscopic findings to predict VFP continues to be problematic. Other studies have shown that 25% to 40% of patients had LEMG findings that were not predicted by their laryngoscopic examination.[2,3] Although vibratory asymmetry is fairly predictive of VFP (83% of cases in our study), determining the distribution (left/right/bilateral) of the paresis is very poorly predictive.

Interpretation of videostroboscopic examinations is by nature subjective. We have observed that vibratory asymmetry can sometimes be difficult to detect on routine stroboscopy. The best method of accentuating asymmetry is to have the patient phonate at a modal or low fundamental frequency at a high intensity. In addition, extinguishing any secondary supraglottic muscular tension seems to be beneficial, as this allows for the differential tension of the true vocal folds to be observed. Last, recording the examination and playing it back in slow motion, or performing frame-by-frame analysis, is yet another method to aid in the detection of vibratory asymmetry.

CONCLUSIONS

The videostroboscopic finding of vibratory asymmetry in mobile vocal folds is a reliable predictor of VFP in most cases. However, the ability of expert reviewers to determine the distribution (left/right/bilateral) of the paresis using videostroboscopic findings is poor. This finding highlights the value of LEMG in arriving at a correct diagnosis in this clinical situation.

REFERENCES

1. Merati AL, Shemirami N, Smith TL, Toohill RJ. Changing trends in the nature of vocal fold motion impairment. Am J Otolaryngol 2006;27:106-8.

2. Heman-Ackah YD, Barr A. Mild vocal fold paresis: understanding clinical presentation and electromyographic findings. J Voice 2006;20:269-81.

3. Koufman JA, Postma GN, Cummins MM, Blalock DP. Vocal fold paresis. Otolaryngol Head Neck Surg 2000;122:537-41.

4. Simpson CB, Cheung EJ, Jackson CJ. Vocal fold paresis: clinical and electrophysiologic features in a tertiary laryngology practice. J Voice 2009;23:396-8.

5. Altman KW. Laryngeal asymmetry on indirect laryngoscopy in a symptomatic patient should be evaluated with electromyography. Arch Otolaryngol Head Neck Surg 2005;131:356-60.

6. Rubin AD, Praneetvatakul V, Heman-Ackah Y, Moyer CA, Mandel S, Sataloff RT. Repetitive phonatory tasks for identifying vocal fold paresis. J Voice 2005;19:679-86.

7. Sulica L, Blitzer A. Vocal fold paresis: evidence and controversies. Curr Opin Otolaryngol Head Neck Surgery 2007;15:159-62.

8. Koufman JA. Evaluation of laryngeal biomechanics by fiberoptic laryngoscopy. In: Rubin JA, Sataloff RT, Korovin GS, Gould WJ, eds. Diagnosis and treatment of voice disorders. New York, NY: Igaku-Shoin, 1995:122-34.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2014; 123(4):271-278.

Article

Annals of Otology, Rhinology & Laryngology
2014, Vol. 123(4) 271–278
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489414525022
aor.sagepub.com
SAGE

Reliability of Clinical Office-Based Laryngeal Electromyography in Vocally Healthy Adults

Daniel J. Croake, MM, MS[1], Joseph C. Stemple, PhD[1], Timothy Uhl, PhD[1], Sanford Archer, MD, FACS, FAAOA[2], and Richard D. Andreatta, PhD[1]

Abstract

Objective: This study aimed to conduct a 3-session reliability assessment of the laryngeal electromyography (LEMG) signal in healthy participants during intensity controlled vocalization tasks. We hypothesized that vocal intensity level and testing session would affect LEMG measures.

Methods: This prospective study used a 2-factor repeated measures design. Seven participants underwent bipolar needle LEMG of the right thyroarytenoid muscle. Data were collected over 3 testing sessions using vocalization tasks performed with visually guided intensity feedback targets (65 and 75 dB SPL). Root mean square amplitudes in microvolts were analyzed for within-session and between-session reliability.

Results: The main effect for intensity was found to approach significance (F = 5.71, P = .054). However, intraclass correlation coefficients (ICCs) using a 2-factor mixed random effect model indicated poor to fair signal reliability between testing sessions (ICC = 0.56 at 65 dB, 0.40 at 70 dB). Intraclass correlation coefficients for within-session data indicated excellent reliability for all testing conditions (0.84–0.98).

Conclusion: Using a quantitative analysis protocol to inform an essentially qualitative technique, our results indicated that there was generally poor to fair reliability in the LEMG signal over testing sessions. Vocal intensity was an important variable that affected LEMG signal reliability. Standardization of LEMG protocols using vocal control parameters and quantitative analyses may help improve LEMG reliability in clinical settings.

Keywords

LEMG, motor units, neuromuscular, thyroarytenoid

Introduction

Laryngeal electromyography (LEMG) is commonly used for the assessment of neuromuscular disorders of the larynx.[1] Laryngeal electromyography is the only direct measure of laryngeal muscle activity, and although it provides general information about the function of the laryngeal musculature, its in-office clinical utility beyond general appreciation of gross neuromuscular function is debatable.[2,3] Studies regarding clinical usefulness of LEMG have not addressed the question of LEMG reliability both within a patient and across clinical testing sessions. Nonclinical and experimental investigations using LEMG have the advantage of signal processing software, control of environmental conditions to reduce electromagnetic field noise, tasks that provide graded control of laryngeal muscle recruitment, and less constrained time frames under which to perform LEMG. Because many variables may affect LEMG reliability, it remains unknown whether clinical office-based LEMG, often performed qualitatively, without vocalization control parameters and under less than optimal recording conditions, can be considered a reliable and clinically meaningful diagnostic tool.

Laryngeal electromyography has been used for research and assessment of the function of the intrinsic laryngeal muscles for more than 60 years and has been shown to be useful in revealing the function of the laryngeal musculature and demonstrating the dynamic control of this musculature during voicing.[4-7] Laryngeal electromyography is frequently used in conjunction with stroboscopic/

[1]College of Health Sciences, University of Kentucky, Lexington, Kentucky, USA
[2]College of Medicine, University of Kentucky, Lexington, Kentucky, USA

Corresponding Author:
Daniel J. Croake, MM, MS, Department of Rehabilitation Sciences, College of Health Sciences, University of Kentucky, Room 120H Wethington Building, 900 South Limestone Street, Lexington, KY 40536-0200, USA.
Email: djcroa2@uky.edu

laryngoscopic evaluation of vocal fold mobility to assess a variety of laryngeal disorders such as dystonia and vocal fold paresis and paralysis, to differentiate among various neurological disorders,[5] and for guiding the placement of botulinum toxin for the treatment of spasmodic dysphonia.[1,8] Laryngeal electromyography may also be potentially useful as a tool for the prognosis of laryngeal nerve disorders.[9,10]

In general, electromyography (EMG) recordings are affected by multiple confounding variables including electrode type and placement, level of muscle activation, left and right side dominance, artifact from electrode movement, and so on, which may all, or in part, compromise the accuracy of the data necessary for diagnostic evaluation.[11,12] Electromyographic investigations of between-session and intra-session reliability for some limb muscles have revealed high reliability for both between- and within-session measurements.[13,14] However, similar data for laryngeal-based EMG are absent and cannot be directly interpolated from limb studies due to significant differences in anatomical structure and the ability to control for muscle length and loading.

Unlike most limb muscles that have skeletal support and firm attachment points, the larynx is suspended in the neck, surrounded by soft muscle tissue, a series of membranes, and a somewhat yielding cartilaginous framework. Distinct muscle force and leverage points are difficult to determine in the laryngeal complex because of these flexible attachment points. Placing a consistent and measureable isotonic load on laryngeal muscles for accurate and reliable activation is difficult, complicating replication of motor unit activation in these muscles.

Another factor to be considered regarding the reliability of clinical in-office LEMG is that phonation is an emergent behavior, arising through the complex interaction of respiratory, phonatory, and resonance subsystems of the vocal tract. These vocal subsystems function synergistically, integrating properties of tissue elasticity, muscle activation, and aerodynamics toward normal vocal function. A change in any subsystem's dimension will potentially alter vocal output. These additional confounding variables have the potential to further complicate in-office LEMG interpretation.[15]

In general, reliable LEMG measurements are dependent on consistent muscle activation tasks. These tasks must be carefully controlled and performed for measurement reliability. For example, to describe relative recruitment of motor unit potentials for the thyroarytenoid muscle (TA), maximum voluntary contraction (MVC) strategies have been used for comparison. Typically, a maximal voluntary contraction is assigned a 100% possible recruitment value whereby subsequent muscle contractions during voicing tasks are given a percentage of decreased recruitment. Maximum voluntary contraction in the laryngeal system is typically accomplished through performance of a Valsalva maneuver (hard breath hold).[8] However, it has been shown that vocal fold closure is not consistently accomplished during Valsalva maneuvers up to 14% of the time, potentially leading to significant diagnostic error.[16] Other qualitative ratings such as decreased recruitment scales are not comparable across offices due to their highly subjective nature and lack of standardized between-office collection protocols.[17]

Another commonly used clinical LEMG technique is comparison of recruitment against the contralateral muscle. Unfortunately, this technique does not take into account the notion that the contralateral TA muscle is dependent on the co-contraction of neighboring intrinsic muscles. Thus, TA contraction may be altered in the presence of a contralateral paresis or paralysis. In this scenario, compensatory muscle activation is a likely confounder.[18] In addition, a large-scale retrospective study reported unexpected contralateral neuropathy in 26% of patients with laryngeal movement disorders.[9] Electromyography studies of limb muscle also indicate significant contralateral differences in motor unit recruitment even during simultaneously controlled muscle contractions.[12]

Because raw EMG signals are quasi-random in nature, they cannot be directly compared. Thus, a principle goal of this study was to use quantitative methodology with the addition of control parameters and measures, to characterize the reliability of a primarily qualitative clinical evaluation. One such measure was quantification of the LEMG signal via calculation of the root mean square (RMS). Root mean square is considered to be the current "gold standard" for quantitative electromyographic analysis[11,12] and allows for rapid quantitative comparisons among groups of signals. Root mean square was chosen as a measurement metric because it provides an indication of mean muscle activity and signal power and is the analog to voltage output. Because RMS is also considered a data smoothing technique, it is not well suited for visualization of waveform transients and morphology characteristics such as polyphasic or nascent potentials; however, it is useful to quantify and compare LEMG across samples in terms of signal voltage and power. Because EMG is a time-varying signal containing positive and negative values, RMS is an ideal quantitative measure that can be easily calculated post hoc or in real time with many commercially available data acquisition software programs.

Determining LEMG data reliability within the context of an in-office clinical environment is important to make careful and useful clinical interpretations and to potentially improve clinical protocols. To our knowledge, in-office clinical LEMG reliability has not been systematically investigated in a cohort of vocally healthy adults. As such, the purpose of this study was to prospectively investigate LEMG signal reliability recorded from the thyroarytenoid muscle over multiple testing sessions using a common in-office clinical routine. We modeled our basic methodology

after the University of Iowa head and neck protocol for LEMG diagnostics.[21] Laryngeal electromyography was performed on the right thyroarytenoid muscle of 7 participants with normal vocal function on 3 separate occasions with the application of an additional behavioral control for vocal intensity. We hypothesized that time of testing and vocal intensity would significantly alter quantitative measures of thyroarytenoid EMG signals.

Materials and Methods

Seven participants, ages 18 to 40 years (3 male and 4 female), were recruited and volunteered for this study. Ten participants were initially recruited based on power analysis for the repeated measures design. Two participants did not complete the entire study protocol, and 1 participant had unusable data. All were nonsmokers, English speaking, and free of laryngeal pathology and movement disorders as judged by stroboscopic examination of the larynx. Exclusion criteria were professional voice training, diagnosed bleeding disorder, prior head/neck/spinal surgery, and/or intubation within the past 5 years. All procedures were approved by the University of Kentucky Institutional Review Board, and informed consent was obtained from all participants. All assessments were performed at the University of Kentucky Laryngeal & Speech Dynamics Laboratory.

Prior to data collection, all participants received a laryngeal exam via videostroboscopy to ensure normal vocal function and structure (Kay Elemetrics Rhino-Laryngeal Stroboscope [Model RLS 9100 B], Kay Elemetrics 70 degree rigid endoscope [Model SN 1541]; PENTAX Medical, Montvale, New Jersey, USA). This examination was performed by a certified/licensed speech-language pathologist and reviewed by a board certified otolaryngologist.

After imaging was completed, electromyographic biopotentials (μV) were acquired from the right TA muscle via a 25-mm, 30-gauge concentric bipolar needle electrode (XLTEK 101468; Natus Medical, San Carlos, California, USA). The raw EMG signal was routed to a biopotential amplifier (Grass Model 15A54; Natus Neurology, Warwick, Rhode Island, USA) and serially coupled to a 16-bit analog-to-digital converter sampling at 10 kHz (PowerLab 1630; ADInstruments, Inc, Colorado Springs, Colorado, USA). Prior to digitization, analog signals were band-pass filtered (30 Hz–3 kHz @ –3 dB), with an in-line notch filter applied to reduce 60 Hz line contamination. The LEMG analog output was paralleled to a stereo amplifier and played over loud speakers to provide the otolaryngologist with auditory feedback of muscle activity during needle insertion. Audio (volts) and sound intensity levels (dB SPL) from the participant's vocalizations were recorded by a lapel-style microphone (Sony ECM44B; Sony Corporation, New York, New York, USA) and a commercially available sound level meter (REED ST-8850; REED Instruments, Sainte-Anne-De-Bellvue, Quebec, Canada). Both signals were also digitized by the A/D system (audio sampling rate = 2 kHz; sound pressure level = 1 kHz). All signals were recorded in calibrated units using a proprietary 2-point interpolation method found in our digitization software package (LabChart 7; ADInstruments, Inc). Post-acquisition LEMG signal processing was completed using custom-coded LabChart routines.

Laryngeal electromyography testing was performed in a custom-built Faraday booth to reduce electromagnetic field effects. Participants were seated in an examination chair, reclined to approximately 60 degrees, with their heads comfortably supported by a neck pillow. No sedation or anesthetic was used during the needle insertion and recording procedure. A ground electrode was placed on the participant's neck, below the mastoid process. Thyroarytenoid needle insertion by an otolaryngologist with 10 years of experience performing clinical LEMG procedures was accomplished using a para-medial approach percutaneously with the electrode directed in a superolateral direction through the cricothyroid ligament with the muscle entered submucosally. Needle electrode placement was confirmed using the following behavioral tasks: normal rest breathing, phonation on sustained /i/, sniff, and sustained phonation of falsetto /i/. Upon completion of the study, the needle electrode was removed by the physician and participants were monitored for 15 minutes after the study in case of complications.

Digitized raw LEMG signals were full-wave rectified and RMS signal amplitude values were calculated online with the LabChart software package. Root mean square was calculated as the square root of the mean of a series of squared LEMG amplitude values. Root mean square amplitudes were used in all analyses to determine if significant variance in the LEMG signal existed as a function of vocal intensity and time of data acquisition.

Experimental Protocol

Electromyography of the right TA was performed on each participant on 3 different occasions with a minimum of 1 month between adjacent procedures (mean duration = 2.5 months) to allow for tissue healing. Each session took place at approximately the same time of day and lasted no more than 30 minutes (including videostroboscopy). After needle electrode placement was confirmed within the TA, LEMG signals were recorded under 2 task conditions. First, a confirmation condition was completed by recording baseline LEMG signals while the participant was instructed to relax and breathe normally, sustain the vowel /i/ at a comfortable modal pitch, gently sniff through the nose, and sustain a falsetto /i/. Second, a feedback condition was completed with the participant producing a sustained /i/ at 2 different

intensity targets (65 and 75 dB [± 3 dB]). These intensity targets were chosen to represent a typical healthy vocalization intensity range present in everyday speech. In this condition, participants received visual feedback of their intensity level by monitoring a dB sound level meter placed 16 inches from the individual's mouth. The participant was asked to hold the intensity constant for a minimum of 3 seconds. Trials less than 3 seconds were not accepted for analysis. A total of 10 trials of each task condition were recorded. The initial and final 2 trials were discarded, leaving 6 trials for post-hoc data analysis.

Habitual fundamental frequency (F_o) for the vocalization tasks was initiated by the participant without prompting from the investigators. The F_o chosen by the participant was recorded and played back via an auto-tuner before data collection so that the participant could remain in an acceptable modal pitch range. Before all data collection procedures and to ensure performance consistency, each participant was trained and given time to practice all tasks before data collection.

Data Analysis

A 2-factor repeated measures analysis of variance (RM-ANOVA) was used to compare the effect of the 2 independent variables among participants. The independent variables included (1) LEMG data recording sessions (3 sessions) and (2) the task-related vocal intensity levels (2 dB levels). To evaluate the reliability of LEMG signals across testing times, intraclass correlation coefficients (ICCs) using a 2-way mixed random effects model were calculated. For between-session measurements, standard error of the measurement (SEM) was calculated to determine the minimum detectable change (MDC) in microvolts with 70% and 95% confidence boundaries. Intraclass correlations were also calculated to measure intra-session reliability. All statistical calculations were performed using SPSS version 18 (IBM, Armonk, New York, USA).

Data Selection and Calculations

A 1-second window from the mid-portion of each 3-second task recording window was selected for detailed analysis. Root mean square amplitude values were automatically calculated for each of the 1-second windows using a custom sub-routine in LabChart (ADInstruments, Inc). To be included for further processing and analysis, digitized samples had to fall within +/– 1 dB of the desired intensity levels (65 or 75 dB). In certain cases, the RMS amplitude was not stable for 1 second at +/– 1 dB. In these cases, a shorter window was averaged to eliminate portions of the signal with poor signal quality. The inclusion of smaller analysis windows for these cases was preferable to averaging poor signal quality. Smaller sampling windows do not significantly affect analyses, as comparisons of EMG sample windows using RMS have been shown to produce moderate to excellent ICC reliability data from 1000 ms down to 100 ms window sizes.[13]

Table 1. Means for Intensity Level at 65 dB.

Variable	Session	Mean	SD	Min	Max
Root mean square (μV)	1	80.2	34.8	44.7	151.5
	2	91.8	28.4	45.4	117.6
	3	73.08	19.64	46.31	98.48

Table 2. Means for Intensity Level at 75 dB.

Variable	Session	Mean	SD	Min	Max
Root mean square (μV)	1	84.0	37.4	46.0	155.5
	2	106.5	29.4	69.2	136.9
	3	85.04	16.82	55.42	101.56

Table 3. Repeated Measures Analysis of Variance for Baseline Laryngeal Electromyography Across Time.

Source	*df*	*F* Value	*P* Value
Participant	6	0.85	.555
Time	2	0.02	.978

Table 4. Repeated Measures Analysis of Variance for Root Mean Square Values.

Source	*df*	*F* Value	*P* Value
Participant	6	1.91	.160
Time	2	1.38	.289
Intensity	1	5.71	.054
Time*Intensity	2	2.23	.150

Results

Statistical Analysis

Group data, including means and standard deviations for the dependent variable (RMS) for intensities at 65 dB and 75 dB, are presented in Tables 1 and 2, respectively. Results of the 2-factor RM-ANOVA during baseline and the task conditions are presented in Tables 3 and 4. Repeated measures ANOVA for the baseline data did not reveal a significant main or interaction effect for testing time (session) versus RMS, indicating stable LEMG baseline activity across recording sessions. Results of the RM-ANOVA for the feedback condition were not significant for the main effect of testing time ($df = 2$, $F = 1.38$, $P = .289$). However, the main effect for intensity level closely approached our a

Table 5. Intraclass Correlation Coefficients (ICCs) Between Sessions, Standard Error of Measurement (SEM), and Minimum Detectable Change (MDC) in μV Necessary to Detect True Change Between Measurements.

ICC Sessions 1–3		SEM Sessions 1–3	MDC in μV @ 70% Confidence	MDC in μV @ 95% Confidence
65 dB	0.56	18.57	26.26	51.47
75 dB	0.40	23.0	32.5	63.7

Table 6. Intraclass Correlation Coefficients (ICCs) for Within-Session Data.

Within-Session ICC	Session 1	Session 2	Session 3
65 dB	0.95	0.93	0.84
75 dB	0.88	0.84	0.98

priori significance level of α = .05 (df = 1, F = 5.71, P = .054). This indicated that vocal intensity may play a role in LEMG signal reliability. No interaction effects between intensity and time factors were found.

To further evaluate the reliability of LEMG signals across the 3 testing days, ICCs using a 2-factor mixed random effect model were calculated. Average measures from 6 sample trials per participant were compared across the 3 testing sessions to evaluate reliability. The ICCs across Sessions 1 through 3 were 0.56 for the 65 dB condition and 0.40 for the 75 dB condition. These between-session ICCs were low, suggesting an interpretation of poor to, at best, fair reliability across testing sessions. Although not standardized, typical calculated values for ICC interpretation are as follows: less than 0.40 = poor reliability, 0.40 to 0.75 = fair to good reliability, and greater than 0.75 = excellent reliability.[19] In addition, the SEM was calculated. The SEM was then used to determine the MDC in microvolts necessary to demonstrate a true difference if the muscle was tested on multiple days with 95% confidence boundaries. Results indicated that a change of 51μV would be necessary to determine a true difference in LEMG activity between testing sessions. Numerical results of ICC, SEM, and MDC are presented in Table 5. Intraclass correlations for within-session data revealed strong reliability among participants ranging from 0.84 to 0.95 and from 0.88 to 0.98 for the 65 dB and 75 dB conditions, respectively. Comparisons of within-session data are located in Table 6.

Discussion

The use of clinical in-office LEMG has been incorporated into the diagnostic routine for the evaluation and treatment of voice disorders in many practices across the country.[8] Although some evidence supports LEMG use in the diagnosis and prognosis of certain neuromuscular disorders, the general reliability of the LEMG signal in normal participants has not been carefully evaluated. This study measured the reliability of the LEMG signal in normal, vocally healthy participants over time with the central aim of determining if significant LEMG signal variance occurred as a function of multiple testing sessions. Our second aim was to determine if vocalization intensity affected the LEMG signal. In limb studies, control over the degree of muscle contraction is necessary to achieve results that are comparable within and across participants. Both maximal and submaximal contractions have been shown to demonstrate strong reliability in limb muscle.[20] We used vocal intensity as a method to control laryngeal muscle contraction levels among participants. Our results indicated that between-session LEMG reliability was poor to fair and that control of vocal intensity may be an important performance variable to help improve the reliability of these measurements.

This study mirrored the University of Iowa head and neck protocol for LEMG diagnostics.[21] In addition to this basic protocol, we used vocal intensity control and a Faraday booth to reduce ambient electrical noise to improve the fidelity of the data and provide the optimal set of circumstances under which to perform our LEMG clinical evaluation. The intent of this study was not to quantify LEMG precisely but rather to use quantitative means to measure LEMG in an ideal environment to test the hypothesis that clinical LEMG data are variable across testing sessions even with added control parameters in place.

Repeated measures analysis of variance indicated a nonsignificant effect for time of testing, suggesting that LEMG signals for pooled data did not vary significantly across testing sessions. Intraclass correlation coefficient analysis for within-session reliability was considered excellent for both intensity conditions ranging from 0.84 to 0.95 and from 0.88 to 0.89 for the 65 dB and 75 dB conditions, respectively. However, the between-session ICC revealed poor to fair reliability for both intensity conditions. It should be noted that the most qualitatively consistent data from our study were collected when the participants vocalized at 65 dB, indicating a less reliable measure at greater loudness levels. Data from Sessions 1 and 2 at 65 dB represented the strongest reliability association with an R^2 value of 0.048. This indicated poor reliability even across the most consistent recording sessions (see Figures 1 and 2).

Reliability debates concerning the clinical usefulness of LEMG for diagnostic and prognostic applications have been raised.[1] According to a recent evidence-based review and clinical recommendations, LEMG data have been considered questionable for clinical uses such as diagnosing paresis/paralysis from joint fixation, for accuracy diagnosing diseases of the neuromuscular junction, and for providing accurate diagnostic information of neuropathic and myopathic disorders.[1,8] To address these questions, it has

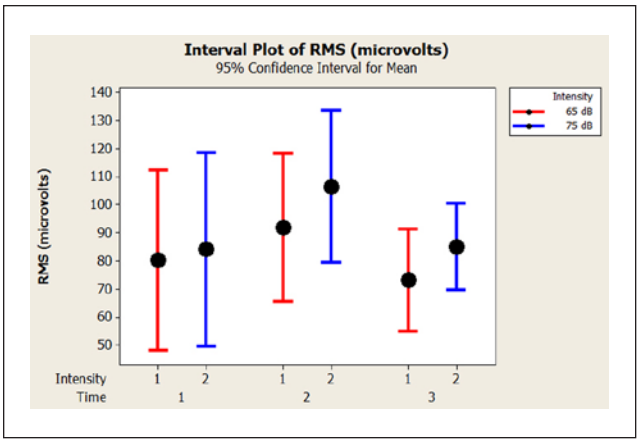

Figure 1. Interval plot with the mean of each data set presented with 95% confidence bars. The x-axis is scaled with regard to intensity and testing session. The y-axis for root mean square (RMS) is scaled in microvolts.

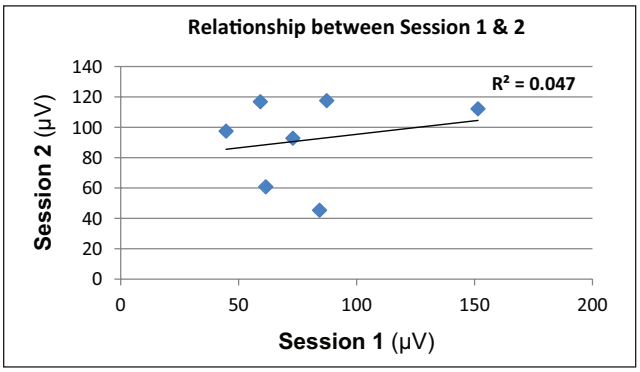

Figure 2. Scatter plot for voice task at the 65 dB target between sessions 1 and 2. These sessions represent the most consistent data from this study. Data points are mean root mean square (RMS) values in μV for each participant. The R^2 value does not indicate a strong association between the data for the 2 sessions. Approximately 5% of the data from Session 1 can be explained by Session 2.

been suggested that additional evidence-based research concerning LEMG methodology and validity be conducted.[8] Data from this study suggest that variables such as data collection time (multiple sessions) and possibly vocal intensity may play a role in the outcomes of LEMG assessment, suggesting methodological limitations of LEMG in terms of its clinical accuracy.

Our results indicated that to be 95% confident that a true detectable change could be observed between testing sessions, a change of 51 μV RMS was necessary with intensity level held constant. It is likely, then, that uncontrolled vocal intensity during LEMG procedures may operate as a confounding variable. Careful regulation of vocal intensity during LEMG may be necessary if the clinical utility of LEMG is to be determined with any degree of accuracy. In fact, intensity would be expected to contribute to changes in the RMS value of the LEMG signal. A near linear relationship between muscle force and EMG activity has been found in classic EMG studies.[22-24] Although EMG does not measure muscle force directly, vocal intensity can be viewed as a global indicator of performance effort and muscle loading on the vocal apparatus. It was not surprising, then, that controlling for intensity revealed changes in our calculated RMS values.

It has been previously demonstrated that both intensity and vocal frequency contribute to variability in quantitative LEMG output with frequency being the greater factor in TA recruitment variability, suggesting the need for control of both parameters for improved clinical assessment.[25,26] In this study, participants were generally able to maintain and regulate the intensity of their vocalization constant at 65 dB across all trials. Although participants were trained to reach the 75 dB target, many could not produce this intensity level consistently for 1 second with the LEMG needle in place. In the 75 dB condition, intensity levels actually ranged from 66.70 dB to 75.22 dB with a mean value of 70.42 dB. Thus, it can be inferred that not only does intensity play a role in LEMG signal stability but relatively small changes in intensity level (approximately 5 dB) can strongly affect RMS values, further arguing for the need and importance of regulating vocal intensity during LEMG diagnostics.

Limitations

The small sample size of 7 participants in this repeated measures study limits the ability to generalize our results to a larger clinical population. Changes in vocal intensity were limited to a 10 dB interval. Larger intensity intervals and additional participant data may better demonstrate differences in mean RMS values across testing conditions. The standard deviations of the RMS values in this study were large. This is an inherent problem with attempting to quantify LEMG because it is difficult to determine which variable(s), such as ambient noise, movement artifact, interpersonal differences in phonation, and so on, may be causing deviations in the signal.[11,27] Needle electrodes, as used in this study, have been shown to demonstrate greater artifact at greater intensities.[28] Movement/vibration artifact cannot be alleviated but is a concern because of the unsteadiness of the needle electrode and the vibration of the vocal fold mucosa. A solution to this problem may be to consider the use of hooked wire electrodes in clinical LEMG studies to ameliorate these concerns.

Clinical Relevance and Future Directions

The results of this study demonstrate that even during controlled laboratory conditions, the LEMG signal appears

significantly variable across testing sessions. It is reasonable to expect greater variability in an office setting without these control parameters. Laryngeal electromyography holds much potential to be a useful clinical tool available for diagnosing movement disorders of the larynx. To obtain the maximum benefit from clinical LEMG, a universal standardized protocol that is feasible within a typical in-office setting should be developed. Further prospective research studies should consider the evaluation of (1) hooked wire electrode use for contralateral comparisons and (2) vocal frequency and intensity controls to regulate muscle recruitment to maximize the interpretability of LEMG measures. Quantitative LEMG techniques that may be used for clinical application are now feasible and relatively inexpensive and should be explored. Laryngeal electromyography follow-up diagnostics would also be useful for data comparisons. A recent article by Sataloff et al stated that follow-up LEMG is not performed in up to 90% of cases if visual inspection of the larynx demonstrates improved vocal fold mobility.[29] Follow-up testing could provide useful reliability data for LEMG as performed in the clinic.

Data from this study offer insight into the importance of using loading controls (control of intensity and frequency) in order to obtain the most accurate data from clinical LEMG. Methodologically, the use of hooked wire LEMG may be a good alternative to needle electrodes for several reasons, including freeing the clinician to direct the patient to control pitch and loudness levels, allowing for simultaneous measures, and reducing the possibility of artifact from needle electrode movement. With today's technology, control of vocal frequency and intensity can be easily accomplished through visual feedback to the patient using an inexpensive headset microphone connected to a laptop computer or other mobile device running commercially available sound intensity applications. In addition, quantitative signal processing tools are becoming more accessible and less expensive, allowing for the real-time use of quantitative techniques such as RMS to improve the quality of in-office assessments and the use of rise-time functions to confirm optimal electrode placement.[30]

In closing, LEMG is a clinical assessment tool that has not yet reached its full potential. The means to make LEMG a more quantitative and reliable assessment method are available and ready for usage to improve the clinical reliability and usefulness of this potentially important diagnostic method.

Acknowledgments

The authors would like to thank the participants for undergoing multiple LEMG procedures for this study, and the University of Kentucky College of Health Sciences Office of Research for supporting this project through pilot funding.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The University of Kentucky College of Health Sciences Office of Research supported this project through pilot funding.

References

1. Sataloff RT, Mandel S, Mann EA, Ludlow CL. Practice parameter: laryngeal electromyography (an evidence-based review). *J Voice*. 2004;18(2):261-274.
2. Ludlow CL, Yeh J, Cohen LG, Van Pelt F, Rhew K, Hallett M. Limitations of electromyography and magnetic stimulation for assessing laryngeal muscle control. *Ann Otol Rhinol Laryngol*. 1994;103(1):16-27.
3. Woo P. Laryngeal electromyography is a cost-effective clinically useful tool in the evaluation of vocal fold function. *Arch Otolaryngol Head Neck Surg*. 1998;124(4):472-475.
4. Heman-Ackah YD, Mandel S, Manon-Espaillat R, Abaza MM, Sataloff RT. Laryngeal electromyography. *Otolaryngol Clin North Am*. 2007;40(5):1003-1023.
5. Sataloff R, Mandel S, Heman-Ackah Y, Manon-Espaillat R, Abaza M. *Laryngeal Electromyography*. 2nd ed. San Diego, CA: Plural Publishing; 2006.
6. Basmajian J. *Muscles Alive: Their Functions Revealed by Electromyography*. 3rd ed. Baltimore, MD: Williams and Wilkins Company; 1974.
7. Hillel AD. The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *Laryngoscope*. 2001;111(4 Pt 2 Suppl 97):1-47.
8. Blitzer A, Crumley RL, Dailey SH, et al. Recommendations of the Neurolaryngology Study Group on laryngeal electromyography. *Otolaryngol Head Neck Surg*. 2009;140(6): 782-793.e786.
9. Koufman JA, Postma GN, Whang CS, et al. Diagnostic laryngeal electromyography: The Wake Forest experience 1995-1999. *Otolaryngol Head Neck Surg*. 2001;124(6): 603-606.
10. Sittel C, Stennert E, Thumfart WF, Dapunt U, Eckel HE. Prognostic value of laryngeal electromyography in vocal fold paralysis. *Arch Otolaryngol Head Neck Surg*. 2001;127(2):155-160.
11. Basmajian J, De Luca C. *Muscles Alive: Their Functions Revealed by Electromyography*. 5th ed. Baltimore, MD: Williams & Wilkins; 1985.
12. De Luca CJ. *Surface Electromyography: Detection and Recording*. Natick, MA: Delsys Inc.; 2002. http://www.delsys.com/Attachments_pdf/WP_SEMGintro.pdf. Accessed November 2008.
13. Bamman MM, Ingram SG, Caruso JF, Greenisen MC. Evaluation of surface electromyography during maximal

voluntary contraction. *Journal of Strength & Conditioning Research*. 1997;11(2):68-72.
14. Yang JF, Winter DA. Electromyography reliability in maximal and submaximal isometric contractions. *Arch Phys Med Rehabil*. 1983;64(9):417-420.
15. Huber JE, Stathopoulos ET. Respiratory and laryngeal responses to an oral air pressure bleed during speech. *J Speech Lang Hear Res*. 2003;46(5):1207-1220.
16. Donzelli J, Brady S. The effects of breath-holding on vocal fold adduction: implications for safe swallowing. *Arch Otolaryngol Head Neck Surg*. 2004;130(2):208-210.
17. Smith LJ, Rosen CA, Niyonkuru C, Munin MC. Quantitative electromyography improves prediction in vocal fold paralysis. *Laryngoscope*. 2012;122(4):854-859.
18. Rubin AD, Sataloff RT. Vocal fold paresis and paralysis. *Otolaryngol Clin North Am*. 2007;40(5):1109-1131.
19. Rosner B. *Fundamentals of Biostatistics*. Belmont, CA: Duxbury Press; 2005.
20. Burden A. How should we normalize electromyograms obtained from healthy participants? What we have learned from over 25 years of research. *J Electromyogr Kinesiol*. 2010;20(6):1023-1035.
21. University of Iowa Health Care. Iowa head and neck protocols: laryngeal electromyography. http://www.healthcare.uiowa.edu/otolaryngology/protocols//Parts/Part1/Part1C/P1C4.htm. Accessed July 14, 2008.
22. Abbott BC, Bigland B. The effects of force and speed changes on the rate of oxygen consumption during negative work. *J Physiol*. 1953;120(3):319-325.
23. Bigland B, Lippold OCJ. The relation between force, velocity and integrated electrical activity in human muscles. *J Physiol*. 1954;123(1):214-224.
24. Bigland B, Lippold OCJ. Motor unit activity in the voluntary contraction of human muscle. *J Physiol*. 1954;125(2):322-335.
25. Lindestad PÅ, Fritzell B, Persson A. Quantitative analysis of laryngeal EMG in normal subjects. *Acta Otolaryngol*. 1991;111(6):1146-1152.
26. Statham MM, Rosen CA, Nandedkar SD, Munin MC. Quantitative laryngeal electromyography: turns and amplitude analysis. *Laryngoscope*. 2010;120(10):2036-2041.
27. Lindestad PÅ, Persson A. Quantitative analysis of EMG interference pattern in patients with laryngeal paresis. *Acta Otolaryngol*. 1994;114(1):91-97.
28. Hirano M, Vennard W, Ohala J. Regulation of register, pitch and intensity of voice. An electromyographic investigation of intrinsic laryngeal muscles. *Folia Phoniatr*. 1970;22(1):1-20.
29. Sataloff RT, Praneetvatakul P, Heuer RJ, et al. Laryngeal electromyography: clinical application. *J Voice*. 2010;24(2):228-234.
30. Statham MM, Rosen CA, Nandedkar SD, Munin MC. Quantitative laryngeal electromyography: turns and amplitude analysis. *Laryngoscope*. 2010;120(10):2036-2041.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(6):1075-1081.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–Head and Neck Surgery
2015, Vol. 152(6) 1075–1081
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599815577104
http://otojournal.org
SAGE

Office-Based Photoangiolytic Laser Treatment of Reinke's Edema: Safety and Voice Outcomes

Ian J. Koszewski, MD[1], Matthew R. Hoffman, PhD[1], W. Greg Young, MD[1], Ying-Ta Lai, MD[1], and Seth H. Dailey, MD[1]

No sponsorships or competing interests have been disclosed for this article.

Abstract

Objective. To evaluate the safety, tolerability, and voice outcomes of office-based photoangiolytic laser treatment of Reinke's edema.

Study Design. Case series with chart review.

Setting. Academic medical center.

Subjects and Methods. We performed a retrospective analysis of patients undergoing office-based laser treatment of endoscopy-proven Reinke's edema. Safety and tolerability were evaluated by reviewing complications. Voice outcomes were analyzed by comparing pre- and postprocedural acoustic, aerodynamic, and Voice Handicap Index measurements. Complete data sets were not available for all subjects; sample size for each parameter is reported with the corresponding result.

Results. Nineteen patients met inclusion criteria. There were no minor or major complications. Five procedures were truncated due to patient intolerance. Phonatory frequency range increased (n = 12, $P = .003$), while percent jitter decreased (n = 12, $P = .004$). Phonation threshold pressure decreased after treatment (n = 4, $P = .049$). Voice Handicap Index also decreased (n = 14, $P < .001$).

Conclusion. This study represents the largest series of patients undergoing office-based photoangiolytic laser treatment specifically for Reinke's edema. Our data suggest that this is a safe and effective modality to treat dysphonia associated with Reinke's edema, although patient intolerance of the procedure may represent a barrier.

Keywords

Reinke's edema, photoangiolytic laser, office-based treatment, voice

Received October 24, 2014; revised January 14, 2015; accepted February 20, 2015.

Reinke's edema is a benign disease of the true vocal folds, characterized by edema and vascular congestion within Reinke's space, leading to diffuse polypoid degeneration. Patients are generally middle-aged women, and they have voice complaints of abnormally low speaking pitch and a rough voice quality.[1] Multiple risk factors have been identified, including tobacco use, vocal abuse, and laryngopharyngeal reflux.[2-5] Critical to preventing progression of the disease and managing the dysphonia are nonsurgical strategies, such as smoking cessation, voice therapy, and treatment of underlying laryngopharyngeal reflux.[1,6] When a nonsurgical approach is inadequate, surgery is employed. Traditional surgical options have focused on mucosal microflap elevation with removal of tissue,[7,8] microdebridement,[9] carbon dioxide laser ablation,[10,11] and cold knife cytoreduction.[12]

Recently, photoangiolytic laser treatment has been proposed as an alternative treatment.[13-16] Unlike traditional methods that involve the physical removal of tissue, the potassium titanyl phosphate (KTP) laser and pulsed dye laser (PDL) target oxyhemoglobin and are thought to address the vascular congestion characteristic of Reinke's edema. Importantly, both the KTP and the PDL have flexible fibers that can be passed through a flexible endoscope for office-based interventions, thus avoiding risks associated with general anesthesia.[17] Performing procedures in the office rather than the operating room has several notable advantages, including decreased cost[18] and avoidance of the potential complications of microlaryngoscopy, such as dental injury and dysgeusia.[19]

Office-based use of lasers for the treatment of Reinke's edema has gained popularity with the advent of improved instrumentation and evidence that the procedures are safe and cost-effective. Koufman et al and Sheu et al reported large series demonstrating support for the use of photoangiolytic lasers in the treatment of an array of laryngeal

[1]Department of Surgery, Division of Otolaryngology–Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA

This article was presented at the 2014 AAO-HNSF Annual Meeting & OTO EXPO; September 21-24, 2014; Orlando, Florida.

Corresponding Author:
Seth H. Dailey, MD, Division of Otolaryngology–Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, 600 Highland Ave, Clinical Science Center-H4, Madison, WI 53792-3284, USA.
Email: dailey@surgery.wisc.edu

lesions, with the latter demonstrating preserved or improved mucosal wave and glottic closure after treatment.[14,15] Pitman et al evaluated the effect of office-based KTP laser treatment in 7 patients with Reinke's edema and found significant improvement in subjective measures for voice quality as well as a trend toward improvement in objective measures.[13]

Although these studies provided important initial support for the use of office-based photoangiolytic laser therapy in the management of Reinke's edema, they are limited by a modest sample size and measurement of few objective voice parameters. As such, larger studies addressing the safety, tolerability, and voice outcomes for this approach are necessary. We report on 19 patients who underwent office-based photoangiolytic laser treatment of Reinke's edema. We hypothesized that no complications would occur, patients would tolerate the procedure, Voice Handicap Index would decrease, and objective voice parameters would move toward the normal ranges. Specifically, we hypothesized that posttreatment assessment would show increased fundamental frequency range, decreased jitter, decreased phonation threshold pressure, and increased maximum phonation time (MPT).

Materials and Methods

Subjects

Approval for this study was obtained from the University of Wisconsin Health Sciences Institutional Review Board. The study was designed as a retrospective case series of patients treated at the University of Wisconsin–Madison. Patient data were obtained from the University of Wisconsin–Madison Voice and Swallow Outcomes Database. Appropriate patients were identified by the University of Wisconsin–Madison Voice and Swallow Outcomes database manager by cross-matching diagnosis with the existence of a procedure, and data extraction was performed by a separate researcher. All patients presented to the University of Wisconsin Hospital and Clinics between January 2007 and November 2013 and underwent voice analysis by a speech-language pathologist, as well as evaluation and treatment by an otolaryngologist. All patients with documented Reinke's edema who underwent at least 1 office-based treatment were considered for inclusion. Patients were excluded if they were <18 years old, had a history of laryngeal malignancy, had a history of a neurolaryngologic disorder (eg, cerebrovascular accident, amyotrophic lateral sclerosis, vocal tremor, or recurrent laryngeal or superior laryngeal nerve injury), had a history of airway stenosis, or were unable to provide consent.

Treatment

All procedures were performed in a clinic setting without sedation. Patients are positioned sitting upright. Local nasal anesthesia is obtained by placing 4% lidocaine and 0.05% oxymetazoline-soaked sponges in the nasal cavities. A flexible endoscope is passed through the nasal cavity for indirect visualization of the endolarynx. Laryngeal anesthesia is obtained by instilling 3 aliquots of 0.5 mL of 4% lidocaine through the working channel of the endoscope during sustained phonation (the ''laryngeal gargle''). The laser fiber is passed through the working channel and advanced until the tip of the fiber is visualized. Laser energy is then applied to the involved tissues. Of note, tissue ablation is not desired during these procedures; rather, enough energy is applied to blanch the superficial tissues (**Figure 1**). Importantly, no immediate reduction is tissue size is desired. Twelve procedures were performed using the KTP laser, and 13 were performed using the PDL.

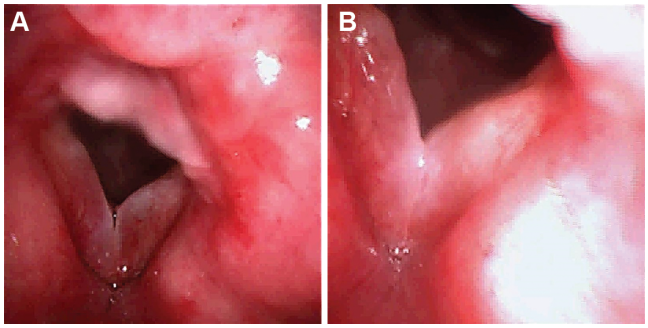

Figure 1. Endoscopic view of glottis immediately before (A) and immediately after (B) treatment. Note the superficial blanching without reduction in tissue mass immediately following application of laser energy.

Experimental Data

Patient demographics, chief complaint, and social history were collected. Endoscopic findings and physician impression were recorded to ensure diagnosis, as well as to document unilateral versus bilateral involvement. Treatment of laryngopharyngeal reflux disease was also documented. Reported complications and patient tolerance data were also collected by extracting physician documentation from the University of Wisconsin–Madison Voice and Swallow Outcomes database. This included whether a procedure was truncated, as well as the number of procedures performed for each patient and whether operative interventions were ultimately required. Minor complications included nasal or pharyngeal pain, minor nosebleed, and vasovagal events. Major complications included need for emergent airway intervention, hospitalization or presentation to the emergency department after treatment, airway bleeding, airway stenosis, reported myocardial infarction or cerebrovascular accident, and extralaryngeal tissue injury. Patients were requested to report complications at follow-up visits; however, our data set did not allow for standardized follow-up of all possible complications.

Objective voice measures were obtained as part of a standard comprehensive voice assessment. Aerodynamic parameters included MPT, mean airflow rate, laryngeal resistance, phonation threshold pressure, and mean peak air pressure. Aerodynamic parameters were measured using the Phonatory Aerodynamic System (model 6600, KayPENTAX, Montvale, New Jersey); the device was calibrated prior to each use according to manufacturer specifications. For MPT, patients produced a sustained /a/ at a modal pitch for as long as possible; this was repeated 3 times, and the longest trial taken as the

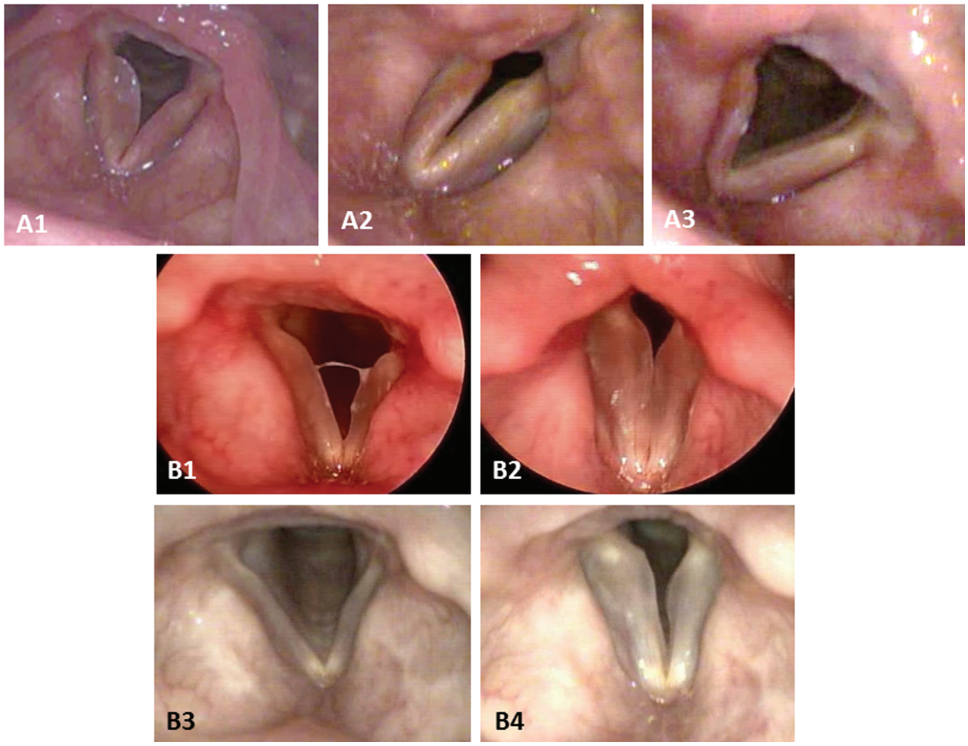

Figure 2. Sample pre- and posttreatment images from 2 patients. A, subject 1: A1, pretreatment, normal inspiration; A2, posttreatment, normal inspiration; A3, posttreatment, vocal fold abduction. B, subject 2: B1, pretreatment, vocal fold abduction; B2, pretreatment, normal inspiration; B3, posttreatment, vocal fold abduction; B4, posttreatment, normal inspiration.

MPT. For mean airflow rate, laryngeal resistance, and peak air pressure, patients produced 3 /pa/ syllable trains at comfortable pitch and loudness; the first and last /pa/ token within each train were removed and the average values computed. For phonation threshold pressure, patients produced 3 /pi/ syllable trains, each beginning with a soft whisper, followed by incremental increases in subglottal pressure until comfortable voicing was achieved; the lowest subglottal pressure at which phonation occurred was recorded as the phonation threshold pressure. Acoustic parameters included minimum and maximum fundamental frequency, phonatory frequency range, and percent jitter. Acoustic data were recorded using the Computerized Speech Lab (model 4150B, KayPENTAX) and Multi-Dimensional Voice Program (model 5105, KayPENTAX). For fundamental frequency values, patients performed ascending and descending glides on the vowel /a/ and were instructed to achieve the lowest and highest frequencies possible, inclusive of falsetto phonation; this was repeated 3 times and the extrema recorded. Phonatory frequency range was calculated as the difference between maximum and minimum fundamental frequency. For percent jitter, patients produced a stable /a/. Dysphonia severity index was calculated as described by Wuyts et al.[20] Subjective voice changes were quantified using pre- and postprocedural scores from the Voice Handicap Index.[21] This instrument measures the impact of one's voice in 3 separate subcategories: functional, physical, and emotional. Finally, total energy delivered was collected.

Although a standardized clinical protocol is followed for collection of voice measures at clinical visits, occasionally not all voice measures are available in the database. Accordingly, a complete data set including all voice parameters was not available for every subject. Analyses were performed using the data that were available, and the number of subjects included in each analysis has been specified. A subject was not included in a given analysis of efficacy if he or she did not have a measurement of that parameter before and after the procedure.

Statistical Analysis

Evaluation of treatment efficacy was performed using paired t tests. If data did not meet assumptions for parametric testing, a Wilcoxon-Mann-Whitney matched pairs signed-rank test was performed. All tests were 2-tailed with a significance level of $\alpha = 0.05$. As complete data sets were not available for all subjects, sample size for each parameter is reported with the corresponding result.

Results

Subject Characteristics

Nineteen patients underwent 25 in-office endoscopic laser treatments of Reinke's edema between January 2007 and November 2013. All but 1 patient was a woman, and all were smokers at the time of presentation. Average age at presentation was 53.9 $\pm$ 7.7 years (range, 43-67 years). All but 1 patient had bilateral involvement, with 1 demonstrating polypoid change affecting only 1 vocal fold. Sample pre- and posttreatment images are provided in **Figure 2**.

Table 1. Voice Outcome Data.[a]

Parameter	Pretreatment	Posttreatment	No.	*P* Value
Dysphonia severity index	−7.0 ± 3.3	−3.0 ± 2.6	12	.003
Acoustic				
Maximum F_0	290 ± 53	482 ± 272	12	<.001
Minimum F_0	110 ± 35	119 ± 95	12	.147
Frequency range	180 ± 67	363 ± 295	12	.003
Percent jitter	4.05 ± 2.83	1.66 ± 1.10	12	.004
Aerodynamic				
Maximum phonation time	8.77 ± 4.28	9.29 ± 3.71	13	.674
Phonation threshold pressure	8.21 ± 3.10	6.69 ± 2.59	4	.049
Mean airflow rate	0.30 ± 0.07	0.27 ± 0.13	4	.536
Laryngeal resistance	47.36 ± 16.97	46.46 ± 24.29	4	.918
Peak pressure	14.04 ± 4.58	10.92 ± 4.07	4	.069
Voice handicap index				
Functional	18 ± 10	12 ± 9	14	<.001
Physical	21 ± 8	15 ± 10	14	.001
Emotional	17 ± 10	11 ± 10	14	.005
Total	56 ± 26	37 ± 27	14	<.001

Abbreviation: F_0, fundamental frequency.
[a]Data are presented as mean ± standard deviation. Complete data sets with measurements of all parameters were not available for every subject; sample size is therefore variable.

Safety and Tolerability

There were no reported minor or major complications. Five procedures had to be truncated due to patient intolerance.

Voice Outcomes

Summary data are presented in **Table 1**. After treatment, dysphonia severity index changed significantly, with a move toward normal voice (P = .003). Phonatory frequency range increased (P = .003), and percent jitter decreased (P = .004). Phonation threshold pressure decreased (P = .049), but there were no significant changes in MPT, mean airflow rate, or laryngeal resistance. Total VHI ($P < .001$) as well as each component of the VHI decreased significantly after treatment (functional: $P < .001$; physical: P = .001; emotional: P = .005; **Figure 3**).

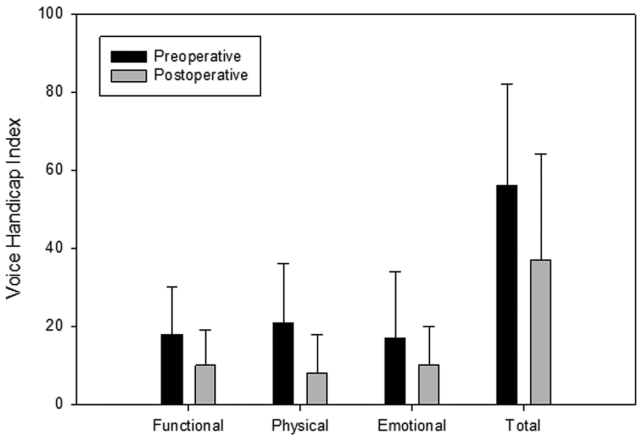

Figure 3. Each component, as well as the total Voice Handicap Index, decreased significantly after treatment. Bar height represents average reported voice handicap; error bars represent standard deviation.

Energy Delivered

Energy delivery data were available on 21 procedures performed to treat bilateral disease. Average energy delivered per procedure was 132 ± 68 J (range, 23-268 J). There was no meaningful difference between the amounts of energy delivered with each laser. For KTP procedures, 126 ± 63 J (range, 47-246 J) were applied; for PDL procedures, 128 ± 75 J (range, 23-268 J) were applied. In 2 procedures for unilateral disease, 108 and 45 J were delivered with the KTP and PDL, respectively.

Discussion

We present a retrospective case series of patients who underwent office-based laser treatment of Reinke's edema. To our knowledge, this study is the largest such series to date.

The increasingly common use of lasers in otolaryngology reflects a general trend toward rendering treatment in the office rather than the operating suite. Office-based treatments offer several advantages. In addition to avoiding the risks of general anesthesia, including myocardial infarction and stroke, unsedated office-based treatment of patients with airway limitations allows the patient to remain in control of his or her own airway throughout the procedure, reducing the risk of airway compromise during induction of general anesthesia. Office procedures cost less,[18] require less time, and avoid the potential complications of microlaryngoscopy, such as dental injury and dysgeusia.[19] Moreover, attempting

an office procedure does not preclude subsequent operative intervention. Still, patients with advanced airway compromise or concerning medical comorbidities are not appropriate for treatment in an office setting, and some patients will demonstrate recalcitrant anxiety to these procedures. In our study, all patients without an obvious airway concern were at least offered an office-based procedure; we do not, however, include patients in the present study who were not amenable to office treatment due to the above limitations.

Surgical lasers fall into 2 broad categories: cutting/ablating lasers and photoangiolytic lasers.[22] Photoangiolytic lasers, including KTP and PDL, selectively target hemoglobin and are therefore most often used to manage highly vascular lesions. Reinke's edema is characterized in part by vascular congestion and stasis within the superficial lamina propria.[2,23] While the exact mechanism of the laser-tissue interaction in benign lesions remains under investigation, it is theorized that photoangiolytic laser energy is effective in improving polypoid degeneration by ablating damaged microvasculature within the SLP, ultimately inducing regression of nonvascular pathologic tissue.[24] It has been proposed that localized energy delivery causes a nonspecific inflammatory response, leading to selective and time-dependent expression of inflammatory cytokines such as transforming growth factor beta 1 and cyclooxygenase 2,[25,26] as well as procollagen/collagenase genes such as matrix metalloproteinases.[26,27] These changes are thought to result in favorable alterations in tissue remodeling. As such, in contrast to classical surgical interventions designed to physically remove excessive tissue, laser therapy is thought to induce a favorable biochemical shift—a biological solution for a biological problem.

In our procedures, energy delivery is titrated to a point of superficial blanching of tissues. No immediate reduction of tissue mass is seen; instead, functional improvement is expected after a period of tissue remodeling. In our study, an average of 132 J was delivered per procedure; however, optimal laser settings and energy titration end points remain undefined. Efforts are underway to characterize these parameters. A recent study examined outcomes for Reinke's edema as a function of laser parameters and initial treatment effects; the average energy applied was 157 J delivered over a 0.369-second exposure time, and voice outcomes were favorable.[28] In an effort to standardize measurement, a validated classification schema was recently proposed to establish a consistent means for measuring response to the KTP laser.[24,29] The present study adds to this growing body of work beginning to evaluate the relationship between amount of energy delivered and treatment outcome.

The patients in our series underwent comprehensive voice analyses before and after completing an intervention, allowing for detailed evaluation of treatment effect. Acoustic measures improved significantly; patients demonstrated improved frequency range due to a higher posttreatment maximum fundamental frequency, representing an improvement in the classic ''low pitched voice'' reported by many patients. Percent jitter also improved after treatment, perhaps reflecting improved vocal fold symmetry after tissue remodeling. Changes in aerodynamic parameters were less pronounced. MPT, laryngeal resistance, mean airflow rate, and peak pressure did not significantly improve following treatment. This may in part be influenced by selection bias, as patients with significant airway compromise—and thus, likely, the most abnormal pretreatment aerodynamic profiles—were not offered office procedures. Phonation threshold pressure did, however, improve after treatment. Finally, our patients demonstrated improvement in all subcategories of the Voice Handicap Index—functional, physical, and emotional. This perhaps more than other measures suggests the utility of these procedures.

Some aspects of the present study may require clarification. First, some individuals showed worsening of certain voice measures after treatment. For example, 1 patient showed increased phonation threshold pressure and airway resistance after a second laser treatment, and 3 patients had decreased MPT after treatment. Also, note that 6 patients underwent multiple procedures. The decision for repeat treatment was based on clinical assessment of recurrent or persistent Reinke's edema with ongoing dysphonia rather than on objective voice data.

While unsedated endoscopic procedures are possible in the majority of patients, anatomic and physiologic limitations as well as anxiety-related factors will represent a barrier in some patients. Of the 25 procedures presented here, 5 were truncated due to patient intolerance. All patients were active smokers at the time of treatment; it is possible that reactive airway physiology contributed to this high rate of intolerance. Our database did not include which patients ultimately underwent operative interventions, but it is likely that some did. Given this limitation, we are unable to assess voice changes related to subsequent surgical intervention and therefore cannot comment on voice outcomes in these patients. Importantly, no patients required emergent airway intervention during or immediately after the procedure, had significant bleeding, or required hospitalization immediately following the procedure. As such, office-based laser treatments in our series were safe.

The present study has several important limitations. As a retrospective analysis without a control group, we cannot determine whether the changes in voice parameters observed after treatment were actually due to the intervention or simply reflect normal temporal variation of the disease. Second, although this is the largest series of patients undergoing photoangiolytic laser therapy for Reinke's edema, our sample size is still modest. Further, complete data sets were not available for all patients. This reduced our effective sample size for the pre- and posttreatment analyses and precluded detailed analysis of parameters over longer periods. Finally, our data set did not provide a standardized means for follow-up; as such, patients who developed complications following the conclusion of their procedure may not be included.

As office-based procedures become increasingly more common, there are many points for further study. Definition of laser settings to optimize tissue remodeling remains an important and active area of investigation. Prospective

studies with larger numbers will help define effect of treatment and clarify the appropriate frequency for follow-up evaluation. To reduce potential cofounders, further studies may benefit from tightly controlling for untreated voice abuse and LPR. Additionally, analyzing differences in outcomes between patients who continue to smoke and those who have quit may define the impact of continued tobacco use on Reinke's edema and healing following photoangiolytic laser intervention. Finally, attention toward procedural details to optimize comfort and minimize anxiety may improve procedure tolerability in a broader population.

Author Contributions

Ian J. Koszewski, study design; data interpretation; data analysis; drafting manuscript; final approval of manuscript; accountable for all aspects of work; **Matthew R. Hoffman**, study design; data interpretation; data analysis; drafting manuscript; final approval of manuscript; accountable for all aspects of work; **W. Greg Young**, study design; data acquisition; data analysis; critical revision of manuscript; final approval of manuscript; accountable for all aspects of work; **Ying-Ta Lai**, study design; data acquisition; data analysis; critical revision of manuscript; final approval of manuscript; accountable for all aspects of work; **Seth H. Dailey**, study design; performed surgical procedures; data interpretation; critical revision of manuscript; final approval of manuscript; accountable for all aspects of work.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Dursun G, Ozgursoy OB, Kemal O, Coruh I. One-year follow-up results of combined use of CO_2 laser and cold instrumentation for Reinke's edema surgery in professional voice users. *Eur Arch Otorhinolaryngol*. 2007;264:1027-1032.
2. Marcotullio D, Magliulo G, Pezone T. Reinke's edema and risk factors: clinical and histopathologic aspects. *Am J Otolaryngol*. 2002;23:81-84.
3. Chung JH, Tae K, Lee YS, et al. The significance of laryngopharyngeal reflux in benign vocal mucosal lesions. *Otolaryngol Head Neck Surg*. 2009;141:369-373.
4. Cipriani NA, Martin DE, Corey JP, et al. The clinicopathologic spectrum of benign mass lesions of the vocal fold due to vocal abuse. *Int J Surg Path*. 2011;19:583-587.
5. Kamargiannis N, Gouveris H, Katsinelos P, et al. Chronic pharyngitis is associated with severe acidic laryngopharyngeal reflux in patients with Reinke's edema. *Ann Otol Rhinol Laryngol*. 2011;120:722-726.
6. Hojslet PE, Moesgaard-Nielsen V, Karlsmore M. Smoking cessation in chronic Reinke's oedema. *J Laryngol Otol*. 1990; 104:626-628.
7. Hirano M, Shin T, Morio M, et al. An improvement in surgical treatment for polypoid vocal cord: sucking technique. *Otologia (Fukuoka)*. 1976;22:583-589.
8. Tan NCW, Pittore B, Puxeddu R. The ''M'' shaped microflap for treatment of complex Reinke's space oedema of the vocal cords. *Acta Otorhinolaryngologica Italica*. 2010;30:259-263.
9. Sant'Anna GD, Mauri M.Use of the microdebrider for Reinke's edema surgery. *Laryngoscope*. 2000;110:2114-2116.
10. Yates A, Dedo HH.Carbon dioxide laser enucleation of polypoid vocal cords. *Laryngoscope*. 1984;94:731-736.
11. Remacle M, Lawson G, Watelet JB. Carbon dioxide laser microsurgery of benign vocal fold lesions: indications, techniques, and results in 251 patients. *Ann Otol Rhinol Laryngol*. 1999;108:156-164.
12. Lumpkin SM, Bishop SG, Bennett S. Comparison of surgical techniques in the treatment of laryngeal polypoid corditis. *Ann Otol Rhinol Laryngol*. 1987;96:254-257.
13. Pitmann MJ, Lebowitz-Cooper A, Iacob C, Tan M. Effect of the 532nm pulsed KTP laser in the treatment of Reinke's edema. *Laryngoscope*. 2012;122:2786-2792.
14. Sheu M, Sridharan S, Kuhn M, et al. Multi-institutional experience with the in-office potassium titanyl phosphate laser for laryngeal lesions. *J Voice*. 2012;26:806-810.
15. Koufman JA, Rees CJ, Frazier WD, et al. Office-based laryngeal laser surgery: a review of 443 cases using three wavelengths. *Otolaryngol Head Neck Surg*. 2007;137:146-151.
16. Rees CJ, Halum S, Wijewickrama RC, Koufman JA, Postma GN. Patient tolerance of in-office pulsed dye laser treatments to the upper aerodigestive tract. *Otolaryngol Head Neck Surg*. 2006;134:1023-1027.
17. Zeitels SM, Burns JA. Office-based laryngeal laser surgery with the 532-nm pulsed-potassium-titanyl-phosphate laser. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:394-400.
18. Rees CJ, Postma GN, Koufman JA. Cost savings of unsedated office-based laser surgery for laryngeal papillomas. *Ann Otol Rhinol Laryngol*. 2007;116:45-48.
19. Klussmann JP, Knoedgen R, Wittekindt C, Damm M, Eckel HE. Complications of suspension laryngoscopy. *Ann Otol Rhinol Laryngol*. 2002;111:972-976.
20. Wuyts FL, de Bodt MS, Molenberghs G, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res*. 2000;43: 796-809.
21. Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI): development and validation. *Amer J Speech Lang Pathol*. 1997;6:66-70.
22. Yan Y, Olszewski AE, Hoffman MR, et al. Use of lasers in laryngeal surgery. *J Voice*. 2010;24:102-109.
23. Martins RH, Dominques MA, Fabro AT, Dias NH, Santana MF. Reinkes' edema: immunoexpression study of fibronectin, laminin and collagen IV in 60 cases by immunohistochemical techniques. *Braz J Otorhinolaryngol*. 2009;75:821-825.
24. Mallur PS, Tajudeen BA, Aaronson N, Branski RC, Amin MR. Quantification of benign lesion regression as a function of 532-nm pulsed potassium titanyl phosphate laser parameter selection. *Laryngoscope*. 2011;121:590-595.
25. Branski RC, Barbieri SS, Weksler BB, et al. Effects of transforming growth factor-beta1 on human vocal fold fibroblasts. *Ann Otol Rhinol Laryngol*. 2009;118:218-226.

26. Lin Y, Yamashita M, Zhang J, Ling C, Welham NV. Pulsed dye laser-induced inflammatory response and extracellular matrix turnover in rat vocal folds and vocal fold fibroblasts. *Lasers Surg Med*. 2009;41:585-594.
27. Mallur PS, Branski RC, Amin MR. 532-nanometer potassium titanyl phosphate (KTP) laser-induced expression of selective matrix metalloproteinases (MMP) in the rat larynx. *Laryngoscope*. 2011; 121:320-324.
28. Young VN, Mallur PS, Wong AW, et al. Analysis of potassium titanyl phosphate laser settings and voice outcomes in the treatment of Reinke's edema. *Ann Otol Rhinol Laryngol*. 2015; 124:216-220.
29. Mallur PS, Johns MM 3rd, Amin MR, Rosen CA. Proposed classification system for reporting 532-nm pulsed potassium titanyl phosphate laser treatment effects on vocal fold lesions. *Laryngoscope*. 2014;124:1170-1175.

Reprinted by permission of Laryngoscope. 2015; 125(4):909-912.

The Laryngoscope
© 2014 The American Laryngological,
Rhinological and Otological Society, Inc.

The Utility of Office-Based Biopsy for Laryngopharyngeal Lesions: Comparison with Surgical Evaluation

Amanda L. Richards, MBBS, FRACS; Manikandan Sugumaran, MD; Jonathan E. Aviv, MD; Peak Woo, MD; Kenneth W. Altman, MD, PhD

Objectives/Hypothesis: Advances in flexible endoscopy with working-channel biopsy forceps have led to excellent visualization of laryngopharyngeal lesions with capability for in-office awake biopsy. Potential benefits include prompt diagnosis without risk of general anesthesia, preoperative counseling, and avoiding an anesthetic should the lesion return benign. We evaluate the accuracy of these biopsies in order to determine their role and diagnostic value.

Study Design: Retrospective chart review.

Methods: Medical records were reviewed from January 1, 2010, through July 31, 2013, of patients who underwent office-based current procedural terminology code 31576 and were taken to the operating room for direct microlaryngoscopy with biopsy/excision. Clinical diagnoses and pathology reports were reviewed. For statistical analysis, we considered three groups: 1) malignant and premalignant, 2) lesions of uncertain significance, and 3) benign lesions.

Results: In the study period, 76 patients with an office biopsy had a clinical picture to warrant direct microlaryngoscopy and biopsy/excision. Kendall's coefficient for each group indicated moderate correlation only. When groups 1 and 2 were considered together, there was a substantial and statistically significant correlation. For malignant and premalignant lesions, the office biopsy analysis was as follows: sensitivity = 60%, specificity = 87%, positive predictive value = 78%, and negative predictive value = 74%.

Conclusion: Office biopsy may offer early direction and avoid operative intervention in some cases; however, for suspected dysplastic or malignant lesions, direct microlaryngoscopy should be the standard of care to ensure adequate full-thickness sampling and staging. For benign pathology, office biopsy is a safe and viable alternative to direct microlaryngoscopy and biopsy/excision.

Key Words: Office biopsy, lesion, leukoplakia, dysplasia, microlaryngoscopy, medical decision making.

Level of Evidence: 4.

***Laryngoscope*, 125:909–912, 2015**

INTRODUCTION

Advances in flexible laryngoscopy, imaging technology, instrument miniaturization, and changes to procedure reimbursement have led to an increase in office-based management in laryngology. Since the introduction of the fiber optic laryngoscope in 1976, there have been steady advances in the quality of lighting and imaging for office laryngeal examinations from fiber optic to distal chip endoscopes.[1] Also, adaptations in the design of the flexible scopes have allowed for the use of a side channel port or disposable sheath for passage of a cupped laryngeal biopsy forceps.[2] The combination of these forceps with optimal imaging has provided an option to obtain tissue for pathology during an outpatient office visit with topical anesthesia. Traditionally, these patients would require a visit to the operating room (OR) with general anesthesia for a direct microlaryngoscopy and biopsy or excision of the lesion. However, regardless of technique, all biopsies need to provide a representative sample of the lesion to demonstrate cell morphology. In addition, sample depth is also important, particularly in cases of dysplasia for which deeper levels may determine a different diagnosis and prognosis.[3]

There are a proposed number of conditions for which office biopsy alone has been proposed as sufficient: 1) confirmed diagnosis of carcinoma when clinically suspected; 2) complete excision of a lesion at the time of office biopsy; 3) benign pathology and resolution of the lesion with treatment; 4) evidence for keratosis, papilloma, or mild dysplasia with stable clinical examination; and 5) the risks of surgical evaluation with general anesthesia outweigh the potential diagnostic or therapeutic benefits of the procedure. Potential benefits include the following: 1) avoiding the risk of general anesthesia, 2) reduced duration from clinical suspicion to histologic confirmation, 3) negating patient anatomic limitations, and 4) avoiding the costs of general anesthesia and the OR.[4]

From the Department of Otolaryngology–Head and Neck Surgery, Icahn School of Medicine, Mount Sinai Hospital, New York, New York, U.S.A.

Editor's Note: This Manuscript was accepted for publication October 7, 2014.

Dr. Aviv is shareholder in Vision Sciences. The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Kenneth W. Altman, MD, PhD, FACS, Professor and Vice Chair for Clinical Affairs, Director, The Institute for Voice and Swallowing at BCM, Bobby R. Alford Department of Otolaryngology - HNS, Baylor College of Medicine, 6501 Fannin Street, Room NA 102, Houston, TX 77030. E-mail: Kenneth.Altman@bcm.edu

DOI: 10.1002/lary.25005

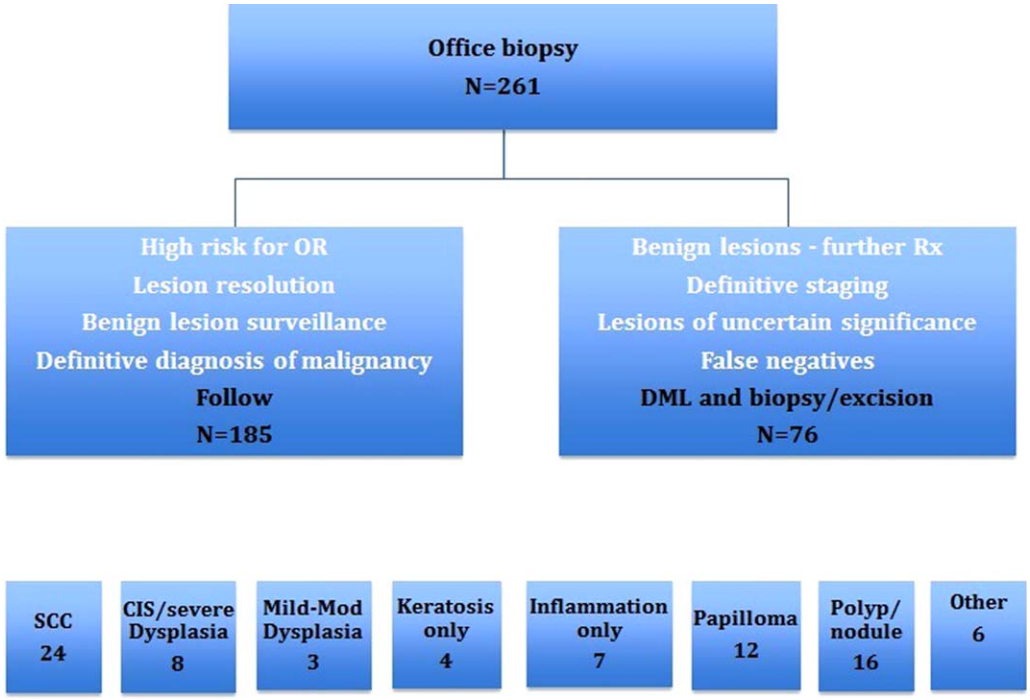

Fig. 1. Flow of office biopsy patients. (Operating room biopsy diagnoses are listed in the last row). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Traditionally, patients warranting direct microlaryngoscopy after office biopsy are those with limited tissue obtained during attempted office biopsy, a concern regarding false-negative office biopsy results, a requirement for disease volume reduction to avoid respiratory or swallowing impairment, and a need for excision of the lesion to improve the voice. Other advantages of direct microlaryngoscopy include a more detailed examination of the extent of a tumor, more accurate biopsy capabilities, and the option for definitive treatment by excision for many lesions.

Despite the popularity of office biopsy, there is a paucity of data in the literature evaluating the accuracy compared to histologic diagnosis at operation. The goals of this study are to determine the accuracy of office biopsies when compared to direct microlaryngoscopy and to evaluate its role and diagnostic value.

MATERIALS AND METHODS

A retrospective medical chart review was performed from January 1, 2010, to July 31, 2013, after receiving approval from the Institutional Review Board Human Subjects Committee. This review identified 261 patients in the clinical practices of the authors who underwent office biopsy (current procedural terminology code 31576) for laryngeal and pharyngeal lesions. Patients' records were then reviewed to determine those who underwent direct microlaryngoscopy with biopsy.

Patients who had resolution of the lesion following biopsy, surveillance of a previously histologically proven benign diagnosis, and a definitive diagnosis of cancer who proceeded to nonsurgical definitive treatment were excluded from the study. We also excluded current anticoagulation, anterior commissure lesions, submucosal lesions, and anatomically obstructive pathology. Patients with brush biopsy alone were also excluded. The pathology reports were reviewed for consistency between office and surgical specimens and compared to clinical diagnoses. The flow of the patients is summarized in Figure 1.

Office biopsies were performed using distal chip video endoscopes (ENT-5000, Vision Sciences, Inc. or VNL-1570STK, KayPENTAX Montvale, NJ) in conjunction with a 2-mm channel endosheath and 1.8-mm nonserrated cup biopsy forceps. The nasal cavity was anesthetized with aerosolized 4% lidocaine with epinephrine 1:100,000 or 4% lidocaine with phenylephrine hydrochloride. The channel-sheathed video endoscope was then passed transnasally into the laryngopharynx. Topical laryngopharyngeal anesthesia was achieved by delivering 0.5 cc of plain 4% lidocaine to the laryngeal surface of the epiglottis. Once supraglottic anesthesia was achieved, 1 to 2 cc of plain 4% lidocaine was then delivered topically to the glottis. The 1.8-mm biopsy forcep was then passed under videoendoscopic guidance and biopsies were performed.

Direct microlaryngoscopy with biopsy was performed under general anesthesia, and lesions were visualized with a zero-degree telescope and binocular microscope. Lesions were excised or sampled for pathologic evaluation using phonosurgical instruments. The procedures included a submucosal dissection in order to obtain epithelial basement membrane in the specimen.

Office biopsy results were divided into clinically relevant groups that would normally used to direct patient care algorithms. For example, mild to moderate dysplasia was separated from severe dysplasia and carcinoma in situ (CIS)/squamous cell carcinoma (SCC). For statistical analysis, we considered three groups: 1) malignant and premalignant (SCC, CIS, and severe dysplasia); 2) lesions of uncertain significance (mild–moderate dysplasia and hyperkeratosis); and 3) benign lesions. Patients who were noted to have a dual diagnosis on histology (e.g., inflammation with mild dysplasia) were analyzed within the group that would direct their final treatment.

To test interrater reliability, we utilized Kendall's coefficient of concordance for the numerically coded ordinal responses

TABLE I.
Summary of Results.

Office Biopsy Pathology	Office Biopsy N =	Accuracy Compared to Final Pathology	Pathology of Missed Diagnosis (False Negatives)	Rate of False Negatives%
SCC	4	100.0%	N/A	0.0%
Severe dysplasia/ CIS	23	17.4%	SCC	56.5%
			Mild–Moderate dysplasia	8.7%
			Inflammation only	8.7%
			Polyps or nodule	4.3%
			Keratosis	4.3%
Mild to moderate dysplasia	12	25.0%	SCC	25.0%
			Severe dysplasia/CIS	33.3%
			Polyps or nodule	8.3%
			Keratosis	8.3%
Keratosis without dysplasia	7	14.3%	SCC	28.6%
			Severe dysplasia/CIS	14.3%
			Polyps or nodule	28.6%
			Inflammation only	14.3%
Inflammation only	8	12.5%	SCC	12.5%
			Severe dysplasia/CIS	25.0%
			Keratosis	12.5%
			Polyps or nodule	25.0%
			Papilloma	12.5%
Polyp/nodule	10	100.0%	N/A	0.0%
Papilloma	11	81.8%	Inflammation only	9.1%
			Other	9.1%
Other benign	6	0.0%	SCC	16.7%
			Keratosis	33.3%
			Inflammation only	33.3%
			Papilloma	16.7%
Inadequate	4	N/A	SCC	50.0%
			Polyp or nodule	50.0%

CIS = carcinoma in situ; SCC = squamous cell carcinoma.

via the SAS macro MAGREE. Kendall's coefficients do not treat all misclassifications equally. For instance, Kendall's coefficients considers the consequences of misclassifying a perfect (rating = 5) object as bad (rating = 1) as more serious than misclassifying it as very good (rating = 4). For all statistical analysis, data was analyzed using the SAS System software (SAS Institute, Inc., Cary, NC).

RESULTS

Seventy-six subjects underwent evaluation with 81 office biopsies with subsequent direct microlaryngoscopy under general anesthesia. The age range of the subjects was 21 to 84 years, with a median age of 62 and a male to female ratio of 5:1. There were 76 laryngeal biopsies and five oropharyngeal biopsies. The oropharynx subsites included two for tonsil, two for tongue base, and one for soft palate. There were no complications from any of the office or operative procedures performed.

The results of office biopsy and their subsequent direct microlaryngoscopy (DML) and biopsy/excision are summarized in Table I.

When considered as separate groups, Kendall's coefficient for each group was all 0.5, indicating "moderate correlation" only. None of these approached statistical significance ($P = 0.5$). When groups 1 and 2 (i.e., lesions of uncertain significance and premalignant/malignancy) were considered together, the coefficient was 0.64 ($P = 0.029$), indicating "substantial correlation." For malignant/premalignant lesions, the office biopsy analysis was as follows: sensitivity = 60%, specificity = 87%, positive predictive value = 78%, and negative predictive value = 74%.

DISCUSSION

Medical decision making, patient counseling, and surgical planning benefit from understanding the nature of a lesion based on patient demographics, clinical history, the physical examination, and cytologic or pathologic diagnosis. The nature of a laryngeal lesion will affect prioritizing the different surgical goals of the following: 1) confidence in a pathologic diagnosis, 2) control of disease, and 3) voice preservation or improvement. See Table II for proposed office biopsy candidacy.

While office biopsy increases in popularity, there needs to be further clarification regarding its utility. The results of this study highlight the concern of office biopsy being used as a substitute for traditional DML. The small tissue sample, limited depth past the basement membrane, and ability to sample only portions of a suspicious lesion are known disadvantages to this technique, especially with leukoplakia and lesions that have some degree of dysplasia. Serrated or other forceps, however, may favourably influence the sensitivity and specificity of the findings, and appropriate selection of forceps is a major consideration in the management of early laryngeal malignancy by any method. Therefore, just as in the OR under direct laryngoscopic conditions for which the surgeon would choose the forceps for biopsy carefully, the same considerations must be applied when performing transnasal biopsy.

Although it may be acceptable for some screening tests to have a high specificity and lower sensitivity, this is not appropriate for this diagnostic test. Sensitivity for malignancy/premalignancy was only 60%, indicating that it is inadequate as a diagnostic test: clinical suspicion alone in this setting would seem at least equivalent. Only 15% of invasive SCC was identified at office biopsy, and it is evident that any clinically suspicious neoplasm must

TABLE II.
Candidates for Office Biopsy.

Inclusion	Exclusion
Anatomic limitations for DML	Anticoagulation
Voice less critical	Anterior commissure location
Following a known benign diagnosis	Submucosal lesion
High risk for general anesthesia	Lesion associated with obstruction

DML = direct microlaryngoscopy.

be evaluated with direct microlaryngoscopy as possible. Cohen et al. noted a 33% false negative rate of in-office flexible laryngeal biopsy; and in their series, CIS on office biopsy was most often SCC at DML.[5] It should be noted, however, that when an office biopsy showed a diagnosis of SCC, it was correlated with the final histologic diagnosis in 100% of patients. For patients with a suspected malignant lesion who are unable to undergo a general anesthetic, reassurance can be given that a diagnosis of cancer in the office correlates with a diagnosis of cancer in OR. The substantial correlation for malignancy/any degree of dysplasia may be adequate to counsel certain patients, and early diagnosis with office biopsy may offer extra confidence in the treatment paradigm, particularly when multiple treatment options are available. It may also reduce time to definitive treatment. When a histologic diagnosis of severe dysplasia or CIS in a lesion with malignant suspicion is added to the diagnostic/staging capabilities of imaging modalities such as CT scanning, radiation may be employed with increased confidence of the clinical diagnosis and stage.[6] Seventy-eight percent of our patients with severe dysplasia, CIS, or SCC on office biopsy were counseled appropriately in advance of surgery and underwent microlaryngoscopy with CO2 laser cordectomy at the time of initial surgical evaluation.

It is an accepted standard of care that mild to moderate dysplasia may be observed, and so it follows that this diagnosis at office biopsy may reassure the clinician. However, lesions of uncertain significance have only moderate correlation to their final pathologic diagnosis. Mild to moderate dysplasia frequently represents more sinister disease, and it may be false reassurance to rely on office biopsy alone. Similar to brush biopsy,[7] a dysplastic office biopsy result indicates a need for further investigation, without providing a definitive diagnosis.

The diagnosis of hyperkeratosis/parakeratosis (pathologic correlates of leukoplakia) and inflammation also frequently correlate with a final histopathologic diagnosis of malignancy and highlight the limitations on biopsy of the awake patient with the absence of tactile feedback. It is well known that potentially malignant and CIS epithelia are associated with pronounced stromal reaction,[8] which is reflected in the 37% of patients with only an office biopsy of inflammation, however, with severe dysplasia/CIS or SCC on final diagnosis. Clinical judgment is paramount in ensuring that patients with suspected false negative biopsies undergo DML.

When considering benign lesions such as papilloma and vocal fold nodules or polyps, office biopsy may play a more definitive role. A clinical suspicion of papilloma or polyps/nodules correlated well with final diagnosis. When clinically indicated, these lesions are appropriate for office management alone.

In the current environment of financially accountable medicine, consideration should be given to cost. Naidu et al.[9] and workers found an average cost saving of $6,970.56 per patient when office biopsy cost was compared to OR biopsy; however, it is instructive to consider that any patient who undergoes both office biopsy and OR biopsy does so at considerable financial burden.

There are limitations to our study, including selection bias determining the candidacy or need for office biopsy. In addition, the technique and types of cupped forceps used were not compared to other approaches. The utility of serrated forceps or rigid endoscopy with nonflexible forceps should also be considered and may enhance sampling. The ability to obtain high diagnostic yield biopsies in the office is also dependent on the experience of the surgeon; our group obtaining the biopsies has over 5 years experience with this method. Classification of the severity of dysplasia is also variable from different classification systems and pathologists,[2] which can skew the accuracy of office biopsies when different grades of dysplasia are considered. A multicenter study is warranted, especially in the realm of medical decisionmaking in complex laryngeal lesions, as well as outcomes of prioritizing disease eradication over voice outcomes with dysplastic lesions.

CONCLUSION

Our study shows that office biopsy has the highest utility in clinically benign lesions and those with SCC. Laryngopharyngeal biopsies have only a moderate correlation with final pathology, although the potential utility increases in certain clinical scenarios and with careful choice of forceps. Office biopsy has a tendency to underestimate the severity of dysplastic lesions, and any degree of dysplasia should be considered as potentially malignant until, as possible, proven otherwise with operative assessment. However, preoperative patient counseling, surgical planning, and therapy may be positively impacted by information from office biopsies, comparable to management of other head and neck neoplasms. For benign pathology that clinically harbors no suspicion for malignancy, office biopsy is a safe and viable alternative to direct microlaryngoscopy and biopsy/excision.

Acknowledgment

We would like to acknowledge Alan Weinberg, MS, for the statistical analysis, and Laurence T. Blanchard, BA, for data preparation.

BIBLIOGRAPHY

1. Eller R, Ginsburg M, Lurie D, Heman-Ackah Y, Lyons K, Sataloff R. Flexible laryngoscopy: a comparison of fiber optic and distal chip technologies. Part 1: vocal fold masses. *J Voice* 2008;22:746–750.
2. Rosen CA, Amin MR, Sulica L, et al. Advances in office-based diagnosis and treatment in laryngology. *Laryngoscope* 2009;119(suppl 2):S185–S212.
3. Zhang HK, Liu HG. Is severe dysplasia the same lesion as carcinoma in situ? 10-Year follow-up of laryngeal precancerous lesions. *Acta Otolaryngol* 2012;132:325–328.
4. Rees CJ, Postma GN, Koufman JA. Cost savings of unsedated office-based laser surgery for laryngeal papillomas. *Ann Otol Rhinol Laryngol* 2007; 116:45–48.
5. Cohen JT, Safadi A, Fliss DM, Gil Z, Horowitz G. Reliability of a transnasal flexible fiberoptic in-office laryngeal biopsy. *JAMA Otolaryngol Head Neck Surg* 2013;139:341–345.
6. Chu MM, Kositwattanarerk A, Lee DJ, et al. FDG PET with contrast-enhanced CT: a critical imaging tool for laryngeal carcinoma. *Radiographics* 2010;30:1353–1372.
7. Nichols ML, Quinn FB, Jr, Schnadig VJ, et al. Interobserver variability in the interpretation of brush cytologic studies from head and neck lesions. *Arch Otolaryngol Head Neck Surg* 1991;117:1350–1355.
8. Hellquist H, Cardesa A, Gale N, Kambic V, Michaels L. Criteria for grading in the Ljubljana classification of epithelial hyperplastic laryngeal lesions. A study by members of the Working Group on Epithelial Hyperplastic Lesions of the European Society of Pathology. *Histopathology* 1999;34:226–233.
9. Naidu N, Noordzij JP, Samin A, Jalisi S, Grillone GA. Comparison of efficacy, safety, and cost-effectiveness of in-office cup forcep biopsies versus operating room biopsies for laryngopharyngeal tumours. *J Voice* 2012; 26:604–606.

Reprinted by permission of J Voice. 2014; 28(3):382-386.

Office-Based Injection Laryngoplasty for the Management of Unilateral Vocal Fold Paralysis

***Sunil P. Verma and †Seth H. Dailey,** **Irvine, California and †Madison, Wisconsin*

Summary: Objective. Office-based injection laryngoplasty (OBIL) is a common method of addressing glottal insufficiency. This retrospective chart review identifies the demographics, laterality, technique, success rate, injectates, and complications of OBIL performed over a 3-year period at a single institution.

Study Design. Retrospective chart review.

Methods. All OBILs performed for the management of UVFP by the senior author over 3 years (2007–2009) were identified from billing records. The age, gender, laterality, underlying disease process, augmentation material, route of injection, and complications were recorded.

Results. Eighty-two OBILs were attempted on 57 patients. The most common route of access was transoral (85.6%). All OBILs were able to be completed. Injectates used were hyaluronic acid derivatives (57.3%), calcium hydroxyapatite (16%), and Cymmetra (16.5%). Three complications (3.7%) occurred. Thirty percent of patients ultimately elected for thyroplasty or ansa reinnervation, 22% found their condition to self-resolve, 14% died, and 25% were lost to follow-up.

Conclusions. Using a variety of approaches, OBIL is possible in almost all patients. The single surgeon transoral route using a rigid angled telescope and curved injection needle was the most commonly used approach. Multiple injectates can be used and have good safety records. The final disposition of patients may be variable and warrants further investigation.

Key Words: Laryngology–Laryngeal surgery–Office-based–Procedures–Surgery–Vocal fold paralysis–Hoarseness–Thyroplasty–Reinnervation.

INTRODUCTION

Injection laryngoplasty (IL) has been a cornerstone in the management of unilateral vocal fold paralysis (UVFP) since its first description.[1] During the majority of the last century, IL was commonly performed in the operating room (OR). However, with the advent of "chip-tip" endoscopes, refinements in the ability to deliver anesthesia to the larynx[2,3] and the development of numerous injectables,[4,5] there has been a move toward IL performed in the office.[6] Advantages of OBIL include markedly decreased cost, avoidance of the risks of general anesthesia, and the ability titrate injectate delivery for optimized voice outcomes, among others.[7]

As the population ages and grows and as some of the most common causes of UVFP increase,[8] including the number of thyroid cancers,[9] cervical spine surgeries,[10] lung cancer resections, and aortic valve replacements,[11] one may expect the incidence of UVFP to increase as well. As the paradigm of OBIL for UVFP continues to evolve, there are questions which remain to be answered.

The first involves the safety profile of both OBIL and the numerous injectables which are being used for the treatment. UVFP often occurs secondary to malignancy, complications from surgery, or both. As such, patients with UVFP often possess multiple morbidities including general health concerns, cardiopulomonary compromise, need for anticoagulation, among other medical and psychosocial concerns. With this in mind, it is critical to evaluate the safety of OBIL as has been done for other office-based laryngeal surgeries.[12,13] In an effort to avoid general anesthesia, another question to be answered is how often OBIL can actually be completed. Finally, there is an active discussion regarding the ultimate disposition of patients after injection.[14–16]

To answer these questions, a retrospective chart review was performed of all OBILs performed for UVFP over a 3-year period at an academic tertiary care institution.

MATERIALS AND METHODS

After obtaining approval by the institutional review board, all OBILs attempted for UVFP by the senior author over 3 years (2007–2009) were identified from billing records. The age, gender, laterality, underlying disease process, route of injection, procedural success rate, amount and type of augmentation material used, complications, and patient disposition were recorded.

All procedures were performed in the otolaryngology clinic examination suite containing a powered examination chair, video tower with photodocumentation capability. Informed consent was obtained and a procedural "time-out" was performed before each procedure. Patient vital signs were collected before the visit; however, no cardiopulomonary monitoring was performed during the procedure. All injectates were directed toward the paraglottic space musculature. Approaches used were transoral,[17] transcricothyroid membrane,[18] transthyrohyoid membrane,[19] and transthyroid ala.

Accepted for publication October 10, 2013.
Financial Disclosures: None.
Conflicts of interest: None.
Presented at the 2011 UCSF Fall Voice Conference/International Association of Phonosurgery; November 3, 2011; San Francisco, California.

From the *University Voice and Swallowing Center, Department of Otolaryngology-Head and Neck Surgery, University of California, Irvine School of Medicine, California; and the †Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin.

Address correspondence and reprint requests to Seth H. Dailey, Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, University of Wisconsin School of Medicine and Public Health, K4/760 Clinical Science Center, Madison, WI, 53792-7395. E-mail: dailey@surgery.wisc.edu

Journal of Voice, Vol. 28, No. 3, pp. 382-386
0892-1997/$36.00
© 2014 The Voice Foundation
http://dx.doi.org/10.1016/j.jvoice.2013.10.006

For the transoral approach, the oral cavity is first anesthetized with topical lidocaine spray applied with an atomizer. The tonsillar pillars, base of tongue, and posterior pharyngeal wall are sprayed with lidocaine. The patient is asked to assume the "sniffing" position and directed to hold his tongue with gauze. Visualization of the laryngopharynx is obtained with a transoral rigid 70° angled telescope held by the surgeon. The view from the scope is transmitted to a screen on the video tower (Figure 1).

An Abraham cannula attached to a syringe with 4% lidocaine is placed along the patient's lingual sulcus and directed over the larynx. A "laryngeal gargle" is performed with 4% lidocaine dripping lidocaine to the endolarynx during sustained phonation. The surgeon then advances a syringe with injectate attached to an orotracheal injector needle (model # 1650030 and 1650050; Medtronic, Minneapolis, MN) along the patient's lingual sulcus and directs it to the larynx. The needle may be used to lateralize the patient's false vocal fold. The needle is inserted through the superior surface of the vocal fold into its body. Injectate is applied within the paraglottic space with approximately 20% overinjection to account for reabsorption.

The percutaneous techniques are performed with a surgeon and an assistant. The skin is anesthetized with 1% lidocaine. After the nasal cavity is anesthetized, a channeled flexible laryngoscope is advanced into laryngopharynx. A laryngeal gargle is performed by dripping 4% lidocaine to the endolarynx *via* the channel of the laryngoscope during sustained phonation. A 25 gauage 1.25-in needle is passed through the skin into the larynx by the surgeon and is directed into the vocal fold.

RESULTS

Eighty-two OBILs were attempted on 57 patients. Patients injected were aged between 16 and 83 years, with a mean age of 60 years. Thirty-five males and 22 females were treated. UVFP occurred on the left side in 40 patients and on the right side in 17. Tables 1 and 2 list the etiology of paralysis and approach used for injection, respectively. No procedure had to be terminated early and all procedures were able to be performed to the intended completion point. On average, 0.64 mL of injectate was used in each setting. The augmentation material used is listed in Table 3.

Three complications (3.7%) were noted during or after OBIL. One patient had a hypersensitivity reaction to Restylane. One patient had calcium hydroxyapatite injected superficially requiring microdirect laryngoscopy and removal at a later date. One patient experienced vocal fold edema after injection and was observed in the office without incident.

Figure 2 details the disposition of patients after OBIL.

FIGURE 1. Surgeon and patient positioning for transoral vocal fold injection.

TABLE 1.
Etiology of UVFP

Etiology	Percentage of Patients
Thoracic	36
Idiopathic	30
Cervical	21
Cerebral	10
Intubation	3

DISCUSSION

UVFP is an entity often encountered by otolaryngologists-head and neck surgeons. Management options include voice therapy, OBIL, and injection laryngoplasty performed under general anesthesia in the OR, reinnervation, thyroplasty, and arytenoid repositioning maneuvers. Definitive treatment typically is deferred for the first 9 months after onset and during that time, patients' options are observation, voice therapy, or IL.

IL has an important role in the management of glottal insufficiency. It provides immediate treatment of symptoms related to voice and cough. OBIL offers some advantages over IL performed in the OR. OBIL permits an unobstructed view of the vocal folds, allowing the surgeon to clearly visualize the change in configuration during injection.[7] There is room for immediate analysis of results permitting simultaneous modification if necessary.[20] Performing the procedure under local anesthesia not only reduces the risks associated with general anesthesia but also allows patients to return to normal activities immediately, preventing lost time from work.

Another advantage of OBIL is cost savings. Grant et al estimated increased charges of $8250 for IL performed in the OR compared with the office.[21] Similarly, other authors have noted significant financial savings associated with performance of IL in office as opposed to the OR.[22,23]

Surgeon preference for performance of IL in the OR versus the office for management of UVFP varies tremendously. A recent multi-institution analysis reported equal numbers of IL performed in the OR and in the office.[6] Recent reports of UVFP management show IL performed entirely in the office[24,25] and entirely in the OR.[26] Rationale beyond surgeon preference drives the decision of where to perform IL, including

TABLE 2.
Approach Used for OBIL

Approach	Number of Times (Percent of Total)
Transoral	71 (86.6)
Transcricothyroid membrane	8 (9.8)
Transthyrohyoid membrane	2 (2.4)
Transthyroid ala	1 (1.2)

TABLE 3.
Injectate Used During OBIL

Injectate	Number of Times Used (Percent of Total)
Hyaluronic acid (Hylaform, Allergan-Inamed Crop, Irvine, CA)	33 (40.2)
Calicium hydroxyapetite (Radiesse Voice, BioForm Medical, San Mateo, CA)	20 (24.4)
Micronized dermis (Cymetra, LifeCell Corp, Branchburgh, NJ)	14 (17.1)
Hyaluronic acid gel (Juvederm Ultra Plus, Allergan, Santa Barbara, CA)	8 (9.8)
Hyaluronic acid (Restylane, Q Med, Uppsala, Sweden)	6 (7.3)
Teflon	1 (1.2)

access to resources. In this series, all patients were treated in office. One reason for this is the fact that University of Wisconsin Clinics is a hospital-based practice in which injectables may be billed to the insurance. In a stand-alone clinic, patients are responsible for cost of the injectate, which causes many to elect for procedures in the OR. Additionally, the office laryngeal surgery suite is located within the hospital building, allowing both inpatients and outpatients to be examined and treated using the same setup.

The average age of patients treated in this series was 60 years which is similar to other reports.[25,27] The left vocal fold was affected more often, which is also consistent with large studies.[8] The most common etiology of paralysis was thoracic which included injury to the recurrent laryngeal nerve (RLN) from mass effect of benign and malignant disease or complications after chest surgery. All patients in this series were able to be injected to the intended completion point using a transoral, transcricothyroid membrane, transthyrohyoid membrane, or transthyroid ala approach. The transoral approach was preferred by the authors as it can be performed by one surgeon, without the need for an assistant. It also allows for the entirety of the needle to be visualized during the injection.

The average amount of injectate applied in this population was 0.64 mL. Mau and Courey[28] demonstrated that on average 0.62 and 0.41 mL of calcium hydroxyapatite were necessary to medialize a cadaveric vocal fold *via* a lateral injection. The increase may be a result of the overinejction necessary to account for reabsorption of injectate. Numerous injectates were used in this study, which were tolerated well by most patients. However, two complications noted in this study were related to the injectate used. The first was a hypersensitivity reaction to Restylane. A study of rabbit vocal folds injected with Restylane revealed that at 1 week and 3 months after injection, the vocal folds experienced "low fibrinogenesis," "a slight inflammatory reaction and absence of necrosis," and "granuloma formation and low fibrinogenesis."[29] However, within the Dermatology literature, injection site inflammation resulting in transient redness and edema of the injected site immediately after injection has been noted in 0.02% of individuals who underwent injection of hyaluronic acid gel for soft tissue augmentation.[30] Additionally, hypersensitivity and inflammatory reactions to hyaluronic acid gel have been noted after cutaneous injections for management of facial rhytids.[30–32] It is very possible that the patient treated in this series experienced a similar reaction in the vocal fold after injection.

The other complication resulted from an injection of calcium hydroxyapatite into the superficial lamina propria, requiring removal under general anesthesia during microlaryngoscopy. This was removed in a manner similar to techniques described by others.[33] Ensuring placement of the injectate into the correct portion of the larynx is paramount in OBIL.

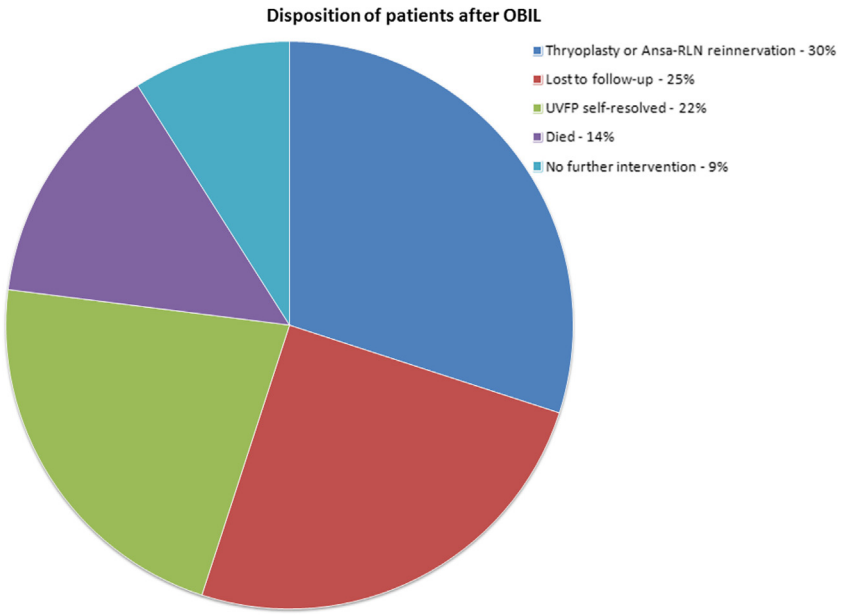

FIGURE 2. Disposition of patients after OBIL.

UVFP often results from malignancy, surgery, and sometimes both. As mentioned earlier, patients may also have general health concerns, cardiopulomonary compromise, anticoagulation needs, as well as psychosocial stressors. As such, the safety of any intervention for this patient population must critically be evaluated. These data, in combination with other data sets, confirm the notion that OBIL is a safe procedure for patients with UVFP.[24,25] One patient had a complication in which vocal fold edema was noted and that the procedure was terminated without incident. There were no complications requiring hospital admission. Patients who were on aspirin prophylactically to prevent cardiac events were asked to stop taking medication 1 week before injection. However, those patients who were taking anticoagulants for therapeutic treatments did not stop taking medications for IL. No complications with hematoma or airway compromise occurred with this approach. For most patients who had injection performed transorally, IL was performed using a 27 gauge needle, in which little, if any, bleeding was noted even if patients were anticoagulated. For this reason, it was deemed safe to continue blood thinners for patients in whom it was medically necessary and do not report any complications with this approach. Others have also shown that procedures performed while a patient is taking anticoagulants are safe.[3,34]

There are risks associated with general anesthesia, which is one of the major motivators to performing office-based laryngeal surgery. Graboyes et al[26] recently published their experience with IL performed under general anesthesia for patients with UVFP after thoracic surgery. Although the majority of their patients did quite well, one of the 20 patients did have intraoperative bile reflux on induction of anesthesia resulting in pneumonitis that may have been avoided with OBIL.

The disposition of patients after injection is shown in Figure 1. Thirty percent of patients sought a definitive intervention in the form of thyroplasty or ansa cervicalis-RLN reinnervation. These results are similar to a study performed by Arviso et al,[16] in which 29% of patients who underwent IL (in the OR or the office) for UVFP required further definitive intervention with medialization thryoplasty. Sixteen percent of the patients treated by Damrose[25] for UVFP required thryoplasty and/or arytenoid adduction after OBIL.

There are multiple reasons why this may have occurred. The concept of laryngeal synkinesis describes abnormal reinnervation of the laryngeal muscles after injury to the RLN.[35–37] After deinnervation of the vocal fold, regeneration of RLN motor axons place the vocal fold in either a favorable or unfavorable position.[37] It has been posited that early medialization of the vocal fold with IL places the vocal fold in a favorable position that is maintained by synkinetic reinnervation.[15] Another consideration is that fibrosis and scarring secondary to IL assist in placing the vocal fold in a permanent medial position.[14,38] Perhaps due to a combination of these reasons, only 30% of the patients in this study required definitive treatment.

In the present study, 22% of individuals had a documented return of function and normal voice noted during stroboscopic examination of the larynx. Fourteen percent of the individuals died, and 9% returned to the office, were noted not to have full recovery of vocal fold motion, and opted for no further intervention. One-quarter of patients did not follow-up. Although this is a sizable number, it is similar to the results of other retrospective studies.[13–16,39] One reason for this is likely due to the large draw of the University of Wisconsin where patients may choose to follow-up with a local otolaryngologist or primary care physician. Some of these patients may have had return of normal or near normal voicing and not found a reason to follow-up. Sulica[40] noted that in idiopathic vocal fold paralysis, which was the second most common reason for UVFP in this series, 52% ± 17% of individuals affected regained complete recovery of voice.

There are limitations to this study which should be recognized. All patients were treated by a single-physician and the data were analyzed in a retrospective fashion. Outcome measures were not obtained in this study, so it is not possible to examine how effective OBIL is. However, other studies have demonstrated improvements in voice quality, swallowing ability, and voice-related quality-of-life after OBIL.[13,25]

From these data, further questions remain to be answered. Multiple injectates were used and it would be interesting to determine which of these is the most durable. The reasons for patients not opting for a more definitive surgery would also be helpful to know.

CONCLUSIONS

OBIL is a safe procedure that is well tolerated in the management of UVFP. Multiple injectates may be used, and familiarity with multiple approaches is beneficial to be able to treat the most number of individuals in the office setting. As noted in this and other studies, a minority of patients who undergo IL require laryngeal framework surgery or a reinnervation procedure.

REFERENCES

1. Bruening W. Uber eine neue Behandlungsmethode der Rekurrenslahmung. *Verhandl Ver Dtsch Laryngol*. 1911;18:93–151. [In German].
2. Bastian RW, Collins SL, Kaniff T, Matz GJ. Indirect videolaryngoscopy versus direct endoscopy for larynx and pharynx cancer staging. Toward elimination of preliminary direct laryngoscopy. *Ann Otol Rhinol Laryngol*. 1989;98:693–698.
3. Bastian RW, Delsupehe KG. Indirect larynx and pharynx surgery: a replacement for direct laryngoscopy. *Laryngoscope*. 1996;106:1280–1286.
4. Rosen CA, Thekdi AA. Vocal fold augmentation with injectable calcium hydroxylapatite: short-term results. *J Voice*. 2004;18:387–391.
5. Ford CN, Bless DM, Loftus JM. Role of injectable collagen in the treatment of glottic insufficiency: a study of 119 patients. *Ann Otol Rhinol Laryngol*. 1992;101:237–247.
6. Sulica L, Rosen CA, Postma GN, et al. Current practice in injection augmentation of the vocal folds: indications, treatment principles, techniques, and complications. *Laryngoscope*. 2010;120:319–325.
7. Rubin HJ. Dysphonia due to unilateral nerve paralysis. Treatment by the intracordal injection of synthetics—a preliminary report. *Calif Med*. 1965; 102:105–109.
8. Rosenthal LH, Benninger MS, Deeb RH. Vocal fold immobility: a longitudinal analysis of etiology over 20 years. *Laryngoscope*. 2007;117: 1864–1870.
9. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973-2002. *JAMA*. 2006;295:2164–2167.

10. Patil PG, Turner DA, Pietrobon R. National trends in surgical procedures for degenerative cervical spine disease: 1990-2000. *Neurosurgery.* 2005; 57:753–758. discussion 753–758.
11. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. *N Engl J Med*. 2011;364: 2128–2137.
12. Yung KC, Courey MS. The Effect of office-based flexible endoscopic surgery on hemodynamic stability. *Laryngoscope*. 2010;120:2231–2236.
13. Mathison CC, Villari CR, Klein AM, Johns MM 3rd. Comparison of outcomes and complications between awake and asleep injection laryngoplasty: a case-control study. *Laryngoscope*. 2009;119:1417–1423.
14. Yung KC, Likhterov I, Courey MS. Effect of temporary vocal fold injection medialization on the rate of permanent medialization laryngoplasty in unilateral vocal fold paralysis patients. *Laryngoscope*. 2011;121: 2191–2194.
15. Friedman AD, Burns JA, Heaton JT, Zeitels SM. Early versus late injection medialization for unilateral vocal cord paralysis. *Laryngoscope*. 2010;120: 2042–2046.
16. Arviso LC, Johns MM 3rd, Mathison CC, Klein AM. Long-term outcomes of injection laryngoplasty in patients with potentially recoverable vocal fold paralysis. *Laryngoscope*. 2010;120:2237–2240.
17. Ford CN, Roy N, Sandage M, Bless DM. Rigid endoscopy for monitoring indirect vocal fold injection. *Laryngoscope*. 1998;108:1584–1586.
18. Ward PH, Hanson DG, Abemayor E. Transcutaneous Teflon injection of the paralyzed vocal cord: a new technique. *Laryngoscope*. 1985;95: 644–649.
19. Amin MR. Thyrohyoid approach for vocal fold augmentation. *Ann Otol Rhinol Laryngol*. 2006;115:699–702.
20. Dedo HH, Urrea RD, Lawson L. Intracordal injection of Teflon in the treatment of 135 patients with dysphonia. *Ann Otol Rhinol Laryngol*. 1973;82: 661–667.
21. Grant JR, Hartemink DA, Patel N, Merati AL. Acute and subacute awake injection laryngoplasty for thoracic surgery patients. *J Voice*. 2008;22: 245–250.
22. Bove MJ, Jabbour N, Krishna P, et al. Operating room versus office-based injection laryngoplasty: a comparative analysis of reimbursement. *Laryngoscope*. 2007;117:226–230.
23. Andrade Filho PA, Carrau RL, Buckmire RA. Safety and cost-effectiveness of intra-office flexible videolaryngoscopy with transoral vocal fold injection in dysphagic patients. *Am J Otolaryngol*. 2006;27: 319–322.
24. Lee SW, Kim JW, Chung CH, et al. Utility of injection laryngoplasty in the management of post-thyroidectomy vocal cord paralysis. *Thyroid*. 2010;20: 513–517.
25. Damrose EJ. Percutaneous injection laryngoplasty in the management of acute vocal fold paralysis. *Laryngoscope*. 2010;120:1582–1590.
26. Graboyes EM, Bradley JP, Meyers BF, Nussenbaum B. Efficacy and safety of acute injection laryngoplasty for vocal cord paralysis following thoracic surgery. *Laryngoscope*. 2011;121:2406–2410.
27. Bhattacharyya N, Kotz T, Shapiro J. Dysphagia and aspiration with unilateral vocal cord immobility: incidence, characterization, and response to surgical treatment. *Ann Otol Rhinol Laryngol*. 2002;111:672–679.
28. Mau T, Courey MS. Influence of gender and injection site on vocal fold augmentation. *Otolaryngol Head Neck Surg*. 2008;138:221–225.
29. Perazzo PS, Duprat Ade C, Lancellotti CL. Histological behavior of the vocal fold after hyaluronic acid injection. *J Voice*. 2009;23:95–98.
30. Friedman PM, Mafong EA, Kauvar AN, Geronemus RG. Safety data of injectable nonanimal stabilized hyaluronic acid gel for soft tissue augmentation. *Dermatol Surg*. 2002;28:491–494.
31. Lupton JR, Alster TS. Cutaneous hypersensitivity reaction to injectable hyaluronic acid gel. *Dermatol Surg*. 2000;26:135–137.
32. Soparkar CN, Patrinely JR. Managing inflammatory reaction to restylane. *Ophthal Plast Reconstr Surg*. 2005;21:151–153.
33. Chheda NN, Rosen CA, Belafsky PC, Simpson CB, Postma GN. Revision laryngeal surgery for the suboptimal injection of calcium hydroxylapatite. *Laryngoscope*. 2008;118:2260–2263.
34. Luu Q, Tsai V, Mangunta V, Berke GS, Chhetri DK. Safety of percutaneous injection of bovine dermal crosslinked collagen for glottic insufficiency. *Otolaryngol Head Neck Surg*. 2007;136:445–449.
35. Crumley RL. Experiments in laryngeal reinnervation. *Laryngoscope*. 1982; 92(9 Pt 2 suppl 30):1–27.
36. Crumley RL. Laryngeal synkinesis: its significance to the laryngologist. *Ann Otol Rhinol Laryngol*. 1989;98:87–92.
37. Crumley RL. Laryngeal synkinesis revisited. *Ann Otol Rhinol Laryngol*. 2000;109:365–371.
38. Carroll TL, Rosen CA. Long-term results of calcium hydroxylapatite for vocal fold augmentation. *Laryngoscope*. 2011;121:313–319.
39. Tan M, Woo P. Injection laryngoplasty with micronized dermis: a 10-year experience with 381 injections in 344 patients. *Laryngoscope*. 2010;120: 2460–2466.
40. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118:1303–1307.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(2):302-307.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–
Head and Neck Surgery
2015, Vol. 152(2) 302–307
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2014
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599814562727
http://otojournal.org
SAGE

Alterations in Extracellular Matrix Composition in the Aging Larynx

Anete Branco, SLP, PhD[1], Alexandre Todorovic Fabro, MD, PhD[2], Tatiana Maria Gonçalves, MD[1], and Regina Helena Garcia Martins, MD, PhD[1]

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract

Objective. To study by immunohistochemistry the alterations of collagens I, III, IV, and V and elastin in the aging process of the human larynx.

Study Design. Cadaver study.

Setting. Universidade Estadual Paulista, Botucatu Medical School, São Paulo State University (UNESP), Brazil.

Subjects and Methods. Thirty vocal folds were obtained at autopsy from 10 adult men (aged 30 to 50 years) and 20 geriatric men (10 aged 60 to 75 years and 10 aged >75 years). Mid membranous vocal fold slides were subjected to immunohistochemical reactions. Digital imaging software (ImageJ) was used to quantify the increase in brownish staining of the lamina propria structures of vocal folds, from superficial to deep layers.

Results. There was an increase of collagen I and III immunoexpression in the elderly larynges, in both layers. Collagens IV and V were immunoexpressed in the vessels endothelium of the lamina propria and in the basement membrane. The immunoexpression of elastin decreased in the elderly larynges, in both lamina propria layers of the vocal folds.

Conclusion. A clear increase of collagens I and III and a decrease of elastic fibers were observed in the lamina propria of vocal folds. The concentration of collagens IV and V was the same across age groups. These findings suggest that as men age, the density of the extracellular matrix increases, brought about by an increase in collagen, while the loss of elastin results in decreased viscoelasticity.

Keywords

collagen, elastin, aged, lamina propria, vocal folds, larynx, immunohistochemistry

Received May 15, 2014; revised October 28, 2014; accepted November 14, 2014.

The physiology of voice production is intimately connected with the characteristics of the vocal fold lamina propria,[1,2] the vocal fold being the main vibrating structure during phonation.[3] In the pursuit to understand the vocal properties typical of age groups, sexes, and diseases, researchers have gone deeper in their morphological studies, trying to elucidate the mechanisms involving the lamina propria microstructures of vocal folds, mainly relating to elastic, collagen, and protein fibers.

Collagen fibers are found across the whole lamina propria, from its superficial to its deepest layers,[4,5] supporting tissue structure through enhanced stability and resistance, relevant to vocal physiology. Gray et al[6] confirmed that collagen and elastic fibers are involved in vocal fold biomechanics, in which the former allow tension and stretching while the latter allow deformation and a quick return to the initial shape.

The composition of the vocal folds extracellular matrix varies significantly between sexes and age groups, implying biomechanical differences that directly influence the vocal properties.[7] The lamina propria trilaminar structure in the adult larynx described by Hirano[8] is not present in the newborn.[3,9,10] It starts to organize only in childhood, from 6 years of age onward. In aging people, anatomical and functional changes occur in the vocal tract, especially in the vocal folds. At this stage of life, structural and functional changes take place in the epithelial covering, the muscle, and the lamina propria of vocal folds, notably, the atrophy of epithelial layers and the vocal muscle, as well as density alterations in collagen and elastic fibers and in hyaluronic acid.[11-13]

Elastic fibers change morphologically in elderly people, relative to adults. Their remodeling with age is reflected by

[1]Ophthalmology, Otorhinolaryngology, Head and Neck Surgery Department, Botucatu Medical School, UNESP–Univ Estadual Paulista, São Paulo, Brazil
[2]Department of Pathology, Botucatu Medical School, UNESP–Univ Estadual Paulista, São Paulo, Brazil

Corresponding Author:
Regina Helena Garcia Martins, SLP, PhD, Botucatu Medical School/UNESP, Ophthalmology, Otorhinolaryngology, Head and Neck Surgery Department, Distrito de Rubião Jr, CEP 18618-970, Botucatu, São Paulo, Brazil.
Email: rmartins@fmb.unesp.br

parameters such as thickness, size, orientation, and location.[14] The ratio of collagen to elastic fibers also evolves, increasing with age, and the predominance of the former explains the rigidity and reduction of the vibratory mucosal wave. Such alterations, combined with loss of elastin and hyaluronic acid, further decrease the viscosity of vocal folds and negatively affect voice emission in the elderly population.[1,4,5,12,15]

Given the complexity and interdependence of vocal fold microstructures, one can expect them to be susceptible to organic changes with aging. Published studies focused more on collagen types I and III, while the rest were little investigated. Thus, the goal of our study was to analyze the concentration of collagen types I, III, IV, and V and of elastin in human larynx senescence by immunohistochemistry.

Methods

The research project was approved by the Ethics Committee for Research on Human Beings of Botucatu Medical School/UNESP, São Paulo, Brazil (reference 3861/2011).

Twenty larynges from elderly men (aged 60-90 years) were included, as well as 10 from male adult controls (aged 30-50 years), collected at autopsy. The elderly group was subdivided into 2 age subgroups: 60 to 75 years (n = 10) and greater than 75 years (n = 10).

The cause of death was retrieved from the autopsy records, and when necessary, additional information was obtained from medical records and family. The exclusion criteria were septicemia; prolonged intubation; systemic infections; persistent dermatologic, autoimmune, or metabolic diseases; smoking habit; and cervical trauma of any kind or other conditions that might compromise the mucosa of vocal folds and invalidate the immunohistochemical analysis.

Fresh larynges were incised at their posterior portion and examined macroscopically to exclude lesions. As a standard procedure, the middle part of the right vocal fold was dissected (**Figure 1**) and immediately embedded in 10% buffered formaldehyde for 24 to 48 hours. Paraffin blocks were prepared for histological slides (hematoxylin and eosin [H&E] stain) and immunohistochemical reactions. The left vocal fold was used in another study.

The following antibodies were used: collagen I (Col1A1, 1:2000 dilution; Dako, Carpinteria, California), collagen III (Col 3A1, dilution 1:1000; Dako), collagen IV (Col4A, 1:40 dilution; Santa Cruz Biotechnology, Santa Cruz, California), collagen V (Col 5A1, 1:100 dilution; Santa Cruz Biotechnology), and elastin (1:200 dilution, Dako). The antigen preparation for immunohistochemical reactions of collagens I, III, and V and of elastin was performed with 1% pepsin at pH 1.8 and incubated for 15 minutes at 60°C and then 30 minutes at 37°C. Blocking was performed with 8% methanol and hydrogen peroxide, followed by 3% Molico milk (Nestlé, Vevey, Switzerland). The secondary complex HRP EnVision (Dako) was added for 1 hour. Diaminobenzidine was added for 5 minutes, and Harris hematoxylin was added for 20 seconds. For collagen IV antigenic sites retrieval, we used a Pascal (Dako) pressure chamber for 3 minutes in the pretreatment solution Trilogy (Cell Marque, Rocklin, California). Following the protocol recommended by our laboratory, the positive control used for collagen I, III, and V and elastin was a fragment of kidney and, for negative control of these same antibodies, only buffered saline, replacing the primary antibody in a series of sections each sampling.

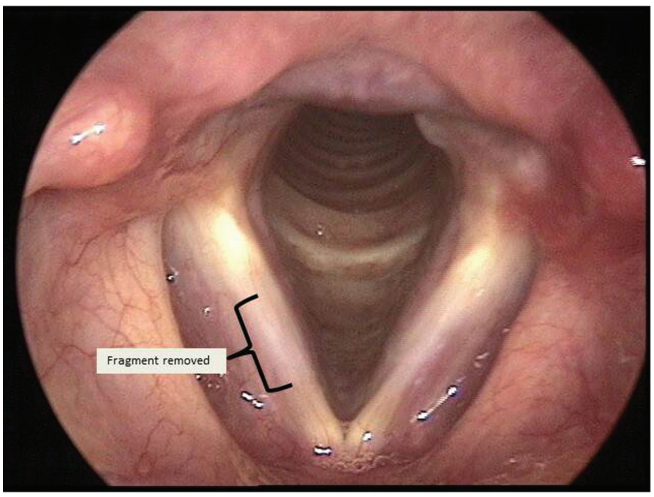

Figure 1. Vocal fold showing the site of the fragment dissection.

Protein Expression Analysis

Reading of histological slides was performed by 3 authors, blindly and randomly. The slides were evaluated at random, without knowledge of the groups and the age of patients, using a light microscope from Zeiss (Axiostar plus, Carl Zeiss do Brasil Ltda), at 40×.

The analyzed locations were the basement membrane, the endothelium vessels, and the lamina propria (superficial and deep layers). Because of the imprecise nature of the boundaries surrounding the intermediate layer of the lamina propria, we decided to divide the lamina propria into 2 layers (superficial and deep) to facilitate presentation of the results. The size of the lamina propria was measured between the basal membrane and muscle fibers. The thickness of the lamina propria was divided into 2 portions subjectively as the superficial layer, corresponding to the upper portion and positioned just below the basement membrane, and the deep layer, corresponding to the lower portion and positioned just above the muscle fibers.

Measurement of antibodies involved quantification of the level of brown staining of the lamina propria and basement membrane structures. The area was set at 2 μm, with percentage used to report results.[16,17] ImageJ software was used in this analysis. We set the polygon as the selection tool for the area of interest, choosing 3 fields at random, and we set the Color Deconvolution plugins to HDAB and Make Binary.

Statistical Analysis

To compare age groups, considering the response profile assessed in 2 layers of the vocal fold (superficial and deep layers), we used a parametric variance analysis combined

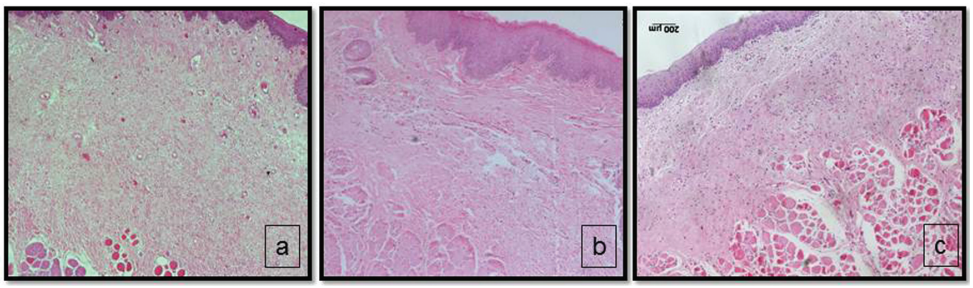

Figure 2. (A) Control group: collagen fibers in the lamina propria, some fibroblasts and vessels. (B, C) Elderly groups: dense collagen in the lamina propria and thickened basement membrane. Hematoxylin and eosin, 10×.

Table 1. Immunoexpression of Antibodies for Collagens I, III, and Elastin, Measured as Area Percentage (%).[a]

	Antibodies								
	Collagen I			Collagen III			Elastin		
Age Group, y	SL	DL	*P* Layers	SL	DL	*P* Layers	SL	DL	*P* Layers
30-50	34.2 (8.5)	3.8 (10.4)	.105	25.6 (21.6-29.4)	23.6 (20.5-29.9)	.046	28.2 (25.7-29.3)	27.5 (23.6- 33.3)	.508
60-75	39.6 (5.1)	44.4 (7.5)	.114	31.2 (23.7-35.6)	24.6 (20.1-29.7)	.0051	20.8 (19.2-23.7)	23.0 (18.7-28.2)	.114
>75	40.4 (5.7)	46.9 (9.3)	.114	32.4 (24.8-44.4)	29.9 (22.1-40.1)	.333	15.9 (10.8-24.3)	18.2 (14.3-25.9)	.285
P ages	.002*			.0061*			.00001*		

[a]Mean (standard deviation) or median (minimum and maximum) according to age group and depth of vocal folds lamina propria. DL, deep layer; SL, superficial layer.
**P* with statistical significance.

with the respective multiple comparisons tests and the nonparametric model when necessary (median, minima, and maxima values). Means and standard deviations of each group were presented when the parametric model was used. For the nonparametric model, medians and their minimum and maximum values were presented, considering the 5% significance level ($P < .05$) within each age range.

Results

The starting point in our study was to analyze H&E slides to check the syntopy of the lamina propria structures in vocal folds and then perform the immunohistochemical study. In these images (**Figure 2**), it was possible to observe in the control group adults that collagen fibers are arranged as a loose fabric between some fibroblasts and vessels in the entire vocal fold lamina propria. In the 60- to 75-year-old group subjects, it was possible to identify a dense collagen thickness deposited on the subepithelial layer (**Figure 2B**). Finally, in the elderly group (greater than 75 years of age), dense collagen uniformly occupied the whole lamina propria (**Figure 2C**).

The immunoexpression of antibodies for collagens I, III, and elastin as well as collagens IV and V is presented in **Tables 1** and **2**, respectively, and shown in **Figures 3** to **7**. We observed that, for collagen I and III, there was an increase of labeling density in the elderly larynges, without any statistical difference between elderly subgroups. Regarding collagens IV and V, a greater labeling density was observed in the endothelium of vessels of the lamina propria and basement membrane. Finally, regarding elastin, we observed a decrease of this antibody as the age increased in both elderly groups.

Table 2. Immunoexpression of Antibodies for Collagens IV and V, Measured as Area Percentage (%).[a]

	Collagen IV		Collagen V	
Age Group, y	SL	DL	SL	DL
30-50	20.9 (2.9)	17.6 (3.7)	38.5 (14.4)	36.6 (8.5)
60-75	20.7 (4.4)	19.3 (4.2)	36.2 (8.9)	35.0 (5.6)
>75	21.3 (3.9)	20.4 (3.5)	31.8 (5.6)	31.3 (6.0)
P ages	.573		.2597	
P Layers	.0034		.397	

[a]Mean (standard deviation) according to age group and depth of vocal folds lamina propria. DL, deep layer; SL, superficial layer.
**P* with statistical significance.

Discussion

Presbyphonia is a physiological effect of the senescence process that takes place in the whole body. Trying to

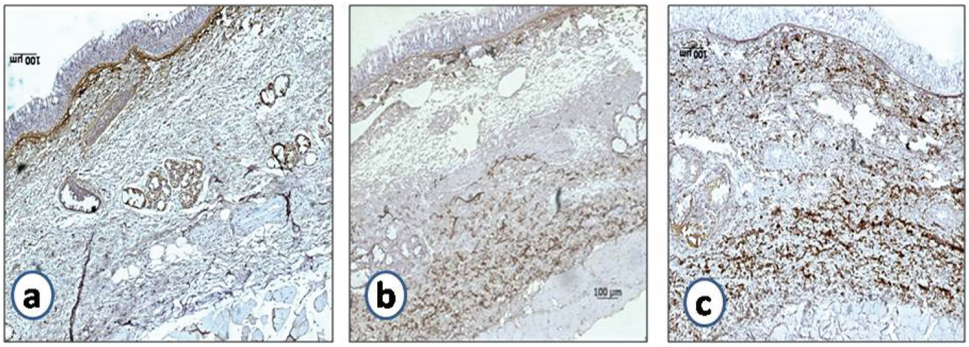

Figure 3. Vocal folds. (A) Control group: collagen I staining the basement membrane and the superficial layer. (B, C) Elderly groups: dense collagen I staining the superficial and deep layers of the lamina propria. Immunohistochemical reaction, 20×.

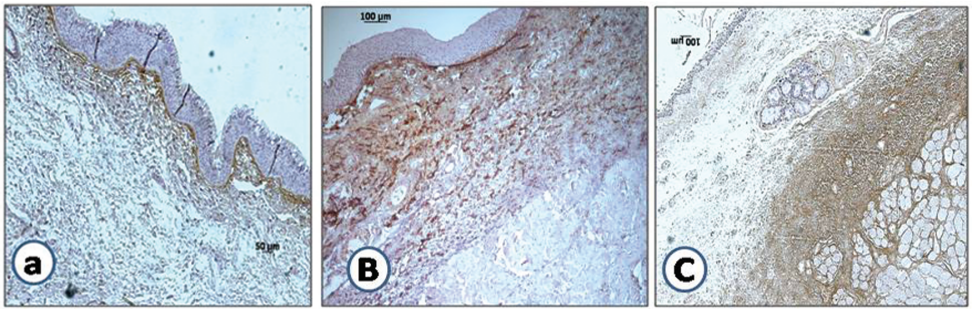

Figure 4. Vocal folds. (A) Control group: collagen III staining the basement membrane and the superficial layer. (B, C) Elderly groups: collagen III staining the superficial (B) and deep layers (C) of the lamina propria. Immunohistochemical reaction, 20×.

correlate the alterations in the vocal folds of the elderly population with the rest of the body's epithelial covering, Ximenes Filho et al[18] performed the simultaneous histomorphometric analysis of vocal folds and inguinal skin from 20 elderly cadavers (10 male and 10 female), finding similar alterations in both locations such as lamina propria and epithelial atrophy.

The immunohistochemical staining revealed that this network is mainly formed by collagen I and III (**Table 1**; **Figures 3** and **4**) with a significant decreased in density of elastin in the elderly larynx (**Table 1**; **Figure 7**). These structural changes in the vocal fold cover are responsible for the hardening of the vocal folds,[6,19,20] clinically manifested by symptoms of hoarseness, vocal fatigue, and vocal range restriction, having a direct impact on speech in higher frequencies.[21,22]

Keeping the proportion of elastic and collagen fibers in the lamina propria of the larynx is important for this organ to retain local resistance, provided by collagen fibers, and distensibility, given by the elastic fibers. Studies of the larynx in animals[23] and in humans[24] have confirmed the role of collagen I in resistance and that of collagen III and elastin in flexibility and elasticity. These parameters work as a ''balance,'' whose equilibrium determines the relative participation of the different vocal fold layers in phonation. When collagen fibers start to predominate in the lamina propria, as seen in our study, the vocal folds become more rigid, which negatively affects voice qualities.[19] According to Ohno et al,[25] the collagen increase in the elderly larynx affects the mucosal wave, resulting in decreases of phonatory intensity and fundamental frequency, especially in women. The voice becomes failing and weak, a vocal pattern known as phonasthenia.[26]

In an attempt to better understand what causes the abnormal production of collagen in the elderly larynx, Kosztyła-Hojna et al[27] studied the ultrastructure of vocal folds retrieved at elderly autopsies or from total laryngectomies due to supraglottic larynx carcinomas, without damage to the actual vocal folds. Using transmission electronic microscopy, the authors observed the destruction of epithelial cells, a vacuolar degeneration of the cell's cytoplasm, a considerable increase of collagen fibers, a vacuolar degeneration of fibroblasts, an increase of the endoplasmic reticulum, and an increase of blood vessels. These authors suggest that the increase of collagen fibers is connected to the cytoplasmatic alterations observed in fibroblasts. This possibility is insufficiently supported by the literature and therefore requires additional studies.

Although the mechanisms are not yet understood, the elderly larynx fibroblasts produce collagen in an excessive and irregular way, as well as a lower amount of hyaluronic acid and elastic fibers. In a study of young and elderly rat larynx fibroblast cultures, Hirano et al[12] observed in the latter lower amounts of hyaluronic acid and higher concentrations of collagen I. They observed the opposite when the fibroblasts were exposed to fibroblast growth factor, which was thus considered by the authors a potential therapeutic tool for

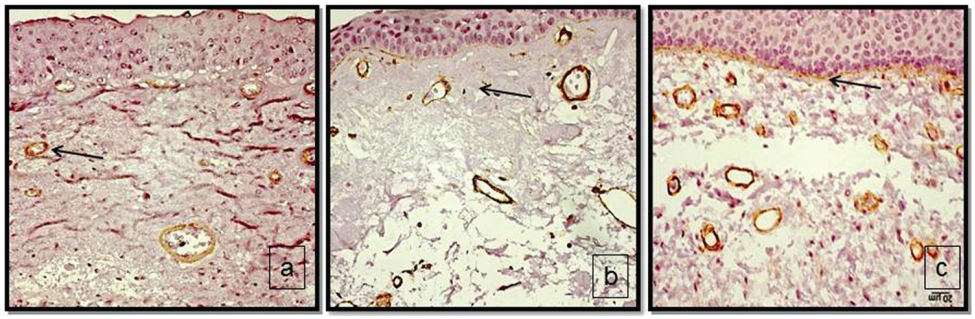

Figure 5. Vocal folds. (A) Control group: collagen IV staining the vessel endothelium (arrow). (B, C) Elderly groups: collagen IV staining the vessel endothelium and the basement membrane (arrows). Immunohistochemical reaction, 20×.

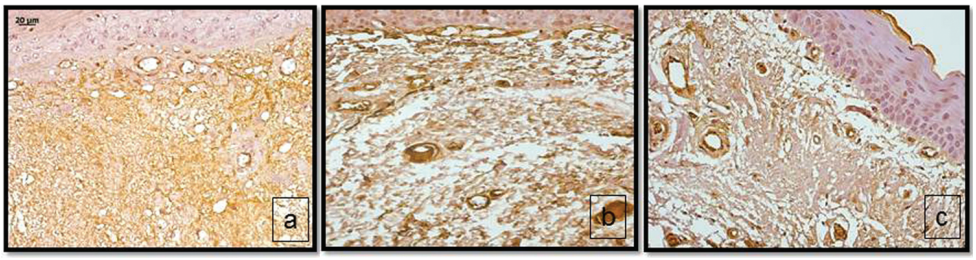

Figure 6. Vocal folds. (A) Control group: collagen V staining the vessel endothelium (arrow). (B, C) Elderly groups: collagen V staining the vessel endothelium and the basement membrane (arrows). Immunohistochemical reaction, 20×.

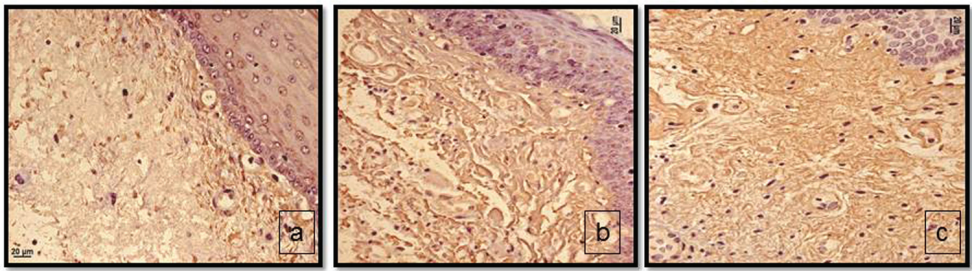

Figure 7. Vocal folds. (A) Control group: uniform elastin antibody staining the lamina propria. (B, C) Elderly groups: sparse distribution of elastin in the lamina propria. Imunohistochemical reaction, 20×.

lamina propria collagen modulation, opening new perspectives for presbyphonia and vocal fold atrophy treatments.[25,28]

The basement membrane and the vessels endothelium having a structural function, they also contain collagen types IV and V, which contribute to support the lamina propria.[29] As in our results, it was observed that these antibodies, when quantified in said locations, also increased with age (**Table 2**). Similar results have been reported in larynges with chronic inflammation in morphological studies.[30] The increase of said collagens is probably related to the simultaneous increase of blood vessels, which is observed with both persistent inflammation and aging.

Elastic fibers in the elderly larynx not only are found in lower amounts but also present structural alterations. Sato and Hirano[15] observed that with aging, these fibers lose elasticity because of an increase in amorphous substance, a decrease in microfibrils, and their metabolic alteration. Thus, we can infer that synchronicity of the phonatory system can depend on biomechanic changes, due to the physiological remodeling of the extracellular matrix.[31]

We can therefore stress that the vibration mechanism of vocal folds is governed by the laryngeal tissue biomechanics. The observed alterations in the elderly larynx lamina propria components are interpreted in the endoscopic examinations as vocal fold atrophy and spindle chink, giving the voice different degrees of hoarseness, asthenia, and breathiness.[32] These ''new'' micro- and macroanatomic configurations present in the elderly larynx may mirror age-dependent physiological remodeling; that is, the structural changes in the vocal folds lamina propria that occur with aging make it necessary for the extracellular matrix and its components to be remodeled to remain functional.

Conclusions

In our methodological conditions, we observed in the vocal folds lamina propria a clear increase of collagens I and III,

as well as a decrease of elastic fibers. The concentration of collagens IV and V did not change according to age group. These findings suggest that as men age, the density of the extracellular matrix increases, brought about by an increase in collagen, while the loss of elastin results in decreased viscoelasticity.

Author Contributions

Anete Branco, interpretation of data for the work, drafting the work, final approval of the work, agreement to be accountable for all aspects of the work; **Alexandre Todorovic Fabro**, interpretation of data for the work, drafting the work, final approval of the work, agreement to be accountable for all aspects of the work, data analysis; **Tatiana Maria Gonçalves**, interpretation of data for the work, drafting the work, final approval of the work, agreement to be accountable for all aspects of the work; **Regina Helena Garcia Martins**, study design, interpretation of the data, writing and approval of the manuscript, critical revision, final approval, agreement to be accountable for all aspects of the work.

Disclosures

Competing interests: None.

Sponsorships: Fapesp and Fundunesp, financial support to immunohistochemical analysis and English version.

Funding source: None

References

1. Butler JE, Hammond TH, Gray SD. Gender-related differences of hyaluronic acid distribution in the human vocal fold. *Laryngoscope*. 2001;111:907-911.
2. Ramos HV, Simões MJ, Pontes PA, et al. Immunohistochemistry as a method to study elastic fibers of human vocal fold. *Braz J Otorhinolaryngol*. 2005;71:486-491.
3. Ishii K, Yamashita K, Akita M, Hirose H. Age-related development of the arrangement of connective tissue fibers in the lamina propria of the human vocal fold. *Ann Otol Rhinol Laryngol*. 2000;109:1055-1064.
4. Hammond T, Gray S, Butler J. Age and gender related collagen distributions in human vocal folds. *Ann Otol Rhinol Laryngol*. 2000;109:913-920.
5. Sato K, Hirano M, Nakashima T. Age-related changes of collagenous. *Ann Otol Rhinol Laryngol*. 2002;111:15-20.
6. Gray SD, Titze IR, Alipour F, Hammond TH. Biomechanical and histologic observations of vocal fold fibrous proteins. *Ann Otol Rhinol Laryngol*. 2000;109:77-85.
7. Kutty JK, Webb K. Tissue engineering therapies for the vocal fold lamina propria. *Tissue Eng Part B Rev*. 2009;15:249-262.
8. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr*. 1974;26:89-94.
9. Hirano M, Kurita S, Nakashima T. Growth, development and aging of human vocal folds. In: Bless DM, Abbs JH, eds. *Vocal Fold Physiology*. San Diego, CA: College-Hill; 1983:22-43.
10. Sato K, Hirano M, Nakashima T. Fine structure of the human newborn and infant vocal fold mucosae. *Ann Otol Rhinol Laryngol*. 2001;110:417-424.
11. Gray SD. Cellular physiology of the vocal folds. *Otolaryngol Clin North Am*. 2000;33:679-697.12
12. Hirano S, Bless DM, Río AM, Connor NP, Ford CN. Therapeutic potential of growth factors for aging voice. *Laryngoscope*. 2004;114:2161-2167.
13. Pontes P, Yamasaki R, Behlau M. Morphological and functional aspects of the senile larynx. *Folia Phoniatr Logop*. 2006;58:151-158.
14. Hammond TH, Gray SD, Butler J, Zhou R, Hammond E. Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngol Head Neck Surg*. 1998;119:314-322.
15. Sato K, Hirano M. Age-related changes of elastic fibers in the superficial layer of the lamina propria of vocal volds. *Ann Otol Rhinol Laryngol*. 1997;106:44-48.
16. Collins TJ. ImageJ for microscopy. *BioTechniques*. 2007;43: S25-S30.
17. Miot HA, Brianezi G. Morphometric analysis of dermal collagen by color clusters segmentation. *An Bras Dermatol*. 2010; 85:361-364.
18. Ximenes Filho JA, Tsuji DH, Nascimento PHS, Sennes LU. Histologic changes in human vocal folds correlated with aging: a histomorphometric study. *Ann Otol Rhinol Laryngol*. 2003;112:894-898.
19. Kendall K. Presbyphonia: a review. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:137-140.
20. Paulsen F, Kimpel M, Lockemann U, Tillmann B. Effects of ageing on the insertion zones of the human vocal fold. *J Anat*. 2000;196:41-54.
21. Gorhan-Rowan MM, Laures-Gore J. Acoustic-perceptual correlates of voice quality in elderly men and women. *J Commun Disord*. 2006;39:171-184.
22. Gregory ND, Chandran S, Lurie D, Sataloff RT. Voice disorders in the elderly. *J Voice*. 2012;26:254-258.
23. Kolachala VL, Torres-Gonzalez E, Mwangi S, et al. A senescence accelerated mouse model to study aging in the larynx. *Otolaryngol Head Neck Surg*. 2010;142:879-885.
24. Tateya T, Tateya I, Bless DM. Collagen subtypes in human vocal folds. *Ann Otol Rhinol Laryngol*. 2006;116:469-476.
25. Ohno T, Hirano S, Rousseau B. Age-associated changes in the expression and deposition of vocal fold collagen and hyaluronan. *Ann Otol Rhinol Laryngol*. 2009;118:735-741.
26. Behlau M, Azevedo R, Pontes P. Conceito de voz normal e classificação das disfonias. In: Behlau M, org. *Voz: o livro do especialista*. Rio de Janeiro, Brazil: Revinter; 2001:53-79.
27. Kosztyła-Hojna B, Andrzejewska A, Rogowski M, Pepinski W. Evaluation of the ultrastructure of the vocal folds mucosa in patients with presbyphonia. *Otolaryngol Pol*. 2002;56:239-243.
28. Ford C. Voice restoration in presbyphonia. *Arch Otolaryngol Head Neck Surg*. 2004;130:1117.
29. Prades JM, Dumollard JM, Duband S, et al. Lamina propria of the human vocal fold: histomorphometric study of collagen fibers. *Surg Radiol Anat*. 2010;32:377-382.
30. Baltaziak M, Kosztyła-Hojna B, Moskal D, Falkowski D. Morphological conditions of dysphonia in patients with vocal folds edema. *Pol Merkur Lekarski*. 2010;29:181-186.
31. Roberts T, Morton R, Al-Ali S. Microstructure of the vocal fold in elderly humans. *Clin Anat*. 2011;24:544-551.
32. Martins RHG, Gonçalvez TM, Pessin AB, Branco A. Aging voice: presbyphonia. *Aging Clin Exp Res*. 2014;26:1-5.

Reprinted by permission of JAMA Otolaryngol Head Neck Surg. 2013; 139(12):1291-1295.

Original Investigation

Influence of Age on Treatment With Proton Pump Inhibitors in Patients With Laryngopharyngeal Reflux Disease
A Prospective Multicenter Study

Young Chan Lee, MD; Jun Seok Lee, MD; Seung Woo Kim, MD; Kee Hwan Kwon, MD, PhD; Young Gyu Eun, MD, PhD

IMPORTANCE Several trials on the predictors of response to proton pump inhibitor (PPI) treatment of laryngopharyngeal reflux (LPR) have shown conflicting results. Furthermore, the influence of age in disease severity and response to PPI therapy is unclear.

OBJECTIVE To assess the difference in disease severity and response to PPI therapy according to age in patients with LPR.

DESIGN, SETTING, AND PARTICIPANTS Prospective multicenter study at 3 tertiary medical centers of 264 consecutive patients with LPR who were referred to the otolaryngology clinic from November 2010 to February 2012.

INTERVENTIONS Participants were prescribed 15 mg of lansoprazole (PPI) twice daily for 3 months.

MAIN OUTCOMES AND MEASURES Reflux Symptom Index (RSI), Reflux Finding Score (RFS), and laryngopharyngeal reflux–health-related quality of life (LPR-HRQOL) were collected at baseline and at 1 and 3 months postbaseline.

RESULTS After 3 months, 35 patients were lost to follow-up and excluded; the remaining 229 patients included 135 men and 94 women. The oldest group (60-79 years; n = 111) showed higher baseline RSI ($P < .001$) and LPR-HRQOL ($P < .001$) scores than the 18- to 39-year-old (n = 35) and 40- to 59-year-old (n = 83) groups. However, baseline RFS scores showed no significant difference among age groups ($P = .44$). Within each age group, the RSI, RFS, and LPR-HRQOL improved significantly with PPI therapy (all $P < .001$); however, no significant difference in improvement of RSI ($P = .59$), RFS ($P = .50$), or LPR-HRQOL ($P = .09$) was seen among the groups. At 3-month follow-up, significantly more responders, defined as those whose RSI score improved by more than 50%, were found in the 18- to 39-year-old and 40- to 59-year-old groups (86% and 75%, respectively) than in the oldest group (57%) ($P = .002$), but there was no significant difference in proportion of responders among age groups at 1-month follow-up ($P = .69$).

CONCLUSIONS AND RELEVANCE In patients with LPR, age seems to affect the subjective symptoms and resulting impact on quality of life but not the laryngeal findings. Furthermore, older patients are more likely not to respond to PPI therapy than younger patients.

Author Affiliations: Department of Otolaryngology–Head and Neck Surgery, Kyung Hee University School of Medicine, Seoul, Korea (Y.C. Lee, J.S. Lee, Eun); Department of Otolaryngology–Head and Neck Surgery, Seoul Veterans Hospital, Seoul, Korea (Kim); Department of Otorhinolaryngology–Head and Neck Surgery, Kangdong Sacred Heart Hospital, Hallym University School of Medicine, Seoul, Korea (Kwon).

Corresponding Author: Young Gyu Eun, MD, PhD, Department of Otolaryngology–Head and Neck Surgery, Kyung Hee University School of Medicine, 1 Hoegi-dong, dongdaemun-gu, Seoul 130-702, Korea (ygeun@hanmail.net).

JAMA Otolaryngol Head Neck Surg. 2013;139(12):1291-1295. doi:10.1001/jamaoto.2013.5556
Published online November 7, 2013.

Copyright 2013 American Medical Association. All rights reserved.

Laryngopharyngeal reflux (LPR) refers to the retrograde flow of stomach contents into the throat and larynx, which leads to symptoms such as chronic dysphonia, throat clearing, cough, globus sensation, and sore throat.[1] In otolaryngologic practice, approximately 10% of patients presenting to outpatient clinics and more than 50% of patients with voice problems receive a diagnosis of LPR.[2] Laryngopharyngeal reflux is a gastrointestinal and otolaryngological condition related to but distinct from gastroesophageal reflux disease (GERD). The reflux of gastric contents is at the core of both LPR and GERD, but the mechanism and symptoms of the 2 disorders are different.[3] The laryngeal mucosa are vulnerable to exposure to acidic substances, so patients often present with laryngopharyngeal symptoms without heartburn or regurgitation.[4] Several studies have suggested that the frequency of GERD complications (esophagitis, Barrett esophagus, stricture) is higher in geriatric patients because of many potential aggravating factors[5-7]; however, the influence of age in LPR is unclear. Although a 3-month empirical trial of proton pump inhibitor (PPI) treatment is generally regarded as a cost-effective modality for the initial management of LPR, the difference in response to PPI therapy according to age is not established. Garrigues et al[8] suggested that response to therapy was associated with younger age and shorter duration of laryngeal symptoms, but the response could not consistently be predicted in patients with chronic posterior laryngitis.

Thus, we conducted a prospective cohort study to assess the difference in symptom severity according to age in patients with LPR through the Reflux Symptom Index (RSI),[9] Reflux Finding Score (RFS),[10] and laryngopharyngeal reflux-health-related quality of life (LPR-HRQOL)[11] score. Furthermore, we evaluated the influence of age on response to PPI therapy.

Methods

Subjects and Study Design

Patients with suspected LPR who were referred to 3 different otolaryngology clinics from November 2010 to February 2012 were assessed for eligibility for the study. All patients underwent otolaryngologic evaluation by 1 of us (S.W.K., K.H.K., Y.G.E.), including laryngoscopy and video strobolaryngoscopy. A diagnosis of LPR was made on the basis of the presence of at least 1 of the following symptoms: hoarseness, chronic cough, throat irritation, laryngospasm, chronic throat clearing, and dysphasia. Diagnosis of LPR was also based on confirmed signs such as erythema, vocal cord edema, subglottic edema, posterior pachydermia, laryngeal edema, ventricular obliteration, and thick endolaryngeal mucus and granuloma from the findings of the laryngoscope. Last, diagnosis required that symptoms not be due to laryngitis caused by upper airway infections and/or allergies. A consensus meeting among the 3 otolaryngologists was conducted to improve interrater reliability. The enrolled participants included all the patients who initially received a diagnosis of LPR and had no history of PPI treatment. Patients younger than 18 years, those experiencing GERD symptoms but not LPR symptoms, and those who had a malignant tumor or major psychosis were excluded. The study protocol was reviewed and approved by the institutional review boards of Kyung Hee University Hospital at Gangdong, Seoul Veterans Hospital, and Samsung Changwon Hospital. Written informed consent was obtained from the participants.

In addition to advice about lifestyle modification (avoidance of caffeine, alcohol, smoking, fatty food, and eating close to bedtime), patients with LPR were prescribed 15 mg of lansoprazole 2 times a day for 3 months. Patients were instructed to take the PPI 30 minutes before meals. The disease severity and changes in subjective symptoms were assessed using 2 surveys, the RSI and LPR-HRQOL, for 3 months.[9,11] The surveys were administered 3 times during this period at outpatient department visits: the first visit and follow-up visits at 1 and 3 months. In addition, to evaluate the objective findings of the laryngeal condition, the RFS by Belafsky et al[10] was conducted by an otolaryngologist. A greater than 50% primary RSI improvement from baseline was considered a response to PPI therapy.

The Questionnaires

The RSI is a high-validity survey that not only assesses the level of severity of LPR but also includes 9 questions to estimate the response to treatment. The questionnaire evaluates the level of symptoms and their severity through a 6-point Likert scale, which ranges from 0 to 5. A high score indicates that patients have more severe symptoms, whereas 0 indicates the absence of symptoms.

The LPR-HRQOL, developed by Carru et al,[11] has been shown to be a reliable and valid rating scale for evaluating the quality of life (QOL) of LPR patients. It consists of a simple questionnaire of 43 questions in the 5 categories of hoarseness, cough, throat clearing, swallowing, and overall impact of acid reflux. The questionnaire consists of basic 7-point Likert scale questions in the first 4 categories and concludes with a 10-point Likert scale question regarding the overall impact of acid reflux. A high score indicates that patients have more severe symptoms, whereas 0 indicates the absence of symptoms.

Examination of the Larynx

All enrolled participants underwent laryngoscopy to evaluate objective signs of LPR on the basis of the RFS. An otolaryngologist performed the examination using a strobolaryngoscope, and 70° rigid endoscopes were generally used. When the vocal cords and surrounding structures were not clearly visible in the rigid endoscope, a flexible nasopharyngoscope was used to increase accuracy. Participants were instructed to verbalize "yee" in a high-pitched tone, low-pitched tone, and regular-pitched tone. Through this procedure, the diagnosis of LPR and the RFS, the rating scale of clinical advanced LPR, were assessed. The RFS ranged from 0 (normal state) to 26; a higher score indicates a deteriorated laryngeal condition. A consensus meeting among the 3 clinics was conducted to improve the interrater and intrarater reliability in RFS scoring. This investigation was performed according to standard protocol and scored by observers blinded to the patient's identity.

Copyright 2013 American Medical Association. All rights reserved.

Table 1. Initial Reflux Symptom Index (RSI), Reflux Finding Score (RFS), and LPR–Health-Related Quality of Life (LPR-HRQOL) According to Age Group

Test	Score, Mean (SD)			*P* Value
	18-39 y (n = 35)	40-59 y (n = 83)	60-79 y (n = 111)	
RSI	13.88 (7.68)	12.20 (8.90)	18.45 (10.43)	<.001
RFS	6.78 (4.86)	7.18 (4.55)	7.75 (3.84)	.44
LPR-HRQOL				
Voice	14.48 (18.65)	18.63 (21.36)	29.55 (21.17)	<.001
Cough	6.42 (7.36)	8.96 (7.36)	14.10 (12.40)	<.001
Throat clearing	6.97 (7.99)	6.15 (7.22)	12.78 (11.50)	<.001
Swallowing	6.05 (7.25)	6.40 (6.38)	11.31 (9.30)	<.001
Overall impact of acid reflux	21.28 (18.46)	21.45 (14.72)	34.81 (23.26)	<.001

Table 2. Improvement in Reflux Symptom Index (RSI), Reflux Finding Score (RFS), and LPR–Health-Related Quality of Life (LPR-HRQOL) After Proton Pump Inhibitor Therapy According to Age

Test	Score, Mean (SD)			*P* Value	
	Baseline	1 Month	3 Months	Within Group	Among Groups
RSI					
18-39 y	13.88 (7.68)	8.77 (6.80)	4.62 (5.35)	<.001	.59
40-59 y	12.20 (8.90)	7.57 (7.31)	5.32 (6.28)	<.001	
60-79 y	18.45 (10.43)	12.89 (9.15)	10.81 (9.86)	<.001	
RFS					
18-39 y	6.21 (4.87)	3.89 (3.10)	2.92 (2.59)	<.001	.50
40-59 y	6.46 (4.70)	4.34 (3.36)	2.98 (2.66)	<.001	
60-79 y	7.74 (3.92)	5.76 (3.39)	4.40 (2.67)	<.001	
LPR-HRQOL					
18-39 y	21.50 (19.29)	14.28 (8.38)	11.93 (4.13)	<.001	.09
40-59 y	20.81 (14.50)	17.41 (12.40)	14.02 (8.03)	<.001	
60-79 y	35.20 (23.29)	24.27 (17.57)	23.08 (19.15)	<.001	

Statistical Analysis

For statistical analysis, SPSS, version 18.0 (SPSS), was used, and all of the data are presented as mean (SD). A *t* test and analysis of variance (ANOVA) were used to compare age differences in RSI, RFS, and LPR-HRQOL data. A repeated measure of ANOVA was used to determine which age group showed a greater response to PPI therapy. An ANOVA model in repeated measures at 3 time points was used with Bonferroni correction for multiple comparisons. Comparison of the proportion of responders according to age was made using a χ^2 analysis. A difference was considered statistically significant when the *P* value was less than .05.

Results

Study Populations

Of 264 consecutive patients considered for the study, 35 were excluded because of loss of follow-up at 3 months. A total of 229 patients with LPR were enrolled and completed the study without loss to follow-up. There were 135 men (59.0%) and 94 women (41.0%). The mean (SD; range) age of the patients was 55.7 (14.0; 18-79) years. Patients were divided into 3 age groups of 18 to 39, 40 to 59, and 60 to 79 years. The number of patients in each group was 35 (15.3%), 83 (36.2%), and 111 (48.5%), respectively.

Difference of RSI, RFS, and LPR-HQOL According to Age

The oldest patient group (60-79 years) with LPR had significantly higher mean (SD) baseline RSI scores than the 18- to 39-year-old and 40- to 59-year-old patient groups (18.45 [10.43] vs 13.88 [7.68] and 12.20 [8.90], respectively; $P < .001$). However, the RFS score showed no significant difference among age groups. The oldest patient group showed significantly worse results on all domains of the LPR-HRQOL (all $P < .001$) (**Table 1**).

Improvement of RSI, RFS, and LPR-HRQOL After PPI Therapy

Within each age group, scores on all 3 tests improved significantly during the period of PPI therapy; however, there was no significant difference among groups in the amount of improvement (RSI, $P = .59$; RFS, $P = .50$; LPR-HRQOL, $P = .09$) (**Table 2**).

Difference in Proportion of Responders on RSI According to Age

Among the age groups, the proportion of responders, as evaluated by RSI score, showed no significant difference at 1 month; however, responders were significantly more plentiful in the 2 younger groups than the oldest group at 3 months ($P = .002$) (**Table 3**).

Copyright 2013 American Medical Association. All rights reserved.

Table 3. Proportion of Responders as Evaluated by Reflux Symptom Index (RSI) According to Age

Follow-up Period, mo	Responders,[a] No. % 18-39 y (n = 35)	40-59 y (n = 83)	60-79 y (n = 111)	*P* Value
1	11 (31)	26 (31)	41 (37)	.70
3	30 (86)	62 (75)	63 (57)	.002

[a] Responders were defined as those whose RSI score improved by more than 50% after proton pump inhibitor therapy.

Discussion

The major finding of this prospective study is that the subjective severity of LPR is significantly greater in older than in younger patients. In addition, the older patients showed lower response rates after PPI therapy.

In a previous cohort study in 100 patients with no history of voice or laryngeal symptoms, 35% were found to have symptoms of LPR and 64% showed 1 or more physical findings of LPR on laryngoscopic examination.[12] Despite the high prevalence of LPR, there are few data on the influence of age on symptom severity or response to PPI treatment. The present prospective study investigated the influence of age on severity and PPI response in LPR. We assessed the subjective severity through the LPR-HRQOL, which evaluated the QOL of patients with LPR, as well as the RSI in groups stratified according to age. To our knowledge, this is the first report of greater severity of disease and negative impact on QOL in geriatric patients with LPR.

It is known that the incidence of GERD symptoms does not increase with age; however, several studies suggest that the frequency of GERD complications such as esophagitis, stricture, or Barrett esophagus is significantly higher in older people.[5,7,13,14] The most likely reason for the increased severity of GERD in older people is the cumulative injury of acid to the esophageal mucosa over time. In addition, a defective antireflux barrier, abnormal esophageal clearance, altered esophageal mucosal resistance, and delayed gastric emptying could contribute to this phenomenon.[15]

It is not known whether the severity of LPR in older patients is greater than in younger patients. Saruç et al[16] demonstrated that age is not a risk factor for the development of LPR. In our results, older patients with LPR showed a higher score on the RSI. Moreover, LPR symptoms had a significantly greater negative impact on the lives of older patients. In a recent study on the QOL impact of LPR, LPR symptoms had a significant correlation with all tested QOL parameters.[17] However, we could not find any difference in RFS, the objective laryngeal finding, among the groups. Our data suggest that age affects the subjective symptoms and resulting impact on the QOL in LPR but not the laryngeal finding. The difference may be the result of a different perspective on their health status among people of different ages.

Many previous studies agree that PPI therapy is the cornerstone of LPR treatment.[18,19] The current management strategy for patients with LPR is empirical therapy with a twice-daily PPI for 3 months[19]; however, the proportion of patients who respond to PPI therapy varies, ranging from 27% to 83% for 1 month of treatment and 41% to 100% for 3 months of treatment.[20-22] Although several randomized clinical trials demonstrated no significant postintervention difference between groups receiving a PPI vs placebo, in a recent open-label observational study, significant improvement in RSI (primary RSI improvement of >50%) was obtained in 75% of patients after 12 weeks.[23] This is similar to the response rate in the 40- to 59-year-old group in our study. Moreover, we were able to find a difference in response among the groups according to age. This is a noteworthy finding in our trial, although there was no placebo group.[21]

Several trials on the predictors of response to PPI treatment have also shown conflicting results. Park et al[24] demonstrated that pretherapy abnormalities in the interarytenoid mucosa and true vocal fold were associated with a 2-fold increase in symptom response to PPI treatment. Williams et al[25] reported that neither baseline GERD symptoms nor endoscopic findings predicted laryngoscopic or symptomatic response. Another study suggested that baseline anxiety levels and heartburn scores and medication dose might be relevant factors in predicting faster response to PPI treatment in carefully selected patients.[26] In our data, different age groups had different proportions of responders as evaluated by the RSI. The response rate in the oldest patients was significantly lower than in other age groups.

Few published articles have investigated PPI resistance in LPR. Amin et al[27] suggested that incomplete suppression might result from a shorter duration of drug action in unresponsive patients, possibly through increased metabolism of the PPI by the liver. Another explanation for poor response to PPI therapy is low bioavailability of the drug. Ashida et al[28] suggested that decreased plasma levels of PPI in patients with resistant gastric ulcers were due to an increase in gastric emptying time. Several authors have showed that older adults have a significant decrease in the amplitude of peristaltic pressures.[29-31] This is associated with a higher prevalence of diabetes mellitus or rheumatological disorders, which may alter esophageal motility in older persons. Therefore, decreased acid clearance in geriatric patients might be a possible cause of decreased response to PPI therapy.

Limitations of the present study include the lack of a placebo group as control. Moreover, we did not demonstrate the reflux events by means of multichannel impedance or pH monitoring studies. Although the gold standard diagnostic method for LPR is dual-probe 24-hour pH monitoring, it is an invasive test with a high false-negative rate.[21] Also, LPR is a fluctuating condition and there can be substantial day-to-day variation of acid exposure in the hypopharynx.[32] However, the response to PPI therapy in patients with suspected LPR is usually so explicit that empirical PPI therapy in LPR is recommended by both gastroenterology and otolaryngology experts and guidelines.[19,33,34] The present study might have meaningful implications for the difference in the effects of PPI therapy according to age.

Copyright 2013 American Medical Association. All rights reserved.

Although there was no significant difference in the objective findings among the different age groups, the subjective severity of LPR in geriatric patients is significantly greater than in younger patients. Furthermore, older patients are more likely not to respond to PPI therapy than younger patients.

ARTICLE INFORMATION

Submitted for Publication: July 2, 2013; final revision received August 18, 2013; accepted September 7, 2013.

Published Online: November 7, 2013. doi:10.1001/jamaoto.2013.5556.

Author Contributions: Dr Eun had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Y. C. Lee, Kim, Kwon, Eun.
Acquisition of data: Y. C. Lee, J. S. Lee, Kim, Kwon, Eun.
Analysis and interpretation of data: Y. C. Lee, J. S. Lee, Eun.
Drafting of the manuscript: Y. C. Lee, J. S. Lee.
Critical revision of the manuscript for important intellectual content: Kim, Kwon, Eun.
Statistical analysis: Y. C. Lee, J. S. Lee.
Study supervision: Kim, Kwon, Eun.

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Wiener GJ, Koufman JA, Wu WC, Cooper JB, Richter JE, Castell DO. Chronic hoarseness secondary to gastroesophageal reflux disease: documentation with 24-h ambulatory pH monitoring. *Am J Gastroenterol*. 1989;84(12):1503-1508.

2. Fraser AG. Review article: gastro-oesophageal reflux and laryngeal symptoms. *Aliment Pharmacol Ther*. 1994;8(3):265-272.

3. Rees CJ, Belafsky PC. Laryngopharyngeal reflux: current concepts in pathophysiology, diagnosis, and treatment. *Int J Speech Lang Pathol*. 2008;10(4):245-253.

4. Koufman J, Sataloff RT, Toohill R. Laryngopharyngeal reflux: consensus conference report. *J Voice*. 1996;10(3):215-216.

5. Collen MJ, Abdulian JD, Chen YK. Gastroesophageal reflux disease in the elderly: more severe disease that requires aggressive therapy. *Am J Gastroenterol*. 1995;90(7):1053-1057.

6. Brunnen PL, Karmody AM, Needham CD. Severe peptic oesophagitis. *Gut*. 1969;10(10):831-837.

7. el-Serag HB, Sonnenberg A. Associations between different forms of gastro-oesophageal reflux disease. *Gut*. 1997;41(5):594-599.

8. Garrigues V, Gisbert L, Bastida G, et al. Manifestations of gastroesophageal reflux and response to omeprazole therapy in patients with chronic posterior laryngitis: an evaluation based on clinical practice. *Dig Dis Sci*. 2003;48(11):2117-2123.

9. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). *J Voice*. 2002;16(2):274-277.

10. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope*. 2001;111(8):1313-1317.

11. Carrau RL, Khidr A, Gold KF, et al. Validation of a quality-of-life instrument for laryngopharyngeal reflux. *Arch Otolaryngol Head Neck Surg*. 2005;131(4):315-320.

12. Reulbach TR, Belafsky PC, Blalock PD, Koufman JA, Postma GN. Occult laryngeal pathology in a community-based cohort. *Otolaryngol Head Neck Surg*. 2001;124(4):448-450.

13. Locke GR III, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ III. Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota. *Gastroenterology*. 1997;112(5):1448-1456.

14. Räihä IJ, Impivaara O, Seppälä M, Sourander LB. Prevalence and characteristics of symptomatic gastroesophageal reflux disease in the elderly. *J Am Geriatr Soc*. 1992;40(12):1209-1211.

15. Castell DO. Esophageal disorders in the elderly. *Gastroenterol Clin North Am*. 1990;19(2):235-254.

16. Saruç M, Aksoy EA, Vardereli E, et al. Risk factors for laryngopharyngeal reflux. *Eur Arch Otorhinolaryngol*. 2012;269(4):1189-1194.

17. Siupsinskiene N, Adamonis K, Toohill RJ. Quality of life in laryngopharyngeal reflux patients. *Laryngoscope*. 2007;117(3):480-484.

18. Ford CN. Evaluation and management of laryngopharyngeal reflux. *JAMA*. 2005;294(12):1534-1540.

19. Koufman JA, Aviv JE, Casiano RR, Shaw GY. Laryngopharyngeal reflux: position statement of the committee on speech, voice, and swallowing disorders of the American Academy of Otolaryngology–Head and Neck Surgery. *Otolaryngol Head Neck Surg*. 2002;127(1):32-35.

20. Vaezi MF. Reflux-induced laryngitis (laryngopharyngeal reflux). *Curr Treat Options Gastroenterol*. 2006;9(1):69-74.

21. Karkos PD, Wilson JA. Empiric treatment of laryngopharyngeal reflux with proton pump inhibitors: a systematic review. *Laryngoscope*. 2006;116(1):144-148.

22. Ulualp SO, Toohill RJ, Shaker R. Outcomes of acid suppressive therapy in patients with posterior laryngitis. *Otolaryngol Head Neck Surg*. 2001;124(1):16-22.

23. Shin MH, Nam SY, Park YH, Son YI. Open-label observational study for evaluating the short-term benefits of rabeprazole medication on laryngopharyngeal reflux. *Clin Exp Otorhinolaryngol*. 2012;5(1):28-33.

24. Park W, Hicks DM, Khandwala F, et al. Laryngopharyngeal reflux: prospective cohort study evaluating optimal dose of proton-pump inhibitor therapy and pretherapy predictors of response. *Laryngoscope*. 2005;115(7):1230-1238.

25. Williams RB, Szczesniak MM, Maclean JC, Brake HM, Cole IE, Cook IJ. Predictors of outcome in an open label, therapeutic trial of high-dose omeprazole in laryngitis. *Am J Gastroenterol*. 2004;99(5):777-785.

26. Siupsinskiene N, Adamonis K, Toohill RJ, Sereika R. Predictors of response to short-term proton pump inhibitor treatment in laryngopharyngeal reflux patients. *J Laryngol Otol*. 2008;122(11):1206-1212.

27. Amin MR, Postma GN, Johnson P, Digges N, Koufman JA. Proton pump inhibitor resistance in the treatment of laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2001;125(4):374-378.

28. Ashida K, Sakaguchi M, Tanaka M, Takiuchi H, Egashira Y, Katsu K. Clinical study on the pathophysiology and treatment of PPI-resistant ulcers. *J Clin Gastroenterol*. 1995;20(suppl 2):S67-S71.

29. Hollis JB, Castell DO. Esophageal function in elderly men: a new look at "presbyesophagus." *Ann Intern Med*. 1974;(3):371-374.

30. Soergel KH, Zboralske FF, Amberg JR. Presbyesophagus: esophageal motility in nonagenarians. *J Clin Invest*. 1964;43:1472-1479.

31. Khan TA, Shragge BW, Crispin JS, Lind JF. Esophageal motility in the elderly. *Am J Dig Dis*. 1977;22(12):1049-1054.

32. Cheung TK, Lam PK, Wei WI, et al. Quality of life in patients with laryngopharyngeal reflux. *Digestion*. 2009;79(1):52-57.

33. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al; American Gastroenterological Association. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. *Gastroenterology*. 2008;135:1383-1391.e5.

34. Altman KW, Prufer N, Vaezi MF. A review of clinical practice guidelines for reflux disease: toward creating a clinical protocol for the otolaryngologist. *Laryngoscope*. 2011;121(4):717-723.

Copyright 2013 American Medical Association. All rights reserved.

Reprinted by permission of J Voice. 2015; 29(6):755-762.

Vocal Fold Vibration in Vocal Fold Atrophy: Quantitative Analysis With High-Speed Digital Imaging

***Akihito Yamauchi, †Hisayuki Yokonishi, *Hiroshi Imagawa, ‡Ken-Ichi Sakakibara, *Takaharu Nito, §Niro Tayama, and *Tatsuya Yamasoba,** *†§*Tokyo and* ‡*Hokkaido, Japan*

Summary: Introduction. Vocal fold vibrations of vocal fold atrophy (VFA), a rapidly increasing voice disorder owing to worldwide societal aging, have not been clarified by high-speed digital imaging (HSDI).
Methods. The HSDI method was performed on 46 patients (33 males and 13 females) with VFA and 20 vocally healthy subjects (8 males and 12 females), and the obtained data were quantitatively evaluated by frame-by-frame analysis, laryngotopography, single- and multi-line kymography, and glottal area waveform.
Results. Overall, patients with VFA revealed larger open quotients, larger lateral phase difference, larger integral glottal width (the average glottal width over a glottal cycle), and smaller speed index than vocally healthy subjects. Some gender difference was noted: in males, lateral phase difference was not significant; and in females, integral glottal width and speed index were not significant. Correlation study revealed moderate correlations between HSDI-derived parameters and conventional acoustic or aerodynamic parameters.
Conclusions. The combination of multiple HSDI analysis methods was effective in documenting the characteristics of vocal fold vibrations in VFA. The knowledge of general vibratory characteristics and gender difference is beneficial for the appropriate clinical care of VFA.
Key Words: Vocal fold atrophy–Presbyphonia–High-speed digital imaging–Aging–Anti-aging.

INTRODUCTION

Vocal fold atrophy (VFA) is a voice disorder resulting from the atrophied muscle and mucosa in the vocal folds.[1] These structural modifications lead to increased glottal air leakage and breathy, rough voice. Aging is considered to be the most major predisposing factor for VFA, although other risk factors have also been proposed (eg, reflux laryngitis, chronic medical conditions, and vocal abuse).[2] The VFA has increased considerably during the past two decades as a result of the worldwide societal aging, and thus, is attracting clinical attention in the world these days.[3,4]

Laryngoscopically, the vocal fold bowing, prominent vocal process, and spindle-shaped glottal gap are usually observed.[1,2] The vibratory characteristics observed with videostroboscopy include normal or decreased amplitude, either complete closure or glottal gap, normal mucosal wave, and small supraglottal area.[1,2,5,6]

The details of vocal fold vibrations in VFA, however, have not yet been documented by high-speed digital imaging (HSDI), although HSDI is considered to be the better choice than videostroboscopy.[7,8] First, HSDI is capable of observing actual vocal fold vibrations with a high frame rate, and guarantees reliable assessment of intra- and intercycle vibratory behaviors, unlike videostroboscopy that only provides reconstructed, averaged, illusory images. Second, HSDI offers wider application to clinical cases than videostroboscopy because HSDI is free from the problem with synchronization and is applicable to severe dysphonia in which videostroboscopy results in desynchronization. Third, various analysis methods for HSDI are now available, and thus HSDI provides more multifaceted information than videostroboscopy that has relatively limited choices of analysis methods. Furthermore, only little is known about the association between vibratory parameters and acoustic or aerodynamic parameters in VFA,[5] and HSDI data have not been reported on this matter. The connection between HSDI parameters and routinely evaluated vocal function parameters in VFA should be beneficial for better understanding the pathophysiological aspects of this clinical entity.

Hence, the purpose of the present study was to quantitatively elucidate the vibratory characteristics in VFA patients using HSDI, and to clarify the relationship between HSDI parameters and aerodynamic/acoustic measures.

MATERIALS AND METHODS

Subjects

Patients who visited the Voice Outpatient Clinic of the Department of Otolaryngology and Head and Neck Surgery at the University of Tokyo Hospital (Tokyo, Japan) and those who were diagnosed with VFA between 2006 and 2013 were included in this study. The diagnosis of VFA was based on careful history taking, acoustic and aerodynamic evaluation, and laryngostroboscopic findings: Patients with objective dysphonia on acoustic or aerodynamic studies; without signs of other laryngeal diseases such as vocal fold paralysis, vocal fold polyp, laryngeal carcinoma, vocal fold scar, or functional dysphonia; and with the prominent vocal process, bowed vocal fold, spindle-shaped or anterior glottal gap, or increased open phase during phonation

Accepted for publication December 15, 2014.
This research was presented in the Seventh East Asian Conference on Phonosurgery; November 26–27, 2010; Tokyo, Japan.
From the *Department of Otolaryngology, The University of Tokyo Hospital, Tokyo, Japan; †Department of Otolaryngology, The Tokyo Metropolitan Police Hospital, Tokyo, Japan; ‡Department of Communication Disorders, The Health Sciences University of Hokkaido, Hokkaido, Japan; and the §Department of Otolaryngology and Tracheo-Esophagology, The National Center for Global Health and Medicine, Tokyo, Japan.
Address correspondence and reprint requests to Akihito Yamauchi, Department of Otolaryngology, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655 Japan. E-mail: drachilles23@yahoo.co.jp
Journal of Voice, Vol. ■, No. ■, pp. 1-8
0892-1997/$36.00
© 2015 The Voice Foundation
http://dx.doi.org/10.1016/j.jvoice.2014.12.008

were diagnosed as those having VFA. The diagnosis was made by three or four certified otorhinolaryngologists specializing in vocal treatment. As for sulcus vocalis, patients with a type 1 sulcus (physiological sulcus) were included in the category of VFA because type 1 sulcus is superficial and generally causes no or minimal functional vocal impairment.[9] Furthermore, type 1 sulcus is considered to be associated with aging.[9] On the other hand, patients with a type 2 sulcus (sulcus vergeture) or a type 3 sulcus (true sulcus vocalis) were excluded from this study.[9]

As a control group, vocally healthy subjects without vocal complaints, history of laryngeal disorders, or signs of laryngeal abnormality with laryngoendoscopy were recruited. As an exception, however, a small glottal gap was permitted for a control group in the present study because vocally healthy elderly population is known to demonstrate a small glottal gap frequently.[10]

All subjects were required to sign a consent form that was approved by our Institutional Review Board. A total of 46 patients with VFA (13 women and 33 men), with the age range between 60 and 91 years, and 20 vocally healthy subjects (12 women and eight men), with the age range between 65 and 81 years, were enrolled in the present study.

TABLE 1.
Clinical Data of All Participants

Parameter (Units)	Control (N = 20)	VFA (N = 46)	*t* Test
Age (yr)	73 ± 5	72 ± 7	0.513
MPT (s)	18.7 ± 6.6	16.4 ± 8.2	0.284
MFR (mL/s)	136 ± 36	210 ± 101	0.002†
AA-F_0 (Hz)	178 ± 50	179 ± 63	0.940
APQ (%)	3.2 ± 1.5	4.1 ± 2.3	0.135
PPQ (%)	0.31 ± 0.46	0.92 ± 0.61	<0.001‡
HNR (dB)	21.5 ± 3.5	12.6 ± 4.9	<0.001‡
Grade	0.80 ± 0.62	1.33 ± 0.52	<0.001‡
Roughness	0.80 ± 0.62	1.17 ± 0.44	0.006†
Breathiness	0.50 ± 0.51	0.65 ± 0.60	0.330

Abbreviations: SD, standard deviation; VFA, vocal fold atrophy; MPT, maximum phonation time; MFR, mean flow rate; AA-F_0, fundamental frequency in acoustic analysis; APQ, amplitude perturbation quotient; PPQ, period perturbation quotient; HNR, harmonics-to-noise ratio.
Notes: Values signify "mean ± SD." The column for *t* test shows the *P* value of Student's *t* test between control and VFA groups.
† $P < 0.01$.
‡ $P < 0.001$.

Background data

Vocal function and voice quality were evaluated by measuring aerodynamic and acoustic parameters. The aerodynamic parameters including the maximum phonation time and mean flow rate were measured with a Nagashima PE-77E Phonatory Function Analyzer (Nagashima Medical, Inc., Tokyo, Japan). Acoustic parameters included the fundamental frequency (AA–F_0), amplitude perturbation quotient, period perturbation quotient, and harmonic-to-noise ratio, which were measured at the University of Tokyo with a dedicated software program, as well as the subjective rating by the GRBAS scale.

Table 1 summarizes the results of aerodynamic and acoustic studies, in which mean flow rate, period perturbation quotient, harmonics-to-noise ratio, and the grade and roughness of the GRBAS scale revealed significant intergroup differences. The scores of Voice Handicap Index-10 and the voice-related quality of life were 13.7 ± 9.4 and 13.8 ± 8.9, respectively, and the rate of synchronization in VFA with videostroboscopy (LS-3A; Nagashima Medical, Inc.) was 67.3%.

High-speed digital imaging

A high-speed digital camera (FASTCAM-1024PCI; Photron, Tokyo, Japan) was connected to a rigid endoscope (#4450.501; Richard Wolf, Vernon Hills, IL) via an attachment lens ($f = 35$ mm; Nagashima Medical, Inc.). Recording was performed under illumination with a 300-W xenon light source at a frame rate of 4500 frames per second and a spatial resolution of 512 × 400 pixels, with an 8-bit grayscale and a recording duration of 1.86 seconds. High-speed digital images of sustained vowel phonation /i/ at a comfortable frequency with a comfortable intensity were recorded. The image sequence of stable vocal fold vibrations were selected for further analyses.

Aerodynamic and acoustic studies were performed approximately 30 minutes before HSDI recording because simultaneous recording was not available at our institution. Both evaluations were done under as similar conditions as possible to allow comparison between the HSDI parameters and the aerodynamic or acoustic parameters.

HSDI analysis methods

The recorded HSDIs were evaluated by frame-by-frame analysis,[11] laryngotopography (LTG),[12] single-/multi-line digital kymography (SLK and MLK, respectively),[13,14] and glottal area waveform (GAW).[15] The details of analysis by these methods are described elsewhere.[11–15]

The size parameters were normalized by the vocal fold length, labeled by "N_L-" (eg, V_L-amplitude mean). The time parameters were normalized by the glottal cycle, labeled by "N_G-" (eg, N_G-lateral phase difference). The size and time parameters were normalized by both glottal cycle and vocal fold length, labeled by "N_{GL}-" (eg, N_{GL}-lateral phase difference).[13]

In the present study, analysis was focused on selected parameters that were considered to be related with the vibratory characteristics of VFA such as amplitude, mucosal wave, lateral/longitudinal phase difference, open quotient, speed index, integral glottal width (the average glottal width over a glottal cycle),[13] maximal/minimal glottal area, glottal area difference, and glottal outlet (normalized supraglottal area).[6]

Frame-by-frame analysis was performed using an assessment form for HSDI developed by the authors, with which vibratory parameters such as symmetry, periodicity, amplitude, mucosal wave, phase difference, glottal closure, and supraglottal hyperactivity were evaluated by two- or four-point scale.[11] For glottal gaps, the incidence (present or absent) and glottal type (incomplete closure, posterior, spindle-shaped, or anterior) were evaluated.

The LTG is a method using a pixel-wise Fourier transform of time-varying brightness curve for each pixel across images and

allows quantitative evaluation of spatial characteristics of frequency and phase. In the present study, the incidence of diplophonia and lateral/longitudinal phase difference normalized by glottal cycle (N_G-lateral/longitudinal phase difference) were evaluated.[12] Image sequence of 512 frames was evaluated.

The SLK analyzes mediolateral vocal fold movements at a midglottal level and allows evaluation of mediolateral and temporal vibratory characteristics such as amplitude (N_L-amplitude mean), mucosal wave (N_L-mucosal wave magnitude mean), phase (N_G-lateral phase difference and N_G-longitudinal phase difference), open quotient (O_q^{SLK}), speed index (SI^{SLK}), and integral glottal width (N_{GL}-integral glottal width).[13] Image sequence of 400 frames was evaluated.

The MLK involves five different longitudinal levels, and can assess temporal and longitudinal oscillatory features such as open quotient (O_q^{MLK}), speed index (SI^{MLK}), and opening/closing longitudinal phase difference (N_G-opening/closing longitudinal phase difference).[14] Image sequence of 400 frames was evaluated.

The GAW provides information on the general dynamics of the glottal area by tracing the vocal fold edges and displaying temporal changes of the glottal area, with which open quotient (O_q^{GAW}), speed index (SI^{GAW}), minimal glottal area (N_L-minimal glottal area), maximal glottal area (N_L-maximal glottal area), glottal outlet (N_L-glottal outlet), and glottal area difference (glottal area difference index = (N_L-maximal glottal area − N_L-minimal glottal area)/N_L-maximal glottal area) can be evaluated.[15] Five consecutive glottal cycles of the most stable segment were evaluated (100–200 frames).

All the HSDI analyses were performed with custom *MATLAB* software (Version 2011a; Mathworks, Inc., Natick, MA) programmed at our institution . An example of HSDI analysis is shown in Figure 1.

Statistics

The difference of clinical and HSDI parameters between VFA patients and vocally healthy subjects were evaluated by Student's *t* test for normally distributed parameters, or either by the Mann-Whitney *U* test or by chi-squared test for other parameters. To assess the correlations with HSDI parameters and reference data (demographic, aerodynamic/acoustic, or stroboscopic data), or the correlations among HSDI parameters, Pearson's correlation analysis for normally distributed parameters and Spearman's correlation analysis for the other parameters were used. In all analyses, *P* value lower than 0.05 was considered significant. Calculations were performed with a custom *MATLAB* software.

RESULTS

HSDI parameters in the VFA

Frame-by-frame analysis. The glottal gap was observed in 45% of the VFA group and 30% of the control group ($P = 0.235$), and the most frequent type was spindle shaped for the VFA group (47.8% of those with glottal gaps) and anterior for the control group (66.7% of those with glottal gaps). Although no parameters revealed significant differences between the two groups, the VFA group tended to demonstrate

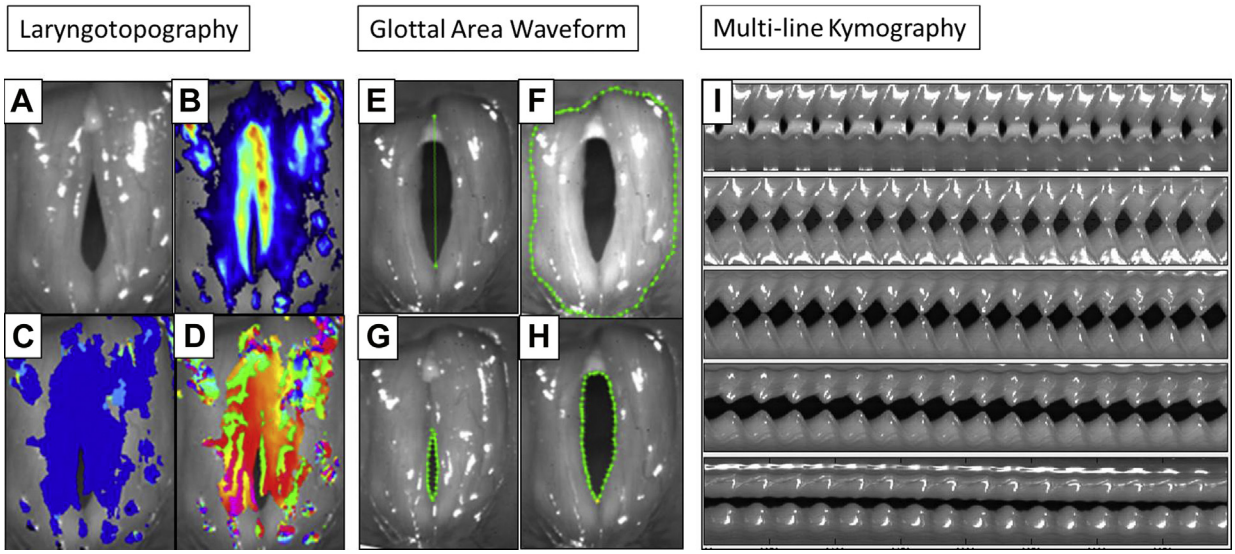

FIGURE 1. An example of the analysis of high-speed digital image is shown. (**A**–**D**) Laryngotopography is shown. (**A**) A static laryngeal image to be superimposed by analyzed topographic data. (**B**–**D**) Spatial distribution of amplitude, frequency, and phase of the maximum amplitude components, respectively. This 71-year-old female patient has a topographic F0 of 225 Hz, left-to-right lateral phase difference (6.3% of a glottal cycle), and anterior-to-posterior longitudinal phase difference (21.9% of a glottal cycle). (**E**–**H**) Glottal area waveform is shown: areas demarcated by a green line in Panels (**E**–**H**) show a minimal glottal area, a maximal glottal area, and a glottal outlet, respectively; and a green line in Panel (**G**) signifies the vocal fold length. This patient has an anterior glottal gap with a N_L-minimal glottal area of 1.2%, a spindle-shaped maximal glottal area with a N_L-maximal glottal area of 11.5%, and no supraglottal hyperfunction with N_L-glottal outlet of 100.2%. (**I**) Multi-line kymography is shown. There are anterior-to-posterior opening and posterior-to-anterior closing longitudinal phase differences (23.8% and 19.0% of a glottal cycle, respectively). O_q^{MLK} and SI^{MLK} are 0.86 and −0.21, respectively.

TABLE 2.
Comparisons of High-Speed Digital Image Parameters Between the Control and VFA Groups

Parameter (Units)	Control (N = 20)	VFA (N = 46)	*t* Test
Laryngotopography			
N_G-lateral phase difference (%)	3.8 ± 4.3	6.3 ± 4.3	0.034*
N_G-longitudinal phase difference (%)	−14.4 ± 16.6	−8.9 ± 13.6	0.175
Single-line digital kymography			
N_L-amplitude mean (%)	8.1 ± 2.7	8.6 ± 3.4	0.544
N_L-mucosal wave magnitude mean (%)	18.6 ± 9.3	16.2 ± 7.9	0.321
O_q^{SLK}	0.57 ± 0.13	0.76 ± 0.17	<0.001†
SI^{SLK}	−0.19 ± 0.18	−0.18 ± 0.18	0.906
N_G-lateral phase difference (%)	9.9 ± 6.5	8.9 ± 10.4	0.689
N_{GL}-integral glottal width (%)	4.5 ± 1.8	7.0 ± 3.5	0.004‡
Multi-line digital kymography			
O_q^{MLK}	0.45 ± 0.11	0.66 ± 0.20	<0.001†
SI^{MLK}	−0.24 ± 0.16	−0.25 ± 0.14	0.908
N_G-opening longitudinal phase difference (%)	−7.4 ± 19.2	−7.3 ± 22.6	0.975
N_G-closing longitudinal phase difference (%)	2.4 ± 19.9	1.6 ± 11.0	0.860
Glottal area waveform			
O_q^{GAW}	0.80 ± 0.18	0.84 ± 0.17	0.908
SI^{GAW}	0.11 ± 0.23	−0.06 ± 0.17	0.002‡
N_L-maximal glottal area (%)	8.8 ± 3.1	16.7 ± 20.3	0.089
N_L-minimal glottal area (%)	0.18 ± 0.53	0.63 ± 1.18	0.104
Glottal area difference index (%)	98.1 ± 5.0	94.1 ± 11.1	0.133
N_L-glottal outlet (%)	61.8 ± 23.6	58.5 ± 21.9	0.592

Abbreviations: VFA, vocal fold atrophy; N_G-, normalized by glottal cycle; N_L-, normalized by vocal fold length; Oq, open quotient; SLK, single-line kymography; SI, speed index; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, multi-line kymography; GAW, glottal area waveform.
Notes: Values for control and VFA columns show "mean ± standard deviation," and the value of *t* test column shows the *P* value of Student's *t* test between all control and VFA groups.
* $P < 0.05$.
† $P < 0.001$.
‡ $P < 0.01$.

smaller amplitude, smaller mucosal wave, greater asymmetry, larger glottal gap, and greater supraglottal hyperactivity than the control group.

Laryngotopography. Three of the patients with VFA (4.9%) revealed diplophonia almost constantly at a comfortable pitch and sound pressure level. The VFA group had larger N_G-lateral phase difference than the control group ($P = 0.034$; Table 2). However, by gender, a significant difference was found only in females (Tables 3 and 4).

Single-line digital kymography. With SLK, patients with VFA revealed significantly greater O_q^{SLK} ($P < 0.001$) and N_{GL}-integral glottal width ($P = 0.004$; Table 2). Although O_q^{SLK}

TABLE 3.
Comparisons of High-Speed Digital Image Parameters Between the Control and VFA Groups in Females

Parameter (Units)	Control (N = 12)	VFA (N = 13)	*t* Test
Laryngotopography			
N_G-lateral phase difference (%)	3.1 ± 4.2	7.3 ± 3.6	0.016*
Single-line digital kymography			
O_q^{SLK}	0.58 ± 0.14	0.77 ± 0.14	0.003†
N_{GL}-integral glottal width (%)	4.4 ± 1.6	5.2 ± 1.8	0.241
Multi-line digital kymography			
O_q^{MLK}	0.46 ± 0.13	0.69 ± 0.21	0.004†
Glottal area waveform			
SI^{GAW}	0.076 ± 0.260	−0.032 ± 0.177	0.261

Abbreviations: VFA, vocal fold atrophy; N_G-, normalized (by glottal cycle); O_q, open quotient; SLK, single-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, multi-line kymography; SI, speed index; GAW, glottal area waveform.
Notes: Values for control and VFA columns show "mean ± standard deviation," and the *t* test column shows the *P* value of Student's *t* test between female control and female VFA groups.
* $P < 0.05$.
† $P < 0.01$.

TABLE 4.
Correlation Coefficients (r) Between High-Speed Digital Image Parameters and Selected Aerodynamic/Acoustic Measures

Parameters	MPT	MFR	AA-F_0	PPQ	HNR	G	R	B
N_L-amplitude mean	−0.06	0.27*	−0.40†	−0.07	0.10	−0.05	−0.05	−0.03
O_q^{SLK}	−0.29†	0.37†	0.02	0.47‡	−0.35†	0.29*	0.29*	0.15
N_{GL}-integral glottal width	−0.20	0.39†	−0.30*	0.19	−0.15	0.18	0.18	0.02
O_q^{MLK}	−0.34†	0.49‡	−0.16	0.34†	−0.31*	0.26*	0.26*	0.13
O_q^{GAW}	−0.28*	0.28*	0.14	0.29*	−0.04	0.24*	0.24*	0.23
SI^{GAW}	0.20	−0.36*	−0.13	−0.28*	0.18	−0.32†	−0.32†	−0.36†
N_L-minimal glottal area	−0.22	0.53‡	−0.43‡	0.12	−0.10	0.20	0.20	0.21
Glottal area difference index	0.17	−0.37†	−0.27*	−0.10	0.12	−0.13	−0.13	−0.09
N_L-glottal outlet	0.37†	−0.11	0.01	−0.09	−0.02	−0.08	−0.08	−0.17

Abbreviations: MPT, maximum phonation time; MFR, mean flow rate; AA-F0, fundamental frequency of acoustic analysis; PPQ, period perturbation quotient; HNR, harmonics-to-noise ratio; G, grade; R, roughness; B, breathiness; NL-, normalized by vocal fold length; O_q, open quotient; SLK, in single-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, in multi-line kymography; GAW, in glottal area waveform; SI, speed index.
* $P < 0.05$.
† $P < 0.01$.
‡ $P < 0.001$.

demonstrated significant differences in both gender, N_{GL}-integral glottal width demonstrated a significant difference only in males (Tables 3 and 5). There were no significant differences in N_L-amplitude mean or N_L-mucosal wave magnitude mean. Although the differences were not statistically significant, N_L-amplitude mean of VFA tended to be greater in males and smaller in females, and N_L-mucosal wave magnitude mean of VFA tended to be smaller in both genders.

Multi-line digital kymography. With MLK, patients with VFA revealed significantly larger O_q^{MLK} ($P < 0.001$; Table 2), and this difference was significant in both genders (Tables 3 and 5).

Glottal area waveform. The VFA group demonstrated significantly smaller SI^{GAW} ($P < 0.001$; Table 2). Although the differences were not significant, the VFA group revealed larger N_L-maximal glottal area and N_L-minimal glottal area (Table 2).

Correlation study

Table 4 summarizes the correlation study between HSDI parameters and background data. Moderate correlations ($0.4 < |r| \leqq 0.7$) were found between mean flow rate and O_q^{MLK}, mean flow rate and N_L-minimal glottal area, AA-F_0 and N_L-amplitude mean, AA-F_0 and N_L-minimal glottal area, and O_q^{SLK} and period perturbation quotient.

Table 6 summarizes the correlation study among the HSDI parameters. A strong correlation ($0.7 < |r|$) was found between N_L-minimal glottal area and glottal area difference index. Otherwise, moderate correlations ($0.4 < |r| \leqq 0.7$) were found in multiple pairs.

DISCUSSION

Videostroboscopy versus HSDI

The use of HSDI offered several advantages to the present study over the previous studies in the literature. First, HSDI enabled the evaluation of time parameters that cannot be assessed by

TABLE 5.
Comparisons of High-Speed Digital Image Parameters Between the Control and VFA Groups in Males were summarized

Parameter (Units)	Control (N = 8)	VFA (N = 33)	*t* Test
Laryngotopography			
N_G-lateral phase difference (%)	4.7 ± 4.4	5.9 ± 4.4	0.511
Single-line digital kymography			
O_q^{SLK}	0.54 ± 0.12	0.75 ± 0.19	0.006*
N_{GL}-integral glottal width (%)	4.6 ± 2.3	7.8 ± 3.8	0.033†
Multi-line digital kymography			
O_q^{MLK}	0.42 ± 0.08	0.65 ± 0.20	0.003*
Glottal area waveform			
SI^{GAW}	0.16 ± 0.17	−0.067 ± 0.166	0.002*

Abbreviations: VFA, vocal fold atrophy; N_G-, normalized (by glottal cycle); Oq, open quotient; SLK, single-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MLK, multi-line kymography; SI, speed index; GAW, glottal area waveform.
Notes: Values for control and VFA columns show "mean ± standard deviation," and the *t* test column shows the *P* value of Student's *t* test between male control and male VFA groups.
* $P < 0.01$.
† $P < 0.05$.

TABLE 6.
Correlation Coefficients (r) Among High-Speed Digital Image Parameters

Parameters	N_L- Minimal GA	SI^{GAW}	$O_q{}^{MLK}$	$O_q{}^{SLK}$	N_L-Amplitude Mean	N_G-O-LPD^{MLK}
GA difference index	−0.90*	−0.46*	−0.65*	−0.54*	−0.34†	−0.15
N_L-minimal GA	1	−0.33†	0.69*	0.47*	−0.24	0.05
SI^{GAW}	—	1	−0.47*	−0.56*	−0.04	−0.39*
$O_q{}^{GAW}$	—	—	0.69*	0.42*	−0.10	0.10
N_G-O-LPD^{LTG}	—	—	0.14	0.19	−0.03	0.50*
$O_q{}^{MLK}$	—	—	1	0.65*	−0.13	0.19
N_{GL}-integral glottal width	—	—	—	0.51*	0.51*	0.28
N_L-MW magnitude mean	—	—	—	—	0.59*	−0.03

Abbreviations: GA, glottal area; N_L-, normalized by vocal fold length; SI, speed index; GAW, in glottal area waveform; O_q, open quotient; N_G-, normalized by glottal cycle; O-LPD, opening longitudinal phase difference; LTG, in laryngotopography; MLK, in multi-line kymography; N_{GL}-, normalized by glottal cycle and vocal fold length; MW, mucosal wave; SLK, in single-line kymography.
* $P < 0.001$.
† $P < 0.01$.

videostroboscopy, which is considered to be meaningful especially because a large part of vibratory characteristics of VFA were reflected in the time parameters (eg, open quotient).

Second, with the use of HSDI, the rate of successful image evaluation increased by 1.5 folds at a rough estimate because videostroboscopic study was successful only in 67.3% because of desynchronization in the present study. The relatively high rate of desynchronization in VFA may be explained by their poor acoustic profile. Patel et al[8] reported that HSDI could be used to augment videostroboscopy for assessment of moderate-to-severe dysphonia, especially in patients with jitter exceeding 0.87%, shimmer exceeding 4.4%, and a signal-to-noise ratio of less than 15.4 dB. In the present study, 21.7% of the VFA group fitted these criteria.

Third, with the application of multiple analysis methods, the present study documented the characteristics of vocal fold vibrations of VFA more extensively and multidirectionally than previous reports, in which vocal fold vibrations were either qualitatively evaluated or quantitatively evaluated with only limited parameters.[5,6]

Although HSDI has disadvantages in comparison with videostroboscopy such as a relatively long time required for analysis (approximately 30 minutes per HSDI at present), a high cost, and the lack of instantaneity (with videostroboscopy, the result of modulation in F_0, sound pressure level, or register can be observed directly and instantaneously),[7,8,16] HSDI is considered to be a good supplementary tool in the assessment of VFA.

Amplitude and integral glottal width

The amplitude mean of VFA was comparable with that of vocally healthy subjects, in this study, which was a consistent result with the previous study.[5] Although not statistically significant, the amplitude of VFA was larger in males and smaller in females than vocally healthy subjects. Various factors can affect amplitude such as amplitude increases as intensity or subglottal pressure increase, or as pitch or stiffness decrease.[6,9,16,17] In female VFA, poor pulmonary function is reported to be frequently associated,[2] which may lead to decreased subglottal pressure and decreased amplitude. In male VFA, a greater glottal flow and lower tension of the thyroarytenoid muscle owing to the muscular atrophy can increase the amplitude.[6,18,19]

On the other hand, N_{GL}-integral glottal width demonstrated a significant difference between the control and VFA groups. The N_{GL}-integral glottal width may be a sensitive parameter than the amplitude *per se* because it has the characteristics of both amplitude and open quotient (Table 6).[13]

Open quotient and speed index

Significant intergroup differences were observed in $O_q{}^{SLK}$ and $O_q{}^{MLK}$ but not in $O_q{}^{GAW}$. This is probably because $O_q{}^{GAW}$ was not a parameter to reflect the size of glottal gap ($O_q{}^{GAW}$ becomes one whether a glottal gap is small or large). Interestingly, the results of $O_q{}^{SLK}$ were comparable with those of $O_q{}^{MLK}$, although $O_q{}^{MLK}$ that assesses the overall glottal area should reflect the pathophysiology of the disease better than $O_q{}^{SLK}$. Perhaps, the midglottal level may represent the vibratory dynamics of overall glottis well enough in VFA, and the information of the glottal ends included in $O_q{}^{MLK}$ may have been less important. Correlation analysis revealed that high $O_q{}^{SLK}$ and $O_q{}^{MLK}$ were associated with poor aerodynamic and acoustic conditions (Table 4). These results seem to stand to reason because weak glottal closure reflected in high open quotient should lead to high glottal flow with high air turbulence.

Speed index of VFA was smaller than that of vocally healthy subjects. Small speed index in the VFA group may originate from the decreased restorative force of the laterally displaced vocal fold toward the medial direction resulting from the disarrangement of collagen fibers or decreased elastin fibers in the lamina propria, the decreased mass or tension of the vocal fold owing to the muscular atrophy.[1,18,19] Contrary to open quotient, SI^{GAW} was more sensitive than SI^{SLK} or SI^{MLK}, probably because SI^{GAW} reflects the general vibratory dynamics than SI^{SLK} or SI^{MLK} (Table 2). Speed index had similar relationships with acoustic and aerodynamic parameters to open quotient (Table 4). These results accord with the findings in the literature, reporting that smaller speed index leads to poorer aerodynamic or acoustic results.[20,21]

Phase difference

With relatively larger N_G-lateral phase difference, vocal fold vibrations in patients with VFA were more asymmetrical than those of vocally healthy subjects. Left-right difference of mass, tension, mucoelasticity of the vocal fold resulting from a different degree of muscular atrophy, and muscular/mucosal degeneration as well as asymmetry of the laryngeal frame may play a role here.[1,18,19]

GAW parameters

The GAW parameters failed to reveal significant intergroup differences although N_L-minimal glottal area and N_L-maximal glottal area were larger, and glottal area difference index was smaller in the VFA group as a trend. This result was consistent with the study of Bloch and Behrman[6] that reported no significant difference in N_L-minimal glottal area between the control and VFA groups. Larger N_L-maximal glottal area found in the present study may be owing to an increased glottal flow in patients with VFA (Table 1), and decreased muscular tension of the vocal fold resulting from the muscular atrophy, leading to a greater lateral excursion of the vocal folds.[18] The smaller glottal area difference index observed in VFA signifies the decreased alternating current of glottal flow, the glottal flow efficiency in other words.

Glottal gap

The result that 30% of elderly vocally healthy subjects had a glottal gap in the present study was consistent with the findings in the literature: Pontes et al[10] reported that the incidence of glottal gap in normal elderly population was 58%, for instance. Strictly speaking, the vocally healthy subjects with a glottal gap in the present study (as well as those in the study of Pontes et al,[10] perhaps) should be termed as "pathological but asymptomatic" rather than "normal," though. Because the preponderant glottal gap was different between the control (anterior) and VFA groups (spindle shaped), the location of the glottal gap may serve as a clue to differentiate VFA from normal aging.

Glottal outlet

No significant intergroup difference of N_L-glottal outlet in the present study was a contradictory result to the report by Bloch and Behrman,[6] who reported significantly smaller N_L-glottal outlet in the VFA group than the normal group. One possible explanation is an interindividual difference of a maladaptive supraglottal hyperactivity as a compensatory strategy for incomplete or decreased glottal closure.[6] Another possible explanation is the posterior displacement of the petiole of epiglottis associated with a descension of the larynx observed in a male low-pitch phonation,[17] or an elevated laryngeal position observed in high-pitch phonation.[22]

Limitations

Overall, the combination of multiple HSDI analysis methods adopted in the present study was effective in the objective documentation of vocal fold vibrations in VFA. Applying the same technique to the evaluation of other laryngeal pathology (eg, vocal fold scar or sulcus vocalis) will be called for in the near future to further validate its utility.

The study design in which the HSDI study and acoustic or aerodynamic studies were performed on separate occasions may be a limitation of the present study, however. Although the effort was made to make the conditions of examination equal as much as possible, there could be a minor variation in F_0 or sound pressure level, leading to relatively low correlations between HSDI parameters and acoustic/aerodynamic parameters. Another limitation may be the use of a rigid endoscope for the HSDI recording, which could yield undesirable laryngeal tension during the study. The short time interval for HSDI analysis as well as the relatively limited subject number (especially of male vocally healthy subjects) may be other limitations. Furthermore, the heterogeneity in the VFA group in the present study may have existed, although the selection of recruited subjects and the diagnosis was based on the agreement of three or four certified otorhinolaryngologists specializing in vocal treatment: Because the differential diagnosis among VFA, sulcus vocalis, and vocal fold scar is not always clear-cut, there is inevitable room for subjectivity.

In the future study, the improvement of the study design by an introduction of simultaneous recording system of HSDI and acoustic signal or aerodynamic data, the introduction of transnasal flexible HSDI, the further refinement of analysis technique with more automation that allows much extended time interval for analysis, and the expansion of subject number will be warranted.

CONCLUSION

The quantitative HSDI analysis of VFA revealed larger open quotients, lateral phase difference and integral glottal width (the average glottal width over a glottal cycle), and smaller speed index than vocally healthy subjects. Gender difference was noted in lateral phase difference, integral glottal width, and speed index. Correlation study revealed mild-to-moderate correlations between HSDI-derived parameters and conventional acoustic or aerodynamic parameters, and moderate-to-strong correlation among HSDI parameters. The combination of multiple HSDI analysis methods was effective in the objective documentation of vocal fold vibrations in VFA.

Acknowledgments

This research was not funded by any organization or grant. There was no conflict of interest to be disclosed.

REFERENCES

1. Kendall K. Presbyphonia: a review. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15:137–140.
2. Yamauchi A, Imagawa H, Sakakibara K-I, et al. Vocal fold atrophy in a Japanese tertiary medical institute: status quo of the most aged country. *J Voice*. 2014;28:231–236.
3. Takano S, Kimura M, Nito T, et al. Clinical analysis of presbylarynx-vocal fold atrophy in the elderly individuals. *Auris Nasus Larynx*. 2010;37:461–467.

4. Davids T, Klein AM, Johns MM. Current dysphonia trends in patients over the age of 65: is vocal atrophy becoming more prevalent? *Laryngoscope*. 2012;122:332–335.
5. Omori K, Slavit DH, Matos C, et al. Vocal fold atrophy: quantitative glottic measurement and vocal function. *Ann Otol Rhinol Laryngol*. 1997;106: 544–551.
6. Bloch I, Behrman A. Quantitative analysis of videostroboscopic images in presbylarynges. *Laryngoscope*. 2001;111:2022–2027.
7. Olthoff A, Woywod C, Kruse E. Stroboscopy versus high-speed glottography: a comparative study. *Laryngoscope*. 2007;117: 1123–1126.
8. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008;117:413–424.
9. Ford CN, Inagi K, Khidr A, et al. Sulcus vocalis: a rational analytical approach to diagnosis and management. *Ann Otol Rhinol Laryngol*. 1996;105:189–200.
10. Pontes P, Yamasaki R, Behlau M. Morphological and functional aspects of the senile larynx. *Folia Phoniatr Logop*. 2006;58:151–158.
11. Yamauchi A, Imagawa H, Yokonishi H, et al. Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: comparative study between healthy young and elderly subjects. *J Voice*. 2012;26: 742–750.
12. Yamauchi A, Imagawa H, Sakakibara K-I, et al. Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography. *J Voice*. 2013;27:39–45.
13. Yamauchi A, Yokonishi H, Imagawa H, et al. Quantitative analysis of digital videokymography: a preliminary study on age- and gender-related difference of vocal fold vibration in normal speakers. *J Voice*. 2014; [E-pub ahead of print].
14. Yamauchi A, Imagawa H, Sakakibara K-I, et al. Characteristics of vocal fold vibrations in vocally healthy subjects: analysis with multi-line kymography. *J Speech Lang Hear Res*. 2014;57:648–657.
15. Yamauchi A, Yokonishi H, Imagawa H, et al. Age- and gender-related difference of vocal fold vibration and glottal configuration in normal speakers: analysis with glottal area waveform. *J Voice*. 2014;28:525–531.
16. Woo P. Quantification of videostrobolaryngoscopic findings—measurements of the normal glottal cycle. *Laryngoscope*. 1996;106:1–27.
17. Timcke R, von Leden H, Moore P. Laryngeal vibrations: measurements of the glottic wave. *Arch Otolaryngol*. 1958;68:1–19.
18. Rodeno MT, Sanchez-Fernandez JM, Rivera-Pomar JM. Histochemical and morphometrical aging changes in humans vocal cord muscles. *Acta Otolaryngol*. 1993;113:445–449.
19. Filho JAX, Tsuji DH, do Nascimento PH, Sennes LU. Histologic changes in human vocal folds correlated with aging: a histomorphometric study. *Ann Otol Rhinol Laryngol*. 2003;112:894–898.
20. Titze IR. Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation. *J Speech Lang Hear Res*. 2006;49:439–447.
21. Mehta DD, Zanartu M, Quatieri TF, Deliyski DD, Hillman RE. Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy. *J Acoust Soc Am*. 2011;130:3999–4009.
22. Guzman M, Barros M, Espinoza F, Herrera A, Parra D, Muñoz D, Lloyd A. Laryngoscopic, acoustic, perceptual, and functional assessment of voice in rock singers. *Folia Phoniatr Logop*. 2013;65:248–256.

Reprinted by permission of Laryngoscope. 2014; 124(8):1869-1876.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Preliminary Data on Two Voice Therapy Interventions in the Treatment of Presbyphonia

Aaron Ziegler, MA (ABD), CCC-SLP; Katherine Verdolini Abbott, PhD, CCC-SLP; Michael Johns, MD; Adam Klein, MD; Edie R. Hapner, PhD, CCC-SLP

Objectives/Hypothesis: Presbyphonia is common among elderly individuals, yet few studies have evaluated behavioral treatment approaches for presbyphonia. The primary aim of this study was to assess the short-term efficacy of two types of voice therapy—vocal function exercises (VFE) and phonation resistance training exercise (PhoRTE) therapy—in the treatment of presbyphonia. The secondary aim was to determine if differences in adherence and treatment satisfaction existed between the two therapy approaches.

Study Design: Prospective, randomized, controlled.

Methods: Preliminary data from 16 elderly participants with presbyphonia randomly assigned to VFE, PhoRTE, or a no-treatment control group (CTL) were analyzed. Before and after a 4-week intervention period, participants completed the *Voice-Related Quality of Life* (V-RQOL) questionnaire and a perceived phonatory effort (PPE) task. Additionally, participants receiving treatment completed weekly practice logs and a posttreatment satisfaction questionnaire.

Results: Preliminary data revealed VFE and PhoRTE groups demonstrated a significant improvement in V-RQOL scores. However, only PhoRTE demonstrated a significant reduction in PPE, as suggested by the study's causal model. The CTL group did not demonstrate significant changes. Numerically, VFE registered slightly greater adherence to home practice recommendations than did PhoRTE, but PhoRTE perceived greater treatment satisfaction than VFE.

Conclusions: Findings provide new evidence regarding the efficacy of voice therapy exercises in the treatment of age-related dysphonia and suggest PhoRTE therapy as another treatment method for improved voice-related quality of life and reduced perceived vocal effort in this population.

Key Words: Aging, presbyphonia, voice disorder, treatment.

Level of Evidence: 2b.

***Laryngoscope*, 124:1869–1876, 2014**

INTRODUCTION

Presbyphonia is a common clinical finding among the elderly and poses a significant barrier to life satisfaction.[1,2] This voice disorder results from age-related laryngeal and respiratory degenerative changes, which lead to glottal incompetence[3] and a decline in inspiratory and expiratory pressures.[4] A deterioration in vocal function in the elderly has been putatively linked to a reduced amount and *intensity* of speech.[5] Interestingly, and analogous to findings of senior athletes,[6] the voice of elderly singers sounds younger, clearer, and louder than the elderly nonsinger's voice.[7] Additionally, both elderly male[8] and female[8,9] singers maintain a stable fundamental frequency throughout the lifespan. Those differences suggest the benefit of increased vocal activity for vocal longevity.

Current Evidence for Behavioral Treatment of Presbyphonia

Over the past decade, eight studies have been conducted on voice therapy for presbyphonia.[10–17] In brief, an overwhelming majority of patients with presbyphonia believe voice therapy is beneficial[15] and exhibit a significant improvement in voice-related quality of life,[13,14] a finding not observed in patients who forego voice therapy.[14] Furthermore, patients with presbyphonia report a significant decrease in phonatory effort after completing voice therapy.[13] Most important, patients with presbyphonia who receive voice therapy exhibit a significant improvement in their functional vocal status.[16]

To date, published prospective studies have only investigated the efficacy of voice therapy approaches for treating individuals with presbyphonia,[10–13,17] but none

From the Emory Voice Center, Department of Otolaryngology—Head and Neck Surgery, Emory University (A.Z., M.J., A.K., E.R.H.), Atlanta, Georgia; the Department of Communication Science and Disorders (K.V.); and the McGowan Institute for Regenerative Medicine (K.V.), University of Pittsburgh; the Center for Neural Basis of Cognition (K.V.), Carnegie-Mellon University and University of Pittsburgh, Pennsylvania; and note Aaron Ziegler is now Acting Assistant Professor in Communication Sciences and Disorders, John A. Burns School of Medicine, University of Hawaii at Manoa (A.Z.), Honolulu, Hawaii, U.S.A

Editor's Note: This Manuscript was accepted for publication November 27, 2013.

Presented at the Voice Foundation's 40th Annual Symposium, Philadelphia, PA, U.S.A., 1–5 June 2011.

Financial Disclosures: Aaron Ziegler, Katherine Verdolini Abbott, Michael Johns III, and Adam Klein-None; Edie Hapner—Plural Publishing. Conflict of Interest: Aaron Ziegler and Edie Hapner—Authors for Plural Publishing. The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Aaron Ziegler, MA (ABD), CCC-SLP, Department of Communication Sciences and Disorders, John A. Burns School of Medicine, University of Hawaii at Manoa, 677 Ala Moana Blvd, Suite 625, Honolulu, HI 96813. E-mail: aszieglerr@gmail.com

DOI: 10.1002/lary.24548

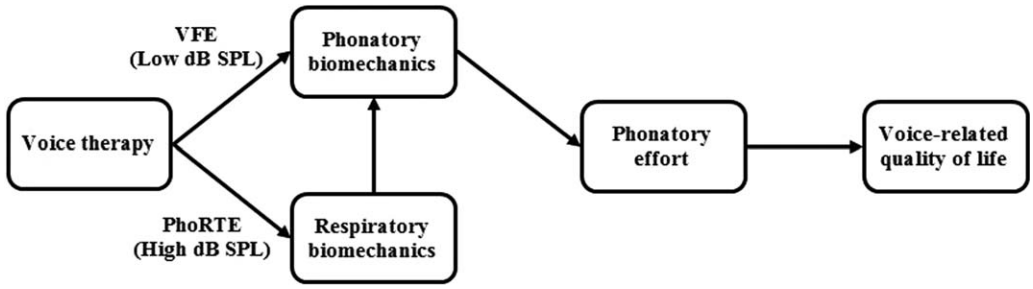

Fig. 1. Proposed flowchart delineating a causal model linking voice therapy to changes in phonatory and respiratory biomechanics, phonatory effort, and voice-related quality of life.

have compared voice therapy techniques to assess the superiority of one approach over another. Furthermore, the literature lacks suggestions for a causal model describing mechanisms of voice change from behavioral treatment of presbyphonia that may assess the potential differential impact of two types of voice therapy. Therefore, a causal model was developed, which stated that targeted voice therapy may affect phonatory biomechanics directly or indirectly through altered respiratory behavior resulting in reduced phonatory effort and lead to an improved voice-related quality of life (Fig. 1).

The causal model suggests that an effective therapeutic approach for presbyphonia will be one that targets the biological bases of the condition, or degenerative respiratory and laryngeal changes as a result of aging. These changes in muscle mass and strength—sarcopenia—are targeted in other parts of the body by engaging in structured exercise that emphasizes an increased level of physical activity to overload the muscle and reverse the sarcopenia process.[18] This type of exercise training—resistance training—has demonstrated positive effects on sarcopenia in older adults by reducing secondary aging effects that occur from muscle atrophy and weakness.[19]

Based on the causal model, it was hypothesized that the intervention groups in this study would result in more positive changes in voice across the experimental period than seen in a no-intervention control group. Furthermore, the causal model suggests that one therapy, a treatment requiring high-vocal intensity phonation and that loads both respiratory and laryngeal musculature, will result in more positive changes than the other therapy, a treatment requiring low vocal intensity phonation.

Study Aims

The purpose of this study was to compare two interventions and no treatment for adults with presbyphonia by using a prospective, randomized, controlled experimental design to assess the short-term efficacy of two voice therapy approaches, as demonstrated by a change in quality of life and perceived phonatory effort. Secondary aims of this study were to examine differences in patient adherence and treatment satisfaction.

MATERIALS AND METHODS

All procedures were approved by the institutional review boards at Emory University and the University of Pittsburgh (IRB #00037045 and #10060268, respectively). The experiment used a prospective, randomized, controlled design.

Participants

Twenty elderly adults aged 60 years and over enrolled in the study (Fig. 2). For this preliminary study, the sample size was selected arbitrarily to generate the necessary results for a power analysis for future studies.

All participants a) reported a current voice problem, including a complaint of reduced vocal loudness or increased vocal effort; b) received a diagnosis of presbyphonia by a fellowship-trained laryngologist[14]; c) received an auditory–perceptual diagnosis of vocal asthenia by a voice-specialized speech-language pathologist (SLP); d) were judged perceptually by a SLP to be free of dysarthria, dysfluency, or language problems; e) passed hearing, cognition, and mood screenings; f) were currently nonsmokers (five years or more); g) reported no progressive neuromuscular diseases affecting voice; h) denied concomitant health problems affecting voice; i) completed menopause, if female; j) reported using current medications for at least one month before participation; k) denied current use of inhaled corticosteroids or prednisone; and l) stated willingness to persist with the 6-week protocol. In addition, participants were included, if stimulable for improved voice quality as assessed by a SLP during the physician's examination visit. Stimulability testing is a routine part of the voice evaluation to determine candidacy for treatment.[20] No participants were excluded based on race, ethnicity, or gender. In accordance with standards on reporting randomized, controlled studies,[21] participant characteristics are provided in Table I.

Procedures

Recruitment, screening, and randomization. Recruitment was performed by a SLP who was part of the multidisciplinary team at the Emory Voice Center. An individual was initially seen for a comprehensive evaluation by a fellowship-trained laryngologist and SLP. Following informed consent, each individual underwent a hearing screening to ensure age-appropriate hearing or adequately managed sensory-neural hearing loss with the use of hearing aids, as evidenced by a response during audiometric testing in a sound-isolated booth at 40 dB HL at 0.5 kHz, 1 kHz, and 2 kHz presented in sound field.[22] Next, each individual underwent a screening to ensure age-appropriate cognitive ability based on results from the *Mini Mental State Examination* (MMSE).[23] A score of ≥ 20 was required for further participation in the study. Then, each individual underwent self-administration of the *Elderly Depression Scale-Short Form* (EDS-SF),[24] and a score of ≤ 5 was required for further participation. Finally, individuals satisfying inclusion criteria were randomized to one of three groups using a

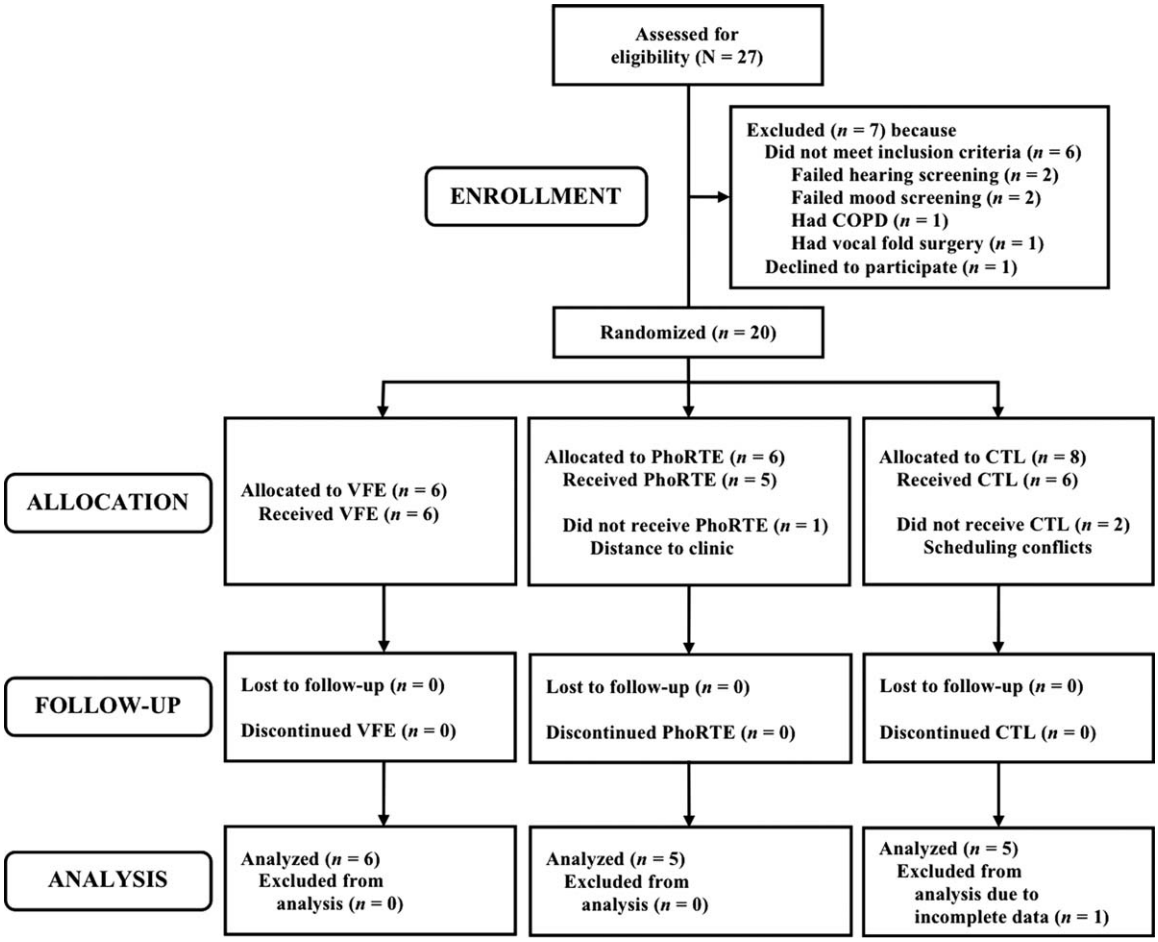

Fig. 2. Flowchart of study procedures.

computer algorithm: vocal function exercises (VFE), phonation resistance training exercise (PhoRTE) therapy, or a no-intervention control group (CTL). All participants were briefly counseled on voice hygiene and given a written copy of a handout that describes vocal hygiene recommendations.

Baseline and follow-up evaluations. At the baseline visit, each participant completed the V-RQOL.[25] Then, the participant was asked to provide an estimation of perceived phonatory effort (PPE). To determine PPE, the participant used a direct magnitude estimation scale[26] on which "100" represented "comfortable effort during phonation," "50" represented "half as much effort as comfortable," "200" represented "two times as much effort as comfortable," and so forth.[27,28]

Participants returned for follow-up measures within one week of completion of the intervention, or 6-weeks postbaseline in the case of the CTL group. At the follow-up visit, each participant completed the V-RQOL[25] and provided a rating of PPE, which were anchored to the participant's baseline ratings to limit drift due to increased awareness of voice. Finally, participants in the VFE and PhoRTE groups completed a post treatment satisfaction questionnaire.[29]

Interventions. Participants receiving an intervention attended four 45-minute treatment sessions—either VFE or PhoRTE—over the course of four weeks, which were provided by one of two participating voice-specialized SLPs. Execution of VFE[30,31] involved four exercises: 1) maximum sustained phonation on /i/ on the pitch F above middle C (males dropped down an octave); 2) an ascending glide over the entire pitch range on /oɬ/; 3) a descending glide over the entire pitch range on /oɬ/; and 4) maximum sustained phonation on the pitches middle C and D, E, F, and G above middle C (males dropped down an octave) on /oɬ/. Participants learned to use low abdominal breathing, a frontal focus with an inverted megaphone mouth shape, and were instructed to complete the exercises as quietly as possible but while maintaining a clear and consistent voice.

PhoRTE[32] (a homophone to the Italian word *forte* meaning loud and strong), adapted from Lee Silverman Voice Treatment (LSVT),[33–35] consisted of four exercises: 1) loud maximum sustained phonation on /a/; 2) loud ascending and descending pitch glides over the entire pitch range on /a/; 3) participant-specific functional phrases using a loud and high voice; and 4) phrases from exercise #3 in a loud and low voice. Low abdominal breathing gestures were encouraged. All feedback thereafter was limited to reminding participants to maintain a "strong" voice. During therapy sessions, participants were expected to maintain a SPL between 80 and 90 dB, as measured by a sound level meter positioned at a microphone-to-mouth distance of 30 cm.

PhoRTE, while derived from the therapeutic studies on LSVT, differed in several ways. First, PhoRTE sessions occurred once weekly as opposed to a more intensive intervention schedule for LSVT (i.e., four days per week for four weeks). Second, PhoRTE incorporated two different manners of producing participant-specific functional phrases (i.e., a loud and high voice and a loud and low voice),[36] Finally, PhoRTE home practice required fewer repetitions than is typically required for

TABLE I.
Summary of Participant Characteristics by Group.

Group/Participant	Sex	Age	Race
VFE			
1	female	83	Caucasian
3	male	66	Caucasian
9	female	74	Caucasian
10	male	78	Caucasian
13	male	78	Caucasian
17	male	60	Caucasian
Mean (SD), $n = 6$	2 females; 4 males	73.2 (8.6)	
PhoRTE			
6	male	79	Caucasian
7	female	78	Caucasian
8	female	72	Caucasian
11	female	80	Caucasian
20	male	71	Asian
Mean (SD), $n = 5$	3 females; 2 males	75.8 (4.0)	
CTL			
2	male	79	Caucasian
4	female	69	Caucasian
5	male	76	African American
14	female	91	Caucasian
15	male	73	Caucasian
Mean (SD), $n = 5$	2 females; 3 males	77.6 (8.4)	
Overall Mean (SD), $N = 16$		75.4 (7.2)1	

CTL = no-treatment control group; PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

patients receiving LSVT (two versus 10 repetitions of each exercise per practice session, respectively).

The PhoRTE exercises were selected because of their high intensity nature that might induce changes to muscle structure and function to reverse the degenerative sarcopenia process.[12] In addition, phonatory–resonatory interaction through a widened mouth and narrow pharynx, as occurs with the use of the vowel /a/, creates an acoustic situation that allows a speaker to shout safely. This megaphone mouth shape at low to medium high pitches raises the first formant frequency to reinforce the fundamental and second harmonic of the source. The resulting phonatory–resonatory interaction helps to recalibrate phonatory effort by assisting vocal fold vibration and maximizing phonatory efficiency. Furthermore, coupling a narrowed epilarynx tube with increased adduction provides maximum power transfer from the glottis to the lips to further increase vocal loudness.[37] Finally, the PhoRTE program subscribes to a task-dependent model of motor control by including functional phrases to help with generalization of voice techniques to conversation.[38]

Home practice program. Participants in both intervention groups were instructed to practice their respective treatments, VFE or PhoRTE, twice daily every day, to perform each exercise twice during each practice session, and to log their practice. Participants were instructed to complete practice logs only for completed exercises. From the practice log, the percent of prescribed exercises completed was computed to measure treatment adherence. The protocols of the two treatments controlled for what was assumed to be equivalent practice durations if the participant was adherent to the twice daily practice sessions. Participants received written instructions on how to complete daily home practice and a compact disc with audio demonstrations of the respective exercises.

RESULTS

Statistical Analysis

Inferential statistical analyses of the preliminary data were used to examine pretreatment to posttreatment changes within groups, and between group differences were examined descriptively for the primary outcome measures (i.e., V-RQOL and PPE). Inferential statistical analyses were also used to investigate between group differences in the secondary outcome measures (i.e., treatment adherence and treatment satisfaction). Due to the preliminary nature of this study and the small sample size, an alpha level of 0.10 was used to minimize the type II error rate in analyzing treatment effects on primary and secondary outcome measures. Of the 20 enrolled participants, only 16 participants were included in the data set for analysis. Of the four who were excluded, three dropped out of the study prior to data collection and one participant in the no-treatment control group had an incomplete data set. Therefore, data from six VFE participants, five PhoRTE participants, and five CTL participants were analyzed.

Participant Characteristics

Participants were seven women (44%) and nine men (56%) aged 60 to 91 years ($M = 75.4$ years, SD = 7.2). Post-hoc analyses using Fisher's exact test and between-subject ANOVAs confirmed the equivalence of groups on gender ($P = .825$, Fisher's Exact Test), age ($F[2, 13] = 0.501$, $P = .617$, $\eta_p^2 = .072$), baseline V-RQOL scores ($F[2, 13] = 0.880$, $P = .438$, $\eta_p^2 = .119$), and baseline PPE ratings ($F[2, 13] = 1.948$, $P = .182$, $\eta_p^2 = .231$) (Tables (I–III)).

V-RQOL

Individual scores, group means and standard deviations, difference scores, and percent change values for the V-RQOL data before and following the 4-week intervention period are displayed in Table II. Results revealed that the VFE and PhoRTE groups experienced a significant improvement in mean pretreatment to posttreatment V-RQOL scores (80.8 to 87.5, $t[5] = 1.964$, $P = .054$, one-tailed, $d = 0.80$ and 88.5 to 95.0, $t[4] = 2.152$, $P = .049$, one-tailed, $d = 0.96$, respectively). The CTL group did not demonstrate a significant change in mean V-RQOL scores (87.5 to 91.5, $t[4] = 1.554$, $P = .195$, $d = 0.70$).

The data were reanalyzed after excluding a PhoRTE participant who commenced therapy without registering quality of life impairment (as evidenced by a score of 100 on the V-RQOL). Removal increased the PhoRTE percent change value (8.03 to 10.66), and it was slightly greater than that of the VFE group (9.30).

TABLE II.
Individual Scores, Mean Pretreatment and Posttreatment Scores, Standard Deviations, Percent Change, and P Values for the VFE, PhoRTE, and CTL Groups on the *Voice-Related Quality of Life*.

Group/Participant	Baseline (Pretreatment)	Follow-Up (Posttreatment)	Absolute Difference	Percent Change	Test Statistic	*P* Value
VFE						
1	80.0	85.0	5.0	6.25		
3	90.0	90.0	0.0	0.00		
9	62.5	85.0	22.5	36.00		
10	90.0	97.5	7.5	8.33		
13	92.5	97.5	5.0	5.41		
17	70.0	70.0	0.0	0.00		
Mean (SD), $n = 6$	80.8 (12.3)	87.5 (10.2)	6.7 (8.3)	9.30 (13.5)	$t = 1.964$**	.054*
PhoRTE						
6	97.5	100.0	2.5	2.56		
7	82.5	97.5	15.0	18.18		
8	75.0	85.0	10.0	13.33		
11	87.5	95.0	7.5	8.57		
20	100.0	97.5	−2.5	−2.50		
Mean (SD), $n = 5$	88.5 (10.4)	95.0 (5.9)	6.5 (6.8)	8.03 (8.25)	$t = 2.152$**	.049*
CTL						
2	90.0	92.5	2.5	2.78		
4	95.0	90.0	−5.0	−5.26		
5	75.0	82.5	7.5	10.00		
14	85.0	95.0	10.0	11.76		
15	92.5	97.5	5.0	5.41		
Mean (SD), $n = 5$	87.5 (7.9)	91.5 (5.8)	4.0 (5.8)	4.94 (6.73)	$t = 1.554$**	.195
Overall Mean (SD), $N = 16$	85.3 (10.4)					

Note. *Significant difference at $P \leq 0.10$ level, one-tailed.
**From repeated-measures *t* test.
CTL = no-treatment control group; PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

PPE

Individual ratings, group means and standard deviations, difference scores and percent change values for PPE ratings before and following the 4-week intervention period are shown in Table III. Results showed that PPE ratings decreased significantly in the PhoRTE group only (144 to 102, $t[4] = -2.370$, $P = .077$, two-tailed, $d = -1.06$). Neither the VFE group nor the CTL group demonstrated a significant difference in PPE ratings (142.5 to 109.2, $t[5] = -1.865$, $P = .121$, two-tailed, $d = -0.76$; 101 to 103, $t[4] = 1.000$, $P = .374$, two-tailed, $d = 0.45$, respectively).

Adherence and Treatment Satisfaction

Participants in the VFE and PhoRTE groups demonstrated adherence to treatment recommendations, and no differences were detected between groups ($P = .411$). One participant in the PhoRTE group practiced significantly less than any other participant and skewed the averaged data for adherence. A post-hoc analysis of the data removing this participant from the PhoRTE data resulted in a more balanced assessment of the practice patterns of the PhoRTE group, 88.2%, nearly equivalent to the average practice of the VFE group (89.3%). Results for treatment satisfaction data revealed no differences in ratings between VFE and PhoRTE on the three questions: extent to which participants a) liked the particular therapy ($P = .285$); b) felt voice changed because of therapy ($P = .227$); and c) felt voice changes were caused by the particular therapy ($P = .550$) (Table IV).

DISCUSSION

The data from this study provide optimism that there may be short-term benefits from two therapy approaches, VFE and PhoRTE, for improvement of voice-related quality of life in elderly individuals with presbyphonia. The causal model tested in this study proposed that therapy-induced changes in laryngeal biomechanics, possibly partly related to changes in respiratory biomechanics, would lead to a reduction in perceived phonatory effort and, ultimately, result in an improvement in voice-related quality of life. Significant pretreatment to posttreatment increases were documented in V-RQOL scores for both intervention groups, in comparison to scores for a no-treatment control group, which did not improve. The magnitude of pretreatment to posttreatment differences on the V-RQOL in each treatment group (VFE and PhoRTE) exceeded changes in an untreated group of elderly individuals with presbyphonia. The improvement of patient-reported outcome measures in a group of elderly individuals with presbyphonia

TABLE III.
Individual and Mean Pretreatment and Posttreatment Ratings, Standard Deviations, Difference Scores, Percent Change, and *P* values for the VFE, PhoRTE, and CTL Groups on Perceived Phonatory Effort.

Group/Participant	Baseline (Pretreatment)	Follow-Up (Posttreatment)	Absolute Difference	Percent Change	Test Statistic	*P* Value
VFE						
1	125	100	−25.0	−20.0		
3	100	100	0.0	00.0		
9	150	100	−50.0	−33.3		
10	200	100	−100.0	−50.0		
13	100	125	−25.0	25.0		
17	180	130	−50.0	−27.8		
Mean (SD), $n = 6$	142.5 (41.7)	109.2 (14.3)	−33.3 (43.8)	−17.7 (26.6)	$t = -1.865$**	.121
PhoRTE						
6	100	100	0.0	00.0		
7	100	50	−50.0	−50.0		
8	200	150	−50.0	−25.0		
11	200	100	−100.0	−50.0		
20	120	110	−10.0	−8.3		
Mean (SD), $n = 5$	144 (51.8)	102 (35.6)	−42.0 (39.6)	−26.7 (23.1)	$t = -2.370$**	.077*
CTL						
2	100	100	0.0	00.0		
4	100	100	0.0	00.0		
5	125	125	0.0	00.0		
14	100	100	0.0	00.0		
15	80	90	10.0	12.5		
Mean (SD), $n = 5$	101 (16.0)	103 (13.0)	2.0 (4.5)	2.5 (5.6)	$t = 1.000$**	.374
Overall Mean (SD), $N = 16$	130 (42.1)					

Note. *Significant difference at $P \leq 0.10$ level, two-tailed.
**From repeated-measures *t* test.
CTL = no-treatment control group; PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

following voice therapy is consistent with results from prior research (Berg et al., 2008; Sauder et al., 2010).

Significant improvement in perceived phonatory effort accompanied voice-related quality of life changes for the PhoRTE group, but not the VFE group, a finding that partially supports the causal model explored in this study, and moreover, that can also be inferred from previous research in a similar cohort.[13] Differences in PPE pretreatment to posttreatment changes between VFE and PhoRTE may be explained by unique vocal tract configurations and their influence on vocal fold vibration. Whereas VFE are characterized by an inverted megaphone-shaped vocal tract, PhoRTE therapy employs a megaphone-shaped vocal tract. Consistent with nonlinear dynamics, rounded vowels such as /o, u/ using a wide open pharynx as in the case of VFE, have been shown to decrease vocal fold adduction. Open vowels such as /a, æ/ using a narrow pharynx and high larynx, as in PhoRTE, have been shown to cause greater vocal fold adduction. In the population of interest, increased adduction is a desired laryngeal target. Perhaps a reduction in the glottal half-width due to increased adduction lowered the required subglottal pressure and resulted in a decrease in perceived phonatory effort.[37,39]

Whereas improvement in V-RQOL scores was accompanied by numerical decreases in PPE in both treatment groups, the no-treatment control group exhibited the opposite finding. For that group, pre- to posttreatment PPE actually increased slightly, even with anchoring the posttreatment estimation of phonatory effort to pretreatment ratings. In light of that finding, elderly individuals who forego therapy seem to employ increased muscle tension at the level of the glottis to achieve phonatory closure during voicing.

Given these preliminary findings, PhoRTE may have a slight advantage over VFE for producing benefit from a physiologic perspective because it demands a higher intensity of effort, which better addresses the overload principle required to induce neuromuscular changes in strength.[40] Increased neuromuscular activity of both the respiratory and laryngeal systems from PhoRTE should lead to even greater improvement in respiratory and laryngeal biomechanics than VFE, ultimately causing a significant reduction in PPE. Furthermore, phonatory efficiency from a megaphone-shaped vocal tract configuration may have also contributed to decreased phonatory effort.[37] Additionally, inclusion of task-specific exercises, as used in PhoRTE, to address the exercise training principle of specificity and promote carryover may result in a greater change in respiratory and laryngeal biomechanics during conversational speech. Consequently, phonatory effort for the PhoRTE group should demonstrate a larger change than VFE.

TABLE IV.
Individual and Group Means, Standard Deviations, and P Values for the VFE and PhoRTE Groups on Weekly Practice Log (% completed) and Posttreatment Satisfaction Questionnaire.

Group/Participant	Adherence Week 1–4	Treatment Satisfaction		
		Like Therapy	Voice Change	Therapy Cause
VFE				
1	78.0	4	4	2
3	79.6	3	3	1
9	100.0	4	5	3
10	95.8	3	4	3
13	87.5	3	4	2
17	94.8	3	4	3
Mean (SD), $n = 6$	89.3 (9.0)	3.3 (.52)	3.9 (.66)	2.3 (.82)
PhoRTE				
6	100.0	3	4	2
7	17.5	3	5	3
8	56.3	4	4	2
11	96.5	5	4	3
20	100.0	4	5	3
Mean (SD), $n = 5$	74.1 (36.6)	3.8 (.84)	4.4 (.55)	2.6 (.55)
Test statistic	t (4.407) = 0.908*	t (9) = −1.137*	t (9) = −1.297*	t (9) = −0.621*
P value, two-tailed	.411	.285	.227	.550

Note. For "like therapy" scale, 1 = not at all; 2 = somewhat; 3 = moderate; 4 = very much; 5 = extremely. For "voice change" scale, 1 = got a lot worse; 2 = got a little worse; 3 = no change; 4 = got a little better; 5 = got a lot better. For "therapy cause" scale, 1 = voice therapy probably irrelevant to voice change; 2 = voice therapy may have caused voice changes; 3 = voice therapy definitely caused voice changes.

*From independent samples *t* test.

PhoRTE = phonation resistance training exercise; SD = standard deviation; VFE = vocal function exercises.

In addition to the foregoing results, this study investigated adherence to home treatment recommendations in this population. Participants in both VFE and PhoRTE appeared to exhibit fairly regular practice of their home programs, a finding that is consistent with published literature.[17] Although self-report may be inaccurate, in the absence of any clear difference in mean practice between the VFE and PhoRTE groups, the most straightforward interpretation is that improvements in V-RQOL are not likely strongly related to treatment adherence.

Accordingly, although not significant, PhoRTE practiced less than VFE and yet consistently perceived greater satisfaction with the therapy they received. This finding supports a model of voice therapy in which treatment efficacy is optimized by a combination of biomechanical, learning, and adherence factors.[41] Specifically, the high intensity component of PhoRTE may necessitate less practice time than VFE to generate neuromuscular changes in muscle strength. Furthermore, the inclusion of functional speech tasks may promote fast learning because it addresses task-specificity and generalization to extra-therapy situations. In addition, practice of functional speech tasks for transfer of therapy techniques to unique communication situations, as well as the emphasis on increased vocal intensity to addresses a key patient concern— reduced loudness— may both increase self-efficacy and lead to improved treatment adherence.

Limitations and Future Aims

This study was designed to develop preliminary data to support the use of voice therapy for a subset of people with voice complaints secondary to presbylaryngeus. It was also designed to support the use of an alternative therapy that was based on high-intensity vocal exercise in the treatment of presbyphonia. Accordingly, one of the aims of the study was to develop an effect size for future research into the therapeutic treatment of presbyphonia. A limitation of this study is thus the small number of participants. Yet another limitation, although a no-treatment control group was included in the experimental design to determine the influence of time, was the lack of an experimental *treatment* control group, which would have provided evidence on whether the perceived change was due to a placebo effect. Additionally, a longitudinal study that follows participants for more than six weeks is necessary to assess maintenance of treatment effects. Future studies should include a larger sample size, incorporate a placebo treatment, and follow participants longitudinally. In addition, future studies should assess differences in vocal load between VFE and PhoRTE, as well as pre- to posttreatment changes in acoustic and aerodynamic parameters.

CONCLUSION

Indications from this study on voice therapy in individuals with presbyphonia are that behavioral

approaches are effective in the management of age-related voice problems. The study provides further preliminary evidence that individuals with presbyphonia may benefit from various therapeutic approaches for which patients express treatment satisfaction. Finally, this study contributes additional support to a previous finding that individuals with presbyphonia regularly practice voice exercises and exhibit good adherence to treatment recommendations.

BIBLIOGRAPHY

1. Golub JS, Chen PH, Otto KJ, Hapner E, Johns III MM. Prevalence of perceived dysphonia in a geriatric population. *J Am Geriatr Soc* 2006;54: 1736–1739.
2. Roy N, Stemple J, Merrill RM, Thomas L. Epidemiology of voice disorders in the elderly: preliminary findings. *Laryngoscope* 2007;117:628–633.
3. Linville SE. Vocal Aging. Clifton Park, NY: Delmar Cengage Learning; 2001.
4. Miller MR. Structural and physiological age-associated changes in aging lungs. *Semin Respir Crit Care Med* 2010;31:521–527.
5. Bastian RW, Thomas JP. Talkativeness and vocal loudness: do they correlate with laryngeal pathology? A study of the vocal overdoer/underdoer continuum. 29th Annual Symposium of the Voice Foundation: Care of the Professional Voice, Philadelphia, PA; 2000.
6. Tanaka H, Seals DR. Invited Review: Dynamic exercise performance in Masters athletes: insight into the effects of primary human aging on physiological functional capacity. *J Appl Physiol* 2003;95:2152–2162.
7. Prakup B. Acoustic measures of the voices of older singers and nonsingers. *J Voice* 2012;26:341–350.
8. Brown Jr WS, Morris RJ, Hollien H, Howell E. Speaking fundamental frequency characteristics as a function of age and professional singing. *J Voice*. 1991;5:310–315.
9. Brown Jr WS, Morris RJ, Hicks DM, Howell E. Phonational profiles of female professional singers and nonsingers. *J Voice* 1993;7:219–226.
10. Gorman S, Weinrich B, Lee L, Stemple JC. Aerodynamic changes as a result of vocal function exercises in elderly men. *Laryngoscope* 2008;118:1900–1903.
11. LaGorio LA, Carnaby-Mann GD, Crary MA. Treatment of vocal fold bowing using neuromuscular electrical stimulation. *Arch Otolaryngol Head Neck Surg* 2010;136:398–403.
12. Ramig LO, Gray S, Baker K, et al. The aging voice: a review, treatment data and familial and genetic perspectives. *Folia Phoniatr Logop* 2001; 53:252–265.
13. Sauder C, Roy N, Tanner K, Houtz DR, Smith ME. Vocal function exercises for presbylaryngis: a multidimensional assessment of treatment outcomes. *Ann Otol Rhinol Laryngol* 2010;119:460–467.
14. .Berg EE, Hapner E, Klein A, Johns III MM. Voice therapy improves quality of life in age-related dysphonia: a case-control study. *J Voice* 2008;22: 70–74.
15. Gartner-Schmidt J, Rosen C. Treatment success for age-related vocal fold atrophy. *Laryngoscope* 2011;121:585–589.
16. Mau T, Jacobson BH, Garrett CG. Factors associated with voice therapy outcomes in the treatment of presbyphonia. *Laryngoscope* 2010;120: 1181–1187.
17. Tay EY, Phyland DJ, Oates J. The effect of vocal function exercises on the voices of aging community choral singers. *J Voice* 2012;26:672 e619–627.
18. Freiberger E, Sieber C, Pfeifer K. Physical activity, exercise, and sarcopenia–future challenges. *Wien Med Wochenschr* 2011;161:416–425.
19. American College of Sports M, Chodzko-Zajko WJ, Proctor DN, et al. American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc* 2009;41: 1510–1530.
20. .ASHA. Voice Evaluation Templates. http://www.asha.org/uploadedFiles/slp/healthcare/AATVoiceEvaluation.pdf. Accessed May 31, 2013.
21. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–694.
22. Cruickshanks KJ, Wiley TL, Tweed TS, et al. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. *Am J Epidemiol* 1998;148:879–886.
23. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
24. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Disabil Rehabil* 1986;5:165–173.
25. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice* 1999;13:557–569.
26. Stevens SS. Psychophysics: introduction to its perceptual, neural, and social prospects. New York, NY: Wiley & Sons; 1975.
27. Colton RH, Brown JWS. Some relationships between vocal effort and intraoral air pressure. *J Acoust Soc Am* 1973;53:296–296.
28. Wright HN, Colton RH. Some parameters of vocal effort. *J Acoust Soc Am* 1972;51:141.
29. Nanjundeswaran C, Li NY, Chan KM, Wong RK, Yiu EM, Verdolini-Abbott K. Preliminary data on prevention and treatment of voice problems in student teachers. *J Voice* 2012;26:816 e811–816, e812.
30. Stemple JC, Lee L, D'Amico B, Pickup B. Efficacy of vocal function exercises as a method of improving voice production. *J Voice* 1994;8:271–278.
31. Wrycza Sabol J, Lee L, Stemple JC. The value of vocal function exercises in the practice regimen of singers. *J Voice* 1995;9:27–36.
32. Ziegler A, Hapner E. Phonation Resistance Training Exercise (PhoRTE) Therapy. In: Behrman A, Haskell J, eds. Exercises for Voice Therapy. 2 ed. San Diego, CA: Plural Publishing; 2013.
33. Ramig LO, Countryman S, Thompson LL, Horii Y. Comparison of two forms of intensive speech treatment for Parkinson Disease. *J Speech Hear Res* 1995;38:1232–1251.
34. Smith ME, Ramig LO, Dromey C, Perez KS, Samandari R. Intensive voice treatment in Parkinson Disease: laryngostroboscopic findings. *J Voice* 1995;9:453–459.
35. Ramig LO, Bonitati CM, Lemke JH, Horii Y. Voice treatment for patients with Parkinson Disease: Development of an approach and preliminary efficacy data *J Med Speech Lang Pathol* 1994;2:191–209.
36. de Swart BJ, Willemse SC, Maassen BA, Horstink MW. Improvement of voicing in patients with Parkinson's Disease by speech therapy. *Neurology* 2003;60:498–500.
37. Titze IR, Verdolini Abbott K. V*ocology: The Science and Practice of Voice Habilitation*. Salt Lake City, UT: National Center for Voice and Speech; 2012.
38. Ziegler W. Review. *Aphasiology* 2003;17:3–36.
39. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am* 1988;83:1536–1552.
40. McArdle WD, Katch FI, Katch VL. *Exercise Phsyiology: Nutrition, Energy, and Human Performance*. 7th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2010.
41. Verdolini-Marston K, Burke MK, Lessac A, Glaze L, Caldwell E. Preliminary study of two methods of treatment for laryngeal nodules. *J Voice* 1995;9:74–85.

Reprinted by permission of Laryngoscope. 2016; 126(1):135-141.

The Laryngoscope
© 2015 The American Laryngological, Rhinological and Otological Society, Inc.

Systematic Review

Cricopharyngeal Dysfunction: A Systematic Review Comparing Outcomes of Dilatation, Botulinum Toxin Injection, and Myotomy

Pelin Kocdor, MD; Eric R. Siegel, MS; Ozlem E. Tulunay-Ugur, MD

Objectives: Cricopharyngeal dysfunction may lead to severe dysphagia and aspiration. The objective of this systematic review was to evaluate the existing studies on the effectiveness of myotomy, dilatation, and botulinum toxin (BoT) injection in the management of cricopharyngeal dysphagia.

Methods: PubMed and Web of Science databases were searched to identify eligible studies by using the terms "cricopharyngeal dysfunction," "cricopharyngeal myotomy," "cricopharyngeal botox," "cricopharyngeal dilation," and their combinations from 1990 to 2013. This was supplemented by hand-searching relevant articles. Eligible articles were independently assessed for quality by two authors. Statistical analysis was performed.

Results: The database search revealed 567 articles. Thirty-two articles met eligibility criteria and were further evaluated. The reported success rates of BoT injections was between 43% and 100% (mean = 76%), dilation 58% and 100% (mean = 81%), and myotomy 25% and 100% (mean = 75%). In logistic regression analysis of the patient-weighted averages, the 78% success rate with myotomy was significantly higher than the 69% success rate with BoT injections ($P = .042$), whereas the intermediate success rate of 73% with dilation was not significantly different from that of either myotomy ($P = .37$) or BoT ($P = .42$). There was a statistically significant difference between endoscopic and open myotomy success rates ($P = .0025$). Endoscopic myotomy had a higher success rate, with a 2.2 odds ratio.

Conclusions: The success rate of myotomy is significantly higher than the success rate of BoT injections in cricopharyngeal dysfunction. Moreover, endoscopic myotomy was found to have a higher success rate compared to open myotomy.

Key Words: Cricopharyngeal dysfunction, cricopharyngeal myotomy, cricopharyngeal botox, cricopharyngeal dilation.

Level of Evidence: NA

***Laryngoscope*, 00:000–000, 2015**

INTRODUCTION

Cricopharyngeal (CP) muscle dysfunction can lead to dysphagia, aspiration, and weight loss, causing significant morbidity and reduced quality of life.[1] Etiologies are numerous and include the general categories of anatomic (cricopharyngeal bar), neuromuscular (central, peripheral, or myogenic), iatrogenic, inflammatory, neoplastic, and idiopathic (Table I).[2] The role of the CP muscle in swallowing has been well established. In 1717, Valsalva first described the anatomy of the cricopharyngeus muscle, which was further clarified by Killian in 1907.[3] CP dysfunction has been attributed mainly to the disordered opening of the CP muscle, which is the main component of the upper esophageal sphincter (UES). The opening of the UES necessitates three factors: neural inhibition of tonic intrinsic sphincter muscle contraction, anterior-superior laryngeal elevation that leads to the mechanical distraction of the UES, and passive stretching of the intrinsic sphincter muscles as the bolus passes.[4,5] A heterogeneous spectrum of disorders can lead to CP dysfunction, including failure of neural inhibition of tonic CP contraction, weakness of pharyngeal muscles with reduced laryngeal elevation and UES opening, as well as decreased compliance of the CP muscle, such as due to radiation fibrosis.

Various preoperative techniques can be used for diagnosis (Table II). The most important component has been a thorough history. In most centers this is followed by a videofluoroscopic swallowing study (VFSS) and manometry. These not only demonstrate the dysfunctional UES, but also demonstrate laryngeal elevation, the strength of the pharyngeal muscles, and laryngeal penetration or aspiration. Although some authors find manometry cumbersome and of limited value,[6,7] others strongly advocate the use of it, especially if coupled with fluoroscopy.[8–11] Manofluoroscopy, which ensures improved sensor placement, also allows assessment of pressures at known sensor locations during swallowing.[10,12,13] It is still

From the Department of Otolaryngology–Head and Neck Surgery (P.K., O.E.T.-U.); and the Department of Biostatistics (E.R.S.), University of Arkansas for Medical Sciences, Little Rock, Arkansas, U.S.A.

Editor's Note: This Manuscript was accepted for publication May 21, 2015.

Presented at American Laryngological Association Annual Meeting, Boston, Massachusetts, U.S.A., April 22–23, 2015.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Ozlem E. Tulunay-Ugur, MD, Associate Professor, Director Division of Laryngology, University of Arkansas for Medical Sciences, Department of Otolaryngology–Head and Neck Surgery, 4301 W. Markham, Slot 543, Little Rock, AR 72205.
E-mail: oetulunayugur@uams.edu

DOI: 10.1002/lary.25447

TABLE I.
Causes of Cricopharyngeal Dysfunction.

Central nervous system
Cerebellar infarct
Brain stem infarct
Parkinsonism
Amyotrophic lateral sclerosis
Base of skull neoplasm
Peripheral nervous system
Peripheral neuropathy
Diabetic neuropathy
Bulbar poliomyelitis
Myasthenia gravis
Neoplasm
Cricopharyngeal muscle
Polymyositis
Oculopharyngeal muscular dystrophy
Hyperthyroidism
Hypothyroidism
Cricopharyngeal disruption
Laryngectomy
Supraglottic laryngectomy
Radical oropharyngeal resections
Pulmonary resections
Cricopharyngeal spasm
Hiatal hernia
Gastroesophageal reflux
Idiopathic cricopharyngeal achalasia

Adapted from Halvorson DJ.[30]

not available and not a part of the workup in many institutions. Poirier et al. advocate the use of manometry to assess the physiological abnormalities at the pharyngoesophageal junction, but do not use it as an indication for surgical treatment.[14] Electromyography has been used by some authors to diagnose swallowing disorders.[15,16]

Numerous treatments exist for CP dysfunction, including swallowing therapy, CP dilation, injection of botulinum toxin, and CP myotomy. The traditional surgical treatment for CP dysfunction has been CP myotomy through a transcervical approach. To minimize the complications of an open approach, endoscopic CP myotomy was introduced using the potassium-titanyl-phosphate laser (wavelength, 532 nm) by Halvorson and Kuhn in 1994.[17] Subsequently, carbon dioxide laser (wavelength, 10,600 nm) gained favor because of its ability to coagulate small vessels and minimize thermal damage.[1]

Blitzer and Brin first presented on the use of in-office botulinum toxin (BoT) injections in 1993 as an alternative to surgery for the treatment of UES dysfunction.[18] In most cases, BoT has been injected under endoscopic visualization and general anesthesia, whereas less has been reported on percutaneous BoT injections under electromyographic guidance and local cutaneous anesthesia.[19] The range of BoT doses reported per injection varies from 10 U to 100 U.[20] Bougienage has been used in the treatment of anatomic esophageal strictures for decades.[21] The commonly used approaches are bougies, wire-guided polyvinyl dilators, air-filled pneumatic dilatation, and water-filled balloon dilatation with or without endoscopy guidance.[22]

CP dysfunction can be challenging diagnostically and in regard to the identification of the best treatment modality for a given patient. The scope of this article was to systematically review the literature regarding CP muscle interventions, specifically myotomy, injection of BoT, and dilation of the CP muscle for the treatment of CP dysfunction in adult patients.

MATERIALS AND METHODS

The literature search was performed according to the guidelines of the Cochrane Collaboration for systematic reviews in PubMed and Web of Science using a time frame from January 1990 until March 2013. Only literature published in English was considered. The search included the following keywords: "cricopharyngeal dysfunction," "cricopharyngeal myotomy," "cricopharyngeal botox," "cricopharyngeal dilation," and their combinations. The inclusion criterion for the studies was for the main focus of the article to be on the success rate and complications of the treatment modality. Bibliographies were manually reviewed to obtain additional articles of relevance. Reviews, editorials, case reports with less than four patients, articles with nonhuman data, duplicate publications, and articles on the pediatric patient population were excluded. Articles describing CP dysfunction attributed directly to Zenker's diverticulum and/or requiring diverticulectomy were also excluded. Articles with one specific etiology (except CP achalasia) as the reason for cricopharyngeal dysfunction were excluded; articles with heterogeneous etiology were included in the study.

The eligible articles were assessed for quality using the modified Downs and Black scale,[23] which is a validated checklist for randomized and nonrandomized studies. Any data extraction or assessment disagreements or inconsistencies were resolved through discussion and consensus.

Statistical Analysis

The average success rate of each procedure was calculated two ways: 1) as the crude (unweighted) average of reported success rates across articles and 2) as the patient-weighted average calculated as the total number of reported successes divided by the total number of treated patients. For logistic regression, the events/trials syntax was used, in which "events" and "trials" respectively represented the number of successes and number of patients in each article; this means that the logistic regression was effectively comparing patient-weighted averages between procedures. Additionally, the procedures were scored for invasiveness as botulinum toxin = low, dilation = medium, and myotomy = high, and the trend in success rate with invasiveness was assessed via the Cochran-Armitage trend test. These analyses assessing success rates were also used for complication rates. SAS version 9.3 (SAS Institute, Cary, NC) was employed for all analyses, and a $P < .05$ significance level was employed for all comparisons.

RESULTS

Study selection identified 567 reference articles; of these 42 met eligibility criteria. An additional five potential relevant reports were identified through scanning reference lists. Ultimately, 32 articles were included in the analysis. Thirteen articles were excluded for the

TABLE II.
Selected Studies of Cricopharyngeal Muscle Intervention for Cricopharyngeal Dysfunction

Author	Year	Study (Level of Evidence)	No. of Patients	Outcome Measures	Type of Treatment	Follow-up	Success Rate (%)	Complication
Schneider[24]	1994	PCS (IV)	7	SR, VFSS, M	BoT-A	4.5–5 months	71	None
Blitzer[18]	1997	RCS (IV)	6	C	BoT-A	Not mentioned	100	Not mentioned
Alberty[4]	2000	PCS (IV)	10	SR, VFSS	BoT-A	1–1.5 months	100	Pharyngeal diffusion
Haapaniemi[34]	2001	RCS (IV)	4	C	BoT-A	2–24 months	75	Urine retention
Shaw[35]	2001	RCS (IV)	12	C	BoT-A	1–14 months	83	Pharyngeal diffusion, pharyngeal tear
Parameswaran[33]	2002	RCS (IV)	12	C, VFSS	BoT-A	3–6 months	92	Neck cellulitis
Zaninotto[25]	2004	PCS (IV)	21	SR, VFSS	BoT-A	12–38 months	43	Death through aspiration
Murry[26]	2005	PCS (IV)	13	C, VFSS	BoT-A	1–9 months	92	Not mentioned
Terre[27]	2008	PCS (IV)	10	C, VFSS, M	BoT-A	12 months	80	None
Lee[36]	2009	RCS (IV)	8	VFSS	BoT-A	0.2–1 month	75	Not mentioned
Alfonsi[28]	2010	PCS (IV)	34	SR	BoT-A	2 months	50	None
Woisard-Bassols[29]	2013	PCS (IV)	11	SR, VFSS	BoT-A	12–48 months	45	None
St. Guily[37]	1994	RCS (IV)	11	C	Myotomy (open)	5–53 months	72	None
Herberhold[38]	1995	RCS (IV)	32	C, VFSS	Myotomy (endoscopic)	Up to 7 years	97	Supraglottic edema, imminent mediastinitis
Poirier[14]	1997	RCS (IV)	40	C, VFSS, M	Myotomy (open)	1–255 months	72.5	Retropharyngeal hematoma
Ali[31]	1997	Cohort study (IIb)	8	C	Myotomy (open)	6 weeks	75	Not mentioned
Halvorson[30]	1998	PCS (IV)	18	C	Myotomy (endoscopic)	Not mentioned	78	Not mentioned
Mason[39]	1998	RCS (IV)	31	C, M	Myotomy (open)	2–48 months	77	Neck hematoma, pulmonary edema, pneumonia
Lawson[40]	2003	RCS (IV)	29	C, VFSS, FEES	Myotomy (endoscopic)	1–36 months	88	None
Zaninotto[25]	2004	PCS (IV)	11	SR, VFSS	Myotomy (open)	6–31 months	73	None
Takes[41]	2005	RCS (IV)	10	SR	Myotomy (endoscopic)	2–24 months	60	None
Dauer[2]	2006	RCS (IV)	22	SR	Myotomy (endoscopic + open)	Not mentioned	58	Fever of unknown etiology, chest pain, pharyngocutaneous fistula, tracheotomy
Munoz[5]	2007	RCS (IV)	14	SR, VFSS	Myotomy (open)	6–10 months	25	Not mentioned
Lawson[42]	2008	RCS (IV)	31	SR, VFSS, FEES	Myotomy (endoscopic)	12–23 months	64.5	None
Kos[9]	2010	RCS (IV)	28	VFSS, M	Myotomy (open)	2.5–203 months	79	Fever, aspiration pneumonia, mucosal perforation
Ozgursoy[10]	2010	RCS (IV)	14	SR, VFSS, MF	Myotomy (endoscopic)	6 months	100	Not mentioned
Bachy[32]	2013	PCS (IV	32	SR	Myotomy (endoscopic)	6–99 months	84	Severe bleeding
Lim[6]	1995	RCS (IV)	40	C, VFSS	Myotomy (endoscopic)	2–22 months	90	Esophageal perforation
Ali[31]	1997	Cohort study (IIb)	12	C	Dilatation (Savary)	6 weeks	58	Not mentioned
Hatlebakk[8]	1998	PCS (IV)	10	SR, M	Dilatation (Savary)	6–20 months	90	Not mentioned
Solt[43]	2001	RCS (IV)	5	C, VFSS, M	Dilatation (balloon)	7–33 months	100	Superficial mucosal splitting
Wang[44]	2005	RCS (IV)	6	SR	Dilatation (balloon + French)	8–27 months	100	None
Clary[21]	2011	RCS (IV)	42	SR	Dilatation (French)	Up to 72 months	64	Partial mucosal tears, laryngospasm
Dou[22]	2012	PCS (IV)	38	SR, VFSS	Dilatation (water-inflated balloon)	3–5 months	76.3	None

BoT-A = botulinum toxin serotype A; C = clinical; FEES = flexible endoscopic evaluation of swallowing; M = manometry; MF = manofluorography; PCS = prospective case series; RCS = retrospective case series; SR = self-rating; VFSS = videofluoroscopic swallowing study.

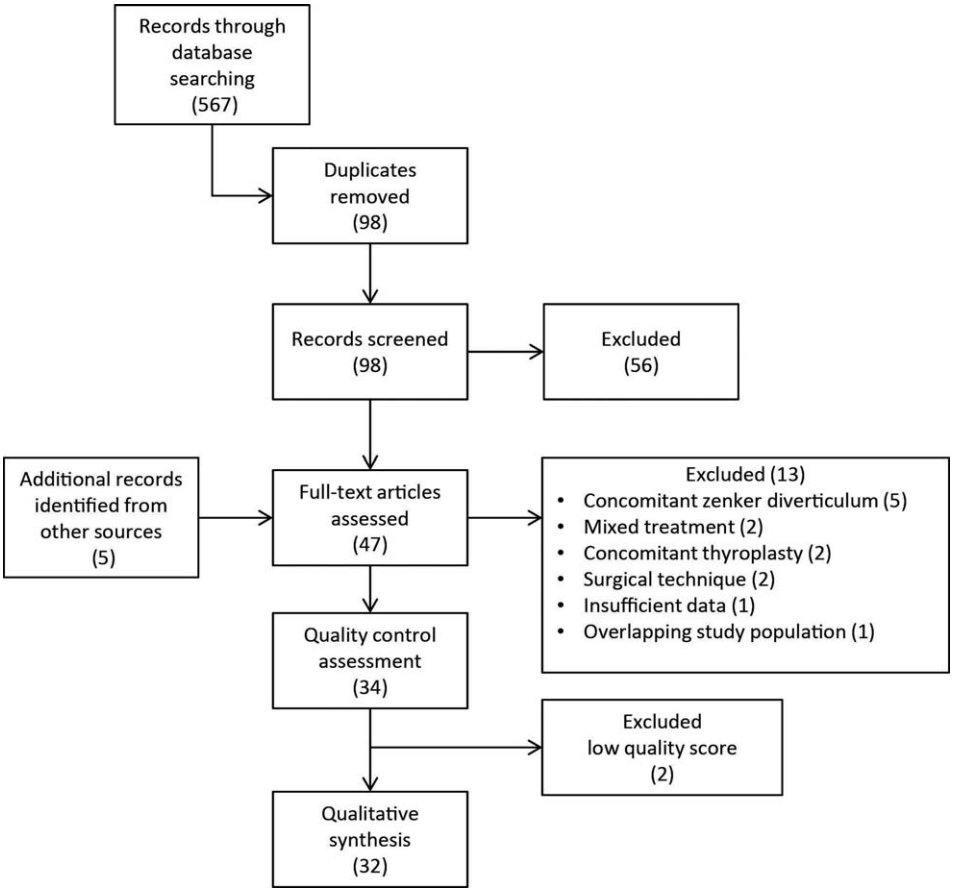

Fig. 1. Flow diagram of the search strategy.

following reasons: 1) surgical technique descriptions (two articles); 2) duplicate and overlapping study populations (one article); 3) insufficient data available to calculate the success rate of the procedure (one article); 4) patients with CP dysfunction besides Zenker's diverticulum over 1.5 cm and/or requiring diverticulectomy (five articles); 5) patients underwent concomitant thyroplasty with BoT injection or myotomy (two articles); 6) patients underwent BoT injection at the same time with myotomy or dilatation (two articles). Studies ranged from 10 to 20 of 25 points on the Downs and Black scale. Two articles receiving a score below 13 were also excluded from the evaluation (Fig. 1).

Twelve studies[4,8,22,24–32] were prospective and 20 [2,5,6,9,10,14,18,21,33–44] were retrospective. All of the publications were observational studies, with a level of evidence of IV; with the exception of one prospective cohort study (IIb) (Table II). All articles except for two dealt with one type of therapeutic procedure; the two exceptions each assessed two procedure types. In these two articles, the authors used one type of procedure for each patient and reported on the success rates and complications of the procedures separately.

Assessment of Success Rates and Complications Between BoT, Dilation, and Myotomy

Of the 32 articles, 12 articles reported on the success rates and complications of BoT injections, six articles on dilation, and 16 articles on myotomy. The range of reported success rates were between 43% and 100% for BoT injections (crude average =76%, patient-weighted average = 69%), between 58% and 100% for dilation (crude average = 81%, patient-weighted average = 73%), and between 25% and 100% for myotomy (crude average = 75%, patient-weighted average = 78%) (Table III).

Patient questionnaires, type of diet tolerated, clinical score of swallowing impairment, and flexible endoscopy had been used for the measurement of success rate in the majority of the articles. In some of the articles, a retrospective review of VFSS had been the choice as an objective tool.

The reported complication rates were between 0% and 25% for BoT injections (crude average = 5%, patient-weighted average = 4%), between 0% and 20% for dilation (crude average = 5%, patient-weighted average = 5%), and between 0% and 39% for myotomy (crude average = 6%, patient-weighted average = 7%) (Table IV). These included pharyngocutaneous fistula, pharyngeal tear, supraglottic edema, imminent mediastinitis, neck cellulitis, retropharyngeal hematoma, neck hematoma, esophageal perforation, laryngospasm, severe bleeding, and death through aspiration.

In logistic regression analysis of the patient-weighted averages, the 78% success rate with myotomy was significantly higher than the 69% success rate with BoT injections ($P = .042$), whereas the success rate of

TABLE III.
Distribution of Success Rates of BoT Injection, Dilation, and Myotomy.

	No. of Articles	Range of Success Rates (Crude Average)	No. of Patients (Sum)	No. of Successes (Sum)	Patient-Weighted Average Success Rate
BoT Injection	12	43%–100% (76%)	148	102	69%
Dilation	6	58%–100% (81%)	113	83	73%
Myotomy	16	25%–100% (75%)	369	286	78%

73% with dilation was not significantly different from that of either myotomy ($P = .37$) or BoT ($P = .42$).

Upon scoring the procedures for invasiveness as BoT injection = low, dilation = medium, and myotomy = high, there was a positive and statistically significant trend favoring increased success rate with increased invasiveness ($P = .039$). In contrast, we found no significant difference in complication rates between procedures via logistic regression, and no significant trend in complication rate with invasiveness via trend analysis.

Subgroup Analysis of Myotomy Procedures

A subgroup analysis was performed to assess the success and complication rates of open versus endoscopic myotomy. For this purpose, one study that used both methods was excluded. There were eight articles reporting outcomes of endoscopic myotomy, whereas seven evaluated open myotomy. Success rates ranged between 60% and 100% with endoscopic myotomy (crude average = 83%, patient-weighted average = 84%) compared to 73% and 79% with open myotomy (crude average = 68%, patient-weighted average = 71%). Comparison of success rates via logistic regression analysis revealed a significant increase in odds of success with the endoscopic procedure (ratio = 2.24, $P = .0025$). Complication rates were reported between 0% and 6% for endoscopic myotomy (crude average = 2%, patient-weighted average = 2%) versus 0% and 39% for open myotomy (crude average = 8%, patient-weighted average = 11%). Comparison of complication rates via logistic regression showed a significant increase in odds of complication with the open procedure (odds ratio = 5.01; $P = .0021$). Brief details of complications were mentioned in Table I.

Subgroup Analysis of BoT Injections

Botulinum toxin units were often reported as a range. We used the midpoint of the BoT unit range in analyzing success and complications rates. Logistic regression analysis indicated that a 20-unit increase in the midpoint BoT dose significantly increased the odds of success (odds ratio = 1.26, $P = .033$) without significantly changing the odds of complication (odds ratio = 0.74, $P = .33$).

DISCUSSION

CP dysfunction can present with various symptoms, often not fitting a common pattern. Patient complaints vary in severity from a lump sensation to complete inability to swallow and life-threatening aspiration. The workup varies among institutions, and there is no agreed on, uniform preoperative or postoperative evaluation technique. Similarly, because outcomes are generally not reported through objective measures, there is continued debate on the best surgical technique and the selection of suitable patients. The aim of this systematic review was to assess the success rates of myotomy, CP dilatation, and botulinum toxin injection in the management of CP dysfunction.

Kaplan is credited for performing the first CP myotomy in 1951 on a patient with bulbar poliomyositis.[45] Varying methods of transcervical myotomy have been described since then as can be seen in Table II. It can be noted that the majority of the articles were on the effectiveness of myotomy (seven papers on open myotomy, eight on endoscopic, and one comparing the two methods) in the management of CP dysfunction. We found the average success rate of myotomy to be 75%, and it was significantly higher than BoT injections ($P = .042$) but not statistically different than dilatation ($P = .37$). The average complication rate of 6% (range = 0%–39%) was not significantly higher than the other methods. Interestingly, myotomy outcomes were significantly better with the endoscopic technique (odds ratio = 2.24), supplemented with the advantage of decreased complication rates ($P = .0021$). Although the risk of mediastinitis and fistula could not be completely excluded by endoscopic laser myotomy, limiting the procedure to the fibers of the cricopharyngeus muscle considerably reduced it.[2,6,10,30,32,38,40–42] Also, any injury to the

TABLE IV.
Distribution of Complications of BoT Injection, Dilation, and Myotomy

	No. of Articles	Range of Complication Rates (Crude Average)	No. of Patients (Sum)	No. of Complications (Sum)	Patient-Weighted Average Complication Rate
BoT Injection	12	0%–25% (5%)	148	6	4%
Dilation	6	0%–20% (5%)	113	6	5%
Myotomy	16	0%–39% (6%)	369	27	7%

recurrent laryngeal nerve is avoided, and the postoperative course is significantly shortened with minimal pain and quick return to swallowing when endoscopic technique can be employed.[30]

The reported articles include patients with various etiologies. Mason et al. reported that the results of myotomy were excellent or good in patients with no discernible (idiopathic) underlying disease, but were not as good in patients with neuropathic or myopathic disease. They also evaluated the role of preoperative manometry and noted that the only factor predicting the success of the procedure, other than the etiology of the disorder, was impaired sphincter opening during manometry (odds ratio = 8.4). They went on to suggest that the most important manometric marker was the absence of the subatmospheric intrasphincteric pressure drop. They concluded that, when combined with an increased intrabolus pressure, the mechanical indicators that the procedure should work are present. Mason et al. also modified the procedure where they divide the sternohyoid and omohyoid muscles (depressors of the hyoid) to improve laryngeal elevation.[39] On the other hand, Poirier et al., in their 40-patient series with a neurogenic origin, reported success if the following criteria were fulfilled: 1) normal voluntary deglutition, 2) adequate tongue movement, 3) intact laryngeal function and phonation, and 4) absence of dysarthria.[14] Kos et al. also report the etiology of the dysphagia to be the most important prognostic factor. The patients with no apparent cause of dysphagia or with non–cancer-related iatrogenic oropharyngeal dysphagia showed 100% improvement. The outcomes in patients with central nervous system damage and extensive head and neck cancer therapy were not as rewarding (25% success rates). Their group also challenged the absence of hypopharyngeal contractions as a contraindication to surgery. In their series, although 71% of the patients with normal constrictor activity showed improvement, 79% with reduced and 71% with absent activity also showed successful outcomes following myotomy.[9] This was also advocated by Ozgursoy and Salassa, and Bammer et al., who reported improved swallowing in patients with weak pharyngeal driving forces.[10,46]

Botulinum toxin injections have been used as a test to determine whether myotomy would be effective.[18] On the other hand, Zaninotto et al. reported success with myotomy even in patients who failed BoT injections, and suggest it should not be used to discriminate between patients who may or may not benefit from surgery.[25]

There is also disagreement between authors on the necessary postoperative studies. Most outcomes are reported on subjective patient improvement. This limits our ability to uniformly compare studies and reported outcomes.

There were fewer studies reporting on the efficacy of CP dilatation. The main advantages include being less invasive and ability to be performed under sedation. This makes it a suitable alternative in patients who cannot undergo general anesthesia along with electromyography-(EMG)-guided in-office BoT injections. Ali et al. performed the only study comparing myotomy and dilatation outcomes. They operated on 20 patients, 12 of whom underwent dilatation and eight myotomy. The patient selection was dictated by clinical circumstances and patient preference, with the exception of patients demonstrating manometric failure of UES relaxations. All of these patients underwent myotomy. They clinically evaluated the patients 6 weeks postoperatively. They had an overall response rate of 65%; 75% of the patients undergoing myotomy and 58% of the patients undergoing dilatation had responded. Unfortunately, when reporting outcomes, they did not differentiate between the two groups.[31] Hatlebakk et al. reported that nine out of the 10 patients remained on an oral diet at 13 months, following dilatation with 18 to 20 mm Savary dilators. On manometry, UES pressures were significantly reduced, and/or the duration and completeness of relaxation increased following dilatation.[8] Solt et al. reported similar improvement in patients without organic stenosis of the UES, with redilatation needed in one patient (out of five) at 21 months.[43] Wang et al. also used dilatation for patients with CP dysfunction that could only be attributed to a CP bar and reported complete response.[44] Clary et al. suggested CP bougie dilatation as a first surgical step. They advocate this two-step approach for two reasons: 1) if dysphagia resolves, the patient can avoid a more morbid myotomy, and 2) if patient experiences no relief, it can suggest a need for further workup to evaluate other causes of dysphagia.[21]

Since the first report of BoT injections for CP dysfunction by Schneider, many have advocated the use of it due to the minimal invasiveness of the procedure, ability to perform in the clinic with EMG guidance, and minimal morbidity.[24] The effective duration varies on the injected site, dosag, and type of disease.[27] Most studies have reported doses between 5 and 50 units[4,18,34] up to 100 units.[20] The maximum duration of the beneficial effects continues to be studied. Terre et al. reported improvement up to a year with a single 100-U injection. They attributed this to the reduction of basal UES pressure, with a subsequent increase in pharyngeal pressure that permitted improvement in sphincter relaxation, as well as the achieved oral diet permitting the strengthening of swallowing musculature.[27] Although Terre et al. recommended BoT injections for patients who had incomplete relaxation of the CP muscle with a certain degree of pharyngeal propulsion, Woisard-Bassols et al. reported good outcomes in patients with CP dysfunction and pharyngolaryngeal weakness.[29] Our review found that BoT injections are not as successful as myotomy, and as the invasiveness of the procedure increased (BoT = low, dilatation = medium, myotomy = high), there was a statistically significant trend favoring increased success rates.

This systematic review has several limitations. Primarily, retrospective chart review studies and prospective cohort studies are subject to selection bias; therefore, the level of evidence provided by this review relies on the strength of the individual articles. The surgeons may select a patient to undergo a particular procedure based on CP dysfunction etiology, patient comorbidities, and surgeon experience. Patients are also allowed to choose the treatment based on recommendations. In CP dysfunction, there is also no universally

agreed on algorithm for management, including preoperative diagnostic testing and patient selection criteria for surgical approach. Due to this, we aimed to only evaluate surgical outcomes. This study is also limited in regard to making recommendations on patient selection for a particular surgical method. Furthermore, the studies reviewed reported outcomes with various methods, relying heavily on self-rating and clinical improvement. Similarly, due to the nature of the disease and infrequency, the largest series in this review included 42 patients. Nevertheless, we believe our data improves our understanding of the surgical management techniques for CP dysfunction and can serve as a starting point for future, well-designed, multicenter prospective trials.

CONCLUSION

In the current systematic review, logistic regression analysis of patient-weighted averages revealed significantly higher success rates with myotomy compared to BoT injections. Although the success rates of dilatation were not found to be significantly different from BoT injections or myotomy, there were also fewer studies assessing myotomy. There was no significant difference in regard to complication rates, and the effectiveness of the procedures improved as the invasiveness increased. As a result, in the well-selected patient, all of these procedures can be employed with good outcomes and minimal morbidity.

BIBLIOGRAPHY

1. Ho AS, Morzaria S, Damrose EJ. Carbon dioxide laser-assisted endoscopic cricopharyngeal myotomy with primary mucosal closure. *Ann Otol Rhinol Laryngol* 2011;120:33–39.
2. Dauer E, Salassa J, Luga L, Kasperbauer J. Endoscopic laser vs open approach for cricopharyngeal myotomy. *Otolaryngol Head Neck Surg* 2006;134:830–835.
3. Lim KG. The mouth of the esophagus. *Laryngoscope* 1907;17:421–428.
4. Alberty J, Oelerich M, Ludwig K, et al. Efficacy of botulinum toxin A for treatment of upper esophageal sphincter dysfunction. *Laryngoscope* 2000;110:1151–1156.
5. Munoz AA, Shapiro J, Cuddy LD, et al. Videofluoroscopic findings in dysphagic patients with cricopharyngeal dysfunction: before and after open cricopharyngeal myotomy. *Ann Otol Rhinol Laryngol* 2007;116:49–56.
6. Lim RY. Endoscopic CO_2 laser cricopharyngeal myotomy. *J Clin Laser Med Surg* 1995;13:241–247.
7. McKenna JA, Dedo HH. Cricopharyngeal myotomy: indications and technique. *Ann Otol Rhinol Laryngol* 1992;101:216–221.
8. Hatlebakk JG, Castell JA, Spiegel J, et al. Dilatation therapy for dysphagia in patients with upper esophageal sphincter dysfunction-manometric and symptomatic response. *Dis Esophagus* 1998;11:254–259.
9. Kos MP, David EF, Klinkenberg-Knol EC, Mahieu HF. Long-term results of external upper esophageal sphincter myotomy for oropharyngeal dysphagia. *Dysphagia* 2010;25:169–176.
10. Ozgursoy OB, Salassa JR. Manofluorographic and functional outcomes after endoscopic laser cricopharyngeal myotomy for cricopharyngeal bar. *Otolaryngol Head Neck Surg* 2010;142:735–740.
11. Olsson R, Nilsson H, Ekberg O. An experimental manometric study simulating upper esophageal sphincter narrowing. *Invest Radiol* 1994; 29:630–635.
12. Salassa JR, DeVault KR, McConnel FM. Proposed catheter standards for pharyngeal manofluorography (videomanometry). *Dysphagia* 1998;3: 105–110.
13. Leonard R, Belafsky PC, Rees CJ. Relationship between fluoroscopic and manometric measures of pharyngeal constriction: the pharyngeal constriction ratio. *Ann Otol Rhinol Laryngol* 2006;115:897–901.
14. Poirier NC, Bonavina L, Taillefer R, et al. Cricopharyngeal myotomy for neurogenic oropharyngeal dysphagia. *J Thorac Cardiovasc Surg* 1997; 113:233–240.
15. Ertekin C, Aydogdu I, Yuceyar N, et al. Electrodiagnostic methods of neurogenic dysphagia. *Electroencephalogr Clin Neurophysiol* 1998;109: 331–340.
16. Elidan J, Shochina M, Gonen B, et al. Electromyography of the inferior constrictor and cricopharyngeal muscles during swallowing. *Ann Otol Rhinol Laryngol* 1990;90:466–469.
17. Halvorson DJ, Kuhn FA. Transmucosal cricopharyngealmyotomy with the potassium-titanyl-phosphate laser in the treatment of cricopharyngeal dysmotility. *Ann Otol Rhinol Laryngol* 1994;103:173–177.
18. Blitzer A, Brin MF. Use of botulinum toxin for diagnosis and management of cricopharyngeal achalasia. *Otolaryngol Head Neck Surg* 1997;116: 328–330.
19. Verhulst J, Ame P, Guatterie M, Barat M, Traissac L. A retrospective study of 91 injections of botulinum toxin into the upper sphincter of the esophagus. *Rev Laryngol Otol Rhinol* 2003;124:315–320.
20. Kim DY, Park CI, Ohn SH, et al. Botulinum toxin type A for poststroke cricopharyngeal muscle dysfunction. *Arch Phys Med Rehabil* 2006;87: 1346–1351.
21. Clary MS, Daniero JJ, Keith SW, et al. Efficacy of large-diameter dilatation in cricopharyngeal dysfunction. *Laryngoscope* 2011;1221:2521–2525.
22. Dou Z, Zu Y, Wen H, et al. The effect of different catheter balloon dilatations modes on cricopharyngeal dysfunction in patients with dysphagia. *Dysphagia* 2012;27:514–520.
23. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377–384.
24. Schneider I, Potatschnig C, Thomfort WF, Eckel HE. Treatment of dysfunction of the CPM with botulinum A toxin: introduction of a new, noninvasive method. *Ann Otol Rhinol Laryngol* 1994;103:31–35.
25. Zaninotto G, Ragona RM, Briani C, et al. The role of botulinum toxin injection and upper esophageal sphincter myotomy in treating oropharyngeal dysphagia. *J Gastrointest Surg* 2004;8:997–1006.
26. Murry T, Wasserman T, Carrau RL, Castillo B. Injection of botulinum toxin A for the treatment of dysfunction of the upper esophageal sphincter. *Am J Otolaryngol* 2005;26:157–162.
27. Terre R, Valles M, Panades A, Mearin F. Long-lasting effect of a single botulinum toxin injection in the treatment of oropharyngeal dysphagia secondary to upper esophageal sphincter dysfunction: a pilot study. *Scand J Gastroenterol* 2008;43:1296–1303.
28. Alfonsi E, Merlo IM, Ponzio M, Montomoli C, et al. An electrophysiological approach to the diagnosis of neurogenic dysphagia: implications for botulinum toxin treatment. *J Neurol Neurosurg Psychiatry* 2010;81:54–60.
29. Woisard-Bassols V, Alshehri S, Simonetta-Moreau. The effects of botulinum toxin injections into the cricopharyngeus muscle of patients with cricopharyngeus dysfunction associated with pharyngo-laryngeal weakness. *Eur Arch Otorhinolaryngol* 2013;270:805–815.
30. Halvorson DJ. The treatment of cricopharyngeal dysmotility with a transmucosal cricopharyngeal myotomy using the potassium-titanly-phosphate (KTP) laser. *Endoscopy* 1998;30:46–50.
31. Ali GN, Wallace KL, Laundl TM, et al. Predictors of outcome following cricopharyngeal disruption for pharyngeal dysphagia. *Dysphagia* 1997; 12:133–139.
32. Bachy V, Matar N, Remacle M, et al. Long-term functional results after endoscopic cricopharyngeal myotomy with CO_2 laser: a retrospective study of 32 cases. *Eur Arch Otorhinolaryngol* 2013;270:965–968.
33. Parameswaran MS, Soliman AMS. Endoscopic botulinum toxin injection for cricopharyngeal dysphagia. *Ann Otol Rhinol Laryngol* 2002;111:871–874.
34. Haapaniemi JJ, Laurikainen EA, Pulkkinen J, Marttila RJ. Botulinum toxin in the treatment of cricopharyngeal dysphagia. *Dysphagia* 2001; 16:171–175.
35. Shaw GY, Searl JP. Botulinum toxin treatment for cricopharyngeal dysfunction. *Dysphagia* 2001;16:161–167.
36. Lee SY, Seo HG, Paik NJ. Botulinum toxin injection for dysphagia: a blinded retrospective videofluoroscopic swallowing study analysis. *Am J Phys Med Rehabil* 2009;88:491–494.
37. St. Guily JL, Perie S, Willig TN, et al. Swallowing disorders in muscular diseases: functional assessment and indications of cricopharyngeal myotomy. *Ear Nose Throat J* 1994;73:34–40.
38. Herberhold C, Walther EK. Endoscopic laser myotomy in cricopharyngeal achalasia. *Adv Otorhinolaryngol* 1995;49:144–147.
39. Mason RJ, Bremner CG, DeMeester TR. Pharyngeal swallowing disorders selection for and outcome after myotomy. *Ann Surg* 1998;228:598–608.
40. Lawson G, Remacle M, Jamart J, Keghian J. Endoscopic CO_2 laser-assisted surgery for cricopharyngeal dysfunction. *Eur Arch Otorhinolaryngol* 2003;260:475–480.
41. Takes RP, Hoogen FJA, Marres HAM. Endoscopic myotomy of the cricopharyngeal muscle with CO_2 laser surgery. *Head Neck* 2005;27: 703–709.
42. Lawson G, Remacle M. Ins and outs of myotomy of the upper esophageal sphincter in swallowing disorders. *B-ENT* 2008;10:83–89.
43. Solt J, Bajor J, Moizs M, et al. Primary cricopharyngeal dysfunction: treatment with balloon catheter dilatation. *Gastrointest Endosc* 2001;6: 767–771.
44. Wang AY, Kadkade R, Kahrilas PJ, Hirano I. Effectiveness of esophageal dilation for symptomatic cricopharyngeal bar. *Gastrointest Endosc* 2005; 61:148–152.
45. Kaplan S. Paralysis of deglutition, a post-poliomyelitis complication treated by section of the cricopharyngeus muscle. *Ann Surg* 1951;133: 572–573.
46. Bammer T, Salassa JR, Klingler PJ. Comparison of methods for determining cricopharyngeal intrabolus pressure in normal patients as possible indicator for cricopharyngeal myotomy. *Otolaryngol Head Neck Surg* 2002;127:299–308.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(3):488-493.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Original Research—Laryngology and Neurolaryngology

Otolaryngology–
Head and Neck Surgery
2015, Vol. 152(3) 488–493
© American Academy of Otolaryngology—Head and Neck Surgery Foundation 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0194599814565599
http://otojournal.org
SAGE

Esophageal Visualization as an Adjunct to the Videofluoroscopic Study of Swallowing

Anna Miles, PhD[1], Jessica McMillan, MSc [1], Katie Ward, MSc[2], and Jacqui Allen, MBChB, FRACS[1,2]

No sponsorships or competing interests have been disclosed for this article.

Received August 13, 2014; revised October 23, 2014; accepted December 4, 2014.

Abstract

Objective. Complaints of dysphagia for solids lead to speech-language pathology (SLP) referral. Yet many of these patients are later diagnosed with esophageal rather than oropharyngeal dysphagia. Fluoroscopic screening involving the oropharynx alone fails to identify these patients. The aim of this study was to investigate the prevalence of esophageal abnormalities in an SLP-led videofluoroscopic study of swallowing (VFSS) clinic.

Study Design. Prospective, observational study.

Setting. Radiology suite, public hospital.

Subjects and Methods. In total, 111 consecutive mixed-etiology patients referred to the clinic by otorhinolaryngology (ORL) (59) or by a speech-language pathologist (52) were recruited. A VFSS was performed according to protocol, and at completion, esophageal visualization (in anterior-posterior plane) was performed by administration of a large liquid barium bolus and a barium capsule. All VFSS recordings were analyzed using objective digital measures of timing and displacement.

Results. Sixty-eight percent of patients had an abnormal esophageal transit. One-third of those referred presented exclusively with esophageal abnormalities, while one-third had both oropharyngeal and esophageal abnormalities. Oral abnormalities, reduced pharyngoesophageal segment maximum opening (PESmax), and increasing age were significantly associated with esophageal abnormalities.

Conclusion. Fluoroscopic evaluation of the pharynx alone, without esophageal review, risks incomplete diagnosis of patients with esophageal disorders. Using esophageal visualization allows timely referral for further investigation by appropriate medical specialties, avoiding incomplete management of patients with dysphagia.

Keywords

deglutition, deglutition disorders, dysphagia, esophageal visualization, speech-language pathology, otorhinolaryngology

Associations between oropharyngeal abnormalities and esophageal abnormalities are poorly understood but well documented.[1-3] Oropharyngeal alterations have been reported in patients with gastroesophageal reflux disease.[4] Neurologic diseases such as Parkinson disease[5,6] and systemic conditions such as scleroderma[7] lead to both oropharyngeal and esophageal abnormalities. In a recent study using high-resolution manometry, O'Rourke and colleagues[8] describe a variety of esophageal alterations during voluntary pharyngeal maneuvers (effortful swallow and Mendelsohn), adding to the theory that changing one point in the swallowing system can lead to positive or negative changes elsewhere. In addition, patient accuracy in locating the level of bolus holdup has been shown to be poor, with patients often indicating the cervical region or levels more proximal than the true site of bolus stasis, particularly when this occurs in the esophagus.[1,9,10] Smith and colleagues[10] reported 57% of respondents located a solid bolus impacted at a distal esophageal ring to the level of the sternal notch.

Complaints of dysphagia for solids regularly lead to speech-language pathology (SLP) referral rather than gastroenterology or otorhinolaryngology (ORL). Traditionally, SLP-led videofluoroscopic study of swallowing (VFSS) has assessed the oropharynx exclusively, even when symptoms might suggest esophageal complaints. This results in failure to identify patients with esophageal problems. These patients are sent home with no diagnosis and either continue to manage their symptoms alone or undergo a variety of other diagnostic tests over a prolonged period of time before reaching correct diagnosis and treatment. An esophageal screen was described and validated by Allen and colleagues in 2012.[11] They compared fluid esophageal screens with full esophagrams in 74 mixed-

[1]The University of Auckland, Auckland, New Zealand
[2]Waitemata District Health Board, Auckland, New Zealand

Corresponding Author:
Anna Miles, PhD, Speech Science, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.
Email: a.miles@auckland.ac.nz

etiology patients. Sensitivity of the screen was 63% with 100% specificity. Sensitivity increased to 71% with the inclusion of a barium capsule. In comparison to the esophagram, the esophageal screen subjects patients to approximately 10 times less radiation dose and is relatively quickly completed at the end of a standard VFSS procedure.[11] The aim of this prospective, observational study was to investigate the prevalence of esophageal abnormalities in an SLP-led VFSS clinic.

Methods

This study received appropriate regional ethics approval (University of Auckland Human Participants Ethics Committee 9263).

Participants

Data from 111 inpatients and outpatients consecutively referred to an SLP-led VFSS clinic (between May and December 2013) were collected. Patients were referred by the ORL service (59) or by another speech-language pathologist (52), with a mean (SD) cohort age of 71 (14.95) years (range, 20-95 years). Fifty-one patients were male (46%). Referrals were all for complaints of dysphagia attributed to mixed etiologies: 36 neurological (32%), 37 dysphagia of unknown cause (33%), 28 otorhinolaryngology (eg, head and neck cancer [25%]), and 10 other (eg, unwell elderly [9%]). Participants were excluded if their VFSS recording had no esophageal visualization or had no timer or no calibration ring, since these were required to complete the timing and displacement measures.

Esophageal Videofluoroscopic Assessment

Studies were performed in a radiology suite using a Videofluoroscope (Toshiba, Tokyo, Japan) and recorded at 30 frames per second onto a USB drive. Timing information was superimposed on the fluoroscopic recording in 100ths of a second using a Horita VS-50 Video Stopwatch (Horita, Capistrano Beach, California). A 19-mm diameter radioopaque ring was taped to the patient's chin (in the lateral plane) and shoulder (in the anterior-posterior plane) to allow calibration for displacement measures. A medical radiation technician (MRT) and a speech-language pathologist were present at all procedures.

For the standard VFSS protocol, the patient was screened in the lateral plane. The patient was presented with 1 mL, 3 mL, and then 20 mL of thin barium (E-Z Paque, E-Z-EM Anjou, Canada; 100% w/v) followed by half a cup of thin barium through a straw. The patient was then given 3 mL of barium paste (E-Z paste, E-Z-EM; 60% w/w). The procedure was truncated if required for patient safety.

The esophageal phase involved the introduction of 2 boluses after completion of the standard VFSS protocol: a 20-mL fluid bolus and a 13-mm barium capsule. The patient was positioned in the anterior-posterior plane, standing whenever possible. The patient was asked to ''swallow all in one go'' to avoid deglutitive inhibition. The MRT followed the bolus from the oral cavity through the lower esophageal sphincter (LES) until clearance into the stomach. Screening was continued for up to 15 seconds. If there was still residue in the esophagus, screening was ceased for 15 seconds, then recommenced. If residue was still present, the patient was asked to take a dry swallow to see if clearance occurred. If residue was still present, the patient was then offered a water swallow as well as being asked if he or she could feel any remaining bolus. A screen shot was used to identify complete clearance.

Data Collected

Age, sex, and comorbidities were recorded for each patient. Each VFSS was analyzed using real-time and frame-by-frame viewing (Quicktime Media Player; Apple, Cupertino, California). Videos were scored for the presence of oral, pharyngeal, and esophageal abnormalities (yes/ no). Prolonged bolus manipulation, anterior spillage from lips, premature spillage into the pharynx, and oral residue were all considered oral abnormalities. Nasal regurgitation, reduced epiglottic deflection, delay in swallowing initiation, pharyngeal residue, penetration, aspiration, and prolonged pharyngeal transit time were considered pharyngeal abnormalities. Esophageal bolus stasis, bolus redirection/intraesophageal reflux, gastroesophageal reflux, esophagopharyngeal reflux, hiatal hernia, prolonged esophageal transit, and pill stasis were all considered esophageal abnormalities. Pharyngeal transit time (PTT) was recorded and translated into a binary measure of (1) within normal limits vs (2) more than 2 standard deviations (SD) outside of normal limits.[12] Esophageal transit time (ETT) was also recorded. A conservative cutoff of over 15 seconds was selected for abnormality. Previous published work has defined normal liquid transit through the esophagus as less than 13 seconds.[11,13,14] Maximum penetration-aspiration scale (PAS) scores were recorded, and scores 6 and above were considered an aspiration event.[15] To explore whether esophageal abnormalities can be predicted by objective pharyngeal measures, we calculated the pharyngeal constriction ratio (PCR)[12] and pharyngoesophageal segment maximum opening (PESmax)[12] using the Universal Desktop Ruler (AVPSoft). These measures were also translated into binary measures of (1) within normal limits vs (2) more than 2 SD outside of normal limits.[12] All measures were taken from the largest fluid bolus ingested.

Data Analysis

Swallow studies were reported by an experienced otolaryngologist, specializing in dysphagia management, and by a speech-language pathologist, trained in quantitative analysis of VFSS using the method developed by Leonard and Kendall.[12] Interrater reliability for all measures was calculated on 30% of videos by a third researcher. Videos were randomly selected by a fourth researcher. The third rater was blinded to the first researcher's scoring and patient etiology. Total agreement across measures was 98%, with a κ coefficient of 0.92. Lack of agreement was found for 1 PAS score (1 vs 2), and although there was slight variance in PESmax (maximum variance .08) and PCR (maximum variance .07), this did not change binary measures of within

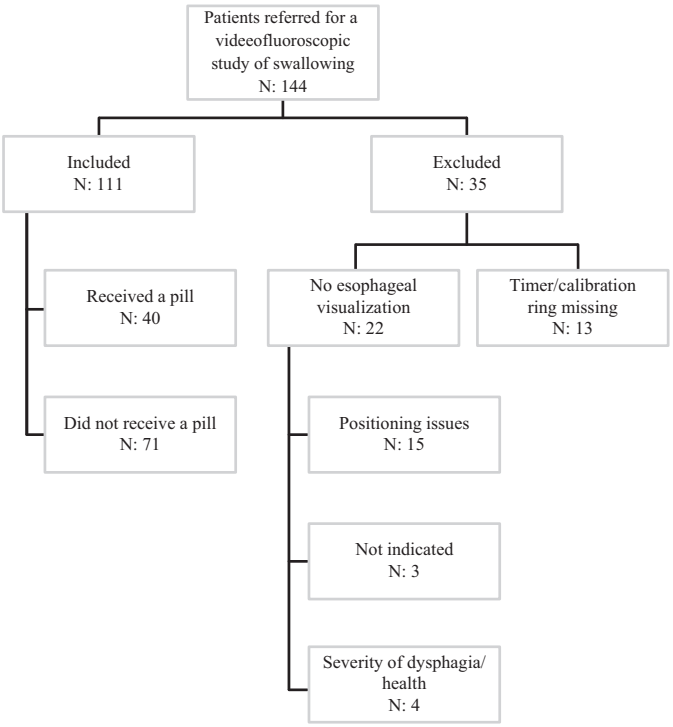

Figure 1. Recruitment inclusion.

Table 1. Frequency of Swallowing Abnormalities in Full Cohort.

Swallowing Indices	Number	% of Cohort
Frequency of oral abnormalities	28	22
Frequency of pharyngeal abnormalities	50	45
Frequency of esophageal abnormalities	76	68
Oral abnormalities alone	4	4
Pharyngeal abnormalities alone	9	8
Esophageal abnormalities alone	34	31
Oral and pharyngeal abnormalities	7	6
Oral and esophageal abnormalities	9	8
Pharyngeal and esophageal abnormalities	15	14
Oral, pharyngeal, and esophageal abnormalities	19	17
No abnormalities	14	13
Frequency of PCR outside 2 SD of norm	10	9
Frequency of PESmax outside 2 SD of norm	20	18
PTT outside 2 SD of norm	19	17
ETT >15 seconds	32	29
ETT pill, >15 seconds (n = 40)	10	25

Abbreviations: ETT, esophageal transit time; PCR, pharyngeal constriction ratio; PESmax, pharyngoesophageal segment maximum opening; PTT, pharyngeal transit time; SD, standard deviation.

normal limits vs outside normal limits. Data were analyzed using SPSS version 20 (SPSS, Inc, an IBM Company, Chicago, Illinois). Descriptive statistics were used to explore the frequency of swallowing abnormalities. Correlation analyses were made using χ^2 for categorical variables and Spearman correlations for continuous variables. Multiple logistic regressions were applied to evaluate the associations between esophageal abnormalities and other clinical indices adjusted for confounding variables based on bivariate analyses (sex, age, etiology, oral abnormalities, pharyngeal abnormalities, PCR, PESmax, PTT, and aspiration event) and 2-way interactions (esophageal abnormalities and PESmax, esophageal abnormalities and oral abnormalities, and esophageal abnormalities and age). First the full model with all confounding factors was fit, and backward selection was used to select the main effect model. The 2-way interactions were then added to the main effect model one by one for the final model.

Excluded Data

Thirty-three videos were excluded from analyses. Reasons for exclusion include inadequate positioning, severe aspiration precluding completion of the study, and missing measurement devices (**Figure 1**). Excluded cases were significantly older ($P < .001$), more likely to be referred by a speech-language pathologist ($P < .001$), and more likely to have a neurologic condition ($P < .01$) than included cases.

Results

Sixty-eight percent of the 111 patients had esophageal abnormalities, with 29% of the total cohort having an esophageal transit time of >15 seconds. Esophageal phase abnormalities were the most common: 68% vs 45% (pharyngeal) and 34% (oral) (**Table 1**). In addition, it was rare for patients to present with isolated oral phase or pharyngeal phase abnormalities (only 4% and 8%, respectively), but one-third of patients demonstrated only esophageal phase abnormalities (**Table 1**).

Referral Sources

There were significant differences in age, sex, etiology, and swallowing indices between referral sources (**Table 2**). However, frequency of esophageal abnormalities was not significantly different (ORL, 73%; SLP, 63%).

Associations with Esophageal Abnormalities

Table 3 displays the associations between esophageal abnormalities and prolonged ETT, with other clinical indices. There was a significant association between esophageal abnormalities and reduced PESmax, with 90% of patients with reduced PESmax opening having concurrent esophageal abnormalities. Esophageal transit time was significantly associated with age ($\rho = .218$, $P < .05$), with an odds ratio of 2.8 for prolonged ETT if a patient was older than 65 years. Oral abnormalities were also significantly associated with prolonged ETT, with 50% of patients with prolonged ETT having oral abnormalities. However, when referral sources were separated, the association was only significant in the SLP cohort (SLP cohort, $P < .05$; ORL cohort, $P = .237$).

Table 2. Comparison of Referral Sources.

Characteristic	Frequency (% within Referral Type)			
	ORL	SLP	χ^2	*P* Value[a]
Number of files	59	52	neurological etiology	neurological etiology
Age (>65 y)	37 (63)	47 (90)	11.50	**.001**
Sex (male)	20 (33)	31 (60)	7.36	**.008**
Etiology—neurological	4 (7)	32 (62)	44.01	**.000**
Dysphagia of unknown cause	27 (46)	10 (19)		
ORL (H&N cancer, GERD)	24 (41)	4 (8)		
Other (unwell elderly, pneumonia)	4 (6)	6 (11)		
Frequency of oral abnormalities	11 (19)	27 (52)	13.60	**.000**
Frequency of pharyngeal abnormalities	10 (17)	40 (77)	40.16	**.000**
Frequency of esophageal abnormalities	43 (73)	33 (63)	1.14	.312
Frequency of PCR outside 2 SD of norm	1 (2)	9 (17)	8.26	**.006**
Frequency of PESmax outside 2 SD of norm	12 (20)	8 (36)	.340	.625
PTT outside 2 SD of norm	7 (12)	12 (23)	2.33	.139
ETT >15 seconds	13 (22)	19 (37)	2.83	.099
Aspiration event occurred (Pen-Asp score 6-8)	2 (3)	12 (23)	9.72	**.003**

Abbreviations: ETT, esophageal transit time; GERD, gastroesophageal reflux disease; H&N, head and neck; ORL, otorhinolaryngology; PCR, pharyngeal constriction ratio; Pen-Asp, penetration-aspiration; PESmax, pharyngoesophageal segment maximum opening; PTT, pharyngeal transit time; SD, standard deviation; SLP, speech-language pathology.
[a]Bolding indicates that the *P* values have reached significance.

Table 3. Associations between Esophageal Abnormalities and Other Clinical Indices.

Characteristic	Esophageal Abnormalities		Esophageal Transit Time >15 Seconds	
	χ^2	*P* Value	χ^2	*P* Value[a]
Age (>65 y)	.000	1.000	3.42	.087
Sex (male)	.006	1.000	.298	.675
Etiology	1.645	.678	2.54	.481
Frequency of oral abnormalities	.089	.826	4.96	**.030**
Frequency of pharyngeal abnormalities	.001	1.000	.446	.504
Frequency of PCR outside 2 SD of norm	.617	.474	.723	.466
Frequency of PESmax outside 2 SD of norm	4.41	**.054**	.029	1.00
PTT outside 2 SD of norm	.047	1.000	1.89	.266
ETT >15 seconds	19.15	**.000**	20.71	**.000**
Aspiration event occurred (Pen-Asp score 6-8)	.009	1.000	.370	.543

Abbreviations: ETT, esophageal transit time; PCR, pharyngeal constriction ratio; Pen-Asp, penetration-aspiration; PESmax, pharyngoesophageal segment maximum opening; PTT, pharyngeal transit time; SD, standard deviation.
[a]Bolding indicates that the *P* values have reached significance.

Discussion

As has been previously reported, there was a high frequency of esophageal abnormalities observed in this cohort of patients referred to an SLP-led VFSS clinic.[1] In fact, esophageal phase abnormalities were the most prevalent finding in all groups regardless of referral diagnosis or source. Esophageal phase abnormalities were not predicted by other phase abnormalities. If fluoroscopic screening had involved the oropharynx alone, one-third of patients would have been sent home with no diagnosis, and one-third of patients would have been treated for oropharyngeal dysphagia without consideration of concurrent esophageal disorders. In agreement with previous manometric data, slower esophageal transit times were found with increasing age.[16] It could, therefore, be hypothesized that, with the inclusion of esophageal visualization, two-thirds of patients potentially had altered clinical recommendations: different diet recommendations, different feeding strategy recommendations, and additional referrals for further investigations and medical specialty input. It may also suggest that esophageal transit time changes with age rather than being a pathologic

finding. Normative data in older adults are required to clarify this and are currently under way.

As one would expect, differences in the primary etiology causing dysphagia were noted between referral sources. However, it is clinically significant that there was no significant difference in the prevalence of esophageal abnormalities between referral sources. These results advocate for esophageal visualization irrespective of the referral source. Esophageal abnormalities appear difficult to predict by etiology. Yet, as has been previously reported, patients with impaired PESmax were more likely to have both pharyngeal[17] and esophageal abnormalities.[1] The association of reduced PESmax and prolonged ETT is hypothesized to represent a compensatory strategy. To limit retrograde flow and minimize the ''threat'' that a retained bolus presents in the esophagus, the upper esophageal sphincter hypertrophies and becomes less compliant.[1,17] This is measured as a reduction in absolute PES opening. Identification of reduction in PES opening therefore should prompt esophageal evaluation.

This study demonstrates that inclusion of esophageal visualization as part of a VFSS protocol can help identify and categorize patients' problems when referred with a symptom of dysphagia. This may allow further investigations to be requested (including formal esophagram) or referral to appropriate medical services to quantitate and characterize the esophageal disorder more thoroughly. Therapeutic recommendations can be refined and targeted to the appropriate service, and the patient receives the most complete information and holistic management.

A short visualization that adds only 2 further swallows (each screened for a maximum of 15 seconds) does not significantly increase radiation exposure or the overall study time. The additional radiation exposure incurred by performing esophageal visualization was recorded as less than 0.1 mSv. Background radiation dose annually exceeds 3 mSv, and therefore the incremental increase in exposure related to esophageal views is very low.[18-21] Compared with formal esophagram, which incorporates several additional views and longer screening time, the simple esophageal visualization may direct referral or management of the patient without exposing him or her to the higher radiation dose of a full esophagram.[11]

Concerns regarding scope of practice with regard to esophageal diagnosis have been raised. Speech-language pathologists are not expected to be esophageal diagnosticians. The visualization provides simple parameters for onward referral, as transit times longer than normal (15 seconds) can be easily measured by automated timer and indicate the need for further review. The referrer, who will decide whether further investigation is warranted and who should perform this, usually directs this. The onus will not fall on the speech-language pathologist performing the test to interpret the clinical significance of any findings. In fact, identification of an esophageal discrepancy likely to produce symptoms, particularly in the absence of other likely causes, may assist the speech-language pathologist in deciding what advice and guidance to give regarding eating strategies, rehabilitative exercises, and body positioning.

Limitations

A proportion of patients (17%) were difficult to screen due to positioning issues, limiting the view of the LES. However, these patients were generally more disabled and often wheelchair bound and may not have tolerated a full esophagram (requiring the ingestion of large quantities of barium in the prone position) either. The barium capsule was used in only 40 procedures, despite evidence that it increases the sensitivity of the screen.[11] This was thought to be due to hesitancy from speech-language pathologist to give patients with dysphagia a capsule. It is likely that capsule use is not safe for all dysphagic patients and that SLP clinical decision making is necessary in evaluating risk in each individual patient. There are no comparative normative data for pill transit times, and this is currently being investigated. PESmax was measured solely in the lateral view. The addition of an anterior-posterior measure would have provided more information regarding the extent of PES opening impairment. A measure of hyoid displacement and/or hyoid-larynx approximation would have added to the study by allowing further analysis of the cause of PES opening impairment. This was not a validation study, and no formal esophagram was performed routinely for comparison. Accuracy of esophageal abnormality detection therefore cannot be confirmed.

Conclusion

Esophageal abnormalities are highly prevalent in patients referred to a VFSS clinic with a symptom of dysphagia. One-third of patients present only with esophageal phase abnormalities. Traditional fluoroscopic screening of the oropharynx alone fails to identify these patients. Esophageal visualization is a useful adjunct to VFSS as it provides preliminary information regarding the esophageal phase of swallowing. It enables appropriate referrals to radiology, ORL, and/or gastroenterology to be made and avoids patients being falsely reassured, misdiagnosed, and mismanaged.

Acknowledgments

The authors thank the Waitemata District Health Board–Speech-Language Therapy Department for access to their clinical database.

Author Contributions

Anna Miles, study design, data collection, analysis and manuscript preparation and final approval; **Jessica McMillan**, study design, data collection, analysis and manuscript preparation and final approval; **Katie Ward**, data collection and analysis and final approval; **Jacqui Allen**, study design, analysis and manuscript preparation and final approval.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Jones B, Ravich W, Donner M, Kramer S, Hendrix T. Pharyngoesophageal interrelationships: observations and working concepts. *Gastrointest Radiol*. 1985;10:225-233.
2. Triadafilopoulos G, Hallstone A, Nelson-Abbott H, Bedinger K. Oropharyngeal and esophageal interrelationships in patients with nonobstructive dysphagia. *Dig Dis Sci*. 1992;37:551-557.
3. Gullung J, Hill E, Castell D, Martin-Harris B. Oropharyngeal and esophageal swallowing impairments: their association and the predictive value of the modified barium swallow impairment profile and combined multichannel intraluminal impedance-esophageal manometry. *Ann Otol Rhinol Laryngol*. 2012;121:738-745.
4. Mendell D, Logemann J. A retrospective analysis of the pharyngeal swallow in patients with a clinical diagnosis of GERD compared with normal controls: a pilot study. *Dysphagia*. 2002;17:220-226.
5. Castell J, Johnston BT, Colcher A, et al. Manometric abnormalities of the oesophagus in patients with Parkinson's disease. *Neurogastroenterol Motil*. 2001;13:361-364.
6. Leopold N, Kagel M. Pharyngo-esophageal dysphagia in Parkinson's disease. *Dysphagia*. 1997;12:11-18.
7. Montesi A, Pesaresi A, Cavalli M, et al. Oropharyngeal and esophageal function in scleroderma. *Dysphagia*. 1991;6:219-223.
8. O'Rourke A, Morgan L, Coss-Adame E, et al. The effect of voluntary pharnygeal swallowing maneuvers on esophageal swallowing physiology. *Dysphagia*. 2014;29:262-268.
9. Edwards D. Discriminatory value of symptoms in the differential diagnosis of dysphagia. *J Clin Gastroenterol*. 1976;5: 49-57.
10. Smith D, Ott D, Gelfand D, Chen M. Lower esophageal mucosal ring: correlation of referred symptoms with radiologic findings using a marshmallow bolus. *AJR Am J Roentgenol*. 1998;171:1361-1365.
11. Allen J, White C, Leonard R, Belafsky P. Comparison of esophageal screen findings on videofluoroscopy with full esophagram results. *Head Neck*. 2012;34:264-269.
12. Leonard R, Kendall K. *Dysphagia Assessment and Treatment Planning: A Team Approach*. 2nd ed.London, UK: Singular Publishing; 2008.
13. Torrico S, Corazziari E, Habib F. Barium studies for detecting esophagopharyngeal reflux events. *Am J Med*. 2003;115:124S-129S.
14. Barloon T, Bergus G, Lu C. Diagnostic imaging in the evaluation of dysphagia. *Am Fam Physician*. 1996;53:535-546.
15. Rosenbek J, Robbins J, Roecker E, Coyle J, Woods J. A penetration-aspiration scale. *Dysphagia*. 1996;11:93-98.
16. Grand L, Lacima G, Ros E, et al. Deterioration of esophageal motility with age: a manometric study of 79 healthy subjects. *Am J Gastroenterol*. 1999;94:1795-1801.
17. Belafsky P, Rees C, Allen J, Leonard R. Pharyngeal dilation in cricopharyngeus muscle dysfunction and zenker diverticulum. *Laryngoscope*. 2010;120:889-894.
18. Herfarth H, Palmer L. Risk of radiation and choice of imaging. *Dig Dis*. 2009;27:278-284.
19. Raelson C, Kanal K, Vavilaia M, et al. Radiation dose and excess risk of cancer in children undergoing neuroangiography. *AJR Am J Roentgenol*. 2009;193:1621-1628.
20. Fazel R, Krumholz H, Wang Y, et al. Exposure to low-dose ionizing radiation from medical imaging procedures. *N Engl J Med*. 2009;361:849-857.
21. Smith-Bindman R, Lipson J, Marcus R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Arch Intern Med*. 2009;169:2078-2086.

Reprinted by permission of Am J Gastroenterol. 2013; 108(3):366-372.

Randomized Controlled Trial Comparing Aerosolized Swallowed Fluticasone to Esomeprazole for Esophageal Eosinophilia

Fouad J. Moawad, MD[1,2], Ganesh R. Veerappan, MD[1,2], Johnny A. Dias, DO[1], Thomas P. Baker, MD[3], Corinne L. Maydonovitch, BS[1] and Roy K.H. Wong, MD[1,2]

OBJECTIVES: **Patients with clinical symptoms of esophageal dysfunction and dense eosinophilic infiltration of the esophageal mucosa are suspected to have eosinophilic esophagitis (EoE). Topical steroids are often used as first-line therapy for EoE, although some patients respond clinically to proton pump inhibitors (PPIs). The purpose of this study was to compare the histological and clinical response of patients with esophageal eosinophilia treated with aerosolized swallowed fluticasone propionate vs. esomeprazole.**

METHODS: **This prospective single-blinded randomized controlled trial enrolled newly diagnosed patients with suspected EoE, defined as having clinical symptoms related to esophageal dysfunction with at least 15 eosinophils/high power field (hpf). Patients underwent 24-h pH/impedance monitoring to establish gastroesophageal reflux disease (GERD). Patients were stratified by the presence of GERD and randomized to receive fluticasone 440 mcg twice daily or esomeprazole 40 mg once daily for 8 weeks followed by repeat endoscopy with biopsies. The primary outcome was histological response of esophageal eosinophilia, defined as <7 eosinophils/hpf. Secondary outcomes included clinical change in symptoms using the validated Mayo dysphagia questionnaire (MDQ) and interval change in endoscopic findings following treatment.**

RESULTS: **Forty-two patients (90% male, 81% white, mean age 38±10 years) were randomized into fluticasone (n=21) and esomeprazole (n=21) treatment arms. In all, 19% (8/42) of patients had coexisting GERD and were equally stratified into each arm (n=4). Overall, there was no significant difference in resolution of esophageal eosinophilia between fluticasone and esomeprazole (19 vs. 33%, P=0.484). In patients with established GERD, resolution of esophageal eosinophilia was noted in 0% (0/4) of the fluticasone group compared with 100% (4/4) of the esomeprazole group (P=0.029). In GERD-negative patients, there was no significant difference in resolution of esophageal eosinophilia between treatment arms with fluticasone and esomeprazole (24 vs.18%, P=1.00). The MDQ score significantly decreased after treatment with esomeprazole (19±21 vs. 1.4±4.5, P<0.001), but not with fluticasone (17±18 vs. 12±16, P=0.162). Improvement in endoscopic findings and other histological markers were similar between treatment groups.**

CONCLUSIONS: **Fluticasone and esomeprazole provide a similar histological response for esophageal eosinophilia. With regard to clinical response, esomeprazole was superior to fluticasone, particularly in patients with established GERD.**

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

Am J Gastroenterol 2013; 108:366–372; doi:10.1038/ajg.2012.443; published online 12 February 2013

INTRODUCTION

Eosinophilic esophagitis (EoE) is a chronic inflammatory immune-mediated condition characterized by symptoms of esophageal dysfunction and the presence of dense eosinophilia on esophageal biopsies (1). Management, most often with topical steroids, is aimed at improving clinical symptoms and reversing the inflammatory changes within the esophagus to prevent tissue remodeling and formation of fibrosis (2).

[1]Department of Medicine, Walter Reed National Military Medical Center, Gastroenterology Service, Bethesda, Maryland, USA; [2]Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA; [3]Department of Pathology, Walter Reed National Military Medical Center, Bethesda, Maryland, USA.
Correspondence: Fouad J. Moawad, MD, Department of Medicine, Walter Reed National Military Medical Center, Gastroenterology Service, 8901 Rockville Pike, Bethesda, Maryland 20889, USA. E-mail: Fouad.Moawad@us.army.mil
Received 16 May 2012; accepted 16 September 2012

Proton pump inhibitors (PPIs) are also used to control symptoms associated with EoE and to treat gastroesophageal reflux disease (GERD), the main differential diagnosis of EoE, which can present with similar clinical symptoms and histopathology. According to the AGA consensus statement published in 2007, administration of PPIs to presumed EoE patients was part of the diagnostic evaluation, primarily to exclude GERD as a cause of esophageal eosinophilia (3). If dense eosinophilia persisted following therapy, then a diagnosis of EoE is made. However, some patients with a phenotypic appearance more suggestive of EoE rather than GERD (i.e. young atopic patient presenting with food impaction with concentric rings on endoscopy and having elevated eosinophils on esophageal biopsies) can respond to PPI therapy (4,5). This phenomenon now recognized by the more recent and updated EoE consensus statement has been termed PPI-responsive esophageal eosinophilia or PPI-responsive EoE (2,6).

The aim of this study was to perform a randomized controlled trial to compare the efficacy of fluticasone propionate (FP) to esomeprazole (ESO) in patients with esophageal eosinophilia. A secondary aim of this study was to determine whether the presence of GERD impacted the response to therapy in each treatment group.

METHODS

Study design and patient population

This is a prospective investigator-blinded randomized study. Adult patients (age≥18 years) seen at Walter Reed Army Medical Center (WRAMC) with esophageal eosinophilia were enrolled from April 2008 to October 2010. All patients had at least one clinical symptom of esophageal dysfunction (dysphagia, food impaction, heartburn) with ≥15 eosinophils/hpf (eos/hpf; high power field) on index endoscopy. Patients who had a history of secondary hypereosinophilic disorders, severe coagulopathy, or who were pregnant were excluded from the study. Patients who were dilated at index endoscopy were not excluded from the study. Baseline demographic data, history of coexisting atopic diathesis (seasonal allergies, food allergies, asthma, and eczema), and data from index endoscopy (concentric rings, longitudinal furrows, white plaques, mucosal tearing/friability, strictures, Schatzki rings, erosive esophagitis) were collected. All patients completed a validated dysphagia questionnaire, known as the 2-week Mayo Dysphagia Questionnaire (MDQ), following index endoscopy once eosinophilic infiltration was established on biopsies (7). This 29-item instrument is scored from 0 to 100 based on the presence and severity of dysphagia and whether patients avoided or had trouble swallowing different foods (oatmeal, banana, apple, ground meat, bread, and fibrous meat) (**Supplementary file**). Informed consent was obtained from all patients. This study was approved by the WRAMC Institutional Review Board (Work Unit number: 08-14045) and the study was registered at www.clinicaltrials.gov (NCT00895817).

GERD diagnosis

Upon enrollment into the study, all patients underwent 24-h pH with impedance studies. Location of the lower esophageal sphincter was determined by esophageal manometry utilizing a Solar Stationary GI motility system (Medical Measurement Systems USA, Dover, NH) and an electrically powered water perfusion pump (Mui Scientific, Ontario, CA). A 24-h pH/impedance catheter was then placed 5 cm above the proximal location of the lower esophageal sphincter. The catheter was connected to a ZepHr-reflux recording system (Sandhill Scientific, Highlands Ranch, CO) to capture pH/impedance, as well as symptom data. Subjects returned to our clinic the following day for analysis of the study. Data was analyzed with Bioview Analysis software (Sandhill Scientific). GERD was defined by the validated Johnson-DeMeester score (8,9). This scoring method takes into account six parameters, which include: total % time pH below 4, % time pH below 4 in the upright position, % time pH below 4 in the supine position, the total number of reflux episodes within a 24-h period, the number of reflux episodes longer than 5 min, and the longest reflux episodes in minutes. A composite score is then calculated with a score of greater than 22 being indicative of GERD. The pH drops without accompanying reflux events on impedance and reflux events during meals were excluded from analysis.

Randomization and drug administration

A computer-generated list of random numbers was used to separate patients into two equal treatment groups (esomeprazole and fluticasone proprionate). Concealed allocation using a sealed opaque envelope containing data on the sequence of randomization was maintained by a research pharmacist. Following data from the 24-h pH study, patients were stratified into GERD-negative or GERD-positive groups. Within each group, subjects were randomized to receive either 40 mg of ESO once daily or 440 mcg of FP twice daily. Patients randomized to ESO were instructed to take the medicine 30–60 min before their first meal. Patients randomized to FP were educated by the research pharmacist on correct delivery of the medication using an inhaler without the use of a spacer and instructed not to drink or eat 30 min following administration. The research pharmacist observed the patients priming the metered dose inhaler and administering at least one puff to ensure correct delivery. Adherence was assessed in the ESO arm by counting the number of pills at the end of the treatment study. For the FP arm, the number of puffs was counted using a specially designed metered dose inhaler, which recorded the number of puffs administered. Patients were considered adherent to treatment if ≥80% of the medication was taken during the study period.

Follow-up

Following 8 weeks of treatment, patients underwent repeat upper endoscopy with esophageal biopsies. A total of eight samples using standard biopsy forceps (Boston Scientific, Natick, MA) were taken from all patients, four from the proximal esophagus, ~15 cm from the gastroesophageal junction, and four from the distal esophagus, ~3 cm above the gastroesophageal junction. All endoscopies were performed with Olympus P160 or 180 endoscopes (Olympus, Tokyo, Japan). Endoscopic data was collected including concentric rings, longitudinal furrows,

© 2013 by the American College of Gastroenterology

white plaques, mucosal tearing/friability, strictures, Schatzki rings, and erosive esophagitis. The MDQ was completed by all patients at the end of the study.

Histopathology

All biopsy samples were embedded in formalin and stained with hematoxylin and eosin. Samples to include slides from index endoscopy and following treatment were reviewed and read by a single-blinded expert gastrointestinal pathologist. Each slide had three separate sections and all were reviewed. Eosinophils were counted in all available fields limited by the size of the biopsy specimens. A high-powered field was considered ×40 magnification on our microscopes, which measured 0.19 mm^2. The total number of eosinophils in all fields was counted and the peak eosinophil count per hpf was reported. Findings of eosinophilic microabscesses, intercellular edema, evidence of eosinophilic degranulation, epithelial basilar hyperplasia, and whether eosinophils were confined to the epithelial basal layer or extended to the epithelial surface (full thickness involvement) were also noted.

End points

The primary end point measured was histological response defined as achieving <7 eos/hpf in both proximal and distal esophageal biopsies following 8 weeks of treatment. Secondary outcomes measured included symptomatic change in dysphagia on the basis of the score from the MDQ and interval change in endoscopic and other histological findings.

Statistical analysis

Sample size estimation was based on the following assumptions: 10% of patients will be GERD positive and respond to PPIs compared with 55% of the patients treated with topical steroids. Controlling the probability of a Type I error at $\alpha = 0.05$, a sample of 38 patients in the treatment groups (19 in each arm) will have 80% power to detect a difference in treatment response of 45%.

Data were collated and analyzed with SPSS 15.0 statistical analysis package (SPSS Inc, Chicago, IL). Categorical data are expressed as frequency and percentage, and continuous data as means and standard deviation (s.d.). Histological response, comparison of histological markers, and endoscopic features between the two treatment groups were analyzed using Fisher's exact test. Within each treatment arm, the change in MDQ score after treatment was compared with Wilcoxon signed rank test. Adherence to treatment was compared using Mann–Whitney *U* test. Spearman rank correlation coefficient (r_s) was used to assess the relationship between reflux impedance episodes in the proximal and distal esophagus to eosinophil counts, as well as the relationship of change in MDQ and eosinophils count post treatment. A *P* value of <0.05 was considered statistically significant. Analysis was performed as intention to treat.

RESULTS

Forty-two patients with esophageal eosinophilia were enrolled into the study. The mean age ± s.d. was 38 years old ±10, 81% were Caucasian, 10% Hispanic, 7% African American, and 1% other. Sixty-two percent of patients had a history of coexisting atopic diathesis (33% seasonal allergies, 29% food allergies, 10% asthma, and 5% eczema). The primary indication for endoscopy was dysphagia (69%), followed by food impaction (19%), heartburn (12%), and other (2%). Nineteen percent (8/42) had GERD by Johnson-DeMeester score and were equally stratified into each treatment arm. Endoscopy revealed Los Angeles Grade A erosive esophagitis in seven patients, all of whom had GERD by Johnson-DeMeester score. There were significantly more acid reflux episodes on 24-h pH monitor in the FP arm compared with the ESO arm (45.8±40.6 vs. 25.5±19.3, $P = 0.045$), as well as impedance reflux episodes (63.6±23.1, 44.8±21.4, $P = 0.012$) (**Table 1**).

There was no significant change in mean eosinophil counts before and after treatment in either arm (FP: 55.9±25 vs. 39.2±29.4, $P = 0.102$; ESO: 42.9±18.9 vs. 30.5±33.7, $P = 0.174$) (**Figure 1**). Histological response was achieved in 33% (7/21) of ESO patients vs. 19% (4/21) of FP patients, $P = 0.484$. Among the eight patients with GERD, all four patients randomized to ESO achieved histological response, whereas none of the four patients randomized to FP achieved a histological response, $P = 0.029$. In the 34 GERD-negative patients, response was achieved in 18% (3/17) of ESO patients vs. 24% (4/17) of FP patients, $P = 1.000$. Among patients with coexisting allergies, response was similar between the two treatment groups, FP: 27% (3/11) vs. ESO: 33% (5/15), $P = 1.000$.

The histological response for FP vs. ESO was similar in the proximal esophagus (29 vs. 55%, $P = 0.118$) and in the distal esophagus (19 vs. 40%, $P = 0.181$). In two patients of the FP arm and four patients of the ESO arm, histological response was achieved in the proximal but not distal esophagus. Improvement in other histological markers of EoE following treatment (basal cell hyperplasia, intercellular edema, eosinophilic microabscess, eosinophilic degranulation, and eosinophilic distribution within the epithelium) was similar between the two treatment arms (**Table 2**).

Frequency and severity of dysphagia were similar at baseline between the two treatment arms. The majority of patients reported moderate to severe dysphagia on question no. 3 of the MDQ with no significant difference between the two groups, (FP 77% vs. ESO 83%, $P = 0.512$). In terms of frequency of dysphagia (question no. 4), the majority of patients reported symptoms from less than once per week to several times per week with no difference between the two treatment arms (FP 89% vs. ESO 83%, $P = 0.646$). Only three patients indicated symptoms with every meal. On baseline questionnaire, 30% of patients randomized to FP avoided fibrous foods (meat, chicken, bread, celery, salad) compared with 42% of patients randomized to ESO, $P = 0.381$.

The MDQ score before and after therapy significantly improved in the ESO group (19±21 vs. 1.4±4.5 $P = 0.001$), but not in the FP group (17±18 vs. 12±16, $P = 0.162$) (**Figure 2**). A similar finding was noted among GERD-negative patients: there was a significant difference in MDQ score before and after treatment with ESO (16±14 to 1.7±5.0, $P = 0.001$) but not with FP (18±19 to 10±16, $P = 0.086$). Overall, there was no significant correlation between the change in symptoms by MDQ and the change in eosinophil count in the proximal ($r_s = 0.001$, $P = 0.996$) and distal ($r_s = 0.101$, $P = 0.558$) esophagus.

Table 1. Baseline patient characteristics

Variables	Fluticasone (*n*=21)	Esomeprazole (*n*=21)	*P* value
Mean age±s.d., years	37.0±11.1	38.0±8.8	0.771
Male, *n* (%)	19 (90.5)	19 (90.5)	1.000
Race, n *(%)*			0.766
Caucasian	17 (81.0)	17 (81.0)	—
Hispanic	3 (14.3)	2 (9.5)	—
African American	1 (4.7)	2 (9.5)	—
Coexisting allergies, n *(%)*			
Any atopic disease	11 (52.4)	15 (71.4)	0.341
Seasonal allergies	7 (33.3)	7 (33.3)	1.000
Asthma	2 (9.5)	2 (9.5)	1.000
Eczema	2 (9.5)	0 (0.0)	0.488
Food allergies	3 (14.3)	9 (42.9)	0.085
Concomitant use of allergy medications, n *(%)*			
Antihistamines	4 (19.0)	4 (19.0)	1.000
Nasal steroid spray	3 (14.3)	1 (4.7)	0.606
Leukotriene antagonist	1 (4.7)	1 (4.7)	1.000
GERD by pH score	4 (19.0)	4 (19.0)	1.000
Erosive esophagitis[a]	3 (14.3)	4 (19.0)	1.000
pH reflux episodes			
No. of episodes	45.8±40.6	25.5±19.3	0.045
% (*x*/*n*) abnormal patients	28.5% (6/21)	9.5% (2/21)	0.238
Impedance reflux episodes			
No. of episodes	63.6±23.1	44.8±21.4	0.012
% (*x*/*n*) abnormal patients	38% (8/21)	14.2% (3/21)	0.159
Mayo dysphagia score	17.1±17.8	19.5±20.7	0.691
Pretreatment eosinophil count			
Proximal biopsy	39.1±33.2	32.9±19.4	0.473
Distal biopsy	38.4±22.3	34.2±25.2	0.593

GERD, gastroesophageal reflux disease.
[a]All cases of erosive esophagitis were Los Angeles Grade A.

With regard to resolution of endoscopic findings, no significant difference was seen between FP and ESO (**Table 2**). Dilation was performed on 15 patients on index endoscopy primarily for a dominant stricture seen or a coexisting Schatzki ring (8 patients taking FP and 7 patients taking ESO). No dilations were performed during the treatment period or on follow-up endoscopy post therapy. There was significant improvement in clinical symptoms based on a mean decrease in the MDQ score seen in both treatment groups who underwent dilation (FP = −10.6±10.5, P=0.027 and ESO = −14.3±14.0, P=0.027). Among patients who did not undergo dilation (n=27) on index endoscopy, there was a significant decrease in MDQ score in the ESO group (−20±24, P=0.005) but not in the FP group (−1.9±21.5, P=0.721).

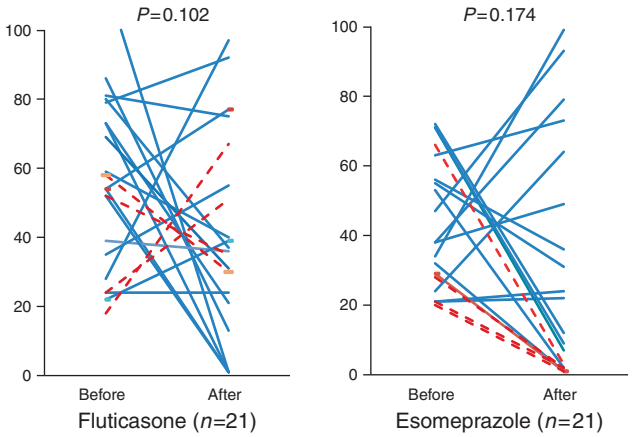

Figure 1. There was no significant change in eosinophil count post treatment in either arm regardless of coexisting gastroesophageal reflux disease (GERD). Dashed lines indicate GERD-positive patients.

To further explore the relationship between reflux and eosinophilia, we examined the association between impedance reflux episodes and eosinophil counts. There was no significant association between impedance reflux episodes and eosinophil counts in the proximal (r_s=0.263, P=0.116) or distal esophagus (r_s=0.162, P=0.359).

In both treatment arms, adherence to therapy was very good. There was no significant difference in adherence to treatment in both groups (FP = 86±24% vs. ESO = 92±10%, P=0.977).

Two patients randomized to the FP arm discontinued treatment during the study period. One patient had worsening of migraine headaches, which he attributed to FP. Another patient had bothersome GERD-related symptoms and discontinued the steroid, and began treatment with a PPIs. Both patients were analyzed as intention to treat. One patient in the fluticasone arm developed esophageal candidiasis. He was asymptomatic during the study period and this was discovered on follow-up endoscopy and confirmed on esophageal biopsies. He was treated with a course of oral fluconazole. No adverse events occurred in the PPI arm.

DISCUSSION

This randomized controlled single-blinded study demonstrated a similar histological response between esomeprazole and fluticasone treatment groups in patients with esophageal eosinophilia. With regard to clinical improvement, based on a validated symptom questionnaire, ESO was significantly better than FP regardless of a concomitant GERD diagnosis.

We defined our patient population who had clinical symptoms of esophageal dysfunction and elevated eosinophil counts on biopsies as having esophageal eosinophilia rather than EoE, based on the most recent updated consensus statement (2). Although swallowed FP is commonly accepted as the first-line treatment for EoE, in this study we examined its efficacy in patients who had phenotypic appearance of EoE and elevated eosinophil counts yet were PPI naive. Our data demonstrated a 19% histological response in patients treated with FP.

© 2013 by the American College of Gastroenterology

Table 2. Data presented for percent (number) of patients with histological and endoscopic findings in pretreatment and posttreatment

	Fluticasone			Esomeprazole		
	Pre	Post	% Improve[a]	Pre	Post	% Improve[a]
Histological findings						
Basal cell hyperplasia	100% (21)	81% (17)	19% (4/21)	100% (21)	52% (11)	43% (9/21)
Intercellular edema	100% (21)	76% (16)	24% (5/21)	86% (18)	57%(12)	35% (7/18)
Eosinophilic microabscess	86% (18)	52% (11)	44% (8/18)	71% (15)	38% (8)	67% (10/15)
Eosinophilic degranulation	95% (20)	52% (11)	50% (10/20)	76% (16)	43% (9)	50% (8/16)
Eosinophilic distribution[b]	95% (20)	76% (16)	25% (5/20)	90% (19)	48%(10)	47% (9/19)
Endoscopic findings						
Stenosis on index endoscopy	24% (5)	14% (3)	80% (4/5)	24% (5)	10% (2)	80% (4/5)
Concentric rings	76% (16)	76% (16)	13% (2/16)	76% (16)	52% (11)	44% (7/16)
Longitudinal furrows	81% (17)	76% (16)	18% (3/17)	81% (17)	52% (11)	41% (7/17)
White plaques	19% (4)	29% (6)	50% (2/4)	24% (5)	0% (0)	100% (5/5)

% improve, % improvement; pre, pretreatment; post, posttreatment.
Percent improvement is among a subgroup of patients who had a pretreatment finding. There was no statistically significant difference in improvement between treatment arms.
[a]Percent improvement is among patient who had a pretreatment finding.
[b]Involving entire thickness of epithelium.

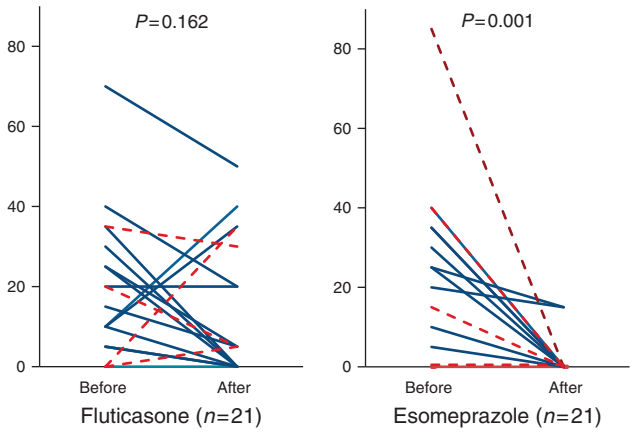

Figure 2. Change in Mayo dysphagia questionnaire (MDQ) score following treatment. There was significant clinical improvement in ESO but not in the FP treatment arm. Dashed lines indicate gastroesophageal reflux disease (GERD)-positive patients.

Response to topical steroids for EoE has varied in the literature with small case series and retrospective studies reporting response rates in up to 80% (1,10). In contrast, data from prospective controlled studies have demonstrated lower response rates. Konikoff *et al.* (11) reported a 50% efficacy (10/20) in a pediatric population with histological response defined as ≤1 eos/hpf. Again, complete histological response was seen in 50% (18/36) of patients after 4 weeks of therapy with FP (12). In another study comparing the efficacy of topical steroids to PPI therapy in adults with EoE, only 15% achieved complete histological response with FP, defined as 5 eos/hpf (13). A recent study by Alexander *et al.* (14) reported a 62% histological response to fluticasone, defined as a more than 90% decrease in mean levels of eosinophils following 6 weeks of therapy. Interestingly, no significant clinical improvement was noted when compared with placebo.

The relatively poor response seen in many randomized studies may at least partially be due to a lack of uniformity in medication dosage, duration of therapy, the definition of response, and delivery method. Additionally, variability in the definition of EoE exists in the literature (15) and in many studies EoE is often synonymous with esophageal eosinophilia. Alexander *et al.* (14) used a higher dose of swallowed fluticasone (880 mcg twice daily), which may reflect the higher response rate seen in their study compared with other randomized controlled studies using topical fluticasone for EoE. Five patients developed esophageal candidiasis, compared with one patient in our study, suggesting this may be dose dependent. Another reason our response rate may have been different when compared with other randomized studies may be due to the definition of response. Variability exists in the literature with some studies using a change in histological grade (12) and even a decrease in percent eosinophils in comparison with baseline (14,16). The majority of the studies have used absolute changes in eosinophil count as an end point to include 1 eos/hpf (11), 5 eos/hpf (13,17), and <7 eos/hpf (18). We chose to use <7 eos/hpf as a target end point. However, we also examined our data using <5 eos/hpf and obtained identical results (19% fluticasone and 33% for esomeprazole).

Another explanation for the low response to FP may be related to the mechanism of drug delivery. FP is an aerosolized medication administered through a metered dose inhaler, which is intended for the airways, and it is unclear how much drug is actually delivered to the esophagus. One would assume that response may be superior in the proximal esophagus than the distal esophagus with this type of delivery. However, when we compared the histological response between the proximal and distal esophagus, there was no

significant difference in eosinophilia between these two locations. It is noteworthy that studies using oral viscous budesonide, which can be administered as a slurry mixed with a sugar substitute or as a swallowed nebulizer, demonstrated higher response rates (72–80%) (17,18). This suggests that the delivery system, and not the steroid type, may be the reason for low response rates. To further study and explore this point, a recent study demonstrated that oral viscous budesonide was more effective in reducing eosinophil counts when compared with nebulized budesonide (19).

Interestingly, there was a greater number of acid and impedance reflux episodes in the FP arm compared to the ESO arm. However, the percentage of patients with abnormal acid (>50) and impedance episodes (>73) was similar between both treatment arms. Therefore, it is unclear whether an increased number of reflux episodes had any role in the response to topical steroids.

The rationale for using PPIs for the treatment of esophageal eosinophilia is that it may help treat underlying GERD, which may contribute to EoE (20). It may be difficult to completely separate these two entities and it appears that many adult EoE patients respond at least clinically to PPI therapy. In one study, 8 of 17 (47%) adult patients presenting with food impaction and dense esophageal eosinophilia (>20 eos/hpf) responded clinically to PPI therapy (5). The mechanism for this is not clear, although PPIs may help heal a disrupted epithelial barrier and therefore reduce further immune activation (2). In our study, the four patients with esophageal eosinophilia and GERD had a complete response to PPI therapy.

The histological response achieved in three patients with PPI therapy despite having a negative 24-h pH study suggests these patients may have PPI-responsive EoE, which stimulates an interesting discussion regarding the pathophysiology of this entity. This described response to esomeprazole in GERD-negative esophageal eosinophilia patients can be attributed to anti-inflammatory properties of PPI independent of acid inhibition. Interestingly, PPIs have been shown to exhibit anti-inflammatory properties by acting directly on principal cytokines (IL-4 and IL-13) involved in the recruitment of eosinophils in the esophagus (21). Additionally, PPI therapy was shown to block the release of eotaxin-3, which has an integral role in the pathogenesis of EoE (22). Other studies have also reported a histological response to PPI therapy in EoE patients even following a negative 24-h pH study. Peterson *et al.* (13) demonstrated a 33% (4/12) histological response to PPI therapy. Similar results have been reported in children as well. In 43 pediatric EoE patients, a subset of patients responded histologically to PPI therapy following a normal pH study (23). These studies coupled with our data suggest that a subset of patients whose clinical presentation suggests EoE and have dense eosinophilia on esophageal biopsies indeed appear to be PPI responsive.

Significant clinical improvement was seen in the majority of patients with esophageal eosinophilia taking PPI even though only a few had histological resolution. Other studies have also reported a poor correlation between clinical remission and histological response. In a randomized placebo controlled study, vomiting was the only symptom that significantly improved in EoE children treated with topical steroids and did not correlate with histological response (11). Similar findings have been described by Molina *et al.* (24) in which clinical remission was achieved in the majority of patients with food impaction or dysphagia despite persistence of eosinophilia. Interestingly, a dissociation was recently reported between clinical response and histological severity in EoE children with the majority of patients reporting continued symptoms despite histological remission (25). These data combined with ours suggest one cannot use clinical response as a surrogate for histological change. However, whether achieving histological response in EoE is an important outcome is yet to be determined. In theory, normalization of the eosinophila may prevent further tissue remodeling, fibrosis, and possibly stricture formation.

Esophageal dilation is one of the most effective treatments for alleviating the symptoms of EoE even though it does not address the underlying disease pathophysiology (26). In our cohort, all of the 15 patients (8 in FP arm and 7 in ESO arm) with dysphagia and esophageal eosinophilia who underwent dilation at index endoscopy had significant improvement in symptoms. Similar clinical improvements have been noted in other studies in EoE (26,27). In contrast, among the 27 patients who did not undergo dilation, only those randomized to ESO had significant improvement in symptoms. As a similar number of patients underwent dilation in each treatment arm and dilation was not performed during the 8-week treatment course or at follow-up endoscopy, it is unlikely that dilation had an effect on the change in MDQ scores before and after therapy.

There are some limitations of our study. A placebo arm would have helped better define the natural history of esophageal eosinophilia. As we were comparing one medication administered as a pill to another given as an inhaler, we found adding a placebo arm to be challenging. Additionally the delivery system of aerosolized fluticasone may have hampered the true assessment of steroid response rates, and using another agent such as oral viscous budesonide may give us a different understanding of the true steroid response. Another limitation of our study was the relatively small sample size, which limited the subgroup analyses. Additionally, we used the 2-week dysphagia score, which may not have accurately reflected the symptoms in some EoE patients, particularly in those with intermittent dysphagia. However, the 2-week MDQ was the only validated questionnaire available at the time of this study. To date, no dysphagia questionnaire developed exclusively for EoE has been published.

In conclusion, histological response between ESO and FP were similar in the treatment of esophageal eosinophilia, with neither drug having overwhelming treatment success. On the other hand, significant improvement in clinical symptoms was demonstrated with PPI therapy. Further larger studies are needed to better define the optimal treatment for patients with esophageal eosinophilia and to better describe the subgroup and natural history of such patients who respond to PPI therapy.

CONFLICT OF INTEREST

Guarantor of the article: Fouad J. Moawad, MD.

Specific author contributions: Study concept and design, subject enrollment, acquisition of data, analysis and interpretation of data,

© 2013 by the American College of Gastroenterology

drafting of the manuscript, critical revision of the manuscript for important intellectual content: Fouad J. Moawad; study concept and design, study enrollment, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content: Ganesh R. Veerappan; data collection, verification and analysis, drafting of the manuscript: Johnny A. Dias; histological analysis, drafting of the manuscript: Thomas P. Baker; study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content: Corinne L. Maydonovitch; study concept and design, drafting of the manuscript, critical revision of the manuscript for important intellectual content: Roy K.H. Wong.

Financial support: None.

Potential competing interests: None.

Disclaimer

The opinions are solely those of the authors and do not represent an endorsement by the Department of Defense. This is US Government work. There are no restrictions on its use.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Topical steroids are commonly used as first line treatment for patients with a clinical presentation suggestive of EoE and who have dense eosinophilic infiltration.
- ✓ PPI therapy is often prescribed in esophageal eosinophilia patients to help treat coexisting GERD and establish diagnosis of EoE.

WHAT IS NEW HERE

- ✓ Topical fluticasone had a lower than expected response rate which may be dose and delivery related.
- ✓ PPIs induce histological response in some patients with esophageal eosinophilia regardless of the presence of GERD.
- ✓ PPIs significantly improved clinical symptoms in patients with esophageal eosinophilia even in the absence of GERD.

REFERENCES

1. Moawad FJ, Veerappan GR, Wong RK. Eosinophilic esophagitis. Dig Dis Sci 2009;54:1818–28.
2. Liacouras CA, Furuta GT, Hirano I *et al.* Eosinophilic esophagitis: updated consensus recommendations for children and adults. J Allergy Clin Immunol 2011;128:3–20.
3. Furuta GT, Liacouras CA, Collins MH *et al.* Eosinophilic esophagitis in children and adults: a systematic review and consensus recommendations for diagnosis and treatment. Gastroenterology 2007;133:1342–63.
4. Ngo P, Furuta GT, Antonioli DA *et al.* Eosinophils in the esophagus-peptic or allergic eosinophilic esophagitis? Case series of three patients with esophageal eosinophilia. Am J Gastroenterol 2006;101:1666–70.
5. Desai TK, Stecevic V, Chang CH *et al.* Association of eosinophilic inflammation with esophageal food impaction in adults. Gastrointest Endosc 2005;61:795–801.
6. Hirano I. Eosinophilic esophagitis and gastroesophageal reflux disease: there and back again. Clin Gastroenterol Hepatol 2011;9:99–101.
7. Grudell AB, Alexander JA, Enders FB *et al.* Validation of the Mayo Dysphagia Questionnaire. Dis Esophagus 2007;20:202–5.
8. Johnson LF, Demeester TR. Twenty-four-hour pH monitoring of the distal esophagus. A quantitative measure of gastroesophageal reflux. Am J Gastroenterol 1974;62:325–32.
9. Johnson LF, DeMeester TR. Development of the 24-hour intraesophageal pH monitoring composite scoring system. J Clin Gastroenterol 1986;8 (Suppl 1): 52–8.
10. Noel RJ, Putnam PE, Collins MH *et al.* Clinical and immunopathologic effects of swallowed fluticasone for eosinophilic esophagitis. Clin Gastroenterol Hepatol 2004;2:568–75.
11. Konikoff MR, Noel RJ, Blanchard C *et al.* A randomized, double-blind, placebo-controlled trial of fluticasone propionate for pediatric eosinophilic esophagitis. Gastroenterology 2006;131:1381–91.
12. Schaefer ET, Fitzgerald JF, Molleston JP *et al.* Comparison of oral prednisone and topical fluticasone in the treatment of eosinophilic esophagitis: a randomized trial in children. Clin Gastroenterol Hepatol 2008;6:165–73.
13. Peterson KA, Thomas KL, Hilden K *et al.* Comparison of esomeprazole to aerosolized, swallowed fluticasone for eosinophilic esophagitis. Dig Dis Sci 2010;55:1313–9.
14. Alexander JA, Jung KW, Arora AS *et al.* Swallowed fluticasone improves histologic but not symptomatic response of adults with eosinophilic esophagitis. Clin Gastroenterol Hepatol 2012;10:742–9.
15. Sperry SL, Shaheen NJ, Dellon ES. Toward uniformity in the diagnosis of eosinophilic esophagitis (EoE): the effect of guidelines on variability of diagnostic criteria for EoE. Am J Gastroenterol 2011; 106:824–32.
16. Straumann A, Conus S, Degen L *et al.* Budesonide is effective in adolescent and adult patients with active eosinophilic esophagitis. Gastroenterology 2010;139:1526–37.
17. Straumann A, Conus S, Degen L *et al.* Long-term budesonide maintenance treatment is partially effective for patients with eosinophilic esophagitis. Clin Gastroenterol Hepatol 2011;9:400–9.
18. Aceves SS, Bastian JF, Newbury RO *et al.* Oral viscous budesonide: a potential new therapy for eosinophilic esophagitis in children. Am J Gastroenterol 2007;102:2271–9.
19. Dellon ES, Sheikh A, Speck O *et al.* Viscous Topical Is More Effective Than Nebulized Steroid Therapy for Patients With Eosinophilic Esophagitis. Gastroenterology 2012;143:321–4.
20. Spechler SJ, Genta RM, Souza RF. Thoughts on the complex relationship between gastroesophageal reflux disease and eosinophilic esophagitis. Am J Gastroenterol 2007;102:1301–6.
21. Cortes JR, Rivas MD, Molina-Infante J *et al.* Omeprazole inhibits IL-4 and IL-13 signaling signal transducer and activator of transcription 6 activation and reduces lung inflammation in murine asthma. J Allergy Clin Immunol 2009;124:607–10.
22. Zhang X CE, Hou X *et al.* In esophageal epithelial cell lines from patients with eosinophilic esophagitis (EoE), omeprazole blocks the stimulated secretion of eotaxin-3: a potential anti-inflammatory effect of omeprazole in EoE that is independent of acid inhibition. Gastroenterology 2010;138 (suppl 1): S–122.
23. Dranove JE, Horn DS, Davis MA *et al.* Predictors of response to proton pump inhibitor therapy among children with significant esophageal eosinophilia. J Pediatr 2009;154:96–100.
24. Molina-Infante J, Ferrando-Lamana L, Ripoll C *et al.* Esophageal eosinophilic infiltration responds to proton pump inhibition in most adults. Clin Gastroenterol Hepatol 2011;9:110–7.
25. Pentiuk S, Putnam PE, Collins MH *et al.* Dissociation between symptoms and histological severity in pediatric eosinophilic esophagitis. J Pediatr Gastroenterol Nutr 2009;48:152–60.
26. Schoepfer AM, Gonsalves N, Bussmann C *et al.* Esophageal dilation in eosinophilic esophagitis: effectiveness, safety, and impact on the underlying inflammation. Am J Gastroenterol 2010;105:1062–70.
27. Straumann A, Spichtin HP, Grize L *et al.* Natural history of primary eosinophilic esophagitis: a follow-up of 30 adult patients for up to 11.5 years. Gastroenterology 2003;125:1660–9.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2015; 124(3):221-226.

Article

Utility of the Transnasal Esophagoscope in the Management of Chemoradiation-Induced Esophageal Stenosis

Annals of Otology, Rhinology & Laryngology
2015, Vol. 124(3) 221–226
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489414550858
aor.sagepub.com
SAGE

Kevin A. Peng, MD[1], Aaron J. Feinstein, MD, MHS[1], Jonathan B. Salinas, MD[1], and Dinesh K. Chhetri, MD[1]

Abstract

Objective: This study aimed to describe management of esophageal stenosis after chemoradiation therapy for head and neck squamous cell carcinoma (HNSCC), with particular emphasis on techniques and outcomes with the use of the transnasal esophagoscope (TNE) in the office as well as operating room settings.

Methods: Retrospective analysis of all patients with esophageal stenosis following head and neck cancer radiation, with or without chemotherapy, and managed with TNE-assisted esophageal dilation over a 5-year period. Preoperative and postoperative swallowing function were assessed objectively with the Functional Outcome Swallowing Scale (FOSS; ranging from score 0, a normal diet, to score 5, complete dependence on nonoral nutrition).

Results: Twenty-five patients met inclusion criteria. The mean pretreatment FOSS score was 4.4, whereas the mean posttreatment FOSS score was 2.7 (Wilcoxon signed-rank test, $P < .001$). Prior to dilation, 16 patients were completely gastrostomy-tube dependent (FOSS 5), of whom 12 (75%) were able to tolerate oral nutrition for a majority of their diet following treatment according to our protocol. No complications were noted.

Conclusion: Dysphagia following chemoradiation therapy for HNSCC is often related to esophageal stenosis. With the aid of TNE, we have developed a successful treatment strategy for esophageal stenosis with improved success rates.

Keywords

chemoradiation, esophageal dilation, esophageal stenosis, head and neck squamous cell carcinoma, transnasal esophagoscopy

Introduction

Squamous cell carcinomas of the head and neck occur frequently, with more than 500000 cases diagnosed worldwide annually.[1] Radiation with concurrent chemotherapy (CRT) is an increasingly used treatment modality for these cancers. As survival rates improve with advances in care, organ preservation—that is, the maintenance of normal mechanisms of breathing, deglutition, and communication—becomes of paramount importance. Following successful treatment of head and neck squamous cell cancer (HNSCC), dysphagia is the most common symptom decreasing quality of life, affecting 50% to 64% of patients after CRT.[2,3]

Whereas early dysphagia is usually temporary, late dysphagia often results from chronic inflammation and fibrosis and is much more difficult to manage.[4-7] This fibrosis may progress to hypopharyngeal or esophageal strictures, which occur in approximately 21% of patients undergoing CRT.[8,9] Risk factors implicated in stricture formation in the general population include reflux, older age, and caustic ingestion; among head and neck cancer patients with HNSCC, additional factors include hypopharyngeal primary site, combined chemoradiation (vs radiation alone), radiation dose, prior neck dissection, female sex, and treatment-induced mucositis.[10]

Objective assessment of dysphagia is essential and comprises 2 complementary tests: the videofluoroscopic swallow study, also known as a modified barium swallow study (MBSS), and the functional endoscopic evaluation of swallow (FEES).[11] Whereas the advantages of FEES include rapidity of the test in an office setting, direct observation of native secretions and swallow anatomy, and lack of radiation for the procedure, MBSS is superior in evaluating the oral and upper esophageal phases.[12] In addition to these tests, flexible transnasal esophagoscopy has seen increasing use in the otolaryngology dysphagia clinic, particularly in

[1]Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

Corresponding Author:
Dinesh K. Chhetri, MD, Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA, 10833 Le Conte Avenue, CHS 62-132, Los Angeles, CA 90095, USA.
Email: dchhetri@mednet.ucla.edu

Table 1. Functional Outcome Swallowing Scale.[a]

Stage	Description
0	Normal function; asymptomatic
1	Normal function; episodic or daily symptoms of dysphagia
2	Compensated abnormal function manifested by significant dietary modifications or prolonged mealtime, without weight loss or aspiration
3	Decompensated abnormal function, with weight loss of 10% or less of body weight over 6 months due to dysphagia, or daily cough, gagging, or aspiration during meals
4	Severely decompensated abnormal function, with weight loss of more than 10% of body weight over 6 months due to dysphagia, or severe aspiration with bronchopulmonary complications; nonoral feeding recommended for most of nutrition
5	Nonoral feeding for all nutrition

[a]Adapted from Salassa.[15]

evaluating the presence, severity, and length of esophageal stenoses.[13]

Following assessment of the stenosis, esophageal dilations with Savary-Gilliard dilators or controlled radial expansion (CRE) balloons can be performed in both operating room and office-based settings with modifications of prior techniques.[14] In this study, we review our management of esophageal stenosis after CRT for HNSCC. We place special emphasis on the use of the transnasal esophagoscope (TNE) to demonstrate that this method has a high success rate with minimal potential for complications. The Functional Outcome Swallowing Scale (FOSS), described by Salassa[15] in 1999, was used to quantify dysphagia prior to and following treatment (Table 1).

Methods

Institutional review board approval was obtained for this study. Inclusion criteria were as follows: history of HNSCC treated with radiation and/or chemotherapy, presence of esophageal stenosis, management of esophageal stenosis by esophagoscopy and dilation, and documentation of swallowing function with instrumental swallow tests (MBSS and/or FEES) both prior to and following dysphagia treatment. Patients with multilevel esophageal stenosis were excluded. Patients who underwent surgery, including tracheostomy, neck dissection, or resection of the primary tumor, were also excluded, with the following exceptions: gastrostomy tube (G-tube) placement, tonsillectomy, or panendoscopy with biopsies.

We managed esophageal stenosis with the following algorithm. After a history and physical examination were performed, FEES was performed in office to assess the current safe diet, and transnasal esophagoscopy was performed if esophageal stenosis was suspected based on MBSS, dysphagia to solid foods, or severe piriform sinus residue. Following a definitive office diagnosis of esophageal stenosis, patients were scheduled for surgery.

In the operating room, suspension direct laryngoscopy was performed under general anesthesia and the rigid operating laryngoscope was placed in the postcricoid space. The TNE was then passed through the laryngoscope into the hypopharynx and advanced into the esophagus. Several dilation scenarios were possible at this point. (1) In a majority of cases, the stricture was seen, and the TNE could be passed atraumatically beyond the stenosis and into the distal esophagus. A CRE balloon was then passed through the stricture under direct visualization and dilation was performed to 18 mm. Alternatively, a Savary-Gilliard dilator guidewire could be passed through the working port of the scope, and dilation could then be performed over the guidewire after retracting the scope completely. (2) The stricture was seen, but the TNE would not pass through the stricture. This indicated that the stenosis diameter was smaller than the diameter of our scope (5.1 mm). At this point, gentle passage of a Savary-Gilliard dilator guidewire was attempted while directly visualizing its passage through the stenotic opening. If this was possible without resistance, dilation was then performed using Savary-Gilliard dilators up to 9 to 10 mm just past the stenosis. Then, the TNE was passed through the stenosis to ensure normal esophageal lumen, after which CRE balloon dilation was performed, typically to 15 mm. (3) A complete stricture was encountered, and the TNE could not pass. In this case, the G-tube was removed and retrograde esophagoscopy was performed. The TNE could be inserted through the G-tube site without dilation of the G-tube tract (Figure 1) and was advanced through the lower esophageal sphincter to the upper esophagus to the stricture site. Anterograde palpation of the esophageal stricture using a blunt instrument such as a rigid esophageal suction tube assisted in identifying the stricture, and under direct retrograde visualization with the TNE, the stricture was punctured. A Savary-Gilliard guidewire was passed through the stenosis with both retrograde and anterograde visualization. This guidewire could be inserted in an anterograde manner under direct visualization of the TNE and dilation performed as in situation 2 above. Topical mitomycin-C (MMC), which inhibits fibroblast proliferation, was applied to the affected region in all cases, using cotton pledgets at a concentration of 0.4 mg/mL for 4 minutes, a technique that has been successfully applied in the treatment of upper aerodigestive tract stenosis.[16-19]

A second dilation was scheduled for 1 to 2 weeks after the first dilation. The second dilation allows an assessment of efficacy of the first dilation, which helps to counsel patients on the anticipated treatment course in regard to repeat dilations and provides an opportunity for the second

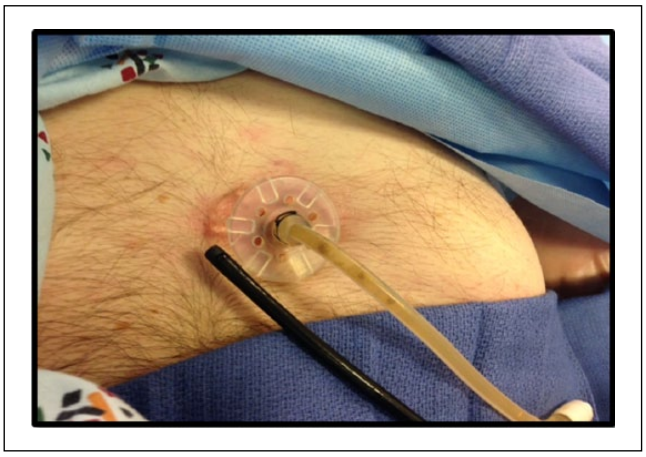

Figure 1. The transnasal esophagoscope (TNE) used in this study pictured adjacent to a standard gastrostomy tube (G-tube), showing similarity of diameters. With the G-tube removed, the TNE can be passed for retrograde esophagoscopy without further dilation of the G-tube site.

application of MMC.[19] Dilation was typically performed to 18 to 20 mm diameter using the CRE balloon. All patients were referred for swallow therapy after the second dilation. Some patients were scheduled for office dilation depending on the degree of stenosis and residual dysphagia. Office dilation was undertaken preferentially as many patients had significant trismus and were high anesthetic risks regarding intubation. Office esophageal dilations were accomplished as follows: bilateral nasal cavities were anesthetized and decongested with topical lidocaine and oxymetazoline. Thereafter, transnasal esophagoscopy was performed via the more patent nasal cavity; once the stenosis was identified, a CRE balloon dilator was passed via the contralateral nasal cavity and to the level of stenosis under direct visualization. Passage of the balloon was sometimes aided by bending the tip slightly to traverse the nasopharyngeal curvature. Dilation was then performed using the CRE balloon, typically to 18 mm. Patients received proton-pump inhibitors for the first 3 months after initial dilation, with further prescriptions based on the presence of reflux symptoms. Esophagoscopy and dilations were performed until the patient's symptoms were alleviated satisfactorily.

To analyze outcomes of our esophageal stenosis treatment algorithm, pretreatment and posttreatment FOSS scores were compared with a Wilcoxon signed-rank test. The number of dilations undergone by each patient was noted, as was the elapsed time between dilations.

Results

Among the 115 patients identified, 81 were excluded due to use of surgery (eg, neck dissection, tumor resection) during their initial treatments and 9 were excluded due to requiring additional surgical procedures at the time of esophageal dilation. There were 25 patients, 21 male and 4 female, who met inclusion criteria (Table 2). The median age was 63 years (range, 40-84 years). The most common primary site was oropharynx (n = 13, 52%), followed by an unknown primary (n = 5, 20%). All patients received combined chemoradiation therapy. Median time from completion of CRT to initiation of esophageal stenosis management was 6.0 months (range, 2 months to 30 years). All patients in the study had single-level stenosis.

Table 2. Patient Characteristics.

Age, y	
Median	63
Range	40-84
Sex	
Male	21
Female	4
Primary site	
Oropharynx	13 (52%)
Unknown primary	5 (20%)
Hypopharynx	3 (12%)
Nasopharynx	2 (8%)
Larynx	1 (4%)
Oral cavity	1 (4%)
Elapsed time between termination of CRT and initiation of esophageal stenosis treatment	
Median	6.0 months
Range	2 months to 30 years
Elapsed time between esophageal dilations	
Median	21 days
Range	6 days to 1.8 years
Number of dilations performed	
Median	2
Range	1-16
Functional Outcome Swallowing Scale score	
Prior to treatment of esophageal stenosis	
Mean	4.36
Range	2-5
Following treatment of esophageal stenosis	
Mean	2.40
Range	1-5

The median number of dilations performed on each patient was 2 (range, 1-16). For patients undergoing multiple dilations, the median time between procedures was 21 days (range, 6 days to 21 months). In 3 patients (12%) who were completely G-tube dependent, retrograde esophagoscopy was performed via the G-tube with the TNE to delineate the esophageal lumen. Mean pretreatment FOSS score for all patients was 4.4 (median, 5; range, 2-5); mean posttreatment FOSS score was 2.7 (median, 3; range, 1-5). A Wilcoxon signed-rank test confirmed a

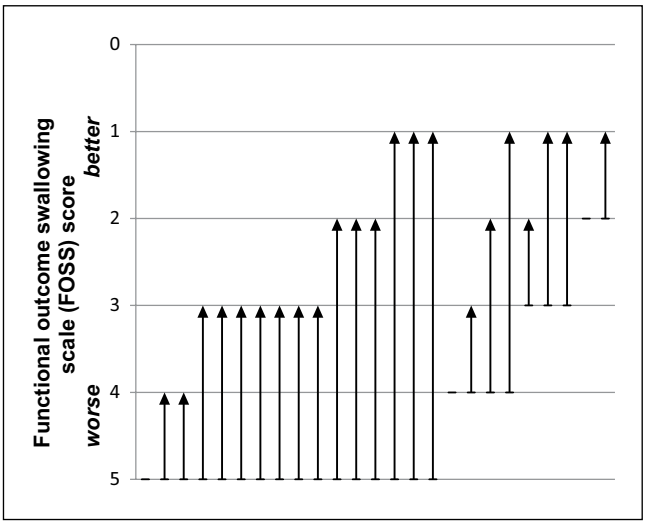

Figure 2. Improvement in Functional Outcome Swallowing Scale (FOSS) score was seen in all but 3 of 25 patients following our esophageal dilation protocol; no patients worsened after therapy. Arrows depict change in FOSS scores following therapy.

statistically significant difference between FOSS scores prior to and following esophageal stenosis treatment ($P <$ .001). The FOSS score did not worsen in any patients (Figure 2).

Prior to treatment, 16 patients (64%) were completely dependent on nonoral nutrition, primarily via G-tube (FOSS score of 5); following treatment, only 2 patients (8%) were completely dependent on nonoral nutrition. Of the 16 patients completely dependent on nonoral nutrition prior to treatment, 12 (75%) transitioned to oral intake for a majority of their nutrition following therapy (FOSS score of 3 or better). Out of all patients studied, 6 (24%) were ultimately on a normal diet following therapy (FOSS score of 0 or 1).

Only 3 patients required 10 or more dilations. Two of these had required initial combined anterograde-retrograde dilations via the gastrostomy, whereas the third received numerous maintenance office dilations. They were typically treated about 3 months apart as they subjectively felt improvements with each office dilation.

Patients who were treated within 6 months after completion of CRT (early dilation) had improved results relative to those treated beyond 6 months (late dilation). Among the 13 patients with early dilation, the mean pretreatment and posttreatment FOSS scores were 4.5 and 2.2, respectively, whereas the 12 patients with late treatment had mean pretreatment and posttreatment FOSS scores of 4.2 and 2.7, respectively. Only 1 of 13 early patients had a posttreatment FOSS score of 4 or 5, as compared to 3 of 12 patients in the late group. There were no documented complications, including zero occurrences of esophageal perforation or mediastinitis.

Discussion

Dysphagia resulting from esophageal stenosis following successful chemoradiation therapy for HNSCC has a significant effect on quality of life.[20] In this setting, optimal treatment is accomplished with the use of serial dilation.[6,21,22] At our institution, we have developed an algorithm to manage esophageal stenosis in the setting of prior CRT, where initial evaluation includes the complementary studies of MBSS, FEES, and transnasal esophagoscopy.

The first dilation occurs in a controlled, operative setting under general anesthesia. The flexible scope is preferred because many of these patients have trismus, friable pharyngeal mucosa, and/or lack of extension precluding rigid esophagoscopy. The otolaryngologist is also more familiar with use of this scope, which has improved maneuverability compared to the regular or even the "ultrathin" but long scope that is typically used in gastroenterology. Following visualization of the stenosis, dilation is performed with CRE balloon or Savary-Gilliard dilators. When using the latter, a guidewire is first passed atraumatically through the stenosis—either parallel to the scope or through the working port of the scope—before the dilator is introduced, thus minimizing the risk of mucosal trauma or extraluminal passage. Retrograde esophagoscopy via the gastrostomy site remains a safe option for patients with complete stenosis. Mitomycin-C can also be applied at this time. The complication risk is very low, and all patients could be discharged to home after recovery from anesthesia. Depending on the severity of stenosis, the timing and the setting of future dilations (office vs operative) are determined.

In our series of patients, we have demonstrated excellent outcomes with our structured management of esophageal stenosis. On Wilcoxon signed-rank test, there was a statistically significant improvement (ie, decrease) in FOSS score, with 6 patients (24%) ultimately tolerating a normal diet (FOSS score of 1). Sixteen patients (64%) were initially G-tube dependent (FOSS score of 5); 12 of these patients (75%) tolerated the oral route for the majority of nutrition (FOSS score of 3 or better) following our therapy.

This compares favorably to previous series: Silvain et al[6] described an early series of 11 patients with esophageal stricture, 9 of whom underwent dilation. This series noted complications in 4 patients, including 1 death, and 4 patients were described to have a semisolid diet after treatment. Dhir et al[23] performed dilations on 21 patients who had undergone radiation with or without surgery and achieved dysphagia relief in 15 of 20 (75%) patients for a median of 14 weeks; however, long-term follow-up was not available. Laurell et al[7] described a similar group who developed moderate to severe esophageal stenosis; their management included both endoscopic dilation and microvascular free flap esophageal reconstruction. In this study, a "nearly normal" diet was achieved in 17 of 22 (78%) patients, although

there was no report of preintervention or postintervention G-tube status. Ahlawat et al[24] performed dilation on 24 patients and reported technical success (endoscopic dilation to 14 mm) in 19 patients and functional success (occasional dysphagia to solid foods) in 18 patients. Again, G-tube status was not available. Our technique improves on these outcomes, however, as the rate of conversion from G-tube dependence to predominantly oral nutrition—75% in our study population—greatly exceeds the success rates reported previously.[6,7,23,24] Furthermore, whereas others have demonstrated good results (81% of patients maintaining weight with oral diet) from dilation of the hypopharynx and upper esophagus,[25] we have achieved these results without complications and with serial dilations in the clinic setting without general anesthesia.

Our use of the TNE accomplishes both diagnostic and therapeutic purposes. Transnasal esophagoscopy is well tolerated in awake patients in the office setting, and we employ the same scope in the operating room, which is beneficial for consistency in assessing the degree of stenosis. Exposure for rigid esophagoscopy may be quite difficult or impossible following CRT, and thus use of the flexible TNE improves our ability to treat challenging cases. Some of the residual esophageal lumens are quite small, and using the 5.1-mm TNE allows successful passage through the stenosis that is not always achieved with the larger gastroscopes. Similarly, the small size allows retrograde passage through the gastrostomy without requiring dilation, thus minimizing morbidity; our results compare very favorably to another series of 45 patients using the retrograde approach reporting G-tube site morbidities in 7 of 63 (11%) procedures.[26] The ability to perform transnasal esophagoscopy and dilation in the office setting confers additional advantages, not in the least that general anesthesia and its concomitant risks are avoided.

Conclusion

Patients with esophageal stenosis after CRT can be successfully managed, with the majority achieving a full oral diet. Transnasal esophagoscopy is an important tool in our armamentarium of management of esophageal stenosis following chemoradiation for head and neck cancer. The versatility of transnasal esophagoscopy as an adjunct to esophageal dilation, with either guidewire or balloon dilators, allows for its use in both operative and office settings. As demonstrated here, our algorithm is well tolerated, highly effective, and associated with little morbidity.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005;55(2):74-108.
2. Francis DO, Weymuller EA Jr, Parvathaneni U, Merati AL, Yueh B. Dysphagia, stricture, and pneumonia in head and neck cancer patients: does treatment modality matter? *Ann Otol Rhinol Laryngol*. 2010;119(6):391-397.
3. Nguyen NP, Sallah S, Karlsson U, Antoine JE. Combined chemotherapy and radiation therapy for head and neck malignancies: quality of life issues. *Cancer*. 2002;94(4):1131-1141.
4. Lazarus CL, Logemann JA, Pauloski BR, et al. Swallowing disorders in head and neck cancer patients treated with radiotherapy and adjuvant chemotherapy. *Laryngoscope*. 1996;106(9, pt 1):1157-1166.
5. Smith RV, Kotz T, Beitler JJ, Wadler S. Long-term swallowing problems after organ preservation therapy with concomitant radiation therapy and intravenous hydroxyurea: initial results. *Arch Otolaryngol Head Neck Surg*. 2000;126(3):384-389.
6. Silvain C, Barrioz T, Besson I, et al. Treatment and long-term outcome of chronic radiation esophagitis after radiation therapy for head and neck tumors. A report of 13 cases. *Dig Dis Sci*. 1993;38(5):927-931.
7. Laurell G, Kraepelien T, Mavroidis P, et al. Stricture of the proximal esophagus in head and neck carcinoma patients after radiotherapy. *Cancer*. 2003;97(7):1693-1700.
8. Hutcheson KA, Lewin JS. Functional outcomes after chemoradiotherapy of laryngeal and pharyngeal cancers. *Curr Oncol Rep*. 2012;14(2):158-165.
9. Lee WT, Akst LM, Adelstein DJ, et al. Risk factors for hypopharyngeal/upper esophageal stricture formation after concurrent chemoradiation. *Head Neck*. 2006;28(9):808-812.
10. Best SR, Ha PK, Blanco RG, et al. Factors associated with pharyngoesophageal stricture in patients treated with concurrent chemotherapy and radiation therapy for oropharyngeal squamous cell carcinoma. *Head Neck*. 2011;33(12):1727-1734.
11. Nguyen NP, Smith HJ, Sallah S. Evaluation and management of swallowing dysfunction following chemoradiation for head and neck cancer. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15(2):130-133.
12. Platteaux N, Dirix P, Dejaeger E, Nuyts S. Dysphagia in head and neck cancer patients treated with chemoradiotherapy. *Dysphagia*. 2010;25(2):139-152.
13. Postma GN, Bach KK, Belafsky PC, Koufman JA. The role of transnasal esophagoscopy in head and neck oncology. *Laryngoscope*. 2002;112(12):2242-2243.
14. Lew RJ, Shah JN, Chalian A, Weber RS, Williams NN, Kochman ML. Technique of endoscopic retrograde puncture and dilatation of total esophageal stenosis in patients with radiation-induced strictures. *Head Neck*. 2004;26(2):179-183.
15. Salassa JR. A functional outcome swallowing scale for staging oropharyngeal dysphagia. *Dig Dis*. 1999;17(4):230-234.

16. Veen EJ, Dikkers FG. Topical use of MMC in the upper aerodigestive tract: a review on the side effects. *Eur Arch Otorhinolaryngol*. 2010;267(3):327-334.
17. Gillespie MB, Day TA, Sharma AK, Brodsky MB, Martin-Harris B. Role of mitomycin in upper digestive tract stricture. *Head Neck*. 2007;29(1):12-17.
18. Rosseneu S, Afzal N, Yerushalmi B, et al. Topical application of mitomycin-C in oesophageal strictures. *J Pediatr Gastroenterol Nutr*. 2007;44(3):336-341.
19. Smith ME, Elstad M. Mitomycin C and the endoscopic treatment of laryngotracheal stenosis: are two applications better than one? *Laryngoscope*. 2009;119(2):272-283.
20. Nguyen NP, Frank C, Moltz CC, et al. Impact of dysphagia on quality of life after treatment of head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2005;61(3):772-778.
21. Ellenhorn JD, Lambroza A, Lindsley KL, LaQuaglia MP. Treatment-related esophageal stricture in pediatric patients with cancer. *Cancer*. 1993;71(12):4084-4090.
22. Moghissi K, Pender D. Management of proximal oesophageal stricture. *Eur J Cardiothorac Surg*. 1989;3(2):93-97.
23. Dhir V, Vege SS, Mohandas KM, Desai DC. Dilation of proximal esophageal strictures following therapy for head and neck cancer: experience with Savary Gilliard dilators. *J Surg Oncol*. 1996;63(3):187-190.
24. Ahlawat SK, Davidson BJ, Al-Kawas FH. Successful use of biliary accessories in antegrade dilation of complex upper esophageal stricture due to chemoradiation and surgery. *Dis Esophagus*. 2008;21(1):86-89.
25. Sullivan CA, Jaklitsch MT, Haddad R, et al. Endoscopic management of hypopharyngeal stenosis after organ sparing therapy for head and neck cancer. *Laryngoscope*. 2004;114(11):1924-1931.
26. Goguen LA, Norris CM, Jaklitsch MT, et al. Combined antegrade and retrograde esophageal dilation for head and neck cancer-related complete esophageal stenosis. *Laryngoscope*. 2010;120(2):261-266.

Reprinted by permission of Laryngoscope. 2014; 124(10):2345-2351.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Correlation Between Reflux and Multichannel Intraluminal Impedance pH Monitoring in Untreated Volunteers

Marie E. Jetté, MS; Eric A. Gaumnitz, MD; Martin A. Birchall, MD; Nathan V. Welham, PhD; Susan L. Thibeault, PhD

Objectives/Hypothesis: Although probable causative agents have been identified (e.g., refluxate components, tobacco smoke), the definitive mechanism for inflammation-related laryngeal mucosal damage remains elusive. Multichannel intraluminal impedance combined with pH monitoring (MII/pH) has emerged as a sensitive tool for diagnosis and characterization of gastroesophageal reflux disease with laryngopharyngeal manifestations. To determine the relationship between laryngeal signs and MII/pH, we examined correlations between Reflux Finding Score (RFS) ratings of videostroboscopic laryngeal examinations and findings from MII/pH.

Study Design: Correlational study.

Methods: Healthy, untreated volunteers (n = 142) underwent reflux diagnosis using data acquired from MII/pH testing. Eight trained clinicians performed RFS ratings of corresponding laryngeal examinations. Averaged RFS ratings were compared to MII/pH data using Pearson correlation coefficients. The relationship between RFS and MII/pH findings and demographic/clinical information (age, sex, smoking status, reflux) was assessed using general linear modeling. Rater reliability was evaluated.

Results: Posterior commissure hypertrophy was negatively correlated with minutes of nonacid refluxate ($R = -0.21$, $P = .0115$). General linear modeling revealed that 28% to 40% of the variance in ratings of ventricular obliteration, erythema/hyperemia, vocal fold edema, diffuse laryngeal edema, posterior commissure hypertrophy, and granulation/granuloma could be explained by main and interaction effects of age, sex, smoking status, and reflux. Intra- and inter-rater reliability for RFS were poor–fair.

Conclusions: These results support the theory that the RFS is not specific for reflux in healthy, untreated volunteers, suggesting there may be alternate explanations for inflammatory clinical signs commonly ascribed to reflux in this population.

Key Words: Impedance monitoring, pH monitoring, gastroesophageal reflux, laryngopharyngeal reflux, laryngopharyngeal reflux diagnosis, Reflux Finding Score.

Level of Evidence: 1b

***Laryngoscope*, 124:2345–2351, 2014**

INTRODUCTION

Chronic laryngitis, one of the most commonly diagnosed dysphonias among healthcare professionals,[1] is characterized by a variety of inflammatory changes observed in patients with an array of symptoms. Gastroesophageal reflux disease (GERD) has been implicated as a probable etiologic factor for chronic laryngitis,[2–4] though treatment with proton pump inhibitors (PPIs), the current standard of care for GERD, demonstrates a nonsignificant benefit over placebo.[5] In spite of lack of efficacy data supporting the use of PPIs, 46.2% of patients with a diagnosis of chronic laryngitis receive medication.[6] Although reflux with laryngeal manifestations (laryngopharyngeal reflux [LPR]) may be an activator of laryngeal inflammation, the extent to which the effects of LPR alone contribute to the clinical picture of chronic laryngitis is unknown.

The Reflux Finding Score (RFS) was developed by Belafsky et al.[7] to document physical LPR findings on a standardized scale, with scores ranging from 0 (no evidence of reflux) to 26 (severe evidence of reflux). To validate this scale, RFS scores from 40 patients with clinically diagnosed LPR documented by esophageal-pharyngeal pH monitoring were compared to scores from 40 age-matched, asymptomatic controls who had not undergone confirmatory pH monitoring, and a statistically significant difference in scores was found.[7] Based on these results, the authors concluded with 95% certainty that a person with RFS >7 has LPR. Other researchers have determined that findings and symptoms ascribed to LPR are not specific to LPR.[8] Milstein et al.[9] found at

From the Department of Surgery, Department of Communication Sciences and Disorders (M.E.J.), Department of Medicine (E.A.G.), Department of Surgery (N.V.W., S.L.T.), University of Wisconsin–Madison, Madison, Wisconsin, U.S.A; and University College London Ear Institute (M.A.B.), The Royal National Throat, Nose and Ear Hospital, London, United Kingdom.

Editor's Note: This Manuscript was accepted for publication April 22, 2014.

Presented at the 2014 Combined Otolaryngological Spring Meetings, American Laryngological Society, Las Vegas, Nevada, U.S.A., May 15, 2014.

This work was supported by grants T32 DC-9401 and R01 DC 9600 from the NIH/NIDCD.

The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Susan L. Thibeault, PhD, 5107 Wisconsin Institutes for Medical Research, 1111 Highland Avenue, Madison, WI 53705. E-mail: thibeaul@surgery.wisc.edu

DOI: 10.1002/lary.24737

least one sign of laryngeal tissue irritation in the majority of volunteers undergoing laryngoscopy with no history of ear-nose-throat complaints or diagnosis of reflux. Similarly, Hicks et al.[10] demonstrated that 86% of normal, healthy, adult volunteers had findings commonly associated with reflux. Moreover, studies examining reliability of subjective laryngoscopic ratings of LPR have revealed mixed results ranging from poor to good.[11]

Ambulatory pH monitoring has been lauded as the gold standard for diagnosis of acid reflux; however, its role in diagnosing LPR remains controversial. In a review of multiple studies, Vaezi et al.[12] revealed that only 54% of patients with laryngoscopic signs of reflux have abnormal esophageal acid exposure on pH probe. They suggest that such low accuracy demonstrates either overdiagnosis of reflux as the cause of laryngeal pathology or lack of sensitivity of pH monitoring in documenting LPR.[12] Other diagnostic tools developed more recently include multichannel intraluminal impedance (MII), pharyngeal pH monitoring,[13] and hypopharyngeal MII (HMII).[14] Impedance monitoring (including MII and HMII) measures both acid and nonacid reflux in liquid and gaseous forms by measuring electrical resistance between different points along the esophagus. Combined with pH monitoring, impedance may offer improved detection of reflux events associated with LPR, though its role in LPR diagnosis has not been established.

The primary goal of this study was to examine correlations between endoscopic findings using RFS and measures acquired from MII with pH (MII/pH) monitoring in healthy, untreated volunteers. Given that the pathophysiology of laryngeal inflammation has not yet been defined and concerns have been published in the literature regarding the specificity of the RFS, we hypothesized that there would be poor correspondence between these sets of variables.

MATERIALS AND METHODS

Participant Selection

Participants aged 21 to 65 years were recruited with newspaper and email advertisements and signs in the clinic and around the University of Wisconsin–Madison. Participants underwent videolaryngostroboscopic examination and 24-hour MII/pH, with each procedure performed on separate dates. The protocol was approved by the institutional review board of University of Wisconsin–Madison, and informed consent was obtained from all participants.

Participants were excluded from the study if they had a history of radiation therapy to the head and neck within the past 5 years, lung or gastroesophageal surgery, chronic sinusitis or rhinitis in the last year, an acute traumatic event near the larynx in the last year, tracheostomy or other significant laryngeal or tracheal surgery, and substance or alcohol abuse in the past year. Consumption of more than 10 (women) and 17 (men) units of alcohol per week (means of United Kingdom and United States recommended weekly limits) excluded participants.[15] Further exclusion criteria included malignancy (except superficial basal cell carcinoma) within the past 5 years; presence of an infectious cause of laryngitis in the past 3 months; need for continuous therapy with diazepam, phenytoin, mephenytoin, warfarin, anticholinergics, antineoplastics, prostaglandin analogs, H2-receptor antagonists, steroids (inhaled, oral, or intravenous), promotility drugs, and sucralfate; use of any PPI or H2 blockers in the past year; theophylline or any other investigational compound or participation in an investigational drug study in the previous 60 days. Women were excluded if pregnant or lactating. Nonsmokers had not smoked during the previous year. Smokers were defined by consumption of a minimum of five cigarettes/5 g of tobacco per day for the duration of 1 or more years, thereby distinguishing them from light smokers.[16,17]

Laryngoscopy

Participants underwent videolaryngostroboscopic examination using rigid or flexible endoscope (Pentax Medical, Lincoln Park, NJ). Topical anesthetic was avoided unless the participant exhibited extreme gag reflex and was unable to tolerate examination. The larynx was visualized during sustained phonation on /i/ and quiet breathing. Digital recordings of laryngoscopic examinations were edited, randomized by clip number (List Randomizer, random.org), and organized into two video montages (iMovie; Apple, Cupertino, CA) representing two randomizations. Sixteen video clips were chosen randomly (List Randomizer, random.org) and included at the end of each video montage to assess intrarater reliability.

Reflux Finding Score

Eight raters provided ratings for this analysis using an adapted RFS (Table I). Raters included clinicians with 55 combined years of experience in voice disorders. A 45-minute training presentation was developed demonstrating published photographic examples of each RFS item[7,18,19] as well as their descriptions. Following training, raters were presented with still images from five examinations and performed group consensus ratings. Notes from the presentation and consensus ratings were saved and raters were able to access these while completing the RFS. Six of eight raters completed the training session with consensus. Two raters that did not attend reviewed the presentation and consensus notes before completing ratings. No demographic or MII/pH data were provided to raters. Raters were also blinded to the purpose of investigation and participant classification.

Combined Multichannel Intraluminal Impedance and 24-Hour pH Probe

After a four-hour fast, participants underwent conventional esophageal manometry (circumferential probe; Medtronic, Shoreview, MN) to locate lower and upper esophageal sphincters (LES and UES, respectively). The MII/pH catheter had two antimony electrodes placed such that proximal sensor was positioned 1 cm below and distal sensor 15 cm below the UES. Impedance was measured through seven sensors placed along a 2.3-mm polyurethane catheter. This catheter was placed transnasally immediately following manometry. Configuration of the catheter allowed recording of changes in intraluminal impedance at 3, 5, 7, 9, 15, and 17 cm above the LES. Data from impedance channels and pH electrodes were transmitted at 50 Hz and stored together on a portable data recorder (Sleuth; Sandhill Scientific Inc., Highlands Ranch, CO) for later synchronization. Participants were monitored for 18–24 hours and encouraged to eat regular meals and participate in routine activities. Change in position (upright and supine) and symptomatic events including heartburn or regurgitation were documented by using buttons on the data recorder. Data were uploaded and analyzed using commercially available software (Bioview Analysis; Sandhill Scientific Inc.).

TABLE I.
Reflux Finding Score Rating Rubric Adapted From Belafsky, Postma, and Koufman.[7]

Reflux Finding Score	
Subglottic edema (pseudosulcus; aka "infraglottic edema")	2 = present, 0 = absent
Ventricular obliteration (false vocal fold edge is indistinct; "complete" refers to the true and false folds appearing to touch)	2 = partial, 4 = complete
Erythema/hyperemia (redness)	2 = arytenoids only, 4 = diffuse
Vocal fold edema (mild is slight swelling, moderate is more perceptible, severe is sessile)	1 = mild, 2 = moderate, 3 = severe, 4 = polypoid
Diffuse laryngeal edema (size of airway relative to size of larynx)	1 = mild, 2 = moderate, 3 = severe, 4 = obstructing
Posterior commissure hypertrophy (pachydermia; mild is mustache-like appearance, moderate is straight line across back of larynx, severe is bulging into airway, and obstructing is airway obliterated)	1 = mild, 2 = moderate, 3 = severe, 4 = obstructing
Granuloma/granulation	2 = present, 0 = absent
Thick endolaryngeal mucus	2 = present, 0 = absent
Total =	

Analysis of pH data. Acid reflux episodes were defined as drops in pH to <4 for at least 5 seconds. Total acid exposure time (%) was calculated as total time of acid reflux episodes divided by monitoring time. Johnson/DeMeester score[20] was obtained using six parameters: 1) total percentage time pH <4.0, 2) percentage time pH <4.0 in an upright position, 3) percentage time pH <4.0 in a recumbent position, 4) total number acid reflux episodes, 5) total number acid reflux episodes longer than 5 minutes, and 6) duration of longest acid reflux episode.

Analysis of MII data. Recorded meal periods were excluded from analysis. On impedance, gas reflux was defined as rapid (>3,000 Ω/s) retrograde moving increase in impedance in at least two impedance sites. Liquid reflux was defined as retrograde moving 40% fall in impedance in two distal impedance sites. Proximal reflux was considered when refluxate reached the 15-cm impedance sensor. Total bolus exposure time (%) was defined as the combination of durations of gas and liquid reflux events divided by total time monitored.

Interpretation of combined dual-channel MII/pH data. Participants were assigned to cohorts—GERD, LPR, normal—based on MII/pH data. GERD was defined by acid exposure percent time of the distal pH probe >4.0, DeMeester score >14.7, and/or bolus exposure percent time of more than 1.4%.[21] LPR was defined by >31 proximal reflux events.[22,23] Normal was defined by the following criteria: acid exposure percent time of the distal pH probe <4.0, DeMeester score <14.7, and <31 proximal reflux events.[22]

Statistical Analysis

To determine inter-rater reliability, intraclass correlation coefficients (ICC) were calculated. Pearson correlation coefficients were used to evaluate intrarater reliability. Average within rater agreement across all eight raters was computed for each RFS item. RFS ratings for each videostroboscopic examination were averaged across all ratings from eight individual raters. Pearson correlation coefficients were used to determine correlations between average RFS ratings and findings on MII/pH and correlations between age and average RFS ratings. General linear models, including repeated measures analysis of variance and analysis of covariance, were fitted to assess main effects of age, cohort, sex, and smoking status, as well as the two-, three-, and four-way interaction effects of age*sex, age*cohort, age*smoking status, cohort*sex, cohort*smoking status, sex*smoking status, age*sex*smoking status, age*cohort*smoking status, age*sex*cohort, cohort*sex*smoking status, and age*cohort*sex*smoker for all RFS ratings. *t*-tests were used to examine differences in variables that could not be accounted for by linear modeling. All analyses were performed with SAS software (SAS Institute Inc., Cary, NC) with type I error set at 0.05.

RESULTS

Clinical and Demographic Characteristics

Of 155 original video clips included in the montages provided to raters, 13 were excluded from rating and analysis due to insufficient views from anterior commissure to posterior pharyngeal wall. Data from 142 participants including videolaryngostroboscopic recordings, MII/pH variables (Table II), and averaged RFS ratings (Table II) were therefore included in the final analysis. Analysis of MII/pH data revealed 38 participants with GERD (27%), 44 with LPR (31%), and 60 normal (42%). Of 142 participants, 116 (82%) had total RFS >7, and 55 (39%) had total RFS >11. Age, sex, smoking, reflux cohort, and total RFS characteristics of these participants are summarized in Table III. Videostroboscopic examination and MII/pH testing were completed with an average of 61 days between each procedure.

RFS Rater Reliability and Agreement

ICC for intrarater reliability ranged from 0.05 to 0.45 (Table IV). Results demonstrate poor to fair reliability for all RFS rating items. Inter-rater reliability was assessed on 256 observations from eight raters. ICC ranged from 0.21 to 0.48 (Table IV), indicating poor to fair inter-rater reliability for all RFS rating items. Average intrarater agreement examines overall levels of rater self-consistency, for each rater and RFS rating. Results are based on repeated ratings of 16 video clips, and indicate that individual raters were 54.8% to 71.7% reliable across all ratings, and that they produced the same value for any individual variable 48.75% to 78.75% of the time (Table V).

TABLE II.
Summary of Reflux Finding Score and Multichannel Intraluminal Impedance pH Monitoring Variables.

	Mean	SD	Minimum	Maximum
RFS variables				
Subglottic edema	0.78	0.74	0	2
Ventricular obliteration	1.77	0.83	0	3.75
Erythema/hyperemia	2.91	0.87	0.5	4
Vocal fold edema	1.25	0.75	0	3.75
Diffuse laryngeal edema	1.05	0.57	0	3
Posterior commissure hypertrophy	1.63	0.66	0	3
Granulation/granuloma	0.38	0.45	0	2
Thick endolaryngeal mucus	1.18	0.72	0	2.25
Total	10.38	3.63	1.75	19.875
MII/pH variables				
Measured by pH monitoring				
% total time pH <4	3.34	7.95	0	80.5
% upright time pH <4	3.69	7.70	0	62.8
% supine time pH <4	2.56	9.41	0	94.9
No. of reflux episodes	19.50	14.20	0	76
No. of reflux episodes ≥5 minutes	1.09	2.64	0	18
Longest reflux episode (min)	13.84	43.69	0	444.8
Johnson/DeMeester score	14.78	31.85	0.8	256.0
Measured by multichannel intraluminal impedance				
Acid refluxate (min)	13.67	15.63	0	102.9
Nonacid refluxate (min)	7.08	9.78	0	102.3
Total % time reflux (min)	1.79	1.66	0.1	12
No. of reflux events	44.3	21.01	7	105
No. of acid reflux events	24.42	17.28	0	91
No. of nonacid reflux events	19.78	11.56	0	62
No. of reflux events that reached the proximal esophagus	23.75	13.97	3	72
No. of acid reflux events that reached the proximal esophagus	14.69	11.32	0	52
No. of nonacid reflux events that reached the proximal esophagus	9.06	6.65	0	38

MII = multichannel intraluminal impedance; RFS = Reflux Finding Score; SD = standard deviation.

Correlations Between RFS and MII/pH

Average RFS ratings for each videostroboscopic examination were compared to individual MII/pH variables resulting in 144 analyzed correlations across 142 participants. There was a single significant correlation between posterior commissure hypertrophy and minutes of nonacid refluxate ($R = -0.21$, $P = .0115$). No other correlations were significant (data not shown).

Effect of Clinical and Demographic Characteristics on RFS

Average RFS ratings for each variable were analyzed relative to clinical and demographic data including cohort, sex, and smoking status. Age was analyzed as a main effect and also included in a separate interaction effects model (Table VI). Interaction effects of cohort, sex, smoking status, and age influenced averaged RFS ratings. General linear modeling, including all variables and their interactions (Table VI, model 2), explained 25% to 40% of the variance observed in many RFS ratings. Although both models tested could not account for variance in ratings of subglottic edema and thick endolaryngeal mucus, further analysis revealed the main effects of sex on both of these variables ($P = .025$, $P = .049$, respectively).

DISCUSSION

The major finding of this study was a single statistically significant correlation between RFS and MII/pH variables in a group of healthy, non–treatment-seeking, untreated volunteers. We found a negative correlation between posterior commissure hypertrophy and duration (minutes) of nonacid reflux ($R = -0.21$, $P = .0115$), suggesting that posterior commissure hypertrophy is decreased with greater duration of nonacid reflux. This result is supported by biological evidence demonstrating less proinflammatory cytokine gene expression with greater acid exposure in biopsies taken from the posterior commissure.[24] Though this correlation coefficient is statistically significant, it is meaningless unless properly interpreted for clinical relevance. Calculating

TABLE III.
Participant Characteristics.

Characteristic	No. (%)	Mean Age, yr
Sex		
Male	64 (45)	40.1
Female	78 (55)	43.5
Cigarette smoking		
Nonsmoker	107 (75)	40.6
Smoker	35 (25)	42.9
Reflux cohort		
GERD	38 (27)	43.4
LPR	44 (31)	37.5
Normal	60 (42)	42.5
Total reflux finding score		
<7	26 (18)	42.0
≥7	116 (82)	41.3
<11	87 (61)	41.4
≥11	55 (39)	41.2

GERD = gastroesophageal reflux disease; LPR = laryngopharyngeal reflux.

coefficient of determination (R^2) yields 0.044, meaning that 4.4% of variation in ratings of posterior commissure hypertrophy can be explained or accounted for by variation in duration of nonacid reflux. This interpretation of the data suggests there are other factors (e.g., demographic characteristics) aside from reflux findings measured by MII/pH that may explain variability in RFS ratings. It is also possible that there is an inherent lack of RFS validity for specific reflux diagnosis.

The primary outcome measures of our study were eight RFS ratings in addition to total RFS averaged across eight trained clinician raters and 16 MII/pH variables. Though averaged RFS ratings were used for analysis, it is worth noting that inter- and intrarater reliability for RFS was poor–fair. In a review of the literature examining reliability for laryngopharyngeal findings in LPR, Powell and Cocks[11] presented a summary from nine publications demonstrating variable reliability ranging from poor–good. They suggested variability might be related to methods of assessment or statistical tests used. Potential explanations for poor intrarater reliability observed in our study relate to the inherent limits of human raters' visual-perceptual systems and the RFS scale itself. Rosen[25] has suggested several limitations and possible errors associated with visual-perceptual ratings of videostroboscopy, including rater fatigue and lack of variability of videos. Additionally, whereas some variables (e.g., subglottic edema) can be scored as 0 (absent) or 2 (present), other variables (e.g., vocal fold edema) are scored on a 5-point scale (0, 1, 2, 3, 4). When data are pooled for statistical calculation of intrarater reliability, the difference between ratings of 0 and 2 is given greater weight than the difference between ratings on a five-point scale. Examining agreement in conjunction with reliability gives an indication of statistical penalties resulting from limits of the scale. For example, upon repeat rating of thick endolaryngeal mucus, clinicians on average agreed with their initial rating 72.4% of the time, whereas intrarater reliability was calculated at $R = 0.12$ ($P = .0001$) indicating poor reliability. Agreement implies that two raters assign identical meanings to each score for each variable, whereas reliability indicates that raters rate variables in parallel fashion, without implying that score values have the same meaning. If the range of scores is restricted (e.g., raters consistently avoid extremes of a scale or scores vary little with respect to variable rated), reliability coefficients may be low, even if raters agree. In this study, it is possible raters avoided severe extremes of the RFS given they were rating images from non–treatment-seeking volunteers as opposed to a pathologic population.

To bolster the clinical relevance of our findings, we used combined MII/pH variables semidiagnostically to categorize our study population into cohorts including LPR, GERD, and normal based on normative data.[21–23] Our study is the first to report on the incidence of

TABLE IV.
Intrarater and Inter-rater Reliability.

	Intrarater Reliability		Inter-rater Reliability	
RFS Variable	*R*	*P* Value	*R*	*P* Value
Subglottic edema	0.05	.06	0.48	<.0001
Ventricular obliteration	0.45	<.0001	0.24	<.0001
Erythema/hyperemia	0.10	.001	0.34	<.0001
Vocal fold edema	0.29	<.0001	0.39	<.0001
Diffuse laryngeal edema	0.17	<.0001	0.29	<.0001
Posterior commissure hypertrophy	0.021	038	0.34	<.0001
Granulation/granuloma	0.20	<.0001	0.21	<.0001
Thick endolaryngeal mucus	0.12	.0001	0.43	<.0001
Total	0.21	.0001	0.48	<.0001

Interpretation of correlation coefficients: <0.40 = poor; 0.40–0.59 = fair; 0.60–0.74 = good; >0.74 = excellent.
RFS = Reflux Finding Score.

TABLE V.
Percent Intrarater Agreement for Each Individual Rater and RFS Rating as Well as Averages Across Raters and RFS Ratings.

	Intrarater % Agreement								Average % Agreement
RFS Variable	R1	R2	R3	R4	R5	R6*	R7	R8*	All Raters
Subglottic edema	70.0	81.8	81.3	72.7	50.0	60.0	90.9	53.8	70.1
Ventricular obliteration	81.8	53.9	100.0	71.4	75.0	88.9	72.7	56.3	75.0
Erythema/hyperemia	73.3	57.2	62.5	66.7	61.5	87.5	61.5	66.7	67.1
Vocal fold edema	76.9	71.4	68.8	58.3	63.6	25.0	30.8	46.7	55.2
Diffuse laryngeal edema	75.0	76.9	18.8	50.0	54.5	50.0	27.3	37.5	48.7
Posterior commissure hypertrophy	42.9	58.3	56.3	58.3	46.2	62.5	66.7	57.2	56.0
Granulation/granuloma	75.0	100.0	81.3	90.0	75.0	77.8	66.7	64.3	78.7
Thick endolaryngeal mucus	78.6	57.1	68.8	68.7	92.9	66.7	92.3	56.3	72.6
Average % agreement	71.7	69.6	67.2	67.0	64.8	64.8	63.6	54.8	

*Indicates rater did not attend training.
RFS = Reflux Finding Score.

GERD and LPR based on MII/pH in untreated, non–treatment-seeking healthy volunteers. Within our participant group, more than half (58%) was categorized as either LPR or GERD, whereas 42% demonstrated normal findings on MII/pH. Similarly, categorization of participants using published thresholds for total RFS of 7[7] and 11[26] yielded 82% and 39%, respectively, categorized as LPR, supporting Hicks et al.'s finding that 86% of normal, healthy, adult volunteers had signs associated with reflux.[10] In a study investigating the diagnostic usefulness of MII/pH in 98 patients with suspected LPR off PPI therapy for at least 2 weeks, Lee et al. found that 54% demonstrated pathologic GERD,[27] a finding consistent with our data in spite of the difference in study populations. It should be noted that in our study, LPR was determined based on impedance and pH findings in the proximal esophagus, not in the hypopharynx, which may have resulted in overestimation of incidence of LPR. Supporting this possibility, an investigation of 34 asymptomatic, untreated research participants using hypopharyngeal MII/pH revealed a single LPR event recorded from one participant (3%), whereas in symptomatic, untreated patients, 24/184 (13%) had at least one LPR event documented.[14] In clinical practice, gastroenterologists use MII/pH to diagnose reflux in patients with persistent symptoms despite acid-suppressive therapy. Diagnosis includes examining symptom association[28] (i.e., determining whether episodes recorded by MII/pH are associated with a corresponding symptom) and comparing MII/pH variables in patients on therapy to normative values.[29] As we were attempting to use MII/pH as the sole objective measure of reflux in a non–treatment-seeking population, symptom association and treatment response were not evaluated within the present research design.

The clinical/demographic interaction and main effects observed within our dataset provide insight into factors that explain some variance in RFS ratings. General linear modeling including main and interaction effects of age, reflux cohort, sex, and smoking status could explain 25% to 40% of the variance observed in all RFS variables except subglottic edema and thick endolaryngeal mucus, suggesting that RFS ratings are influenced by clinical and demographic factors. Inflammatory

TABLE VI.
Summary of Results of Generalized Linear Modeling.*

	Age		Model 1		Model 2 (With Age)	
RFS Variable	R^2	*P*	R^2	*P*	R^2	*P*
Subglottic edema	0.001	0.69	0.11	.19	0.16	.51
Ventricular obliteration	0.08	0.0006	0.29	<.0001	0.40	<.0001
Erythema/hyperemia	0.01	0.19	0.35	<.0001	0.39	<.0001
Vocal fold edema	0.04	0.03	0.33	<.0001	0.39	<.0001
Diffuse laryngeal edema	0.03	0.04	0.28	<.0001	0.37	<.0001
Posterior commissure hypertrophy	0.03	0.04	0.17	.01	0.25	.03
Granulation/granuloma	0.03	0.03	0.05	.80	0.28	.01
Thick endolaryngeal mucus	0.03	0.70	0.10	.20	0.16	.47
Total	0.02	0.09	0.35	<.0001	0.39	<.0001

*Generalized linear modeling demonstrating the total variance (R^2) accounted for by: 1) the main effect of age; 2) the main and interaction effects of cohort, sex, smoking status; and 3) the main and interaction effects of cohort, sex, smoking status, and age for each RFS variable.
RFS = Reflux Finding Score.

signs measured with RFS are in part related to the combinations of sex, smoking status, and age of the larynx being rated as opposed to reflux alone. Subglottic edema, also referred to as pseudosulcus and infraglottic edema, has long been thought to be predictive of,[30] and specific for,[19] LPR; however, our results demonstrate that males receive greater ratings than females on this variable regardless of reflux cohort, smoking status, and age. It seems possible that this finding so commonly ascribed to inflammation from reflux may be a result of anatomic differences between males and females. Males also received greater ratings than females for thick endolaryngeal mucus, suggesting that this finding provides more information about the sex of the person being examined than it does about reflux.

Although attempts were made to eliminate bias, we recognize limitations in our study design that may have prejudiced our results. Of primary consideration is that we examined data from non–treatment-seeking volunteers, a population not representative of a typical clinical population. It would be ideal to repeat the study in treatment-seeking patients for whom laryngeal inflammation impacts vocal function, thereby addressing the role of reflux specific to diagnosis of chronic laryngitis. We also recognize that we persisted in analyzing averaged RFS ratings in spite of poor reliability, though we attempted to avoid this issue by providing raters with training. Finally, we acknowledge that reflux status may have changed in the time between videostroboscopic examination and MII/pH testing. This could be avoided in future studies by completing videostroboscopic examination immediately prior to MII/pH.

CONCLUSION

Our data demonstrate an overall lack of correlation between RFS and MII/pH, supporting the hypothesis that RFS is not specific for reflux in non–treatment-seeking, untreated volunteers. Our findings also illustrate that in spite of training, raters demonstrated poor–fair inter- and intrarater reliability on RFS, consistent with results from other studies. Finally, we suggest that clinical and demographic characteristics, including sex, smoking status, and age, contribute to differences in RFS ratings.

Acknowledgments

The authors thank Dr. Glen Leverson for statistical support.

BIBLIOGRAPHY

1. Cohen SM, Kim J, Roy N, Asche C, Courey M. Prevalence and causes of dysphonia in a large treatment-seeking population. *Laryngoscope* 2012; 122:343–348.
2. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope* 1991;101:1–78.
3. Koufman J, Sataloff RT, Toohill R. Laryngopharyngeal reflux: consensus conference report. *J Voice* 1996;10:215–216.
4. Hanson DG, Jiang JJ. Diagnosis and management of chronic laryngitis associated with reflux. *Am J Med* 2000;108(suppl 4a):112S–119S.
5. Qadeer MA, Phillips CO, Lopez AR, et al. Proton pump inhibitor therapy for suspected GERD-related chronic laryngitis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol* 2006;101:2646–2654.
6. Cohen SM, Kim J, Roy N, Courey M. Assessing factors related to the pharmacologic management of laryngeal diseases and disorders. *Laryngoscope* 2013;123:1763–1769.
7. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope* 2001;111:1313–1317.
8. Barry DW, Vaezi MF. Laryngopharyngeal reflux: More questions than answers. *Cleve Clin J Med* 2010;77:327–334.
9. Milstein CF, Charbel S, Hicks DM, Abelson TI, Richter JE, Vaezi MF. Prevalence of laryngeal irritation signs associated with reflux in asymptomatic volunteers: impact of endoscopic technique (rigid vs. flexible laryngoscope). *Laryngoscope* 2005;115:2256–2261.
10. Hicks DM, Ours TM, Abelson TI, Vaezi MF, Richter JE. The prevalence of hypopharynx findings associated with gastroesophageal reflux in normal volunteers. *J Voice* 2002;16:564–579.
11. Powell J, Cocks HC. Mucosal changes in laryngopharyngeal reflux–prevalence, sensitivity, specificity and assessment. *Laryngoscope* 2013;123: 985–991.
12. Vaezi MF, Hicks DM, Abelson TI, Richter JE. Laryngeal signs and symptoms and gastroesophageal reflux disease (GERD): a critical assessment of cause and effect association. *Clin Gastroenterol Hepatol* 2003;1:333–344.
13. Becker V, Graf S, Schlag C, et al. First agreement analysis and day-to-day comparison of pharyngeal pH monitoring with pH/impedance monitoring in patients with suspected laryngopharyngeal reflux. *J Gastrointest Surg* 2012;16:1096–1101.
14. Hoppo T, Sanz AF, Nason KS, et al. How much pharyngeal exposure is "normal"? Normative data for laryngopharyngeal reflux events using hypopharyngeal multichannel intraluminal impedance (HMII). *J Gastrointest Surg* 2012;16:16–24; discussion 24–25.
15. Furtwaengler NA, de Visser RO. Lack of international consensus in low-risk drinking guidelines. *Drug Alcohol Rev* 2013;32:11–18.
16. Kenford SL, Wetter DW, Welsch SK, Smith SS, Fiore MC, Baker TB. Progression of college-age cigarette samplers: what influences outcome. *Addict Behav* 2005;30:285–294.
17. Husten CG. How should we define light or intermittent smoking? Does it matter? *Nicotine Tob Res* 2009;11:111–121.
18. Belafsky PC, Postma GN, Amin MR, Koufman JA. Symptoms and findings of laryngopharyngeal reflux. *Ear Nose Throat J* 2002;81:10–13.
19. Belafsky PC, Postma GN, Koufman JA. The association between laryngeal pseudosulcus and laryngopharyngeal reflux. *Otolaryngol Head Neck Surg* 2002;126:649–652.
20. Johnson LF, DeMeester TR. Development of the 24-hour intraesophageal pH monitoring composite scoring system. *J Clin Gastroenterol* 1986; 8(suppl 1):52–58.
21. Lee JH, Park SY, Cho SB, et al. Reflux episode reaching the proximal esophagus are associated with chronic cough. *Gut Liver* 2012;6:197–202.
22. Shay S, Tutuian R, Sifrim D, et al. Twenty-four hour ambulatory simultaneous impedance and pH monitoring: a multicenter report of normal values from 60 healthy volunteers. *Am J Gastroenterol* 2004;99:1037–1043.
23. Cho YK. How to interpret esophageal impedance pH monitoring. *J Neurogastroenterol Motil* 2010;16:327–330.
24. Thibeault SL, Smith ME, Peterson K, Ylitalo-Moller R. Gene expression changes of inflammatory mediators in posterior laryngitis due to laryngopharyngeal reflux and evolution with PPI treatment: a preliminary study. *Laryngoscope* 2007;117:2050–2056.
25. Rosen CA. Stroboscopy as a research instrument: development of a perceptual evaluation tool. *Laryngoscope* 2005;115:423–428.
26. Postma GN, Halum SL. Laryngeal and pharyngeal complications of gastroesophageal reflux disease. GI Motility Online. 2006. Available at: http://www.nature.com/gimo/contents/pt1/full/gimo46.html. Accessed on January 31, 2014.
27. Lee BE, Kim GH, Ryu DY, et al. Combined dual channel impedance/pH-metry in patients with suspected laryngopharyngeal reflux. *J Neurogastroenterol Motil* 2010;16:157–165.
28. Weusten BL, Roelofs JM, Akkermans LM, Van Berge-Henegouwen GP, Smout AJ. The symptom-association probability: an improved method for symptom analysis of 24-hour esophageal pH data. *Gastroenterology* 1994;107:1741–1745.
29. Tutuian R, Castell DO. Review article: complete gastro-oesophageal reflux monitoring—combined pH and impedance. *Aliment Pharmacol Ther* 2006;24(suppl 2):27–37.
30. Hickson C, Simpson CB, Falcon R. Laryngeal pseudosulcus as a predictor of laryngopharyngeal reflux. *Laryngoscope* 2001;111:1742–1745.

Reprinted by permission of Laryngoscope. 2013; 123(10):2463-2468.

The Laryngoscope
© 2013 The American Laryngological,
Rhinological and Otological Society, Inc.

The Role of Impedance Monitoring in Patients With Extraesophageal Symptoms

Robert T. Kavitt, MD, MPH; Elif Saritas Yuksel, MD; James C. Slaughter, DrPH; C. Gaelyn Garrett, MD; David Hagaman, MD; Tina Higginbotham, MPA; Michael F. Vaezi, MD, PhD, MSc (Epi)

Objectives/Hypothesis: Ambulatory esophageal impedance monitoring is commonly employed to assess for nonacid reflux in patients with extraesophageal reflux. We aimed to determine if *on* therapy impedance data can be predicted from *off* therapy upper endoscopy, manometry, or pH parameters.

Study Design: Prospective Cohort Study.

Methods: Patients with extraesophageal reflux symptoms and either partial- or nonresponders to twice-daily PPI underwent impedance monitoring *on* twice-daily PPI, as well as manometry, upper endoscopy, and 48-hour wireless pH monitoring *off* acid-suppressive medications for 1 week. Percent time pH < 4 and number of reflux episodes were obtained. Multivariable linear regression was used to determine association between the impedance data *on* therapy and upper endoscopy, manometry, and pH parameters measured *off* therapy.

Results: Seventy-five patients (77% female, median BMI 29, 38% with hiatal hernia, and 19% with esophagitis) were studied both *on* and *off* therapy. Thirty-five percent had abnormal impedance monitoring *on* therapy and 84% had abnormal pH testing *off* therapy. There was no significant ($P = 0.184$) overall correlation between total number of impedance events and the baseline physiologic parameters of hiatal hernia, degree of acid reflux, or manometric findings, with only weak correlation ($r = 0.54$, $P = 0.045$) with % time pH < 4 among patients with esophagitis.

Conclusions: In patients with suspected extraesophageal reflux refractory to PPI therapy, impedance measures *on* therapy cannot be predicted from traditional baseline esophageal physiologic parameters. We recommend caution regarding overinterpretation of impedance data. Laryngoscope, 000:000–000, 2013

Key Words: Impedance, GERD, refractory reflux.

Level of Evidence: 2b.

***Laryngoscope,* 123:2463–2468, 2013**

INTRODUCTION

Gastroesophageal reflux disease (GERD) is a common disorder with increasing prevalence in the Western world.[1] Approximately 40% of adults frequently complain of heartburn,[2] and GERD remains the leading outpatient physician diagnosis for gastrointestinal disorders in the United States.[3]

Extraesophageal reflux (EER) is widely implicated in the etiology of laryngeal, pharyngeal, and pulmonary symptoms, and controversy exists regarding the diagnosis and management of this condition.[4] Currently, most patients with signs and symptoms attributed to EER are empirically treated with proton pump inhibitors (PPIs). However, symptomatic improvement on PPIs is not as consistent compared to those with classic GERD.[5] Patients with presumed EER refractory to initial empiric medical therapy are often referred for further testing. Current guidelines recommend diagnostic testing, which can include the use of upper endoscopy and pH and/or impedance monitoring.[6]

Combined impedance-pH monitoring can detect various types of esophageal refluxate: gas, liquid, acid, or nonacid, and is used to clarify the mechanisms of PPI-refractory symptoms.[7–11] Multicenter studies utilizing impedance-pH testing in patients with PPI-refractory symptoms suggest that approximately one-third of patients exhibit weakly acid or nonacid reflux. In the background of potent acid suppression, the clinical significance of these findings currently remains controversial. While some advocate for the clinical utility of impedance-pH monitoring in assessing the impact of weakly acidic material on patients' persistent symptoms, others are not as enthusiastic. Studies have suggested that 40% to 50% of patients with persistent symptoms on acid-suppressive therapy have no temporal correlation between their symptoms and any type of reflux.[12,13]

While there is no doubt that impedance-pH testing is currently the most accurate and detailed method for detecting reflux of all kinds, the clinical indications for its use are still evolving. Its role in the management of GERD patients awaits further definition, mainly due to

From the Division of Gastroenterology, Hepatology and Nutrition (R.T.K., E.S.Y., T.H., M.F.V.), the Department of Biostatistics (J.C.S.), the Vanderbilt Voice Center; Department of Otolaryngology (C.G.G.), and the Allergy, Sinus and Asthma Program (D.H.), Vanderbilt University Medical Center, Nashville, Tennessee, U.S.A.

Editor's Note: This Manuscript was accepted for publication August 22, 2012.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Michael F. Vaezi, MD, PhD, MSc (Epi), Division of Gastroenterology, Hepatology and Nutrition, Vanderbilt University Medical Center, 1660 TVC, 1301 22nd Ave. South, Nashville, TN 37232-5280. E-mail: Michael.vaezi@vanderbilt.edu

DOI: 10.1002/lary.23734

the paucity of high-quality studies examining the clinical relevance of impedance findings. Thus, the aim of our study was to determine if *on* therapy impedance values assessing for nonacid reflux can be predicted from the traditionally recognized and commonly employed clinical markers of reflux based on *off* therapy upper endoscopy, manometry, or pH monitoring.

MATERIALS AND METHODS

The study was performed in accordance with the Declaration of Helsinki, Good Clinical Practice, and applicable regulatory requirements. The Vanderbilt Institutional Review Board approved this clinical trial (IRB# 090872).

STUDY DESIGN AND PATIENT POPULATION

The study population consisted of patients with suspected reflux-associated extraesophageal symptoms refractory to PPI therapy referred to the Esophageal Motility Center at Vanderbilt University Medical Center for evaluation and treatment. Refractory symptoms were defined as less than 50% improvement in the chief complaint after at least 12 weeks of twice-daily PPI therapy. This group was chosen since they represent the patient population for whom pH and/or impedance monitoring is currently indicated.[14,15] In order to assess the severity and frequency of their extraesophageal symptoms, patients were asked to complete a questionnaire previously described in the literature.[16] The following information were collected for all patients: presence, severity, and frequency of GERD symptoms (heartburn ± regurgitation) and extraesophageal symptoms (cough, hoarseness, throat clearing, sore throat, globus sensation, postnasal drip symptoms, chest pain), current medications, demographic data (age, sex, race, body mass index), history of alcohol and tobacco use, and presence of voice/laryngeal and nasal symptoms.

Patients underwent esophagogastroduodenoscopy (EGD), wireless 48-hour pH monitoring, and esophageal manometry *off* acid suppression to assess the baseline esophageal mucosal integrity, motility, and acid exposure. They also underwent 24-hour impedance/pH monitoring while *on* twice daily PPI therapy to determine the presence of acid and nonacid reflux in the setting of acid suppression. The presence and size of a hiatal hernia were determined at endoscopy; the presence and severity of esophagitis was graded by the Los Angeles Classification; and the presence of Barrett's esophagus was also noted.[17] Inclusion criteria were age greater than 18 years and chronic EER symptoms refractory to PPI therapy. Patients were excluded from the study if they were unwilling to undergo testing, were pregnant, had undergone surgery for reflux or peptic ulcer disease, or if they had a serious illness that would interfere with study participation.

Wireless pH Monitoring

Ambulatory pH monitoring was performed for 48 hours using a wireless monitoring device (Given Imaging, Duluth, GA). Patients were instructed to stop taking all PPIs and H2-receptor antagonists (H2RAs) for at least 7 days prior to undergoing testing. Wireless capsules were calibrated by submersion in buffer solutions at pH 7.0 and pH 1.0, and then activated by magnet removal. Patients underwent EGD with conscious sedation for visual anatomic inspection and distance measurements from the incisors to the squamocolumnar junction (SCJ). Capsules were then placed using the manufacturer's delivery system at 6 cm above the SCJ and attached with vacuum suction. Capsule placement was confirmed with endoscopy. After successful placement, patients were given wireless pH recorders to wear about their waists, or to keep within 3 feet to 5 feet at all times. Recording devices receive pH data sampling transmitted by the capsule at 433 Hz with 6 second sampling intervals. Patients were instructed to perform their normal daily activities and dietary practices. Distal esophageal pH recording was conducted for a total of 48 hours. During this time patients kept diaries of meal times, symptoms, and supine positioning.

After completion of the 48-hour study, data were downloaded from recording devices to dedicated computers. Patient diary information was manually entered into the computer-based record. Measurements of the total, upright, and supine percentage time when esophageal pH was below 4 were determined over day 1 and day 2 of the wireless study. Total acid exposure time (% total time pH less than 4) greater than 5.5% was considered abnormal, while greater than 8.2% was considered abnormal for the upright state and greater than 3.0% was considered abnormal for the supine state.[18]

Esophageal Motility Testing

High-resolution manometry (Given Imaging, Duluth, GA) was used to measure the location of the lower esophageal sphincter prior to placement of the impedance-pH catheter. A solid-state assembly with 36 circumferential sensors spaced at 1-cm intervals (outer diameter 4.2 mm) was used. This device detects pressure over a length of 2.5 mm in each of the 12 radially dispersed sectors of the 36 pressure-sensing elements. The sector pressures are averaged, making the sensors a circumferential pressure detector. Prior to recording the transducers were calibrated at 0 and 100 mm Hg using externally applied pressure. Using this device the lower esophageal sphincter was measured and the proximal location noted for placement of the impedance-pH catheter.

Combined Impedance-pH Monitoring

Patients underwent impedance-pH monitoring while *on* at least twice daily PPI therapy for 1 month prior to evaluation. They were instructed to fast for 4 hours before testing. Each patient's primary symptom complaint was recorded as part of the preprocedure evaluation. Patients were given diaries to record the timing of initiation and completion of meals and position changes (upright or supine) during 24-hour impedance and pH monitoring.

Impedance testing was performed using a combined impedance-pH monitoring device (Sandhill Scientific Inc; Highlands Ranch, CO) comprising a data recorder (Sleuth System; Sandhill Scientific Inc.) and a 2.1-mm diameter

polyvinyl catheter embedded by one pH and six impedance sensors at predefined positions. The pH sensors were calibrated before placement using standardized buffer solutions at pH 4.0 and 7.0, as recommended by the manufacturer. The catheter was placed intranasally so that the esophageal pH sensor was positioned 5 cm above the manometrically defined upper border of the lower esophageal sphincter. Intraluminal impedance was measured at 3, 5, 7, 9, 15, and 17 cm above the lower esophageal sphincter. Data sampling frequency for both impedance and pH sensors was 50 Hz. Studies were performed for 24 hours, after which patients returned to the lab for catheter removal and data review.

Data were downloaded from the recorder and analyzed using BioView Analysis software (Sandhill Scientific Inc). Reflux episodes were identified by computerized detection (Autoscan; Sandhill Scientific Inc.) of proximally directed decreases in impedance. Tracings were then manually reviewed by an experienced investigator (MFV) to confirm accuracy and correct any errors. Total, upright, and supine reflux events were recorded. Acid reflux events were defined as those occurring with pH less than 4.0, and nonacid reflux events were defined as those occurring at pH of 4.0 or greater. For impedance parameters, total number of reflux events greater than 48 was considered abnormal.[19–22]

Statistical Analysis

Data were collected and stored at the secure Web-based Vanderbilt Digestive Disease Center REDCap (Research Electronic Data Capture) (1 UL1 RR024975 NCRR/NIH). REDCap is an application designed to support data capture for research studies providing: 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources. There was strict control and supervision of the data entry and access for this study.

Continuous variables were summarized using the median, 25th, and 75th percentiles. Categorical variables were summarized using percentages. We used separate Kruskal-Wallis tests to determine if the total number of impedance events and the percent total time pH less than four differed by esophagitis and pH groups. Spearman's rank correlation was used to estimate the overall association between impedance events (on therapy) and percent total time pH less than 4 (off therapy). Multivariable linear regression was used to assess the relationship between *on*-therapy impedance parameters and *off*-therapy pH parameters, manometry findings, and upper endoscopy findings. All analyses were conducted using the R statistical program.

RESULTS

Demographics and Endoscopic Findings

Seventy-five patients with suspected extraesophageal reflux underwent testing with 48-hour wireless pH monitoring *off* PPI therapy and 24-hour impedance/pH monitoring *on* PPI therapy between 2005 and 2012. Their chief complaints included: cough (53%), asthma (12%), hoarseness (7%), throat clearing (6%), pulmonary fibrosis (6%), sore throat (6%), postnasal drip (5%), and sinusitis (5%). A total of 58/75 (77%) of patients were female, 71/75 (95%) were Caucasian, with a median (IQR) age of 55 (45–64) years, and median (IQR) body mass index of 29 (25–33). Hiatal hernias greater than 1 cm in size were present in 29/75 (38%), of which 61% were 2 to 3 cm in size, and 39% were 4 cm or greater. Esophagitis was present in 14/75 (19%) of patients, of which 90% were grade A or B by Los Angeles Classification. None of the patients were found to have endoscopic evidence of possible Barrett's esophagus.

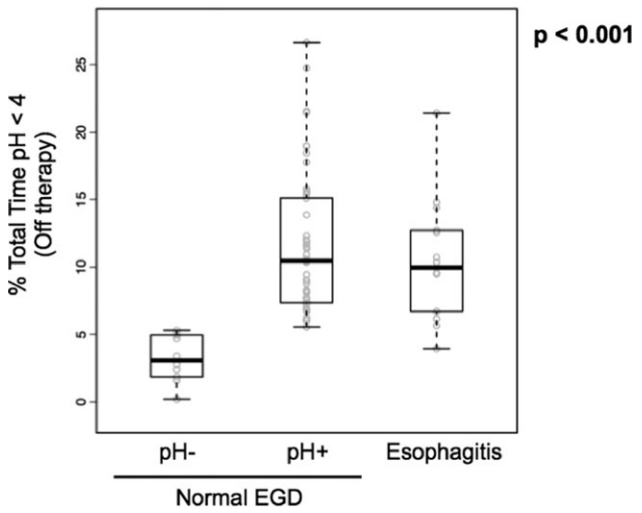

Fig. 1. Individual and group median (IQR) % total time pH < 4 for the three patient subgroups. As expected, acid reflux was significantly less in those with normal endoscopy and pH than those with abnormal pH or esophagitis.

Impedance and pH Parameters

Overall. Median (IQR) total, upright, and supine impedance events *on PPI therapy* were 39 (20–54), 29 (6–12), and 4 (1–8), respectively. The impedance events were predominantly mixed gas and liquid of nonacidic nature (pH > 4) and were abnormal (greater than 48 reflux events) in 35% of tested subjects. Abnormal acid reflux was not present in any of the subjects when tested *on PPI* therapy. Wireless pH testing *off PPI therapy* showed a median (IQR) % time pH < 4 of 9% (7%–13%) in total, 11% (8%–16%) in the upright state, and 2% (1%–8%) in the supine state. Eighty-four percent of patients had abnormal wireless pH testing results *off* therapy.

Subgroups. Patients were divided into three subgroups based on endoscopic and pH findings: Esophagitis, normal endoscopy but abnormal pH (referred to as "pH+" in Figures 1 and 2), and normal endoscopy with normal pH (referred to as "pH−" in Figures 1 and 2). Figures 1 and 2 depict individual and group results on the degree of

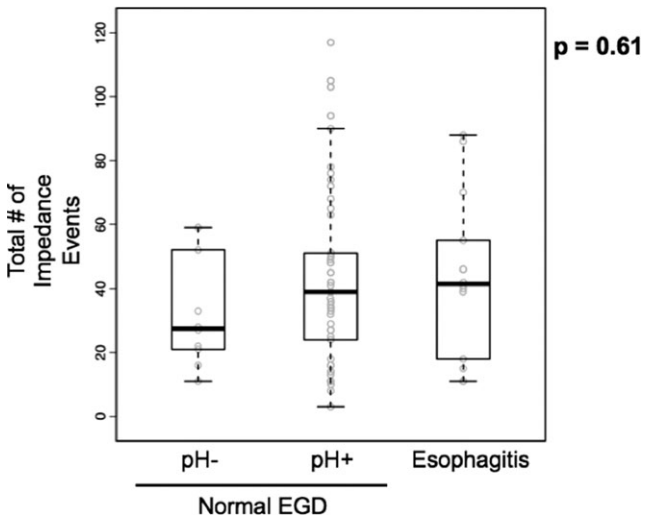

Fig. 2. Individual and group median (IQR) total number of impedance events for the three patient subgroups. Impedance events were similar between the groups.

acid reflux (% total time pH < 4) *off* therapy and total number of impedance events *on* therapy by patient subgroup. As expected, patients with a normal upper endoscopy and normal pH testing exhibited a significantly ($P < 0.001$) lower % total time pH < 4 (3.08% [1.99%–4.89%]), compared to the patients with a normal upper endoscopy and abnormal pH testing (10.45% [7.35%–15.10%]) and those with esophagitis (9.98% [6.70%–12.66%]) (Fig. 1). However, there was no significant difference among the three groups with respect to median (IQR) total number of impedance events on therapy ($P = 0.61$) (27.50 [21.25–47.25];39 [24–51]; 42 [23.25–52.75], respectively) (Fig. 2). Median (IQR) lower esophageal sphincter pressure for the three groups were similar ($P = 0.78$) between the three groups (3.0 [1.0–7.0], 5.5 [0.4–15.5], and 8.0 [1.3–29], respectively).

Predictors of Impedance Findings

No significant overall relationship was observed between the total number of impedance events and the traditionally recognized physiologic parameters for reflux. Specifically, there was no correlation (r = 0.15, $P = 0.2$) between the total number of impedance events *on* therapy and the total % time pH < 4 *off* therapy (Fig. 3A). Presence or absence of hiatal hernia did not change this relationship (Fig. 3B). However, among patients with endoscopic evidence of esophagitis, there was a weak significant correlation (r = 0.54, $P = 0.045$) between the *on* therapy impedance and *off* therapy pH parameters (Fig. 3C). However, the impedance parameters *on* therapy could not be predicted based on severity of reflux parameters at baseline. Number of impedance events *on* therapy were similar ($P = 0.99$) between patients who had no or mild reflux (% time pH < 4 of less than 10%) compared to those with moderate to severe reflux (% time pH < 4 of more than 10%) (Fig. 4).

DISCUSSION

In a unique cohort of patients who had both *off* and *on* therapy testing, we sought to determine if *on* therapy impedance data can be predicted from the more traditionally recognized and commonly employed *off* therapy upper endoscopy, manometry, or pH parameters. In this stepwise diagnostic approach, we found that among patients with extraesophageal reflux refractory to PPI

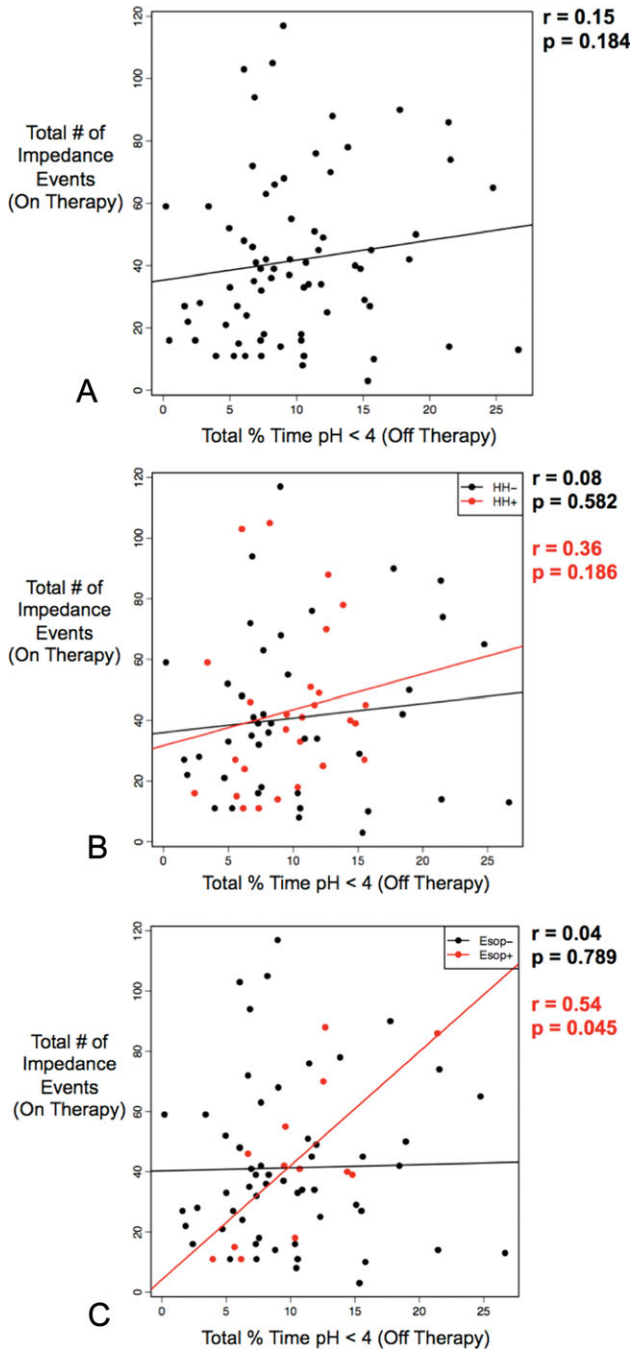

Fig. 3. Correlation between % total time pH < 4 *off therapy* and total number of impedance events *on therapy* for Overall (A), stratified by hiatal hernia (B), and stratified by endoscopic presence of esophagitis (C). HH+ = hiatal hernia present; HH– = no hiatal hernia; Esop– = no esophagitis; Esop+ = esophagitis. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

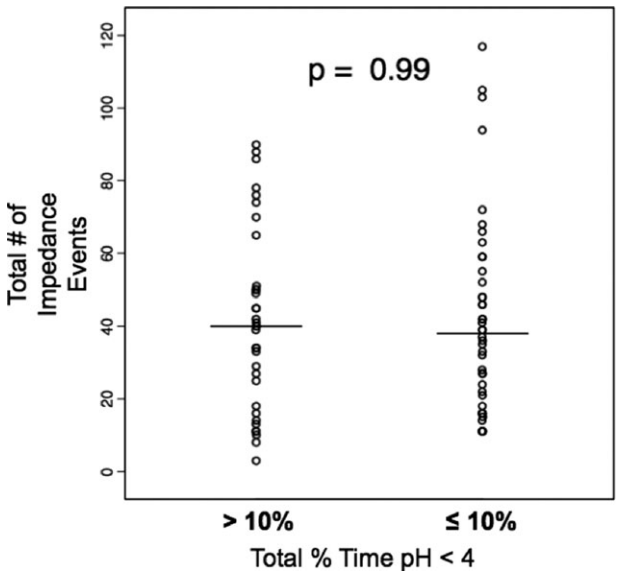

Fig. 4. Impedance parameters dichotomized by severity of baseline acid reflux. The impedance parameters were no different in those with no or mild reflux (defined as % total time pH $<$ 4 of less than or equal to 10%) compared to those with moderate to severe reflux ($>$10% acid reflux).

therapy, impedance testing results *on* therapy cannot be predicted from customary baseline parameters, except among patients with esophagitis in whom the correlation was weak at best, with only 19% of the population having esophagitis. Our data suggests that impedance parameters *on* therapy do not correlate well with any reflux parameters previously employed to assess disease severity. Thus, we urge caution regarding the over-interpretation of impedance parameters, as the clinical relevance of impedance testing remains unclear at this time.

The presence of non- or weakly acidic reflux in patients on PPI therapy is suggested to imply continued reflux and the need for additional therapies.[10,11] In a group of 19 patients who had positive symptom association with acid or nonacid reflux on impedance testing, a retrospective phone interview study suggested 94% fundoplication success.[23] However, two recent prospective trials have questioned the clinical reliability of symptom indices in reflux disease.[24,25] Furthermore, the most recent surgical trial in patients with extraesophageal syndrome showed that impedance parameters *on* therapy did not predict symptom response postfundoplication.[26] In this study, the traditional parameters of increased acid exposure, presence of hiatal hernia, and typical reflux symptoms (heartburn and regurgitation) were predictive of extraesophageal symptom response to fundoplication.

Important controversy in patients with continued symptoms, despite aggressive PPI therapy, is whether to conduct testing *on* or *off* PPI therapy. Employing both impedance-pH monitoring *on* therapy and wireless pH monitoring *off* therapy in the same group of patients with PPI-refractory symptoms, we confirmed that non- or weakly acid reflux may be present in up to 35% of patients; however, continued acid reflux was not seen in any patient. Our data are in agreement with two prior studies; one showing that continued acid reflux is a rarity on twice daily PPI therapy,[27] and the other showing continued nonacid reflux by impedance testing in 37% of patients refractory to PPI therapy.[12] More important, we could not identify any *off* therapy traditionally employed physiologic parameter that could predict the *on* therapy impedance findings. Furthermore, patients with more severe reflux by pH testing defined as % time pH $<$ 4 of greater than 10% had similar impedance parameters than those with no or mild reflux at baseline (Fig. 4). Thus, it appears that the impedance parameters do not correlate with any of the traditionally employed tools in assessing reflux severity. For example, it has been shown that patients with hiatal hernia typically have higher reflux scores compared to those without hiatal hernia,[28] esophagitis severity is expected to correlate with hiatal hernia size and esophageal acid exposure,[29] and % time pH $<$ 4 increases in a graded fashion across the GERD spectrum.[30] Thus, given the lack of any correlation between impedance results and these traditional markers, we urge caution regarding the clinical relevance of impedance testing.

Our study is unique in that the same patient population underwent physiologic testing off and on PPI therapy. However, some limitations of our study should also be highlighted. First, the results from our study underscore the need for larger outcome studies among patients with refractory symptoms and abnormal impedance testing. Second, our present analysis discusses the impedance findings with respect to abnormal number of reflux events in the distal esophagus. We did not evaluate proximal extent and liquid, gas, or mixed nature of the refluxate, as some believe may be important in a subgroup of treatment-resistant patients.[31] Additionally, we have used number of reflux events as the primary measure as opposed to SI or SAP. However, the use of SI and SAP is problematic in this group since patients have already declared lack of clinical response to aggressive acid suppression, and recent studies suggest that these metrics may not be reliable or reproducible.[24,25]

CONCLUSION

In a unique group of patients who had both *off* therapy traditional esophageal physiologic testing and *on* therapy impedance monitoring, our study shows limited correlation between the latter results with the former previously recognized and employed methodologies. There remains uncertainty regarding the clinical utility of impedance testing among patients with extraesophageal symptoms, and we recommend caution in over-interpretation of impedance pH monitoring data.

BIBLIOGRAPHY

1. Dent J, El-Serag HB, Wallander MA, Johansson S. Epidemiology of gastro-oesophageal reflux disease: a systematic review. *Gut* 2005;54: 710–717.
2. Locke GR 3rd, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ 3rd. Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota. *Gastroenterology* 1997;112: 1448–1456.

3. Shaheen NJ, Hansen RA, Morgan DR, et al. The burden of gastrointestinal and liver diseases, 2006. *Am J Gastroenterol* 2006;101: 2128–2138.
4. Frye JW, Vaezi MF. Extraesophageal GERD. *Gastroenterol Clin North Am* 2008;37:845–858, ix.
5. Qadeer MA, Phillips CO, Lopez AR, et al. Proton pump inhibitor therapy for suspected GERD-related chronic laryngitis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol* 2006;101:2646–2654.
6. Kahrilas PJ, Shaheen NJ, Vaezi MF. American Gastroenterological Association Institute technical review on the management of gastroesophageal reflux disease. *Gastroenterology* 2008;135:1392–413, 413 e1–5.
7. Park W, Vaezi MF. Esophageal impedance recording: clinical utility and limitations. *Curr Gastroenterol Rep* 2005;7:182–189.
8. Sifrim D, Blondeau K. Technology insight: the role of impedance testing for esophageal disorders. *Nat Clin Pract Gastroenterol Hepatol* 2006;3: 210–219.
9. Tutuian R, Castell DO. Review article: complete gastro-oesophageal reflux monitoring—combined pH and impedance. *Aliment Pharmacol Ther* 2006;24 (suppl 2):27–37.
10. Agrawal A, Castell DO. Clinical importance of impedance measurements. *J Clin Gastroenterol* 2008;42:579–583.
11. Bredenoord AJ, Tutuian R, Smout AJ, Castell DO. Technology review: esophageal impedance monitoring. *Am J Gastroenterol* 2007;102: 187–194.
12. Mainie I, Tutuian R, Shay S, et al. Acid and non-acid reflux in patients with persistent symptoms despite acid suppressive therapy: a multicentre study using combined ambulatory impedance-pH monitoring. *Gut* 2006;55:1398–1402.
13. Zerbib F, Roman S, Ropert A, et al. Esophageal pH-impedance monitoring and symptom analysis in GERD: a study in patients off and on therapy. *Am J Gastroenterol* 2006;101:1956–1963.
14. DeVault KR, Castell DO. Updated guidelines for the diagnosis and treatment of gastroesophageal reflux disease. *Am J Gastroenterol* 2005;100: 190–200.
15. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. *Gastroenterology* 2008;135:1383–1391, 91 e1–5.
16. Park W, Hicks DM, Khandwala F, et al. Laryngopharyngeal reflux: prospective cohort study evaluating optimal dose of proton-pump inhibitor therapy and pretherapy predictors of response. *Laryngoscope* 2005;115: 1230–1238.
17. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut* 1999;45:172–180.
18. Pandolfino JE, Richter JE, Ours T, Guardino JM, Chapman J, Kahrilas PJ. Ambulatory esophageal pH monitoring using a wireless system. *Am J Gastroenterol* 2003;98:740–749.
19. Zerbib F, des Varannes SB, Roman S, et al. Normal values and day-to-day variability of 24-h ambulatory oesophageal impedance-pH monitoring in a Belgian-French cohort of healthy subjects. *Aliment Pharmacol Ther* 2005;22:1011–1021.
20. Shay S, Tutuian R, Sifrim D, et al. Twenty-four hour ambulatory simultaneous impedance and pH monitoring: a multicenter report of normal values from 60 healthy volunteers. *Am J Gastroenterol* 2004;99:1037–1043.
21. Pritchett JM, Aslam M, Slaughter JC, Ness RM, Garrett CG, Vaezi MF. Efficacy of esophageal impedance/pH monitoring in patients with refractory gastroesophageal reflux disease, on and off therapy. *Clin Gastroenterol Hepatol* 2009;7:743–748.
22. Tutuian R, Mainie I, Agrawal A, et al. Normal values for ambulatory 24-hour combined impedance-pH monitoring on acid suppressive therapy. *Gastroenterology* 2006;130:A171.
23. Mainie I, Tutuian R, Agrawal A, Adams D, Castell DO. Combined multichannel intraluminal impedance-pH monitoring to select patients with persistent gastro-oesophageal reflux for laparoscopic Nissen fundoplication. *Br J Surg* 2006;93:1483–1487.
24. Slaughter JC, Goutte M, Rymer JA, et al. Caution about overinterpretation of symptom indexes in reflux monitoring for refractory gastroesophageal reflux disease. *Clin Gastroenterol Hepatol* 2011;9:868–874.
25. Vaezi MF, Higginbotham T, Kavitt RT, et al. Do patients report symptoms when they occur? Fundamental consequences for interpretation of ambulatory reflux monitoring. *Gastroenterology* 2012;142:S–415.
26. Francis DO, Goutte M, Slaughter JC, et al. Traditional reflux parameters and not impedance monitoring predict outcome after fundoplication in extraesophageal reflux. *Laryngoscope* 2011;121:1902–1909.
27. Charbel S, Khandwala F, Vaezi MF. The role of esophageal pH monitoring in symptomatic patients on PPI therapy. *Am J Gastroenterol* 2005;100: 283–289.
28. Sloan S, Rademaker AW, Kahrilas PJ. Determinants of gastroesophageal junction incompetence: hiatal hernia, lower esophageal sphincter, or both? *Ann Intern Med* 1992;117:977–982.
29. Jones MP, Sloan SS, Rabine JC, Ebert CC, Huang CF, Kahrilas PJ. Hiatal hernia size is the dominant determinant of esophagitis presence and severity in gastroesophageal reflux disease. *Am J Gastroenterol* 2001;96: 1711–1717.
30. Vaezi MF, Richter JE. Role of acid and duodenogastroesophageal reflux in gastroesophageal reflux disease. *Gastroenterology* 1996;111:1192– 1199.
31. Zerbib F, Duriez A, Roman S, Capdepont M, Mion F. Determinants of gastro-oesophageal reflux perception in patients with persistent symptoms despite proton pump inhibitors. *Gut* 2008;57:156–160.

Reprinted by permission of Otolaryngol Head Neck Surg. 2008; 139(3):414-420.

Otolaryngology–Head and Neck Surgery (2008) 139, 414-420

ORIGINAL RESEARCH—LARYNGOLOGY AND NEUROLARYNGOLOGY

Double-blind, placebo-controlled trial with esomeprazole for symptoms and signs associated with laryngopharyngeal reflux

Oliver Reichel, MD, Holger Dressel, MD, Katrin Wiederänders, and **Wolfgang J. Issing, MD, PhD,** Munich, Germany; and Newcastle Upon Tyne, UK

OBJECTIVE: To determine the efficacy of proton pump inhibitor (PPI) therapy with esomeprazole on symptoms and signs associated with laryngopharyngeal reflux (LPR).

STUDY DESIGN AND METHODS: Prospective, double-blind, randomized, placebo-controlled study. Sixty-two patients with a reflux finding score (RFS) > 7 and a reflux symptom index (RSI) > 13 were enrolled and received either esomeprazole 20 mg twice daily or placebo for three months. RSI and RFS were assessed at baseline, after six weeks, and after three months.

RESULTS: Reductions of total RSI and RFS as well as of several subscores were significantly higher in the treatment group compared to placebo after three months ($P < 0.05$ each). The difference between study groups was most pronounced for posterior commissure hypertrophy ($P < 0.01$).

CONCLUSION: In the treatment of LPR-related symptoms a high placebo effect can be observed. However, compared to control, twice-daily PPI treatment for three months demonstrated a significantly greater improvement in laryngeal appearance and LPR symptoms.

© 2008 American Academy of Otolaryngology–Head and Neck Surgery Foundation. All rights reserved.

Laryngopharyngeal reflux (LPR) has had a significantly increasing impact on otolaryngologist office visits in the last decade. In the three-year period 1990-1993, annual visits to otolaryngologists by patients suffering from reflux-related problems averaged 89,000; this average increased to 421,000 annual visits during the three-year time span 1998-2001.[1] The generally recommended treatment in patients with LPR is twice-daily-dose proton pump inhibitor (PPI) therapy for three to six months.[2,3] However, there is some controversy regarding the adequate PPI dosage and even the efficacy of these drugs for LPR treatment.[3,4] Some authors criticize these recommendations for LPR treatment as being based on poor levels of evidence from uncontrolled studies.[5] In addition, several studies in the past could not demonstrate superiority of PPIs over placebo for treatment of suspected LPR.[6-9] Therefore, the primary objective of this prospective, double-blind, randomized, and placebo-controlled study was to evaluate the effect of a three-month treatment with esomeprazole 20 mg twice daily on symptoms and laryngeal signs in patients with suspected LPR. We chose esomeprazole for this study as we could show adequate measurable acid suppression with a once-daily dose of this PPI in a large number of LPR patients in a former trial[4] and because esomeprazole provides greater 24-hour control of intragastric acid than all other available PPIs at standard doses used to treat erosive esophagitis.[10]

MATERIALS AND METHODS

The following procedures were performed in accordance with the Declaration of Helsinki, Good Clinical Practice, and applicable regulatory requirements. The study was approved by the Ethics Committee of the Medical Faculty of Ludwig Maximilians University Munich (project number 265-05). Before initiation of any procedure, signed informed consent was obtained from all patients.

Participants

Between February 2006 and July 2007, 62 consecutive patients (30 women and 32 men; age range, 21-77 years; mean, 48.7 years) were enrolled. They presented to the Department of Otorhinolaryngology–Head and Neck Surgery of Ludwig Maximilians University, Munich, with unspecific otolaryngologic and respiratory disorders such as chronic cough, dysphagia, throat clearing, globus sensation, hoarseness, sore throat, and heartburn. All patients were examined by one otolaryngologist and rigid or flexible laryngoscopy had to reveal mucosal abnormalities consistent with LPR reflected by a reflux finding score (RFS) > 7.[11] In addition, the reflux symptom index (RSI),[12] a self-administered nine-item outcomes instrument for the diagnosis of LPR, had to exceed the value of 13 for inclusion. None of the study patients were treated with PPIs or any other antireflux medication for at least three months. Further

Received April 9, 2008; revised May 27, 2008; accepted June 3, 2008.

0194-5998/$34.00 © 2008 American Academy of Otolaryngology–Head and Neck Surgery Foundation. All rights reserved.
doi:10.1016/j.otohns.2008.06.003

exclusion criteria included age younger than 18 years, a history of laryngeal malignancy or gastrointestinal surgery, and the need for continuous therapy with warfarin, coumarin, or acetylsalicylacid. No patient with a clinically significant condition that could put the patient at risk, affect the patient's ability to participate in the study, influence the trial results, or necessitate surgery during the study was included. Patients with a contraindication to esomeprazole (eg, known or suspected hypersensitivity or allergy to esomeprazole or other PPIs) or participants of another investigational drug study in the past 30 days were excluded from participation. No pregnant or lactating women were enrolled, and in case of childbearing potential, effective contraception had to be maintained. Finally, patients with drug or alcohol addiction or with any psychiatric disease were not allowed to participate in the study.

Study Procedures

Patients passing all inclusion and exclusion criteria were then sequentially randomized in a 1:1 ratio to receive either coated tablets of esomeprazole 20 mg (twice daily, 30 minutes before meals) or a placebo tablet indistinguishable from the esomeprazole tablet (Fig 1). Both esomeprazole and placebo were given for a total of three months and were supplied in four small plastic bottles, each containing 48 tablets. Esomeprazole and placebo were provided by Astra-Zeneca (Wedel, Germany), and Astra-Zeneca also performed the randomization. Patients and investigators were blinded as to the medication randomization. Enrolled patients did not receive any instructions regarding lifestyle modifications to reduce acid reflux like avoidance of fatty meals or caffeine. However, all included subjects were explicitly instructed to take the medication with water at least 30 minutes before morning and evening meals. The RFS was readministered six weeks and three months after the start of treatment by the same otolaryngologist by laryngoscopy and each patient filled in the RSI questionnaire on these two follow-up visits. The primary objective of the study was the comparison of the total RSI and RFS after a treatment period of three months. At the final visit all patients completing the trial were additionally asked if they thought that therapy had completely resolved their complaints (possible answers were "yes" or "no"). In order to check for compliance with treatment, pill counting was performed at each visit. It was considered adequate if 75% or more of the treatment medication was taken. If one or more study medication bottles were not returned, compliance was not calculated.

Statistical Analysis

For sample size calculation we assumed that a typical LPR-associated subscore (eg, posterior commissure hypertrophy) would improve by at least one point in 65% of the PPI group and in 30% of the placebo group. For comparison of those two proportions in independent samples, a sample size of 31 patients per treatment group was calculated (two-tailed z test with alpha = 0.05, power = 80%, and accounting for a 10% drop-out rate). Baseline characteristics of both groups were reported as mean values and standard deviations (SD) or proportions and compared using the Mann-Whitney U test for continuous variables and the χ^2 test for categorical variables (Table 1). Treatment or placebo effects after six weeks and three months were reported as mean differences and standard errors (SEM) from baseline and were tested using Wilcoxon signed rank tests (Tables 2 and 3). Differences in mean changes between esomeprazole and placebo were additionally tested using Mann-Whitney U test (Tables 2 and 3). In order to compare the subjective estimation of the therapeutic drug effect between the two study groups, again the χ^2 test was used. A P value < 0.05 was considered statistically significant. The statistical analyses were done with SPSS 14.0 for windows (version 14.0.1; SPSS Inc, Chicago, IL).

Role of the Funding Source

Astra-Zeneca had no involvement in the study design, collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

RESULTS

Fifty-eight patients completed the study; the drop-out rate was low with 6.5%. Thirty patients received esomeprazole and the control group consisted of 28 individuals (Fig 1, Table 1). Baseline characteristics between the two study groups were comparable. This was also true for all respective subscores of both RSI and RFS.

Compared to baseline, the total RSI and RFS were significantly reduced in both study groups after a treatment period of six weeks (Table 2). However, differences in total scores between the esomeprazole and placebo group were not statistically significant. The only symptom with a statistically significant decrease for the esomeprazole group was heartburn ($P < 0.05$). When the differences in laryngeal findings found after six weeks of treatment were compared for the two study groups, only diffuse laryngeal edema showed a significantly stronger improvement with esomeprazole ($P < 0.05$). For ventricular obliteration and laryngeal erythema a nonsignificant trend for better improvement in the esomeprazole group was found.

At the final visit, patients of both study groups again reported on an improvement of symptoms, reflected by a highly significant reduction from baseline of the total RSI ($P < 0.001$) (Table 3). However, the total improvement was significantly stronger in the esomeprazole group ($P < 0.05$). While in esomeprazole recipients a highly significant reduction for each single item of the RSI was evident, no significant improvement from baseline for swallowing difficulties, coughing after meals, or heartburn was found in the

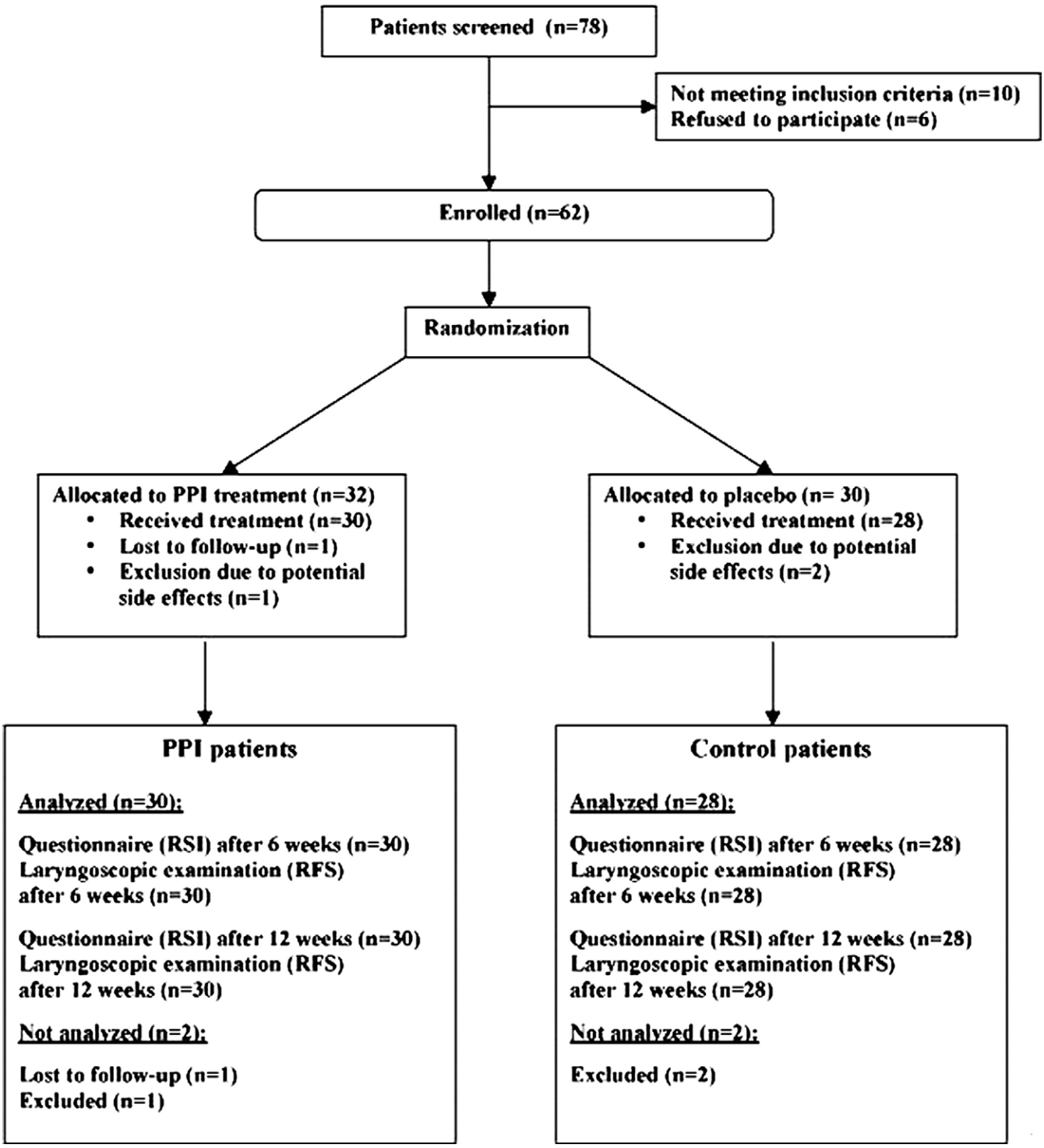

Figure 1 Summary of patient flow throughout study. PPI treatment = esomeprazole 20 mg twice daily.

placebo group. Similar to the result of the first follow-up visit, there was a significantly stronger improvement of the symptom heartburn in patients receiving esomeprazole. The comparison of the total RFS reduction between the two study groups reveals a statistically significant difference in favor of the esomeprazole group ($P < 0.05$). In contrast to esomeprazole recipients, patients receiving placebo did not show any significant improvement of ventricular obliteration or laryngeal erythema. After a treatment period of three months, the final examination revealed a statistically significant stronger improvement of erythema and diffuse laryngeal edema in the esomeprazole group ($P < 0.05$). The most significant difference between the study groups was the much stronger reduction of posterior commissure hypertrophy in the esomeprazole group after three months of therapy ($P < 0.01$).

Table 1
Baseline characteristics*

Variable	Control (n = 28)	PPI (n = 30)	*P* value†
Sex: male	13 (46.4%)	17 (56.7%)	0.160
Smoker	5 (17.9 %)	6 (20.0%)	0.871
Total RSI before therapy	21.79 ± 6.69	23.1 ± 7.45	0.503
Total RFS before therapy	14.89 ± 2.5	14.9 ± 2.75	0.969
Age	47.6 ± 16	49 ± 13.9	0.663

n, number of subjects in each group; *PPI,* proton pump inhibitor (esomeprazole); *RFS,* reflux finding score; *RSI,* reflux symptom index.
*Comparison of baseline characteristics between groups, reported as mean ± standard deviation (SD) or proportion of subjects (%).
†*P* value = Mann-Whitney *U* test or χ^2 test, all nonsignificant.

When asked about persisting complaints at the end of the trial, 11 of 26 placebo recipients (42%) felt themselves to be free of symptoms compared to 22 of 28 patients (>78%) from the esomeprazole group (two patients of each study group could not clearly answer the question with "yes" or "no"). This difference is statistically significant ($P = 0.006$).

All documented adverse events were considered mild and not treatment-related. The most common ones were bronchitis (n = 7), pharyngitis (n = 5), and headache (n = 5). One patient receiving esomeprazole was lost to follow-up after the initial visit. Two female patients complained of abdominal pain two and three weeks after receiving the study medication and were withdrawn from treatment. Their treatment was unblinded, and as they were part of the placebo group their complaints could not be caused by esomeprazole. Another female patient developed a skin rash six weeks after the beginning of therapy. Although this drop-out patient was not a placebo recipient, allergologic testing excluded drug hypersensitivity to esomeprazole (Fig 1).

DISCUSSION

Our study is the second largest to assess the efficacy of a PPI therapy for symptoms and signs associated with LPR compared to placebo and the second one choosing esomeprazole for a controlled trial. The results of this study demonstrate a significant placebo effect in the therapy of LPR-related symptoms as reported by patients especially within the first 6 weeks of PPI treatment. This phenomenon was already observed in former placebo-controlled studies investigating the efficacy of PPI therapy in patients with

Table 2
Change from baseline within each group after six weeks

	Visit I vs visit II		
	Esomeprazole	Placebo	Difference
RSI			
Total	9.87 ± 1.61‡	6.93 ± 1.61‡	2.94 ± 2.27
Hoarseness	1.03 ± 0.31†	1.07 ± 0.35†	−0.04 ± 0.47
Throat clear	1.20 ± 0.31†	1.32 ± 0.32†	−0.12 ± 0.45
Throat mucus	1.10 ± 0.28†	0.75 ± 0.22†	0.35 ± 0.35
Difficulty swallowing	1.23 ± 0.30†	0.57 ± 0.27*	0.66 ± 0.41
Coughing after meals	0.53 ± 0.34	−0.11 ± 0.32	0.64 ± 0.47
Breathing difficulties	0.50 ± 0.29	0.57 ± 0.25*	−0.07 ± 0.39
Annoying cough	0.87 ± 0.39*	0.68 ± 0.27*	0.19 ± 0.48
Throat sensations	1.60 ± 0.32‡	1.25 ± 0.34†	0.35 ± 0.46
Heartburn	1.53 ± 0.32‡	0.75 ± 0.27†	0.78 ± 0.42*
RFS			
Total	3.47 ± 0.47‡	2.46 ± 0.71†	1.00 ± 0.85
Subglottic edema	−0.07 ± 0.07	0.21 ± 0.19	−0.28 ± 0.12
Ventricular	0.60 ± 0.17†	0.36 ± 0.18	0.24 ± 0.25
Erythema	0.77 ± 0.21†	0.43 ± 0.26	0.34 ± 0.33
Vocal fold edema	0.33 ± 0.09†	0.46 ± 0.13†	−0.13 ± 0.16
Diffuse laryngeal edema	0.60 ± 0.13‡	0.21 ± 0.11	0.39 ± 0.16*
Posterior commissure hypertrophy	0.80 ± 0.15‡	0.54 ± 0.13†	0.26 ± 0.20
Granuloma/granulation tissue	0.07 ± 0.07	0.00 ± 0.18	0.07 ± 0.19
Thick endolaryngeal mucus	0.40 ± 0.23	0.29 ± 0.25	0.11 ± 0.33

Data are given as mean differences ± standard error of the mean (SEM). The first two data columns show differences from baseline and the third demonstrates the difference between study groups. *P* values are from Wilcoxon signed rank test and Mann-Whitney *U* test as appropriate.
*$P < 0.05$; †$P < 0.01$; ‡$P < 0.001$.

Table 3
Change from baseline within each group after three months

	Visit I vs visit III		
	Esomeprazole	Placebo	Difference
RSI			
Total	14.27 ± 1.58‡	7.79 ± 1.74‡	6.48 ± 2.34*
Hoarseness	1.37 ± 0.33‡	1.18 ± 0.36†	0.19 ± 0.48
Throat clear	1.80 ± 0.26‡	1.18 ± 0.31†	0.62 ± 0.41
Throat mucus	1.43 ± 0.33†	0.96 ± 0.27†	0.47 ± 0.43
Difficulty swallowing	1.40 ± 0.31‡	0.57 ± 0.28	0.83 ± 0.42
Coughing after meals	1.17 ± 0.37†	0.21 ± 0.33	0.95 ± 0.50
Breathing difficulties	0.87 ± 0.27†	0.57 ± 0.26*	0.30 ± 0.38
Annoying cough	1.60 ± 0.37†	1.07 ± 0.34†	0.53 ± 0.51
Throat sensations	2.27 ± 0.28‡	1.43 ± 0.33†	0.84 ± 0.43
Heartburn	1.97 ± 0.31‡	0.64 ± 0.42	1.32 ± 0.52*
RFS			
Total	4.60 ± 0.63‡	2.32 ± 0.76†	2.28 ± 0.98*
Subglottic edema	0.00 ± 0.10	0.21 ± 0.21	−0.21 ± 0.24
Ventricular	0.53 ± 0.20*	0.36 ± 0.21	0.18 ± 0.28
Erythema	0.77 ± 0.20†	0.07 ± 0.19	0.70 ± 0.27*
Vocal fold edema	0.67 ± 0.13‡	0.46 ± 0.16†	0.20 ± 0.20
Diffuse laryngeal edema	0.83 ± 0.16‡	0.36 ± 0.13*	0.48 ± 0.21*
Posterior commissure hypertrophy	0.97 ± 0.15‡	0.32 ± 0.12*	0.65 ± 0.19**
Granuloma/granulation tissue	0.07 ± 0.07	0.00 ± 0.15	0.07 ± 0.16
Thick endolaryngeal mucus	0.67 ± 0.20†	0.58 ± 0.25*	0.10 ± 0.32

Data are given as mean differences ± standard error of the mean (SEM). The first two data columns show differences from baseline and the third demonstrates the difference between study groups. *P* values are from Wilcoxon signed rank test and Mann-Whitney *U* test as appropriate.
*$P < 0.05$; †$P < 0.01$; ‡$P < 0.001$.

LPR.[7,8,13] After three months of treatment, however, the differences in symptom improvement between the two study groups became more obvious, reflected by a significantly stronger decrease in the total RSI after esomeprazole therapy. As a consequence, even in patients with suspected LPR reporting early symptom relief under PPI medication, the treatment should continue for at least three months.

In contrast to Vaezi et al,[9] who had performed the largest double-blind and placebo-controlled trial evaluating a PPI effect on LPR symptoms and laryngeal signs, we could demonstrate that twice-daily esomeprazole was superior to placebo in improving both LPR symptoms and laryngeal findings. In the mentioned study, Vaezi et al included 145 patients receiving either 40 mg of esomeprazole (n = 95) or placebo (n = 50) twice daily for 16 weeks. From their findings, Vaezi et al concluded that compared with placebo the PPI therapy was of no therapeutic benefit on signs and symptoms associated with LPR.[9] However, they excluded patients with moderate to severe heartburn from their study. Thus, patients with a symptom typical for gastroesophageal reflux disease (GERD) but also relevant for LPR (6% to 43% of LPR patients suffer from heartburn[14,15]) were not part of the large study population. This could have affected the study results. Another reason for this study result differing from our findings might be the fact that Vaezi et al did not use the RFS for control of laryngeal changes due to its lack of external validation.

Four further studies in the past also revealed no statistically significant benefit of a PPI therapy on characteristic LPR symptoms and signs compared to placebo. Havas et al performed a double-blind, placebo-controlled, and randomized study with 15 LPR patients evaluating the therapeutic efficacy of 30 mg lansoprazole twice daily for 12 weeks.[6] From their findings the authors concluded that lansoprazole was not more effective than placebo in the treatment of cervical symptoms of LPR and posterior pharyngolaryngitis. As the study population was small and no regular statistical analysis of the results was performed, the significance of this conclusion remains unclear. In a randomized, double-blind, crossover study by Eherer et al, pantoprazole 40 mg twice daily for three months did not significantly affect symptom or laryngeal scores compared with placebo in 14 LPR patients.[16]

Nevertheless, the authors hypothesized that pantoprazole compared to placebo may have resulted in faster improvement of LPR symptoms. Again, the limitation of this study is a small sample size, which makes it difficult to draw meaningful conclusions from this data. Another double-blind, randomized trial by Steward et al, comparing two-month rabeprazole (20 mg twice daily) to placebo-control PPI treatment, also failed to demonstrate significantly greater improvement in reflux symptoms, health status, or laryngeal appearance.[7] However, for laryngeal symptoms such as hoarseness, dry cough, and throat clearing a statis-

tically significant better improvement was noted for PPI-treated patients. In our opinion, it is not surprising that in this study laryngeal signs of LPR showed no stronger improvement in PPI recipients compared to control within two months, as the physical findings of LPR improve more slowly than the symptoms.[17] In our study we could also find no statistically significant difference in the laryngeal appearance (reflected by the total RFS) between the esomeprazole and placebo group after a short treatment period of six weeks (Table 2). Another critical point for the mentioned study is the fact that patients were not instructed on when to take the medication. This could have seriously affected the results of this study. In a randomized, double-blind, placebo-controlled trial, Wo et al evaluated the efficacy of single-dose pantoprazole 40 mg for 12 weeks in newly diagnosed LPR.[8] The response was similar between the pantoprazole and placebo group. As 60% of the study subjects had additional abnormal distal esophageal reflux and the single-dose PPI was not sufficient to reach a pH-documented measurable suppression of hypopharyngeal acid reflux, the results of this study are not adequately comparable to those of our study with a double-dose PPI design. Noordzij et al performed another prospective, placebo-controlled, randomized, and double-blind study to determine the efficacy of 40 mg omeprazole twice daily for two months in the treatment of LPR.[13] The authors could demonstrate a significant improvement of a composite laryngeal symptom score in the omeprazole group but not the placebo group. Again, the endoscopic laryngeal signs did not change significantly over the course of the study for either treatment group. As mentioned above, the two-month follow-up period may not have been sufficient to detect changes in laryngeal appearance. Another study investigating the therapeutic benefit of lansoprazole 30 mg twice daily for treating LPR, by El-Serag et al, provided evidence that lansoprazole therapy for three months achieved significantly better symptomatic response than placebo.[18] Due to a selected referral study population with a high likelihood of GERD, the authors suggested not to generalize their results to patients with LPR in a primary care setting.

The analysis of the respective RFS subscores in our study revealed a highly significant reduction of posterior commissure hypertrophy in the esomeprazole but not in the placebo group after a treatment period of 12 weeks ($P < 0.01$). This laryngeal sign is supposed to be one of the mucosal alterations most related to LPR.[19] The fact that the most significant difference in laryngeal appearance between the two study groups after 12 weeks could be detected in this special area in our opinion strongly indicates the efficacy of a PPI treatment in patients with symptoms and signs associated with LPR. The improvement of posterior commissure hypertrophy was not significantly better in the esomeprazole group compared to control at the first follow-up. This result underlines the importance of a PPI treatment for at least three months in patients with suspected LPR. Another striking result was that we found heartburn to be the only RSI subscore with a statistically significant difference in improvement between the two study groups after both six weeks and three months. This finding, in our opinion, clearly demonstrates that PPI therapy has achieved an earlier and stronger reduction of distal reflux episodes potentially causing heartburn compared to placebo. This result also reflects the well-known fact that esophagitis caused by reflux shows an earlier resonse to PPI treatment than reflux-induced laryngitis.[3,17]

Several issues should be addressed. As we did not randomize our otolaryngologic evaluations performed by only one examiner, the term double-blind can only be used to describe the medication randomization. Another critical aspect of our trial may be the fact that we did not perform 24-hour pH monitoring to diagnose LPR objectively before patient inclusion. According to reports of other authors, patient tolerance is poor for this procedure.[7] As such, we decided to assess LPR-related symptoms and signs alone with the use of RSI and RFS. Moreover, Vaezi et al, in the above-mentioned study, found that only a small proportion of their patients undergoing pH monitoring had documented pharyngeal acid exposure despite typical LPR symptoms and laryngoscopic signs.[9] From this finding they concluded that sensitivity of pH monitoring for detection of proximal esophageal or hypopharyngeal reflux episodes might not be more than 50%.[9] Another argument against performing pH monitoring before inclusion was the fact that more and more authors doubt that 24-hour pH monitoring, although supposed to be the gold-standard test for LPR, is the preferable initial step in the work-up of most patients with LPR.[9,18,20] A third limitation of our study might be the short follow-up of 12 weeks. The significance of our study results probably would have been higher after a treatment period of six months. However, only Vaezi et al performed a trial with a follow-up of more than three months.[9] Another critical aspect might be the dose of esomeprazole used (20 mg). It can be hypothesized that the differences in RSI and RFS improvement between the study groups would have been even more significant with a dose of 40 mg twice daily. This is the dose generally recommended for treating LPR.[2] However, we chose a dose of 20 mg esomeprazole twice daily as many institutions and general practitioners in Germany prefer to start with the lower dose. Nevertheless, we suppose our results clearly demonstrate a therapeutic effect of PPI treatment for LPR-related symptoms and signs. This estimation is confirmed by the subjective opinion of our study patients concerning the drug effect, with only 42% of placebo recipients and more than 78% of the esomeprazole group being free of symptoms ($P < 0.006$).

CONCLUSION

Especially during the first weeks of PPI therapy, a significant placebo effect appears to exist in the treatment of LPR-related symptoms. However, compared to placebo,

twice-daily esomeprazole treatment with a dose of 20 mg for three months demonstrated a significantly greater improvement in laryngeal appearance and LPR-related symptoms, as reflected by the total score differences of RFS and RSI. The most striking difference between the study groups was the significant reduction of posterior commissure hypertrophy in the PPI group.

AUTHOR INFORMATION

From the Department of Otorhinolaryngology–Head and Neck Surgery (Dr Reichel and Ms Wiederänders) and the Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine (Dr Dressel), Ludwig Maximilians University, Munich, Germany; and the Department of Otolaryngology (Dr Issing), The Freeman Hospital, Newcastle Upon Tyne, UK.

Corresponding author: Oliver Reichel, MD, Department of Otolaryngology–Head and Neck Surgery, Klinikum Großhadern, Marchioninistraße 15, 81377 Munich, Germany.

E-mail address: Oliver.Reichel@med.uni-muenchen.de.

AUTHOR CONTRIBUTIONS

Oliver Reichel, study design, data collection; **Holger Dressel**, data analysis, statistical analysis; **Katrin Wiederänders**, data collection; **Wolfgang J. Issing**, study design.

FINANCIAL DISCLOSURE

This study was partially supported by a grant from Astra-Zeneca, Wedel, Germany. AstraZeneca had no involvement in the study design, collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

REFERENCES

1. Altman KW, Stephens RM, Lyttle CS, et al. Changing impact of gastroesophageal reflux in medical and otolaryngology practice. Laryngoscope 2005;115:1145–53.
2. Koufman JA, Aviv JE, Casiano RR, et al. Laryngopharyngeal reflux: position statement of the Committee on Speech, Voice, and Swallowing Disorders of the American Academy of Otolaryngology–Head and Neck Surgery. Otolaryngol Head Neck Surg 2002;127:32–5.
3. Ford CN. Evaluation and management of laryngopharyngeal reflux. JAMA 2005;294:1534–40.
4. Reichel O, Keller J, Rasp G, et al. Efficacy of once-daily esomeprazole treatment in patients with laryngopharyngeal reflux evaluated by 24-hour pH monitoring. Otolaryngol Head Neck Surg 2007;136:205–10.
5. Karkos PD, Wilson JA. Empiric treatment of laryngopharyngeal reflux with proton pump inhibitors: a systemativ review. Laryngoscope 2006; 116:144–8.
6. Havas T, Huang S, Levy M, et al. Posterior pharyngolaryngitis: double-blind, randomised placebo-controlled trial of proton pump inhibitor therapy. Aust J Otolaryngol 1999;3:243–6.
7. Steward DL, Wilson KM, Kelly DH, et al. Proton pump inhibitor therapy for chronic laryngo-pharyngitis: a randomized placebo-control trial. Otolaryngol Head Neck Surg 2004;131:342–50.
8. Wo JM, Koopman J, Harrell SP, et al. Double-blind, placebo-controlled trial with single-dose pantoprazole for laryngopharyngeal reflux. Am J Gastroenterol 2006;101:1972–8.
9. Vaezi MF, Richter JE, Stasney CR, et al. Treatment of chronic posterior laryngitis with esomeprazole. Laryngoscope 2006;116:254–60.
10. Miner P Jr, Katz PO, Chen Y, et al. Gastric acid control with esomeprazole, lansoprazole, omeprazole, pantoprazole, and rabeprazole: a five-way crossover study. Am J Gastroenterol 2003;98:2616–20.
11. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). Laryngoscope 2001;111:1313–7.
12. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the Reflux Symptom Index (RSI). J Voice 2002;16:274–7.
13. Noordzij JP, Khidr A, Evans BA, et al. Evaluation of omeprazole in the treatment of reflux laryngitis: a prospective, placebo-controlled, randomized, double-blind study. Laryngoscope 2001;111:2147–51.
14. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. Laryngoscope 1991;101(suppl 53):1–78.
15. Kamel PL, Hanson D, Kahrilas PJ. Omeprazole for the treatment of posterior laryngitis. Am J Med 1994;96:321–6.
16. Eherer AJ, Habermann W, Hammer HF, et al. Effect of pantoprazole on the course of reflux-associated laryngitis: a placebo-controlled double-blind crossover study. Scand J Gastroenterol 2003;38:462–7.
17. Belafsky PC, Postma GN, Koufman JA. Laryngopharyngeal reflux symptoms improve before changes in physical findings. Laryngoscope 2001;111:979–81.
18. El-Serag HB, Lee P, Buchner A, et al. Lansoprazole treatment of patients with chronic idiopathic laryngitis: a placebo-controlled trial. Am J Gastroenterol 2001;96:979–83.
19. Book DT, Rhee JS, Toohill RJ, et al. Perspectives in laryngopharyngeal reflux: an international survey. Laryngoscope 2002;112:1399–406.
20. DelGaudio JM, Waring JP. Empiric esomeprazole in the treatment of laryngopharyngeal reflux. Laryngoscope 2003;113:598–601.

Reprinted by permission of Arch Otolaryngol Head Neck Surg. 2010; 136(11):1089-1092.

ORIGINAL ARTICLE

ONLINE FIRST

Treatment of Clinically Diagnosed Laryngopharyngeal Reflux Disease

Tarek Fouad Youssef, MD; Mohamed Rifaat Ahmed, MD

Objectives: To determine the incidence of *Helicobacter pylori* (HP) stool antigen (HPSA) in patients with laryngopharyngeal reflux disease (LPRD), and to make a comparison of 2 treatment regimens that have been used based on the presence or absence of HPSA positivity in patients with LPRD.

Design: Randomized controlled study.

Setting: Suez Canal University Hospital, Ismalia, Egypt.

Patients: A total of 212 patients with symptoms of LPRD.

Intervention: Patients were evaluated by laryngoscopy, ambulatory pH monitoring for 24 hours, and HPSA testing. Esomeprazole magnesium as a monotherapy was evaluated vs triple therapy in patients with HP infection.

Main Outcome Measures: To determine the incidence of HPSA in patients with LPRD, and to make a comparison of 2 treatment regimens that have been used based on the presence or absence of HPSA positivity in patients with LPRD.

Results: Persistent dry cough and a feeling of a lump in the throat (globus sensation) were the most frequent symptoms of LPRD, while posterior laryngeal inflammation was the main laryngoscopic finding. Results from the HPSA test were positive in 57% of the studied group. Patients with negative HPSA were treated with esomeprazole as single modality with a reported improvement score of 96.6%. Patients with positive HPSA test results were divided into 2 groups: 1 received only esomeprazole, with reported improvement in 40%, whereas the second group was treated with esomeprazole, plus amoxicillin sodium and clarithromycin (triple therapy) and reported a 90% incidence of symptom improvement.

Conclusion: The incidence of HP infection in patients with LPRD in our study was 57%. Triple therapy showed a higher cure rate in patients with HPSA-positive test results.

Arch Otolaryngol Head Neck Surg. 2010;136(11):1089-1092.
Published online September 20, 2010.
doi:10.1001/archoto.2010.165

GASTROESOPHAGEAL REflux disease (GERD) is defined as a backward flow of gastric contents into the esophagus.[1] Beaver et al[2] suggested that laryngopharyngeal reflux disease (LPRD) means a backward flow of the stomach contents up to the throat. The clinical symptoms usually occur secondary to a refluxate of hydrochloric acid and pepsin.[3] The gastric refluxate in the larynx might be the causative factor in posterior laryngeal inflammation, laryngeal contact ulcers, and laryngeal granuloma formation.[3,4] It is associated with many otolaryngology disorders, such as reflux laryngitis, cervical dysphagia, globus pharyngeus, chronic cough, laryngeal or tracheal stenosis, and laryngeal carcinoma.[4] The incidence of laryngopharyngeal symptoms is greater than expected.[5]

There is a complex multifactor set of pathophysiologic characteristics of LPRD besides simple acid reflux.[6] *Helicobacter pylori* (HP) is a gram-negative, microaerophilic bacterium that can cause infection of the stomach and is also strongly linked to the development of duodenal and gastric ulcers.[7,8] A relationship between the rates and degree of reflux esophagitis with HP infection has been reported, but to our knowledge, no relationship with reflux laryngitis has been reported.[9,10] The HP stool antigen (HPSA) test is a rapid, noninvasive diagnostic method based on a sandwich enzyme immunoassay with antigen detection, which has a high sensitivity and specificity.[11,12]

No standard guidelines are available for treatment of LPRD; proton pump inhibitors, twice daily for 8 weeks, have been recommended[13] if HP is present. However, clinical guidelines may consider revision to add a triple therapy regimen.

Author Affiliations:
Department of Otolaryngology–Head and Neck Surgery, Faculty of Medicine, Suez Canal University, Ismalia, Egypt.

Table 1. Relationship Between Laryngopharyngeal Reflux Disease (LPRD) Symptoms With HPSA Test Results[a]

LPRD Symptom	Patients With HPSA Test Result, No.	
	Positive (n=122)	Negative (n=90)
Chronic cough	59	46
Feeling of lump in throat	52	46
Frequent throat clearing	40	38
Bad/bitter taste	37	32
Hoarseness	24	30

Abbreviation: HPSA, *Helicobacter pylori* stool antigen.
[a] All comparisons were nonsignificant ($P<.05$ was considered statistically significant).

Table 2. Relationship Between Laryngoscopic Findings and HPSA Test Results[a]

Laryngoscopic Findings	Patients With HPSA Test Result, No.	
	Positive (n=122)	Negative (n=90)
Red, irritated arytenoids	61	55
Swelling of the vocal cords	32	27
Small laryngeal ulcers	11	13
Granulomas in the larynx	2	2

Abbreviation: HPSA, *Helicobacter pylori* stool antigen.
[a] All comparisons were nonsignificant ($P<.05$ was considered statistically significant).

We determined the incidence of HPSA-positive findings among patients diagnosed as having LPRD and compared the efficacy of 2 treatment regimens.

METHODS

We performed a randomized controlled study at Suez Canal University Hospital, Ismalia, Egypt. A total of 212 patients with LPRD symptoms (hoarseness, chronic unexplained cough, frequent throat clearing, a feeling of a lump in the throat [globus sensation], and a bad or bitter taste in the mouth[3]) were included in this study, but we excluded patients with a history of smoking, alcohol intake, chronic rhinosinusitis, or treatment for LPRD.

All patients underwent laryngoscopic examination to confirm reflux signs, then 24 hours of pH monitoring was ordered (the pH test was considered to be positive for LPRD when the pH was lower than 4; HPSA testing was ordered when a fresh stool sample was obtained).[11,12]

Patients with negative HPSA test results received once-daily esomeprazole magnesium, 40 mg, for 4 weeks.[14] Patients with positive HPSA test results were divided into 2 equal randomized groups: one was a control group that received only esomeprazole magnesium, 40 mg, for 4 weeks,[14] and the other was a study group that received triple therapy comprising esomeprazole magnesium, 40 mg, plus amoxicillin sodium, 1 g, and clarithromycin, 500 mg, for the same period.[14]

A senior otolaryngologist (T.F.Y.) who was blind to the treatment protocol performed follow-up evaluation for all patients after the end of medical treatment.

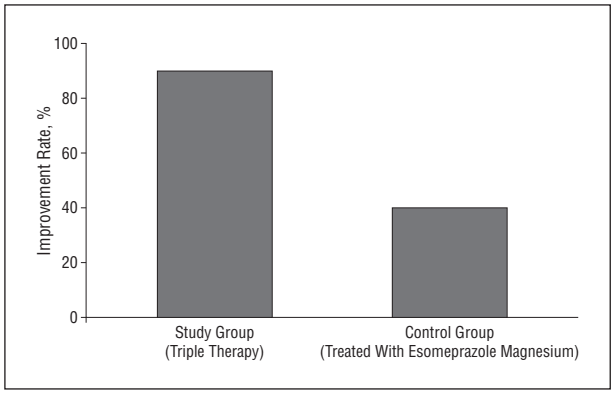

Figure. The clinical improvement in both controls and study patients with laryngopharyngeal reflux disease is seen. Triple therapy comprised esomeprazole plus amoxicillin sodium and clarithromycin. In both groups, treatment was daily for 4 weeks.

RESULTS

The mean age of the 212 patients in the study was 32.4 years. Cough, the main LPRD symptom, was found in 105 patients (49%), followed by globus sensation in 98 patients (46%), frequent throat clearing in 78 patients (36%), a bad or bitter taste in the mouth in 69 patients (32%), and hoarseness in 54 (25%).

Red, irritated arytenoids was the main laryngoscopic finding in 116 patients (54%), followed by swelling of the vocal cords in 59 patients (27%), small laryngeal ulcers in 24 patients (11%), and laryngeal granulomas in 4 patients (2%).

Among the patients in the study, the HPSA test results were positive in 57% of cases, and we found them to be statistically nonsignificant in relation to patient symptoms (**Table 1**) ($P<.05$ was considered statistically significant). Also, they were found to be nonsignificant in relation to the laryngoscopic findings (**Table 2**).

Marked improvement in symptoms occurred in 87 of the 90 patients with negative HPSA test results who received once-daily esomeprazole magnesium, 40 mg, for 4 weeks.[14]

The 122 patients with positive HPSA test results were randomized into 2 equal groups (61 patients each). The control group (61 patients) received only esomeprazole magnesium, 40 mg, for 4 weeks; 23 patients (40%) showed marked improvement in symptoms, partial improvement occurred in 9 patients (16%), while 25 patients (44%) reported no improvement. Four patients discontinued follow-up.

The second study group (61 patients) received triple therapy comprising esomeprazole magnesium, 40 mg, plus amoxicillin sodium, 1 g, and clarithromycin, 500 mg,[14] for the same period. Two patients discontinued follow-up. Fifty-three patients (90%) showed marked improvement in symptoms, partial improvement occurred in 3 patients (5%), and 3 patients (5%) showed no improvement (**Figure**).

COMMENT

In the practice of otolaryngology, it is now common to encounter patients with LPRD symptoms. Most of these pa-

©2010 American Medical Association. All rights reserved.

tients have been seen in thoracic and gastroenterology departments with atypical GERD symptoms. Laryngopharyngeal reflux disease is a diagnostic dilemma given the lack of solid guidelines for diagnosis and management.

In a recent report, Barry and Vaezi state,[15] "more questions than answers" were given, which best describes the current state of knowledge of LPRD. Our current study contributed several more questions.

In 1 limb of the study, a trial was made to associate HP infection with the degree or severity of symptoms and laryngoscopic findings. It was shown clearly based on statistical analysis that HP has no relation with any of the symptoms or signs of HPSA-positive or HP-negative individuals.

The second limb of the study compared the efficacy of proton pump inhibitor monotherapy vs triple therapy, and we have shown in our results that triple therapy gave better results in patients with positive HPSA test results. This study presents as much raw data as possible in compliance with the most recent guidelines to enable future evidence-based meta-analysis.

Gastroesophageal reflux disease is a common acid-related disorder presenting with a broad spectrum of symptoms with or without complications.[3] The incidence of laryngopharyngeal symptoms is greater than expected.[5] There are more complex multifactorial pathophysiologic characteristics of LPRD than simply acid reflux.[6] Laryngopharyngeal reflux disease is considered to be a variant of GERD in which the incidence of throat and laryngeal symptoms is more evident and encountered in practice more often than expected.[16]

A large number of studies have raised the issue of the role of HP infection and its role in the pathophysiologic mechanism of GERD, but the interest in its role in LPRD has not been adequately studied.[17] An estimated prevalence rate of HP infection of 30% among the general population has been given and shows that it is quite common.[18] Various theories and mechanisms have been proposed to clarify its role in GERD.

In our study, 212 patients with symptoms of LPRD and positive results from 24 hours of pH monitoring were evaluated clinically. The most common symptoms were dry, persistent cough (49%) followed by a globus sensation (46%); other studies have also reported a globus sensation or throat-clearing, voice change, persistent sore throat, dysphagia, and cough as the predominant symptoms.[19-21]

The common reported findings of LPRD are in the domain of posterior laryngitis; we reported red, irritated arytenoids in 54% and swollen vocal folds (27%); other reports[20,21] found endoscopic abnormalities in up to 98% of patients with LPRD, including nonspecific hyperemia, usually of the posterior larynx.

In our study, the 57% incidence rate of positive HPSA test scores is higher than that reported by Haruma et al,[21] who mentioned that in Japan there is a relationship between HP infection and LPRD with a reported incidence of 31% to 41%. *Helicobacter pylori* stool antigen testing is a relatively new, noninvasive diagnostic technique with high sensitivity and specificity[11,22]

Several authors suggested a correlation of HP infection and the degree of GERD,[9,10,19,23,24] while others[25] did not find any association between HP positivity and symptoms; the latter is in agreement with our data, which failed to demonstrate such a connection, and this variable report adds more to the dilemma of diagnosing LPRD.

As mentioned in the introductory paragraphs, we did not aim to point to a specific treatment regimen, a task better left for meta-analysis trials, but our raw data showed that patients with LPRD and with negative HPSA test results benefit from esomeprazole magnesium, 40 mg, for 4 weeks, with marked symptom improvement in most cases. While the patients with positive HPSA test results who received only esomeprazole magnesium, 40 mg, for 4 weeks showed a 40% rate of improvement, the second study group of patients with positive HPSA test results receiving triple therapy showed a 90% rate of improvement. Reports of a more successful triple therapy in GERD[26] are in agreement with our results, but still, no clear guidelines for treatment of LPRD are available.

In conclusion, the incidence of the HP infection in patients with LPRD in our study is 57%. Second, HP infection should be considered when treatment is prescribed to patients with LPRD because the standard therapy for GERD might be insufficient. Finally, the use of triple therapy (esomeprazole magnesium, 40 mg, plus amoxicillin sodium, 1 g, and clarithromycin, 500 mg) in the treatment of LPRD with HP infection might result in a higher cure rate.

Submitted for Publication: March 9, 2010; final revision received June 20, 2010; accepted July 22, 2010.
Published Online: September 20, 2010. doi:10.1001/archoto.2010.165
Correspondence: Mohamed Rifaat Ahmed, MD, Department of Otolaryngology–Head and Neck Surgery, Faculty of Medicine, Suez Canal University, Ismalia, Egypt (m_rifaat@hotmail.com).
Author Contributions: Both authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Analysis and interpretation of data:* Youssef and Ahmed. *Critical revision of the manuscript for important intellectual content:* Youssef and Ahmed. *Statistical analysis:* Youssef and Ahmed. *Obtained funding:* Ahmed. *Administrative, technical, and material support:* Ahmed.
Financial Disclosure: None reported.

REFERENCES

1. Koufman JA. Gastroesophageal reflux disease. *Otolaryngol Head Neck Surg*. 1993; 81(9)(suppl 2):2351.
2. Beaver ME, Stasney CR, Weitzel E, et al. Diagnosis of laryngopharyngeal reflux disease with digital imaging. *Otolaryngol Head Neck Surg*. 2003;128(1):103-108.
3. Bough ID Jr, Sataloff RT, Castell DO, Hills JR, Gideon RM, Spiegel JR. Gastroesophageal reflux laryngitis resistant to omeprazole therapy. *J Voice*. 1995; 9(2):205-211.
4. Hanson DG, Kamel PL, Kahrilas PJ. Outcomes of antireflux therapy for the treatment of chronic laryngitis. *Ann Otol Rhinol Laryngol*. 1995;104(7):550-555.
5. Dinis PB, Subtil J. *Helicobacter pylori* and laryngopharyngeal reflux in chronic rhinosinusitis. *Otolarygol Head Neck Surg*. 2006;134(1):67-72.
6. Reimer C, Bytzer P. Management of laryngopharyngeal reflux with proton pump inhibitors. *Ther Clin Risk Manag*. 2008;4(1):225-233.
7. Kim JG. Treatment of *Helicobacter pylori* infection [in Korean]. *Korean J Gastroenterol*. 2005;46(3):172-180.

©2010 American Medical Association. All rights reserved.

8. Yamaoka Y, ed. *Helicobacter pylori: Molecular Genetics and Cellular Biology*. Norwich, England: Caister Academic Press; 2008.
9. Oridate N, Takeda H, Yamamoto J, et al. *Helicobacter pylori* seropositivity predicts outcomes of acid suppression therapy for laryngopharyngeal reflux symptoms. *Laryngoscope*. 2006;116(4):547-553.
10. Tezer MS, Kockar MC, Koçkar O, Celik A. Laryngopharyngeal reflux finding scores correlate with gastroesophageal reflux disease and *Helicobacter pylori* expression. *Acta Otolaryngol*. 2006;126(9):958-961.
11. Gulcan EM, Varol A, Kutlu T, et al. *Helicobacter pylori* stool antigen test. *Indian J Pediatr*. 2005;72(8):675-678.
12. Oderda G, Rapa A, Ronchi B, et al. Detection of *Helicobacter pylori* in stool specimens by non-invasive antigen enzyme immunoassay in children: multicentre Italian study. *BMJ*. 2000;320(7231):347-348.
13. Scott LJ, Dunn CJ, Mallarkey G, Sharpe M. Esomeprazole: a review of its use in the management of acid-related disorders. *Drugs*. 2002;62(10):1503-1538.
14. Ni YH, Lin JT, Huang SF, Yang JC, Chang MH. Accurate diagnosis of *Helicobacter pylori* infection by stool antigen test and 6 other currently available tests in children. *J Pediatr*. 2000;136(6):823-827.
15. Barry DW, Vaezi MF. Laryngopharyngeal reflux: more questions than answers. *Cleve Clin J Med*. 2010;77(5):327-334.
16. Thomson ABR, Barkun AN, Armstrong D, et al. The prevalence of clinically significant endoscopic findings in primary care patients with uninvestigated dyspepsia: the Canadian Adult Dyspepsia Empiric Treatment–Prompt Endoscopy (CADET-PE) study. *Aliment Pharmacol Ther*. 2003;17(12):1481-1491.
17. Issing WJ, Karkos PD, Perreas K, Folwaczny C, Reichel O. Dual-probe 24-hour ambulatory pH monitoring for diagnosis of laryngopharyngeal reflux. *J Laryngol Otol*. 2004;118(11):845-848.
18. Ahmad I, Batch AJ. Acid reflux management: ENT perspective. *J Laryngol Otol*. 2004;118(1):25-30.
19. Pinar E, Oncel IS, Calli C, Atalay M. Detection of gastroesophageal reflux by scintigraphy in patients with laryngopharyngeal symptoms and findings [in Turkish]. *Kulak Burun Bogaz Ihtis Derg*. 2003;10(4):153-158.
20. Yorulmaz I, Ozlugedik S, Kucuk B. Gastroesophageal reflux disease: symptoms versus pH monitoring results. *Otolaryngol Head Neck Surg*. 2003;129(5):582-586.
21. Haruma K, Hamada H, Mihara M, et al. Negative association between *H. pylori* infection and reflux esophagitis in older patients: case-control study in Japan. *Helicobacter*. 2000;5(1):24-29.
22. Rouev P, Chakarski I, Doskov D, Dimov G, Staykova E. Laryngopharyngeal symptoms and gastroesophageal reflux disease. *J Voice*. 2005;19(3):476-480.
23. Tauber S, Gross M, Issing WJ. Association of laryngopharyngeal symptoms with gastroesophageal reflux disease. *Laryngoscope*. 2002;112(5):879-886.
24. Ercan I, Cakir BO, Uzel TS, Sakiz D, Karaca C, Turgut S. The role of gastric Helicobacter pylori infection in laryngopharyngeal reflux disease. *Otolaryngol Head Neck Surg*. 2006;135(1):52-55.
25. Postma GN, Johnson LF, Koufman JA. Treatment of laryngopharyngeal reflux. *Ear Nose Throat J*. 2002;81(9)(suppl 2):24-26.
26. Ching SS, Sabanathan S, Jenkinson LR. Treatment of Helicobacter pylori in surgical practice: a randomised trial of triple versus quadruple therapy in a rural district general hospital. *World J Gastroenterol*. 2008;14(24):3855-3860.

Reprinted by permission of Laryngoscope. 2015; 125(5):1093-1097.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Does Body Mass Index Predict Tracheal Airway Size?

Brian D'Anza, MD; Jesse Knight, MD; J. Scott Greene, MD, FACS

Objectives/Hypothesis: To determine the relationship between body mass index along with other anthropomorphic variables as they relate to tracheal airway dimensions.

Study Design: Retrospective case series.

Methods: This was a radiographic study of 123 consecutive hospitalized patients undergoing tracheotomy over a 4-year period (2007–2011). We measured airway dimensions in axial computed tomography imaging and made comparisons with height, weight, body mass index, gender, and age. Measurements were taken at the first tracheal ring level including anterior-posterior length, width, and calculated area. We expected higher body mass index not to be a good predictor of larger airway dimensions.

Results: The linear regression model showed body mass index was significantly inversely related to tracheal width after controlling for gender and age ($P = .0389$). For every 1 kg/m^2 increase in body mass index, the tracheal width decreased by 0.05 mm. There was a trend for airway area to diminish with increasing body mass index.

Conclusions: These results are consistent with the hypothesis that obese patients do not have larger airways. Our study indicated a trend toward smaller airways as body mass index increased. Specifically, as body mass index increases, tracheal width appears to decrease. This information should help medical professionals avoid the tendency to use a larger tube to secure the airway of an obese patient. Hopefully, this will result in further research into the field and may prevent future airway injuries in a society where obesity has become epidemic.

Key Words: Tracheal airway size, endotracheal tube size, obesity.

Level of Evidence: 4

***Laryngoscope*, 125:1093–1097, 2015**

INTRODUCTION

General guidelines exist for endotracheal tube (ETT) size selection in adults.[1,2] Variations in the choice of tube selection are influenced by factors such as patient age, gender, and body habitus. Generally speaking, a larger diameter tube is used for adult males and those with larger body habitus when compared to females or smaller individuals. Tracheostomy tube selection follows similar decision making algorithms in the adult population.[3] Due to normal anatomical variability, it is difficult to standardize recommendations for endotracheal and tracheostomy tube sizes for adults.[4] On a case-by-case basis, considerations for choosing a tube size might include those mentioned previously. Body dimensions that could be used to predict airway sizes include body mass index (BMI). A literature review shows the lack of a study comparing BMI with airway dimensions.

Larger-than-necessary ETTs are known to cause laryngeal or tracheal trauma and are to be avoided. Common problems can include laryngeal webs, vocal cord ulcerations, vocal cord paralysis or paresis, subglottic stenosis, and tracheal stenosis among others.[5,6] However, our experience has shown the tendency of emergency medical providers and critical care physicians to place a larger tube in a larger patient. Based on our observations at time of tracheotomy, we have found that it is common to find a smaller than anticipated trachea in an obese patient. Based on this experience, we hypothesize that higher BMI is not a good predictor of larger airway size. The purpose of our study was to evaluate the airway dimensions and identify anatomical concerns for the use of relatively oversized ETTs in an obese population. We measured airway dimensions in axial computed tomography (CT) imaging of 123 patients who underwent tracheotomy and made comparisons with their height, weight, BMI, gender, and age.

MATERIALS AND METHODS

The Geisinger Medical Center Office of Research Compliance and Institutional Review Board approved the retrospective electronic chart case series titled Predictors of Airway Size. The study was performed reviewing the information on 123 patients who underwent tracheotomy surgery by the otolaryngology department over a 4-year period (2007–2011) and who also had CT imaging of the trachea within the previous 3 months. Anthropomorphic measurements were taken from the time of

From the Department of Otolaryngology, Head and Neck Surgery, Facial Plastic Surgery (B.D., J.S.G), Geisinger Medical Center, Danville, Pennsylvania; and the ENT Associates of Los Alamos, PLLC (J.K.), Los Alamos, New Mexico, U.S.A.

Editor's Note: This Manuscript was accepted for publication September 3, 2014.

Presented at the 2013 American Academy of Otolaryngology–Head and Neck Surgery National Meeting, Vancouver, British Columbia, Canada, September 30, 2013.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Brian D'Anza, MD, Department of Otolaryngology, Head and Neck Surgery, Facial Plastic Surgery, Geisinger Medical Center, 100 N. Academy Ave., Danville, PA 17822. E-mail: bddanza@geisinger.edu

DOI: 10.1002/lary.24943

TABLE I.
Descriptive Statistics for the Entire Sample Stratified by Gender.

Variable	Entire Sample, N = 123	Female, n = 47	Male, n = 76	*P* Value
Age, yr	63.1 (15.3)	65.1 (14.2)	61.9 (15.9)	.2636
BMI, kg/m^2	32.4 (11.8)	33.5 (12.5)	31.7 (11.3)	.3919
CT AP, mm	20.8 (4.4)	18.8 (4.4)	22.1 (4.0)	<.0001
CT width, mm	17.2 (3.5)	15.6 (2.8)	18.3 (3.6)	.7164
Airway area, mm^2	294.8 (112.8)	240.0 (87.4)	328.8 (113.7)	<.0001

Values for age, BMI, CT AP, CT width, and airway area reflect the average (mean) calculation. Standard deviation is noted in parentheses

AP = anterior-posterior diameter; BMI = body mass index; CT = computed tomography.

tracheotomy in 76 males and 47 females to include height, weight, calculated BMI, age, and gender. Two hundred eighty patients were noted to have undergone tracheotomy over a 4-year period, but 157 patients were excluded due to malignancy involving the tracheal airway or with suboptimal imaging precluding accurate assessment of the tracheal dimensions. This excluded any patient who required placement of an ETT at the time of the CT study, or any other intervention such as a nasogastric tube that might impinge on the membranous tracheal wall dimensions. None of the patients included in this study had a prior tracheostomy or airway procedure. All of the included patients were spontaneously breathing and in a supine position at the time of the CT study. In all cases, the reason for ETT and tracheostomy placement was ventilator-dependent respiratory failure.

We noted the size of the ETT in place at the time of the tracheotomy procedure and documented this for each patient. Axial CT measurements were analyzed at the level of the first tracheal ring, as this was the area most commonly associated with the presence of a balloon cuff and thus the most likely area for tracheal injury.[7–9] We did not measure the level of the cricoid due to the circumferential cartilage ring structure at this level, making the airway less likely to be compressed or narrowed.

Measurements included the anterior-posterior diameter (CT AP) as well as the width (CT width) of the trachea at this point. Cross-sectional airway areas were also calculated and compared. Data are described using means and standard deviation for continuous variables and frequency and percentages for categorical variables. Data are described for the full sample and stratified by gender. Comparison across gender was accomplished using the two-sample t test and Pearson χ^2 tests, as appropriate. Pearson correlation was estimated between the continuous variables. Finally, linear regression was used to identify predictors of airway size. BMI classification was taken from current World Health Organization standards.[10]

TABLE II.
ETT Sizes for the Entire Sample.

ETT Size, mm	Entire Sample	Female	Male
6.0	12 (9.8%)	8 (17.0%)	4 (5.3%)
6.5	3 (2.4%)	2 (4.3%)	1 (1.3%)
7.0	12 (9.8%)	8 (17.0%)	4 (5.3%)
7.5	42 (34.2%)	21 (44.7%)	21 (27.6%)
8.0	52 (42.3%)	8 (17.0%)	44 (57.9%)
8.5	1 (0.8%)	0 (0%)	1 (1.3%)
9.0	1 (0.8%)	0 (0%)	1 (1.3%)

ETT sizes are all standardized numbers based on internal diameter in millimeters. Numbers in columns designate number of patients. Percentages designate the percentage of patients who had the specified size of tube placed.

ETT = endotracheal tube.

TABLE III.
BMI Classification and Mean ETT Sizes.

			Mean ETT Sizes, mm		
BMI, kg/m^2	Males	Females	Males	Females	All
Underweight	4	1	7.5	N/A	7.5
Normal	18	9	7.69	6.88	7.42
Overweight	15	10	7.70	6.95	7.40
Obese	39	27	7.83	7.39	7.65
	Range 14–70	Range 17–75			

The number in each table row for the columns titled Males and Females reflects the number of patients in each category.

Underweight: BMI = <18.50, Normal: BMI 18.50–24.99, Overweight: BMI 25–29.99, Obese: BMI ≥30. Mean ETT sizes are all based on internal diameter in millimeters.

BMI = body mass index; ETT = endotracheal tube.

RESULTS

Descriptive statistics for the entire sample size including age, BMI, CT AP, and CT width are shown in Table I, and the ETT sizes placed at the time of tracheostomy are shown in Table II. These statistics are stratified by gender. The BMI range for males was 14 to 70 and was 17 to 75 for females. Using these outlier patients as examples, the male with a BMI of 14 had an anterior-posterior tracheal airway measurement of 26 mm and a width of 21 mm. The male with a BMI of 70 had an anterior-posterior tracheal airway measurement of 24 mm and a width of 20 mm for comparison. The female with the BMI of 17 had an anterior-posterior tracheal airway measurement of 16 mm and a width of 15 mm. The female with the BMI of 75 had an anterior-posterior tracheal airway measurement of 14 mm and a width of 10 mm.

We have also included the mean ETT sizes calculated for the entire sample stratified by BMI classification. We found that for both obese men and women, the higher the BMI, the higher the average ETT size. Specifically, compared with normal sized patients, men and women who were obese had on average 0.14 and 0.51 mm-larger ETTs placed, respectively. This is demonstrated in Table III.

Most importantly, findings from the linear regression models yielded significant associations between BMI and tracheal airway dimensions. Specifically, BMI was inversely related to tracheal width after controlling for gender and age ($P = .0389$). For every 1 kg/m^2 increase in BMI, the tracheal width decreased by 0.05 mm. In addition, although not statistically significant, on linear regression analysis there was a trend for the airway area to decrease as the BMI increased.

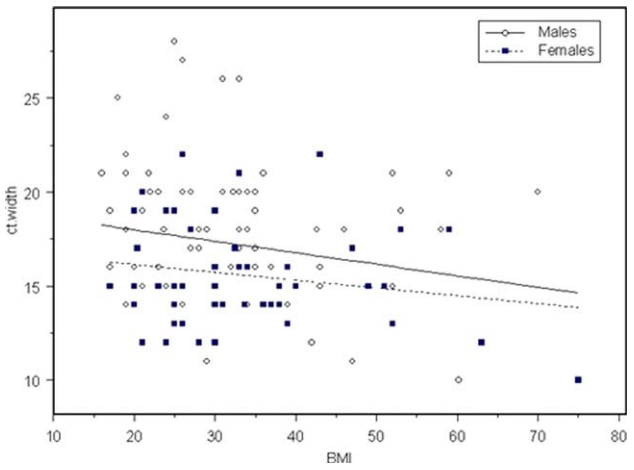

Fig. 1. Airway computed tomography (CT) width (mm) versus body mass index (BMI). The CT width on the y-axis is measured in millimeters. BMI on the x-axis is measured using weight in kilograms and height in meters and is calculated by dividing the subject's weight by the square of his/her height (kg/m^2). The white circle represents one individual male subject. The solid square represents one individual female subject. The solid line represents the linear regression analysis for all male subjects. The dotted line represents the linear regression analysis for all female subjects. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

Figures 1 and 2 provide graphical representations of the linear regression analyses.

Linear regression models also yielded significant associations for gender, with males having increased size on both tracheal CT AP diameter and width measurements on CT imaging compared to females. After controlling for BMI and age, there was an increase in CT AP diameter and CT width for males of 3.34 and 2.59 mm, respectively (Table IV).

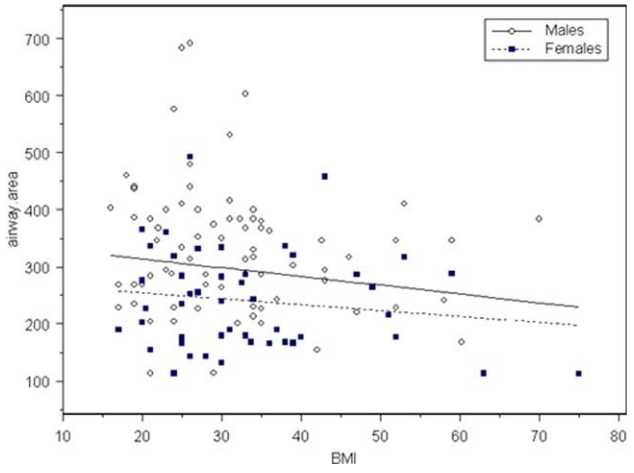

Fig. 2. Airway area (mm^2) versus body mass index (BMI). The airway area on the y-axis is measured in mm^2. BMI on the x-axis is measured using weight in kilograms and height in meters and is calculated by dividing the subject's weight by the square of his/her height (kg/m^2). The white circle represents one individual male subject. The solid square represents one individual female subject. The solid line represents the linear regression analysis for all male subjects. The dotted line represents the linear regression analysis for all female subjects. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

TABLE IV.
Linear Regression Results.

	CT AP, mm		CT Width, mm	
Variable	Estimate (SE)	*P* Value	Estimate (SE)	*P* Value
Male	3.34 (0.78)	<.0001	2.59 (0.61)	<.0001
BMI	−0.02 (0.03)	.5040	−0.05 (0.025)	.0389
Age	0.03 (0.02)	.2208	−0.002 (0.019)	.9359

Estimate is the score on change in millimeters for a given variable based on the linear regression best fit line from the scatter plot.
AP = anterior-posterior diameter; BMI = body mass index; SE = standard error; CT = computed tomography.

We compared each anthropomorphic variable for the strongest effect by using standardized regression coefficients. These were estimated from a linear regression model after the risk factors had been rescaled. Among males, the strongest predictor of CT AP diameter and airway area was height, and for CT width it was BMI. Among females, the strongest predictor of CT AP diameter and airway area was BMI, and for tracheal width it was height. However, none of the variables was significant. Therefore, no further modeling was performed.

Among both males and females, airway area correlated directly with height. Those findings are shown in Figure 3.

DISCUSSION

Our results support BMI to be inversely related to tracheal width on CT imaging. As shown in Figure 1, for every 1 kg/m^2 increase in BMI, the CT width decreased by 0.05 mm ($P = .0389$). All anthropomorphic measurements showed a trend for decreased airway dimensions (area, AP diameter, width) with increasing BMI based on linear regression of scatter plots. This would suggest consistency with the trend we have clinically observed in smaller tracheal airway sizes in patients noted to have much larger body habitus. We speculate that this could be a secondary effect from numerous factors, one being increased pressure on the trachea due to increased adiposity in these patients.

Animal studies have revealed a relationship between the natural caudal traction of the trachea by the thoracic contents and airway patency. Prior research has shown there is an influence of thoracic volumes on upper airway obstruction and compression.[11–13] Studies have shown that obese patients have problems with lower and upper airway compression due to increased weight and adiposity.[14,15] Specifically, abdominal obesity is suggested to negatively influence upper airway function during sleep.[14] It is believed that increased abdominal adiposity causes diaphragmatic compression of intrathoracic contents, which results in their cephalic deviation. As a result, the natural caudal traction of intrathoracic contents via the trachea is reduced and thus increases the distensibility of the airway. This is

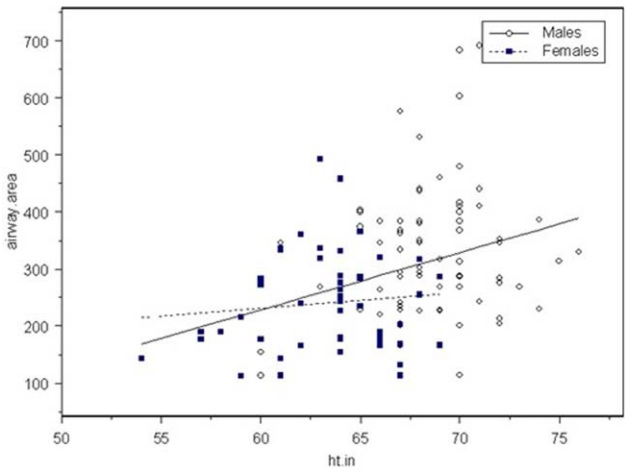

Fig. 3. Airway area (mm^2) versus height (inches). Airway area on the y-axis is measured in $mm.^2$ Height (ht) in on x-axis is measured in inches. The white circle represents one individual male subject. The solid square represents one individual female subject. The solid line represents the linear regression analysis for all male subjects. The dotted line represents the linear regression analysis for all female subjects. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

thought to compromise upper airway patency, especially when obese patients are supine.

Obese subjects have been shown to have increased positive end expiratory pressures due to extrinsic compression of surrounding tissues.[15] Based on the current study results, these relationships may also extend to the tracheal airway. The cartilage framework is more resistant to external thoracic pressures than the soft tissues of the lungs, oropharynx, or hypopharynx; however, the lack of caudal traction may be playing a role in the increased collapsibility and resultant decreased caliber of the airway. This may be especially pronounced as it relates to the membranous portion of the trachea as evidenced in the significantly decreased airway width of the obese study subjects. Other factors, such as genetic differences in cartilage composition, strength of surrounding musculature, inherent conditions of the trachea, such as tracheomalacia, could all be contributing. Further studies are needed to better examine these relationships in the trachea.

It is worth mentioning that when compared to females, males were found to have larger tracheal dimensions at the first tracheal ring. This is consistent with previous reports.[3]

As demonstrated in Table I, the study population had average tracheal airway dimensions that were slightly smaller than published normative values.[16,17] A large radiographic study performed by Breatnach et al. showed that the average tracheal airway width and AP diameter were 25 and 27 mm in men and 21 and 23 mm in women, respectively. These are generally accepted to be average airway dimensions for a presumably normal-sized population. Conversely, our patient sample had a much higher proportion of obese patients compared to the general population, comprising well over half the cohort. According to the Centers for Disease Control, the percentage of the US population that was obese was 35.5% and 35.8% for men and women, respectively.[18] This would seem to further corroborate the evidence that there may be an inverse correlation between obesity and airway size.

A patient with a higher body mass index might be presumed to have a larger tracheal airway. Anecdotally, we have found it is not uncommon to see an inappropriately sized ETT chosen for use in such a patient. Our data corroborate this contention, as obese patients were found on average to have a larger ETT in place, as shown in Table III. More importantly, the results of our study suggest that further research is needed in this area, as it is incorrect to assume a larger patient deserves a larger tube based on body size.

There are several limitations to our data. First, our study represents a retrospective case series of a limited population without randomization. Other limitations of this study include the inability to account for any chronic airway disease, such as tracheomalacia, which could lead to dynamic airway changes. Patients with acquired tracheomalacia can have more dynamic changes of the trachea, which has been noted on CT imaging studies previously.[19] Even cross-sectional changes during breathing and coughing as a result of changes in head and neck position and intrathoracic pressure can be noted.[4] One could make the argument that a larger ETT is required for obese patients to counteract the compressive forces on the airway. This must be balanced against the concern for further airway trauma and resultant complications such as stenosis, webbing, and ulcerations. The study population is also nearly exclusively from the intensive care unit, as the overwhelming majority of the tracheostomies performed at our institution are due to ventilator-dependent respiratory failure. This may bias the results as it relates to the type and size of tubes chosen. However, it is precisely this population of severely ill patients who require long-term intubation and who we are concerned about developing complications of tracheal trauma.

CONCLUSION

The population in this study demonstrated significant decreases in airway size with increasing BMI. Obese patients demonstrated radiographic evidence of a significant decrease in width and area of the airway. There was a trend for larger ETTs being placed in patients with a higher BMI. These findings have implications for airway management in obese populations. This study does not confirm the need to place a smaller tube in obese patients, but does suggest that due to smaller tracheal airway sizes, there may be a higher risk of airway injury in obese populations. Larger tubes are not anatomically indicated simply because of a greater BMI. More research in this population is needed to address airway management of obese patients. Currently, there is a lack of literature addressing this topic. This becomes more important, as population rates of obesity have become epidemic.

Acknowledgments
The authors thank H. Lester Kirchner, PhD, for his invaluable contributions to the data analysis.

BIBLIOGRAPHY

1. Allen MS. Surgical anatomy of the trachea. *Chest Surg Clin N Am* 1996;6: 627–635.
2. Miller RD. *Miller's Anesthesia*. 7th ed. London, United Kingdom: Churchill Livingstone; 2009: Chapter 50: Airway Management in Adult: 1573-1610.
3. Randestad A, Lindholm CE, Fabian P. Dimensions of the cricoid cartilage and the trachea. *Laryngoscope* 2001;110:1957–1961.
4. Gray H, Williams PL, Warwick R. In: *Gray's Anatomy*. 36th ed. Philadelphia, PA: Saunders; 2000:1246.
5. Deeb ZE, Williams JB, Campbell TE. Early diagnosis and treatment of laryngeal injuries from prolonged intubation in adults. *Otolaryngol Head Neck Surg* 1999;120:25–29.
6. Gaynor EB, Greenberg SB. Untoward sequelae of prolonged intubation. *Laryngoscope* 1985;95:1461–1467.
7. Sarper S, Ayten A, Eser I, Ozbudak O, Demircan A. Tracheal stenosis after tracheostomy or intubation. *Tex Heart Inst J* 2005;32:154–158.
8. Stauffer JL, Olson PE, Petty TL. Complications and consequences of endotracheal intubation and tracheostomy. A prospective study of 150 critically ill adult patients. *A J Med* 1981;70:65–76.
9. Pearson FG, Andrews MJ. Detection and management of tracheal stenosis following cuffed tube tracheostomy. *Annals Thoracic Surg* 1971:12:359–374.
10. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. WHO Technical Report Series 894. Geneva, Switzerland: World Health Organization; 2000.
11. Van de Graaff WB. Thoracic influence on upper airway patency. *J Appl Physiol* 1988;65:2124–2131.
12. Series F, Marc I. Influence of lung volume dependence of upper airway resistance during continuous negative airway pressure. *J Appl Physiol* 1994;77:840–844.
13. Series F, Cormier Y, Desmeules M. Influence of passive changes of lung volume on upper airways. *J Appl Physiol* 1990;68:2159–2164.
14. Pankow W, Podszus T, Gutheil T, Penzel T, Peter JH, Von Wichert P. Expiratory flow limitation and intrinsic positive end expiratory pressure in obesity. *J Appl Physiol* 1998:85:1236–1243.
15. Ferretti A, Giampiccolo P, Cavalli A, Milic-Emili J, Tantucci C. Expiratory flow limitation and orthopnea in massively obese subjects. *Chest* 2001; 119:1401–1408.
16. Brodsky JB, Macario A, Mark JBD. Tracheal diameter predicts double-lumen tube size: a method for selecting left double lumen tubes. *Anesthesia Analg* 1996;82:861–864.
17. Breatnach E, Abbott GC, Fraser RG. Dimensions of the normal human trachea. *Am J Roentgenol* 1984;142:903–906.
18. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of obesity in the United States, 2009–2010. NCHS Data Brief. 2012;(82):1–8.
19. Aquino S, Shepard J, Ginns L, Moore R, Halpern E, Grillo H. Acquired tracheomalacia: detection by expiratory CT Scan. *J Comput Assist Tomogr* 2001;25:3:394–399.

Reprinted by permission of Laryngoscope. 2015; 125(5):1137-1143.

The Laryngoscope
© 2014 The American Laryngological, Rhinological and Otological Society, Inc.

Causes and Consequences of Adult Laryngotracheal Stenosis

Alexander Gelbard, MD; David O. Francis, MD, MS; Vlad C. Sandulache, MD, PhD; John C. Simmons, MD; Donald T. Donovan, MD; Julina Ongkasuwan, MD

Objectives/Hypothesis: Laryngotracheal stenosis (LTS) is largely considered a structural entity, defined on anatomic terms (i.e., percent stenosis, distance from vocal folds, overall length). This has significant implications for identifying at-risk populations, devising systems-based preventive strategies, and promoting patient-centered treatment. The present study was undertaken to test the hypothesis that LTS is heterogeneous with regard to etiology, natural history, and clinical outcome.

Study Design: Retrospective cohort study of consecutive adult tracheal stenosis patients from 1998 to 2013.

Methods: Subjects diagnosed with laryngotracheal stenosis (ICD-9: 478.74, 519.19) between January 1, 1998, and January 1, 2013, were identified. Patient characteristics (age, gender, race, follow-up duration) and comorbidities were extracted. Records were reviewed for etiology of stenosis, treatment approach, and surgical dates. Stenosis morphology was derived from intraoperative measurements. The presence of tracheostomy at last follow-up was recorded.

Results: One hundred and fifty patients met inclusion criteria. A total of 54.7% had an iatrogenic etiology, followed by idiopathic (18.5%), autoimmune (18.5%), and traumatic (8%). Tracheostomy dependence differed based on etiology ($P < 0.001$). Significantly more patients with iatrogenic (66%) and autoimmune (54%) etiologies remained tracheostomy-dependent compared to traumatic (33%) or idiopathic (0%) groups. On multivariate regression analysis, each additional point on Charlson Comorbidity Index was associated with a 67% increased odds of tracheostomy dependence (odds ratio 1.67; 95% confidence interval 1.04–2.69; $P = 0.04$).

Conclusions: Laryngotracheal stenosis is not a homogeneous clinical entity. It has multiple distinct etiologies that demonstrate disparate rates of long-term tracheostomy dependence. Understanding the mechanism of injury and contribution of comorbid illnesses is critical to systems-based preventive strategies and patient-centered treatment.

Key Words: Tracheal stenosis, subglottic, laryngotracheal stenosis, intubation, tracheostomy.

Level of Evidence: 4.

***Laryngoscope*, 125:1137–1143, 2015**

INTRODUCTION

Laryngotracheal stenosis (LTS) is a life-threatening, fixed, extrathoracic restriction in pulmonary ventilation. LTS is an umbrella term, encompassing luminal compromise at the level of the larynx, subglottis, or trachea, which exists in a watershed of specialty care. Diagnosis is frequently delayed as patients rapidly transition from acute inpatient care to outpatient facilities. The majority of patients are not definitively diagnosed until outpatient specialty evaluation.[1] Many specialists (e.g., intensivists, otolaryngologists, interventional pulmonologists, thoracic surgeons) initially interact with this population, which makes it difficult to establish the natural history of the disease, define universal predictors of disease outcome, and create cogent personalized plans of care. Additionally, long-term sequelae of intensive respiratory support (endotracheal intubation and elective tracheostomy) do not develop on a timescale necessary for recognition by practitioners providing acute care, impeding quality-driven improvement efforts.[2]

Moreover, LTS is generally described in terms of its structural characteristics, defined in anatomic terms (i.e., percent stenosis, distance from vocal folds, overall length). This neglects the unique biology driving luminal compromise in heterogeneous patient populations and has significant implications for identifying at-risk populations, devising systems-based preventive strategies, and promoting patient-centered treatment directed at the diverse pathophysiology driving airway injury. The present study was undertaken to test the hypothesis that LTS is heterogeneous with regard to etiology, natural history, and clinical outcome.

PATIENTS AND METHODS

This study was performed in accordance with the Declaration of Helsinki, Good Clinical Practice, and was approved by the Baylor College of Medicine Institutional Review Board (IRB No. H33195).

Patients

Subjects diagnosed with laryngotracheal stenosis (ICD-9: 478.74, 519.19) between January 1, 1998, and January 1, 2013, were

From the Department of Otolaryngology, Vanderbilt School of Medicine (A.G., D.O.F.), Nashville, Tennessee; and the Bobby R. Alford Department of Otolaryngology–Head and Neck Surgery Baylor College of Medicine (V.C.S., J.C.S., D.T.D., J.O.), Houston, Texas, U.S.A.

Editor's Note: This Manuscript was accepted for publication September 8, 2014.

All authors have completed and submitted the International Committee of Medical Journal Editors Form for Disclosure of Potential Conflicts of Interest. The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Alexander Gelbard, MD, Assistant Professor, Department of Otolaryngology, Vanderbilt School of Medicine, Medical Center East, S. Tower, 1215 21st Ave. South, Suite 7302, Nashville, TN 37232-8783. E-mail: alexander.gelbard@vanderbilt.edu

DOI: 10.1002/lary.24956

TABLE I.
Definitions of LTS Etiology of Injury Utilized in This Study.

Idiopathic	No history of significant laryngotracheal injury. No significant history of endotracheal intubation or tracheotomy within 2 years of the presentation. No thyroid or major anterior neck surgery. No neck irradiation. No caustic or thermal injuries to the laryngotracheal complex. No history of vasculitis. Negative titers for angiotensin-converting enzyme and antinuclear cytoplasmic antibody. The lesion must involve the subglottis.
Autoimmune	Patients with documented clinical, along with serologic and/or histologic, diagnosis of Wegener's granulomatosis, relapsing polychondritis, systemic lupus erythematous, rheumatoid arthritis, epidermolysis bullosa, sarcoidosis, or amyloidosis
Polytrauma	Patients presenting with laryngotracheal stenosis following documented traumatic injuries involving multiple organ systems
Iatrogenic	Patients who developed subglottic or tracheal stenosis following tracheostomy *or* subglottic or tracheal stenosis developing within 2 years of intubation

identified. Those with a history of tracheal malignancy or isolated laryngeal stenosis were excluded. Laryngeal and tracheal stenosis both share an association with prolonged endotracheal intubation, as well as many of the same comorbid medical risk factors. However, isolated laryngeal stenosis remains a distinct anatomic and structural injury with a unique treatment algorithm that merits dedicated independent study and is not discussed in the present work. Patients meeting inclusion were grouped into four categories based on stenosis etiology: 1) idiopathic, 2) iatrogenic, 3) autoimmune, and 4) polytrauma (Table I).

Data Collected

Patient characteristics (age, gender, race, follow-up duration) and comorbidities were extracted. Records were reviewed for etiology of stenosis, treatment approach (i.e., endoscopic, open), and surgical dates. Stenosis morphology (% luminal obstruction, distance from glottis [cm], and overall length [cm]) and tracheomalacia were derived from intraoperative findings. Patients were staged with the established Cotton-Myer, Lano, and McCaffrey classification systems, as previously described[3–5] (Table II). The number and frequency between repeat procedures were captured.

Procedures

Treatments for tracheal stenosis included: 1) endoscopic dilations of the stenotic trachea,[6] 2) open surgical resection of the diseased tracheal segment with end-to-end anastomosis,[7] and 3) permanent tracheostomy. The treatment algorithm consisted of initial endoscopic dilation for all patients. In patients who required multiple dilation procedures, rigorous selection criteria were applied for consideration of open surgical reconstruction. Patients less than 45 years old, without type 2 diabetes or connective tissue disease, and with stenosis 2 cm or more below the glottis and less than 2 cm in length were offered open surgical reconstruction.

TABLE II.
Definitions of Clinical LTS Classification Systems.

Cotton-Myer	I	<70% obstruction
	II	70%–90% obstruction
	III	>90% obstruction
	IV	Complete obstruction
Lano	I	One subsite* involvement
	II	Two subsite involvement
	III	Three subsite involvement
McCaffrey	I	Subglottis or trachea < 1 cm
	II	Subglottis > 1 cm
	III	Subglottis and trachea >1 cm
	IV	Any lesion involving glottis

*Subsites defined as glottis, subglottis, and trachea.

Outcomes

Presence of a tracheostomy at last follow-up was the primary outcome. This represented failure of surgical management to correct airway narrowing.

Statistical Analysis

All data management and analysis were done using STATA/MP version 12.1 software (STATACorp, College Station, Texas). Univariate analyses were performed using analysis of variance, Pearson's chi-squared tests, and Fisher's exact tests, as appropriate. Stepwise multivariate logistic regression analysis was used to identify independent risk factors for tracheostomy. A significance level of $P < 0.20$ on univariate analysis was used as the criterion for inclusion in the multivariate model. As per convention, $P < 0.05$ was required for statistical significance in the model.

RESULTS

A total of 340 patients with a diagnosis of tracheal or laryngeal stenosis were identified. Excluded were those with tracheal malignancy (N = 9) and isolated bilateral vocal-fold immobility (N = 181). In all, 150 patients met inclusion criteria. The most common etiology was iatrogenic (54.7%), followed by idiopathic (18.5%), autoimmune (18.5%), and traumatic (8%: Table III). Mean follow-up was 39.3 months (95% confidence interval [CI], 31.9–46.6), but varied significantly by etiology ($P < 0.001$; Table III).

Univariate Analysis

Patient Characteristics. Age at presentation differed significantly by strata ($P = 0.002$) with those in the traumatic group being significantly younger than all others (34.4 years, CI 23.5–45.3; Table III). Gender distribution also differed based on etiology ($P < 0.002$; Table III). In order, the idiopathic group had a significantly higher percentage of females (93%) than autoimmune (68%), iatrogenic (62%), or traumatic (33%) LTS patients. Charlson Comorbidity Index (CCI) varied between groups ($P < 0.001$). Iatrogenic and autoimmune strata had significantly higher indices than either idiopathic or traumatic strata (Table III).

TABLE III.
Demographics, Stenosis Characteristics, Comorbidities, and Treatment of LTS Grouped by Etiology of Injury.

Patient Characteristics	Idiopathic (n = 28)	Polytrauma (n = 12)	Autoimmune (n = 28)	Iatrogenic (n = 82)	Significance (*P*)
Demographics					
Follow-up (Mean months, 95% CI)	56.07	12.3	69.1	27.05	<0.001
	(41.5–70.6)	(7.2–17.5)	(39.7–98.6)	(20.9–33.1)	
Age (Mean years, 95% CI)	50.35	35.7	45.1	51	0.002
	(45.9–54.8)	(24.1–47.4)	(39.7–50.4)	(48.0–54.7)	
Sex (% female)	93	33	68	62	0.002
Race (%)					
Caucasian	89	50	71	63	0.330
African American	7	17	14	16	
Asian	0	8	0	2	
Hispanic	4	17	14	17	
Disease Morphology					
% Stenosis (Mean %, 95% CI)	57.86	69.6	68.5	72.8	0.010
	(52.3–63.4)	(55.1– 84.1)	(60.6–76.4)	(68.1–77.6)	
Distance below glottis (Mean cm, 95% CI)	1.289	2.17	1.94	1.77	0.110
	(1.0–1.6)	(1.29–3.05)	(1.38–2.51)	(1.5–2.02)	
Stenosis Length (Mean cm, 95% CI)	1.657	1.95	2.12	2.167	0.440
	(1.3−1.99)	(0.99–2.9)	(1.62–2.62)	(1.91–2.42)	
Comorbidities					
Charlson Index (Mean, 95% CI)	0.07	0.00	1.28	1.32	<0.001
	(0–0.16)	(0)	(0.99 −1.58)	(0.94–1.7)	
DMII (%)	0	0	11	39	<0.001
MI (%)	0	0	3.6	28	<0.001
CHF (%)	0	0	0	13	0.027
CVA (%)	0	0	0	7	0.008
COPD (%)	4	0	7	13	0.390
Connective tissue (%)	0	0	100	0	<0.001
GERD (%)	18	8	21	20	0.859
Treatment					
No. procedures/year (Mean, 95% CI)	1.75	3.41	1.8	2.65	0.490
	(0.8–2.6)	(1.6–5.2)	(0.9–2.7)	(1.7–3.6)	

CHF = congestive heart failure; CI = confidence interval; COPD = chronic obstructive pulmonary disease; CVA = cerebrovascular accident; DMII = diabetes mellitus type 2; GERD = gastroesophageal reflux disease; MI = myocardial infarction.

Examination of the individual components of the CCI showed cardiovascular comorbidities (i.e., myocardial infarction, congestive heart failure, peripheral vascular disease, and cerebrovascular disease) and diabetes mellitus type 2 (DMII) were significantly more prevalent in the iatrogenic strata than in other etiologies (Fig. 1A; Table III). There were not significant differences in the rate of gastroesophageal reflux disease (GERD) between strata (Table III).

Disease Morphology. Degree of stenosis differed between etiologic strata ($P = 0.01$). Idiopathic LTS involved less of the tracheal lumen (mean 57%; CI 52%–63%) than those in the autoimmune or iatrogenic groups (Table III). There were no differences in the mean distance from the glottis ($P = 0.11$) or the length of stenoses between strata ($P = 0.44$). In the iatrogenic group, LTS occurred in the subglottis (1.5 cm from the glottis) in 59% of patients (49/82) (Fig. 1B). Even in those patients presenting with iatrogenic LTS following tracheostomy, 41% (16/39) had subglottic injuries on intraoperative examination.

Treatment. There was no difference in number of surgeries per year of follow-up ($P = 0.49$) or the types of surgeries performed by etiologic strata ($P = 0.14$; Table III). Most patients were treated with tracheal dilation (84%), followed by T-tube placement (8%), resection (6%), and no treatment (2%).

Tracheostomy Dependence. Tracheostomy dependence differed based on etiologic strata ($P < 0.001$; Fig. 1C). Significantly more patients in the iatrogenic (66%) and autoimmune (54%) groups were tracheostomy-dependent at last follow-up compared to those in either the traumatic (33%) or idiopathic (0%) groups. Tracheostomy dependence also differed based on established staging systems (Fig. 2A). When stratified via Cotton-Myer staging (based on the degree of luminal

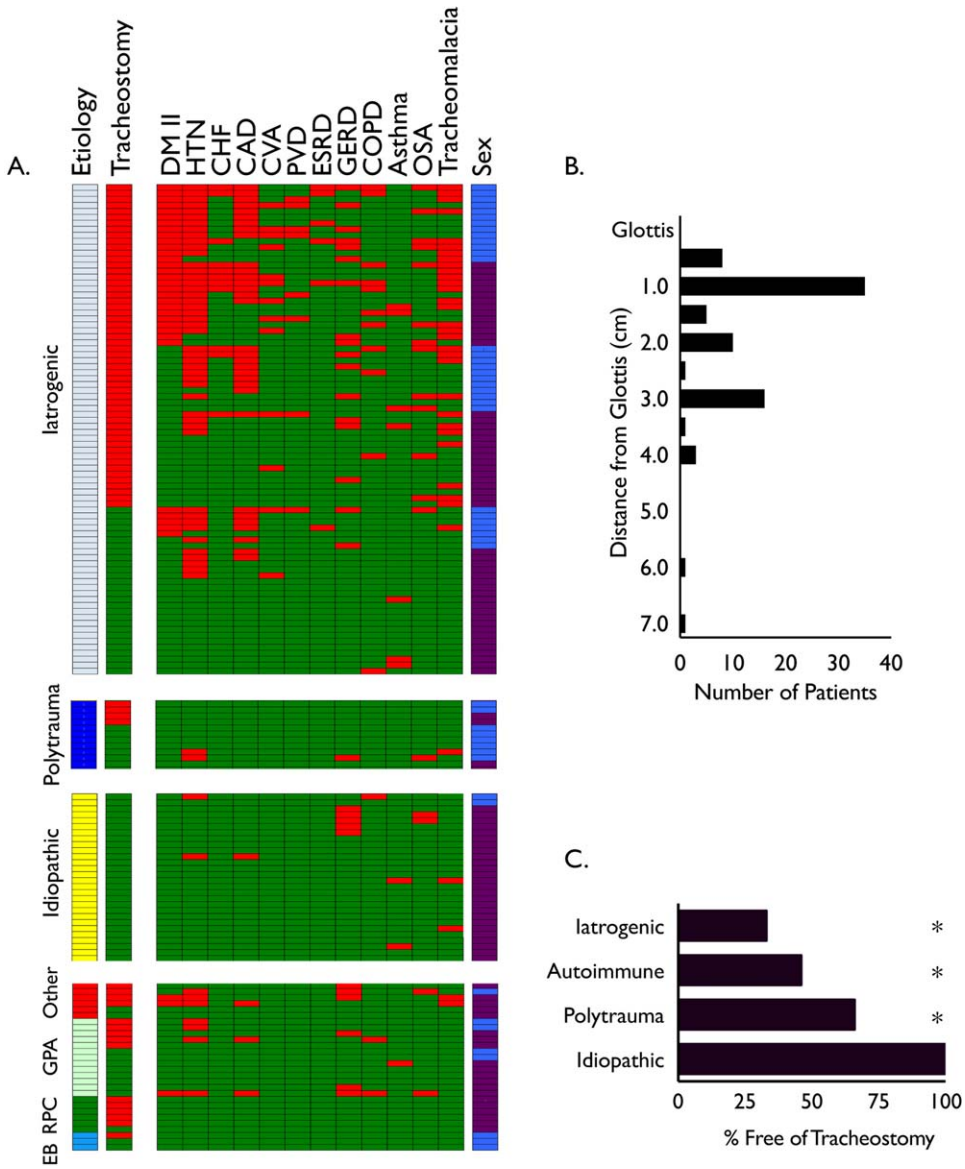

Fig. 1. Heatmap grouped by different etiologies of stenosis. Each line represents an individual patient. Tracheostomy status (red indicating tracheostomy), medical comorbidities (presence highlighted in red), and sex (blue indicating male, purple indicating female). In autoimmune subgroup: GPA (granulomatosis with polyangitis, i.e., Wegener's granulomatosis), RPC (relapsing polychondritis), EB (epidermolysis bullosa) (A). Location of tracheal stenosis in iatrogenic injuries. Histogram showing location of stenotic lesion in iatrogenic subgroup in relation to distance from glottis (B). Tracheostomy status of different etiologies at last follow-up. Asterisk denotes statistical significance from idiopathic group (C).

obstruction), significantly more patients with grade III (90%) and grade IV (90%) lesions were tracheostomy-dependent at last follow-up compared to those in either the grade II (38%) or grade I (32%) groups ($P < 0.001$; Fig. 2A). When stratified by the Lano classification (based on the stenosis location), increasing subsite involvement was significantly associated with a higher rate of tracheostomy ($P < 0.001$; Fig. 2A). When staged according to the McCaffrey classification system (based on both stenosis location and length), increased stage was associated with progressively increased risk of tracheostomy ($P < 0.01$; Fig. 2A).

All three of the established, adult LTS staging systems accurately stratified patients' outcomes based on the severity of their structural injury. Overall (consistent with prior reports), patients in our series with more severe luminal compromise, those with longer stenosis, and those with lesions spanning multiple subsites (glottis, subglottis, and/or trachea) had a much higher incidence of tracheostomy. However, this observation did not hold when patients were stratified by etiology of injury (Table IV.) No patients in the idiopathic group required tracheostomy (even those with lengthy, severe stenosis involving multiple subsites). Conversely, patients with iatrogenic injuries had a significantly higher rate of tracheostomy, even when matched at lower stenosis grades when compared with the other etiologic strata. The non-uniform rate of tracheostomy observed in different

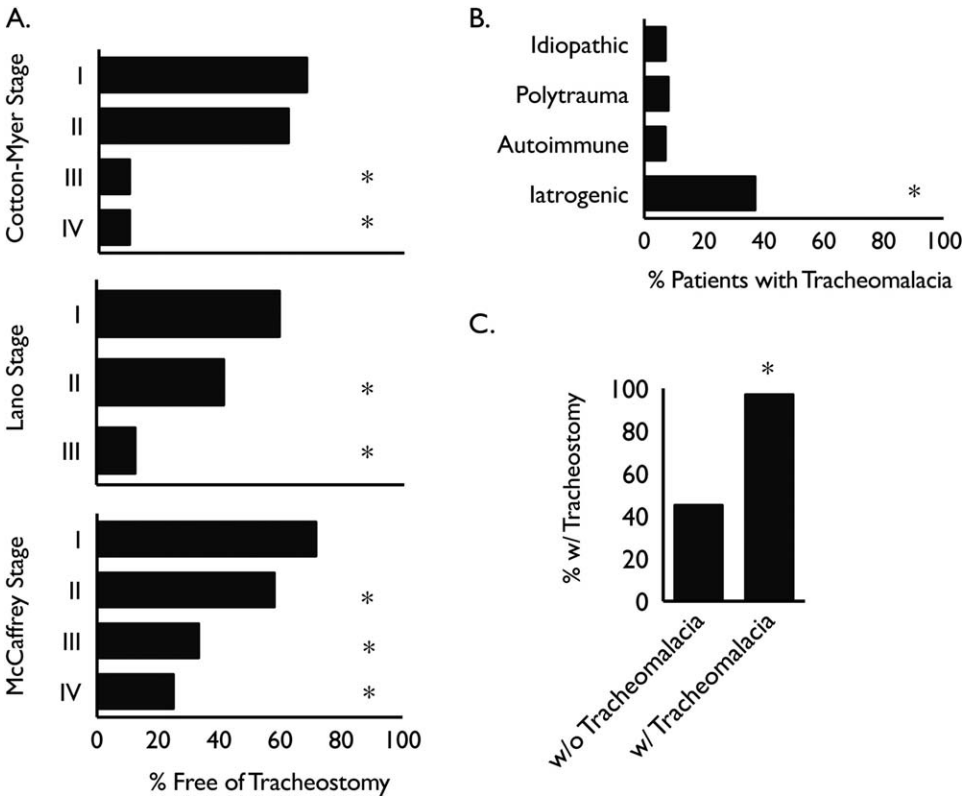

Fig. 2. Tracheostomy status of different Cotton-Myer, Lano, and McCaffrey stages at last follow-up. For Cotton-Myer staging, asterisk denotes statistical significance between grade I and II vs. grade III and IV (A). Diagnosis of tracheomalacia stratified by etiology. Asterisk denotes statistical significance between iatrogenic etiology and all other groups (B). Rate of tracheostomy in iatrogenic etiology patients with and without a diagnosis of tracheomalacia. Asterisk denotes statistical significance (C).

etiologic groups was seen in all three established LTS classification systems (Table IV).

Tracheal Structural Instability. Patients with iatrogenic injuries had a significantly higher rate of tracheomalacia observed on bronchoscopic evaluation (37% vs. 8%; $P < 0.001$; Fig. 2B). Given the retrospective nature of this work, it is not possible to establish a causative relationship between the initial injury and the loss of structural integrity associated with tracheomalacia. However, it is interesting that among the iatrogenic group, 45% of patients without malacia required tracheostomy, whereas 97% of those with malacia necessitated long-term tracheostomy ($P < 0.001$; Fig. 2C).

Multivariate Analysis

Multivariate regression analysis was performed to determine independent predictors of ultimate tracheostomy dependence. Each additional point on CCI was associated with a 67% increased odds of tracheostomy dependence (odds ratio [OR] 1.67; 95% CI 1.04–2.69; $P = 0.04$). Moreover, there was a 3% increased odds of tracheostomy dependence for each additional percentage of airway compromise (OR 1.03, 95% CI 1.01–1.06; $P = 0.001$). LTS patient characteristics (etiology, age, sex, race) were not significantly associated with odds of tracheostomy dependency.

DISCUSSION

Although most airway stenosis appears similar on anatomic imaging and clinical examination, we present data supporting the hypothesis that different mechanisms of injury are associated with differing rates of long-term tracheostomy dependence. The relationships between the anatomic stenosis characteristics (% stenosis, location, and length) and endoscopic or open surgical "success" have been established through pioneering work in children[8,9] and adults.[7] In advanced centers, procedural intervention for LTS offers a high rate of long-term tracheostomy free survival.[4,10,11] However, success in these large published series remains critically dependent on patient selection. With our consecutive series of both inpatient and outpatient consultations, we believe that this study captured a more representative cross-section of symptomatic LTS patients than many prior adult surgical case series. In the "real world," those patients deemed poor operative candidates (e.g., sicker patients) are often left with limited therapeutic options regardless of the structural morphology of their stenosis.

Endotracheal intubation and tracheostomy can be lifesaving but should not be considered benign procedures. They harbor significant long-term risks to communication,[12] swallowing,[13] and breathing,[14] particularly in the subset of patients with comorbid illness.[15] Ironically,

TABLE IV.
Percentage of LTS Patients With Tracheostomy by Cotton-Myer, Lano, and McCaffrey Stage, Grouped by Etiology of Injury.

	Cotton-Myer			
	I	II	III	IV
Idiopathic	0	0	0	n/a
Polytrauma	0	25	100	100
Autoimmune	36	50	100	100
Iatrogenic	57	44	92	88

	Lano		
	I	II	III
Idiopathic	0	0	n/a
Polytrauma	27	100	n/a
Autoimmune	50	42	83
Iatrogenic	60	75	100

	McCaffrey			
	I	II	III	IV
Idiopathic	0	0	0	0
Polytrauma	0	60	50	n/a
Autoimmune	63	50	50	n/a
Iatrogenic	36	65	82	80

n/a refers to an absence of patients within a given stage.

this is also the population that more frequently requires intensive respiratory support. In our series, each additional point on CCI was associated with a 67% increased odds of tracheostomy dependence. Although this association does not appear surprising, we believe that it is powerful. It demonstrates the suitability of the CCI to serve as a systems-based protocol to identify patients who mandate a heightened awareness of complications from these procedures.

Consistent with previously published series,[4,16,17] despite many risk factors for iatrogenic injury being clarified over the past 40 years,[15,18–20] more than half the LTS burden in our cohort was potentially preventable. Overall, 59% of iatrogenic injuries occurred within the subglottis; therefore, they are attributable to intubation. In a post hoc analysis, 83% (15/18) of the "healthy" patients (those without DMII or cardiovascular disease) with iatrogenic LTS were women. This previously reported observation[21] suggests that endotracheal tube size may contribute to tracheal injury and should be carefully considered in the smaller female trachea.[22]

As has been consistently shown across other large series,[15] patients with DMII are particularly vulnerable to airway injury and have a higher likelihood of long-term tracheostomy dependence when injury occurs. Interestingly, the rate of GERD was not significantly different between the etiologic subgroups. Although other investigators have suggested a tight relationship between GERD and adult idiopathic LTS, this was not seen in our patient population. The limits of retrospective review prevent us from direct comparison of the objective data on the frequency and severity of reflux episodes between individuals and subgroups. Increased body mass index also has a suggested association, with increased risk of tracheal injury with intubation and worse response to procedural intervention. Our series lacked the biometric data to address this concept. Additionally, the limits of our tertiary care referral center (with limited out-of-network medical records) prevented us from exploring the relationship between the length of intubation or type of tracheostomy procedure (open vs. percutaneous) and the ultimate injury severity or treatment outcome.

A strong association between the degree of stenosis and ultimate decannulation has previously been reported in children.[23] Our series supports these prior observations in the pediatric population and now extends them to adults. As previously reported in adults, the location of injury and the length of stenosis are also essential components to predict long-term tracheostomy dependence. Critically, we now also offer data supporting an additional relationship between the cause of upper airway injury and its ultimate response to therapy. This relationship had been assumed; we offer the first formal demonstration.

Anatomic staging systems are numerous,[3–5,24–28] yet the ideal system in adult LTS remains unresolved. The most established allow some degree of prognosis, promote individualized treatment planning, and facilitate multi-institutional comparison. In this work, we utilized three separate, established LTS classification systems. As expected, they all effectively stratify the patient's risk of long-term tracheostomy. Of interest, however, in adult LTS it appears that the McCaffrey and Lano systems offer more precision than does the Cotton-Myer scale.

In general, although those patients in our series with more severe luminal compromise, longer stenoses, and lesions spanning multiple subsites had a much higher incidence of tracheostomy, this observation did not hold in the idiopathic group (patients who never required tracheostomy), suggesting a unique injury. Conversely, whereas lower LTS stages (in all 3 systems) overall had a lower rate of tracheostomy, patients with iatrogenic injuries had a significantly higher rate, even when matched at lower stenosis grades (identically in all 3 systems). Grouping LTS patients solely by an anatomic classification of their injury neglects a critical component of the heterogeneous biology responsible for tracheal scar.

Patients with iatrogenic stenosis appear to possess unique medical comorbidities. Their disease ultimately behaves differently, as evidenced by their disparate rate of long-term tracheostomy dependence, even when matched for similar degree of luminal compromise. These separate subgroups likely merit tailored treatment strategies.

The finding of the high rate of tracheomalacia in the subgroup with iatrogenic injuries, and the significant association between tracheomalacia and long-term tracheostomy dependence in this subgroup, raises questions regarding the relative contributions of mucosal injury

and cartilaginous injury in LTS. Ultimately, we believe the degree of tracheal wall injury (what we term "superstructure instability") may have significant prognostic power for overall response to therapy. However, this is difficult to quantify at present with our current diagnostic modalities and remains an area of active research.

Our study represents one of the largest published adult LTS series in the scientific literature. The data supports the hypothesis that laryngotracheal stenosis is a common endpoint to multiple pathophysiologic processes. Although different mechanisms of airway injury physiologically affect the patient in similar ways, we show that they occur in unique populations and have divergent responses to therapy. Management and prevention strategies should carefully consider this heterogeneous pathophysiology. This difference is not reflected in staging systems limiting themselves to an anatomic description of the tracheal scar.

CONCLUSION

Relief through endoscopic dilation, or open tracheal resection, is attainable in some cases of LTS; however, treatment is not universally successful. It is incumbent on the scientific community to move beyond viewing LTS as a purely anatomic problem, remedied only through surgical reconstruction. Rather, the management of airway stenosis should transition to increasingly personalized plans of care based on early recognition of at-risk populations, and an understanding of the divergent pathophysiology affecting the unique subgroups with LTS.

Acknowledgments

Study concept and design: Drs. A. Gelbard and V.C. Sandulache; acquisition of data: Drs. A. Gelbard and J.C. Simmons; analysis and interpretation of data: Drs. A. Gelbard and D.O. Francis; drafting of the manuscript: Drs. A. Gelbard and D.O. Francis; critical revision of the manuscript: Drs. J. Ongkasuwan and D.O. Francis; statistical analysis: Dr. D.O. Francis.

Administrative, technical, or material support: Dr. D.T. Donovan; study supervision: Drs. D.T. Donovan and J. Ongkasuwan.

BIBLIOGRAPHY

1. Nouraei SA, Singh A, Patel A, Ferguson C, Howard DJ, Sandhu GS. Early endoscopic treatment of acute inflammatory airway lesions improves the outcome of postintubation airway stenosis. *Laryngoscope* 2006;116:1417–1421.
2. Institute of Medicine (U.S.). Committee on Quality of Health Care in America. *Crossing the Quality Chasm : A New Health System for the 21st century*. Washington, DC: National Academy Press; 2001.
3. Cotton RT. Pediatric laryngotracheal stenosis. *J Pediatr Surg* 1984;19: 699–704.
4. Lano CF Jr, Duncavage JA, Reinisch L, Ossoff RH, Courey MS, Netterville JL. Laryngotracheal reconstruction in the adult: a ten year experience. *Ann Otol Rhinol Laryngol* 1998;107:92–97.
5. McCaffrey TV. Classification of laryngotracheal stenosis. *Laryngoscope* 1992;102:1335–1340.
6. Baugnee PE, Marquette CH, Ramon P, Darras J, Wurtz A. [Endoscopic treatment of post-intubation tracheal stenosis. *Apropos of 58 cases]. Rev Mal Respir* 1995;12:585–592.
7. Grillo HC, Donahue DM, Mathisen DJ, Wain JC, Wright CD. Postintubation tracheal stenosis: treatment and results. *J Thorac Cardiovasc Surg* 1995;109:486–493.
8. Hartley BE, Rutter MJ, Cotton RT. Cricotracheal resection as a primary procedure for laryngotracheal stenosis in children. *Int J Pediatr Otorhinolaryngol* 2000;54:133–136.
9. Manning PB, Rutter MJ, Border WL. Slide tracheoplasty in infants and children: risk factors for prolonged postoperative ventilatory support. *Ann Thorac Surg* 2008;85:1187–1191; discussion 1191–1192.
10. Grillo HC. Primary reconstruction of airway after resection of subglottic laryngeal and upper tracheal stenosis. *Ann Thorac Surg* 1982; 33: 3–18.
11. Laccourreye O, Brasnu D, Seckin S, Hans S, Biacabe B, Laccourreye H. Cricotracheal anastomosis for assisted ventilation-induced stenosis. *Arch Otolaryngol Head Neck Surg* 1997;123:1074–1077.
12. Whited RE. Posterior commissure stenosis post long-term intubation. *Laryngoscope* 1983;93:1314–1318.
13. Macht M, Wimbish T, Clark BJ, et al. Postextubation dysphagia is persistent and associated with poor outcomes in survivors of critical illness. *Crit Care* 2011;15:R231.
14. Streitz JM Jr, Shapshay SM. Airway injury after tracheotomy and endotracheal intubation. *Surg Clin North Am* 1991;71:1211–1230.
15. Wu X, Su ZZ, Hu LJ, et al. [Analysis of the risk factors causing tracheal stenosis after tracheotomy for mechanical ventilation in 560 patients]. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 2007;42:839–842.
16. Kleiss IJ, Verhagen AF, Honings J, Schuurbiers OC, van der Heijden HF, Marres HA. Tracheal surgery for benign tracheal stenosis: our experience in sixty three patients. *Clin Otolaryngol* 2013;38:343–347.
17. Mutrie CJ, Eldaif SM, Rutledge CW, et al. Cervical tracheal resection: new lessons learned. *Ann Thorac Surg* 2011;91:1101–1106; discussion 1106.
18. Mathias DB, Wedley JR. The effects of cuffed endotracheal tubes on the tracheal wall. *Br J Anaesth* 1974;46:849–852.
19. Seegobin RD, van Hasselt GL. Endotracheal cuff pressure and tracheal mucosal blood flow: endoscopic study of effects of four large volume cuffs. *Br Med J (Clin Res Ed)* 1984;288:965–968.
20. Andrews MJ. The incidence and pathogenesis of tracheal injury following tracheostomy with cuffed tube and assisted ventilation. Analysis of a 3-year prospective study. *Br J Surg* 1971;58:749–755.
21. Zias N, Chroneou A, Tabba MK, et al. Post tracheostomy and post intubation tracheal stenosis: report of 31 cases and review of the literature. *BMC Pulm Med* 2008;8:18.
22. Griscom NT, Wohl ME. Dimensions of the growing trachea related to age and gender. *AJR Am J Roentgenol* 1986;146:233–237.
23. Hartnick CJ, Hartley BE, Lacy PD, et al. Surgery for pediatric subglottic stenosis: disease-specific outcomes. *Ann Otol Rhinol Laryngol* 2001;110: 1109–1113.
24. Anand VK, Alemar G, Warren ET. Surgical considerations in tracheal stenosis. *Laryngoscope* 1992;102:237–243.
25. Brichet A, Verkindre C, Dupont J, et al. Multidisciplinary approach to management of postintubation tracheal stenoses. *Eur Respir J* 1999;13: 888–893.
26. Freitag L, Ernst A, Unger M, Kovitz K, Marquette CH. A proposed classification system of central airway stenosis. *Eur Respir J* 2007;30: 7–12.
27. Grillo HC, Mark EJ, Mathisen DJ, Wain JC. Idiopathic laryngotracheal stenosis and its management. *Ann Thorac Surg* 1993;56:80–87.
28. Grundfast KM, Morris MS, Bernsley C. Subglottic stenosis: retrospective analysis and proposal for standard reporting system. *Ann Otol Rhinol Laryngol* 1987;96:101–105.

Reprinted by permission of Am J Surg. 2014; 208(5):770-774.

The American Journal of Surgery (2014) 208, 770-774

The American Journal of Surgery®

Clinical Science

Incidence of overall complications and symptomatic tracheal stenosis is equivalent following open and percutaneous tracheostomy in the trauma patient

William W. Kettunen, M.D.[a], Stephen D. Helmer, Ph.D.[a], James M. Haan, M.D.[a,b,*]

[a]*Department of Surgery, The University of Kansas School of Medicine – Wichita, 929 N. Saint Francis Street, Room 3082, Wichita, KS 67214, USA;* [b]*Department of Trauma Services, Via Christi Hospital, Saint Francis Campus, Wichita, KS, USA*

KEYWORDS:
Tracheal stenosis;
Percutaneous tracheostomy;
Open tracheostomy

Abstract

BACKGROUND: While percutaneous tracheostomy (PT) is becoming the procedure of choice for elective tracheostomy, there is little late complication data. This study compared incidence of, and factors contributing to, tracheal stenosis following PT or open tracheostomy (OT).

METHODS: A 10-year review was conducted of trauma patients undergoing tracheostomy. Data on demographics, injury severity, tracheostomy type, complications, and outcomes were compared between patients receiving PT or OT and for those with or without tracheal stenosis.

RESULTS: Of 616 patients, 265 underwent OT and 351 underwent PT. Median injury severity score was higher for PT (26 vs 24, $P = .010$). Overall complication rate was not different (PT = 2.3% vs OT = 2.6%, $P = .773$). There were 9 tracheal stenosis, 4 (1.1%) from the PT group and 5 (1.9%) from the OT group ($P = .509$). Mortality was higher in OT patients (15.5% vs 9.7%, $P = .030$). Patients developing tracheal stenosis were younger (29.8 vs 45.2 years, $P = .021$) and had a longer intensive care unit length of stay (28.3 vs 18.9 days, $P = .036$).

CONCLUSION: Risk of tracheal stenosis should not impact the decision to perform an OT or PT.
© 2014 Elsevier Inc. All rights reserved.

Percutaneous tracheostomy (PT) is becoming the procedure of choice for elective tracheostomy in trauma patients. Many studies have proven this more prevalent technique to be safe, and possibly more cost-effective, than the traditional open tracheostomy (OT).[1–3] Most of the literature consists of observational data or small prospective studies, therefore debate still continues as to which method is preferred.

The literature is less clear on late complications, specifically tracheal stenosis. The exact incidence of tracheal stenosis following tracheostomy is difficult to quantify because many patients are critically ill and may die before decannulation, are lost to follow-up after being dismissed from a level-I trauma center, or are asymptomatic.[4–6] With a shortage of evidence, some postulate that the percutaneous technique predisposes patients to tracheal stenosis, more so than the open technique. They cite that the

No current or previous support was received from industry or organizations that might have influenced this work.

Presented at the 71st Annual Meeting of the American Association for the Surgery of Trauma, September 12 to 15, 2012, Kauai, Hawaii.

* Corresponding author. Tel.: +1-316-268-5538; fax: +1-316-291-7892.

E-mail address: James.Haan.Research@viachristi.org

Manuscript received May 22, 2013; revised manuscript December 4, 2013

0002-9610/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amjsurg.2013.12.036

ostomy is often times placed higher on the trachea percutaneously than it is when performed open and there is more trauma and granulation tissue to the trachea when passing dilators percutaneously.[7]

The purpose of this investigation was to compare outcomes and complications between OT and PT. All major complications, including tracheal stenosis, were recorded to determine the incidence of, and any risk factors for, tracheal stenosis.

Patients and Methods

A retrospective review of all trauma patients who received a tracheostomy from August 1, 2001 to July 31, 2011 was conducted. Patients were identified using the trauma registry of an established American College of Surgeons-verified level-1 trauma center. Patient demographics, mechanism of injury, injury severity score (ISS), Glasgow coma scale (GCS) score, time from injury to tracheostomy creation, method of performing tracheostomy (open vs percutaneous), complications associated with tracheostomy (tracheo-innominate artery fistula, tracheal stenosis, scar and excess granulation tissue requiring surgical scar revision, loss of airway requiring conversion to open, and bleeding requiring conversion to open), intensive care unit (ICU) length of stay (LOS), mechanical ventilator days, overall LOS, and patient disposition were collected using the trauma registry and patient records. Tracheal stenosis was identified based on clinical symptoms (ie, difficulty with decannulation or shortness of breath with exertion). Complications were defined as being early, those occurring within the first 48 hours of tracheostomy, or late, those occurring more than 48 hours post-tracheostomy. Outcomes and complication data were collected from the in-hospital stay and from rehospitalizations. Study subjects were not contacted for long-term follow-up.

Analyses were conducted using IBM SPSS Statistics for Windows, Version 19.0. (IBM Corp, Armonk, NY). Data were initially summarized. Primary analyses were conducted comparing outcomes between patients based on the method of tracheostomy creation (OT vs PT). Secondary analyses were conducted comparing outcomes between patients who developed tracheal stenosis and patients who did not develop tracheal stenosis. Quantitative data were analyzed using the Student t test. If heterogeneity of variance was identified, the Mann–Whitney test was used. Comparisons of ordinal data were analyzed with the Mann–Whitney test. Qualitative data were analyzed with chi-square analysis or the Fisher's exact test in instances where cell size was 5 or less observations. All analyses were conducted as 2-tailed tests and statistical significance was defined as $P < .05$.

This study was reviewed and approved for implementation by the Institutional Review Board of Via Christi Hospitals Wichita, Inc.

Results

During the 10-year study period, 629 tracheostomies were performed on trauma patients. We excluded 13 patients who had an emergency cricothyroidotomy or whose LOS was for more than 1 day. Of the remaining 616 patients, the average age was 45.0 ± 20.6 years, the majority were male (n = 458, 74.4%), white (n = 534, 86.7%), and median ISS and GCS scores were 25 (25th and 75th percentiles = 17 and 33) and 5 (25th and 75th percentiles = 3 and 14), respectively. Forty-three percent (n = 265) had an OT and 57% (n = 351) had a PT. There were no significant differences in age, sex, GCS score, mechanism of injury, interval from admission to tracheostomy formation, ICU LOS, ventilator days, or hospital LOS between the 2 groups (Table 1). There was a significant

Table 1 Comparison of demographics, injury severity, mechanism of injury, and hospitalization details for patients who received a tracheostomy through an open or percutaneous procedure

Parameter	Open procedure	Percutaneous procedure	P value
No. of subjects (%)	265 (43.0%)	351 (57.0%)	
Age (years)*	45.0 ± 21.3	44.9 ± 20.1	.932
Sex (male)	204 (77.0%)	254 (72.4%)	.194
Injury severity score†	24.0 (17.0, 30.0)	26.0 (18.0, 34.0)	.010
Glasgow coma scale score†	6.0 (3.0, 15.0)	3.0 (3.0, 14.0)	.116
Mechanism of injury			.068
Blunt	244 (92.1%)	337 (96.0%)	
Penetrating	18 (6.8%)	13 (3.7%)	
Drowning	1 (.4%)	1 (.3%)	
Burn	2 (.8%)	0 (.0%)	
Admission to tracheostomy interval (days)	7.0 ± 5.4	7.0 ± 4.7	.988
Intensive care unit days*	19.3 ± 15.2	18.9 ± 11.8	.223
Mechanical ventilation days*	16.7 ± 12.9	15.8 ± 11.2	.945
Hospital length of stay (days)*	27.6 ± 19.9	26.7 ± 29.2	.643

*Mean ± standard deviation.
†Median (25th and 75th percentile).

Table 2 Comparison of complication, disposition, and death data for patients who received a tracheostomy through an open or percutaneous procedure

Parameter	Open procedure Number (%)	Percutaneous procedure Number (%)	*P* value
Complication	7 (2.6%)	8 (2.3%)	.773
Tracheal stenosis	5 (1.9%)	4 (1.1%)	.509
Other major complications	2 (.8%)	4 (1.1%)	.704
Disposition			.007
Home/home with home health care/jail/mental health facility	44 (16.6%)	35 (10.0%)	
Rehabilitation center/select specialty hospital acute care/other acute hospitals	165 (62.3%)	262 (75.1%)	
Skilled nursing unit/nursing home	14 (5.3%)	14 (4.0%)	
Hospice/death	42 (15.8%)	38 (10.9%)	
Death	41 (15.5%)	34 (9.7%)	.030

difference in ISS between the 2 groups with the percutaneous group having a higher median score than the open group (24 vs 26, $P = .007$).

The overall complication rate was similar between the OT and PT groups (Table 2). The incidence of tracheal stenosis was also similar when comparing the open group with the percutaneous group (1.9% vs 1.1%, $P = .509$). The open group had an incidence of major complications other than tracheal stenosis of .8%. These complications involved scar and excess granulation tissue requiring surgical scar revision ($n = 2$). The percutaneous group had an incidence of other major complications of 1.1% ($n = 4$). These included tracheo-innominate artery fistula ($n = 1$), loss of airway requiring conversion to open ($n = 2$), and bleeding requiring conversion to open ($n = 1$). The patient with a tracheo-innominate artery fistula hemorrhaged while on the floor. The hemorrhage was occluded manually while the patient was taken to the operating room for repair, but the patient exsanguinated before repair could be accomplished. All conversions to an open procedure occurred during the initial hospitalization.

Four of the 9 patients were immediately diagnosed with tracheal stenosis after a failed decannulation attempt; however, the other 5 patients presented in a delayed fashion after being decannulated. The delay ranged from 3 to 12 months, with patients presenting with shortness of breath with exertion ($n = 4$) and with trouble extubating after elective laparoscopic cholecystectomy ($n = 1$). All of the 9 patients underwent some form of treatment for their stenosis. Five of these 9 patients underwent bronchoscopy with tracheal balloon dilation, while 4 of the 9 patients underwent tracheal resection. In our study, the risk of tracheal stenosis requiring invasive intervention following tracheostomy was 1.5%.

There was a significant difference in patient disposition between the 2 groups (Table 2). The open group was discharged home more often and to a rehabilitation center less often when compared with the percutaneous group ($P = .007$); however, mortality rate was higher in the open group (15.5% vs 9.7%, $P = .030$).

We also conducted analyses comparing those patients with tracheal stenosis with those with no tracheal stenosis, independent of which method of tracheostomy was performed (Table 3). Patients who developed tracheal stenosis were younger (29.8 vs 45.2 years of age, $P = .021$), had a longer ICU LOS (28.3 vs 18.9 days, $P = .036$), and tended to require mechanical ventilation for a longer interval (26.7 vs 16.1 days, $P = .055$) compared with those who did not develop tracheal stenosis. There were, however, no differences between the groups in regard to sex, ISS, GCS score, mechanism of injury, interval between admission and tracheostomy formation, hospital LOS, disposition, or mortality.

Comments

While there is support in the literature of equivalent early complication rates between open and percutaneous techniques,[8,9] there is less evidence about their equivalency with regard to late complications such as tracheal stenosis. For this reason, there is still debate about which method provides superior patient outcomes. The incidence of symptomatic tracheal stenosis following OT or PT ranges in the literature from 0% to 10%.[4–6] The true incidence of tracheal stenosis is difficult to ascertain because it is often subclinical in nature. In our study, tracheal stenosis was identified based on clinical symptoms. Our study was similar to these published results, demonstrating equivalent symptomatic tracheal stenosis rates for OT and PT (1.9% vs 1.1%, respectively).

As stated earlier, several studies demonstrate complication rates that are equivalent for PT and OT. Our study supports the literature in this regard with an overall complication rate of 2.3% and 3.3%, respectively. The types of complications encountered during tracheostomy creation have been described in the literature and include peristomal bleeding, peristomal infection, loss of airway during procedure, surgical scar contracture, and tracheo-innominate artery fistula.[2,10] The complications reported in our study are in line with those previously described. Major complications in our study were defined as need for surgical intervention or death. Both of the reoperations in the

Table 3 Comparison of demographics, injury severity, mechanism of injury, hospitalization details, disposition, and death for patients with tracheal stenosis versus patients without tracheal stenosis

	Tracheal stenosis	No tracheal stenosis	
Parameter	Value	Value	*P* value
No. of subjects	9 (1.5%)	607 (98.5%)	
Age (years)*	29.8 ± 11.8	45.2 ± 20.6	.021
Sex (male)	6 (66.7%)	452 (74.5%)	.701
Injury severity score†	30.0 (19.5, 37.0)	25.0 (17.0, 33.0)	.175
Glasgow coma scale score†	3.0 (3.0, 13.0)	5.0 (3.0, 14.0)	.659
Mechanism of injury			1.000
Blunt	9 (100.0%)	572 (94.2%)	
Penetrating	0 (.0%)	31 (5.1%)	
Drowning	0 (.0%)	2 (.3%)	
Burn	0 (.0%)	2 (.3%)	
Admission to tracheostomy interval (days)	9.3 ± 7.9	7.0 ± 4.9	.175
Intensive care unit days*	28.3 ± 18.8	18.9 ± 13.3	.036
Mechanical ventilation days*	26.7 ± 21.7	16.1 ± 11.7	.055
Hospital length of stay (days)*	40.0 ± 21.7	26.9 ± 25.6	.127
Disposition			.604
Home/home with home health care/jail/mental health facility	2 (22.2%)	77 (12.7%)	
Rehabilitation center/select specialty hospital acute care/ other acute hospitals	7 (77.8%)	42 (69.4%)	
Skilled nursing unit/nursing home	0 (.0%)	28 (4.6%)	
Hospice/death	0 (.0%)	80 (13.2%)	
Death	0 (.0%)	75 (12.4%)	.610

*Mean ± standard deviation.
†Median (25th and 75th percentile).

open group were for surgical scar revision. The other major complications observed in our PT patients included loss of airway, bleeding, and tracheo-innominate artery fistula, all well-known risks of PT.

As we did not find the route of tracheostomy formation to influence the incidence of tracheal stenosis, we attempted to identify factors that may have contributed to stenosis formation. Numerous and variable risk factors for tracheal stenosis following intubation have been suggested in the literature and include trauma and inflammation at the endotracheal tube cuff site, excess granulation tissue around the tracheal stoma site or over a fractured cartilage, high tracheostomy site, prolonged intubation, traumatic intubation, or previous intubation or tracheostomy.[1] Both groups in this study received tracheostomy within 7 days of admission as per American College of Surgeons recommendations. Our study demonstrated that patients who developed tracheal stenosis tended to have longer mechanical ventilator requirements (26.7 vs 16.1 days, $P = .055$), with patients developing stenosis being on the ventilator on average 11 additional days. It could be hypothesized that additional ventilator days meant more time with an inflated tracheal cuff causing tracheal ischemia and stenosis. We did identify that younger age and longer length of ICU stay were associated with increased rate of tracheal stenosis; however, the reason for these findings is unclear and these findings were not observed in similar studies.

There are several limitations to our study. It is retrospective in nature with a relatively small sample size. Also, the study time frame included the widespread introduction of the percutaneous technique at our institution. The associated learning curve for a new procedure may have influenced the results or influenced which technique was used to create the tracheostomy in specific patients. Also, the percentage of patients seen after dismissal from the hospital was unknown.

Common shortcomings of other investigations into this topic are length and reliability of follow-up, thereby calling into question the accuracy of the reported incidence of tracheal stenosis. We have a unique practice environment in which there are 2 level-1 trauma centers serving the entire population center with extremely rural surroundings. This leads to an isolated trauma population for study. These centers inform each other of any complications or admissions from each other's population. Additionally, the next closest trauma center for follow-up for tracheal symptoms is 200 miles away. Because of this, we do not believe that any patients were lost to follow-up or transferred to the other trauma center with a complication of tracheal stenosis. Also, in our city, if a patient presented to one of the other large multispecialty groups, then that patient would be redirected to our clinic. Evidence for this rests in the fact that 5 of the 9 tracheal stenosis patients presented to our clinic in a delayed fashion well after hospital discharge. These factors help distinguish our follow-up results as compared with other studies. This said, there still exists the possibility that late occurring and/or subclinical tracheal stenoses may have

been missed as we did not contact study subjects for late-term follow-up or measure tracheal circumference in this retrospective study.

PT has been proven to be a safe and cost-effective method for elective tracheostomy. Our study demonstrates that the complication rate, particularly that of tracheal stenosis, after PT is equivalent to that observed after the traditional OT procedure.

References

1. Freeman BD, Isabella K, Lin N, et al. A meta-analysis of prospective trials comparing percutaneous and surgical tracheostomy in critically ill patients. Chest 2000;118:1412–8.
2. Friedman Y, Fildes J, Mizock B, et al. Comparison of percutaneous and surgical tracheostomies. Chest 1996;110:480–5.
3. Holdgaard HO, Pedersen J, Jensen RH, et al. Percutaneous dilational tracheostomy versus conventional surgical tracheostomy. A clinical randomized study. Acta Anaesthesiol Scand 1998;42:545–50.
4. Norwood S, Vallina VL, Short K, et al. Incidence of tracheal stenosis and other late complications after percutaneous tracheostomy. Ann Surg 2000;232:233–41.
5. Kornblith LZ, Burlew CC, Moore EE, et al. One thousand bedside percutaneous tracheostomies in the surgical intensive care unit: time to change the gold standard. J Am Coll Surg 2011;212:163–70.
6. Jackson LS, Davis JW, Kaups KL, et al. Percutaneous tracheostomy: to bronch or not to bronch–that is the question. J Trauma 2011;71:1553–6.
7. Raghuraman G, Rajan S, Marzouk JK, et al. Is tracheal stenosis caused by percutaneous tracheostomy different from that by surgical tracheostomy? Chest 2005;127:879–85.
8. Ciaglia P, Graniero KD. Percutaneous dilatational tracheostomy. Results and long-term follow-up. Chest 1992;101:464–7.
9. Cobean R, Beals M, Moss C, et al. Percutaneous dilatational tracheostomy. A safe, cost-effective bedside procedure. Arch Surg 1996;131:265–71.
10. Wood DE. Tracheostomy. Chest Surg Clin N Am 1996;6:749–64.

Reprinted by permission of Ann Otol Rhinol Laryngol. 2015; 124(2):137-142.

Article

Using Pulmonary Function Data to Assess Outcomes in the Endoscopic Management of Subglottic Stenosis

Annals of Otology, Rhinology & Laryngology
2015, Vol. 124(2) 137–142
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0003489414548915
aor.sagepub.com
SAGE

Shannon M. Kraft, MD[1], Kevin Sykes, PhD, MPH[1], Andrew Palmer, CCC-SLP[2], and Joshua Schindler, MD[2]

Abstract

Objective: This study aimed to examine the authors' experience with endoscopic management of idiopathic subglottic stenosis (iSGS), and to identify pulmonary function test (PFT) values that can be used to quantify outcomes.

Methods: Retrospective review.

Results: Twenty-five patients with a new diagnosis of iSGS were seen between 2006 and 2012. Median age at surgery was 45.3 years (interquartile range [IQR], 38.5-67.0), and median body mass index was 28.7 kg/m^2 (IQR, 23.5-32.1). Forty-five procedures were performed. Median preoperative stenosis was 56.8% (Cotton-Myer grade 2). The typical stenosis began 15 mm below the true vocal folds and was 12 mm long. Median follow-up was 21.4 months (IQR, 5.1-43.1). For patients receiving multiple dilations, median time between procedures was 23.7 months. Four PFT parameters demonstrated significant improvement after intervention: (1) PEF (absolute change = 2.54 L/s), (2) PIF (absolute change = 1.57 L/s), (3) FEV1/PEF (absolute change = 0.44), and (4) FIF50% (absolute change = 1.71 L/s). PIF was the only parameter affected by using a larger balloon (P = .047).

Conclusion: PEF, PIF, FEV1/PEF, and FIF50% improved significantly after endoscopic incision and dilation of iSGS, and this could potentially be used as a metric by which to evaluate outcomes in the endoscopic management of subglottic stenosis.

Keywords

endoscopic, idiopathic subglottic stenosis, pulmonary function test

Introduction

Etiology/Pathophysiology

Idiopathic subglottic stenosis (iSGS) is a nonspecific fibroinflammatory process that results in progressive narrowing of the airway at the level of the cricoid and proximal trachea. Investigators have proposed a multitude of theories as to the etiology of the process since its first description in 1972.[1] The disorder occurs almost exclusively in women, typically presenting in the fourth or fifth decade.[2] This female preponderance led to the hypothesis that estrogen might play a role in the pathogenesis of iSGS, although the evidence for this has not borne out.[3,4]

Current theories regarding the etiology of iSGS have focused on non-gender-specific causes. Subglottic injury from laryngopharyngeal reflux (LPR) has been implicated as a possible cause.[5] Autoimmune disease, specifically a limited form of seronegative polyangiitis with granulomatosis (GPA), has also been proposed as a mechanism for iSGS.[6] Other theories on the etiology of the disorder include repetitive microtrauma from cough[7] and hereditary factors.[8] Ultimately a disease of the subglottic lamina propria, no single etiologic factor has been identified.

Endoscopic Management of iSGS

Although segmental resection of the diseased portion of the airway is considered the gold standard for obtaining long-term improvement in the airway in iSGS, it is a potentially morbid procedure.[9] Endoscopic surgery provides an attractive alternative that is less invasive, albeit less definitive. As surgeon experience has grown, several variations in technique have been explored. Dedo and

[1]Department of Otolaryngology, University of Kansas, Kansas City, Kansas, USA
[2]Department of Otolaryngology, Oregon Health and Science University, Portland, Oregon, USA

Corresponding Author:
Shannon M. Kraft, MD, Department of Otolaryngology–Head and Neck Surgery, University of Kansas Medical Center, 3901 Rainbow Boulevard, MS 3010, Kansas City, KS 66160, USA.
Email: skraft3@kumc.edu

Catten[3] described a microflap technique in which a CO_2 laser was used to make a mucosal flap followed by resection or ablation of the aberrant soft tissue between the flap and the cricoid. Shapshay et al[10] reported on the use of radial incisions with a CO_2 laser, followed by dilation with a rigid bronchoscope. Some surgeons favor a cold technique over the laser for making radial incisions.[11] Balloon dilation was introduced in the management of subglottic stenosis as a theoretically less traumatic alternative to passing rigid dilators.[12]

In addition to the different surgical techniques, there have been several investigations evaluating wound-healing modulators as adjunctive therapies. Most notable, inhaled, systemic, and locally injected steroids have been used extensively in an attempt to slow restenosis after dilation.[13] Mitomycin C, an alkylating agent, has been proven to prevent fibroblast proliferation[14] and has been used with varying degrees of success in endoscopic airway surgery.[15] Halofuginone, an inhibitor of collagen 1-α synthesis,[16] and 5-fluorouracil, an antimetabolite that inhibits fibroblast activity,[17] are also being investigated in animal models as potential adjunctive therapies.

Given the variety of options in the endoscopic management of iSGS, it has become obvious that an objective means by which to quantify operative outcomes and to compare the efficacy of different techniques or adjunctive therapies is sorely needed.

Assessing Outcomes in the Endoscopic Management of Subglottic Stenosis

In the late 1960s and early 1970s, there was a great deal of interest in using pulmonary function tests (PFTs) to aid in the diagnosis of upper airway obstruction (UAO). Initial investigations focused on identifying values that could differentiate UAO from lower airway disease.[18,19] It was ultimately Hyatt's[20] description of the flow-volume loop and the different patterns of obstruction (variable intrathoracic, variable extrathoracic, and fixed) that provided clinicians with a powerful tool for diagnosing and classifying UAO. There has been a renewed interest in the past 2 decades in using PFT data not only as a diagnostic tool but also as a means of quantifying the results of interventions in UAO.[21,22]

This retrospective study is designed to review our experience with using pulmonary function data in the management of patients with iSGS. Specifically, we aim to (1) describe our experience with iSGS, (2) identify which PFT parameters change following endoscopic intervention, (3) quantify the degree of improvement in airflow postoperatively using PFT data, and (4) determine if PFTs change in a predictable manner postoperatively as restenosis invariably occurs.

Methods

Inclusion/Exclusion Criteria

This study was approved by the institutional review board at Oregon Health and Science University. A retrospective chart review was performed, examining a single surgeon's experience with iSGS. Records for all newly diagnosed adult patients with iSGS referred to the senior author between January 1, 2006, and December 31, 2012, were reviewed. At least 1 standard endoscopic intervention (described below) and 1 pulmonary function test was required for inclusion in the study. Patients with a history of previous airway surgery, airway trauma, tracheotomy, positive serology (c-ANCA or ACE level), or intubation within the previous 12 months were excluded from the analysis.

Management of Stenosis

All patients included in the study were managed endoscopically in a standardized fashion. After confirming that the patient could be mask ventilated, anesthesia was induced and maintained using a total intravenous technique. A plastic tooth guard was used to protect the maxillary dentition. The airway was exposed using an Osshoff-Pilling laryngoscope, and the patient was suspended from the Mustard table. Ventilation was conducted via jet technique. Standard laser precautions were used throughout the procedure, to include placing moist gauze pads over the eyes and wet towels over any exposed skin.

A 4-mm rod lens telescope was used to evaluate and measure the stenosis. If the initial values were obtained by the resident or fellow, the senior author (J.S.) repeated and confirmed the measurements. The stenosis was described by its distance below the true vocal folds and by its total length. Once measurements were complete, 2-mm cup forceps were used to take a biopsy. The stenosis subsequently was infiltrated with 1 mL of triamcinolone suspension (40 mg/mL). A CO_2 laser was then used to make 3 to 5 radial incisions in the stenosis. Dilation of the stenosis was performed using a constant radial expansion (CRE) balloon. After confirming hemostasis and obtaining postoperative photo documentation, a small, cuffed endotracheal tube (5.0) was placed through the laryngoscope. The patient was taken out of suspension and the laryngoscope removed. The patient was then remanded to the care of the anesthesia team for recovery. Patients were discharged home from the postanesthesia care unit that same day with a fluticasone inhaler (220 mcg) and instructed to use 2 puffs twice daily until the inhaler was empty.

Data Collection and Statistical Analysis

After obtaining approval from the institutional review board, the senior surgeon's operative record was reviewed

Table 1. Clinical Presentation.

	Median	Interquartile Range
Race		
Caucasian	23	
Latina	2	
Age at surgery, mo	45.3	38.5-67.0
Weight, kg	70	61.5-85.5
Height, m	1.6	1.55-1.65
Body mass index, kg/m^2	28.7	23.5-32.1
Stenosis, %	58.60	38.9-78.4
Cotton-Myer grade	2	1-3
Distance below cords, mm	15	11.0-16.5
Length of stenosis, mm	12	9.0-17.0
Follow-up, mo	21.4	5.1-43.1
No. of procedures	2	1-3
Time between first & second surgeries, mo	23.7	15.6-31.8

to identify patients with a preoperative diagnosis of subglottic stenosis. Clinical records were then further screened to select patients who met the inclusion criteria. All data points, to include demographic, clinical, operative, laboratory, radiographic, and pulmonary functional data, were entered in a Microsoft Excel spreadsheet. Data were imported into and analyzed via SPSS version 20.0 (SPSS Inc, Chicago, Illinois, USA). A P value < .05 was considered to be significant.

As we could not assume a normal distribution for our small sample size, a Wilcoxon signed rank test was used to look for differences in the pre-dilation and post-dilation differences in pulmonary function data. Subsequently, a Kruskal-Wallis test was performed on those PFT parameters that improved significantly to determine the contribution of dilation size on the change. For those measures that changed significantly, postoperative PFT values were plotted as a function of time to assess the rate of change after intervention.

Results

Clinical Presentation and Evaluation of Stenosis

A total of 25 new patients with iSGS were seen between 2006 and 2011 at our clinic. As anticipated, all patients were female. Ninety-two percent (N = 23) were white and 8% (n = 2) Hispanic. The median age at the time of the first surgical intervention was 45.3 years (interquartile range [IQR], 38.5-67.0), with a median body mass index of 28.7 kg/m^2 (IQR, 23.5-32.1) (Table 1).

A total of 45 procedures were performed. Twelve patients had 1 procedure, 9 patients had 2 procedures, 2 patients had 3 procedures, 1 patient had 4 procedures, and 1 patient had a total of 7 procedures (median = 2). At the time of initial presentation, the typical stenosis was described as beginning 15 mm below the true vocal folds and measuring 12 mm long. Fifteen patients had preoperative computed tomography (CT) scans. The median degree of stenosis as determined by CT was 56.8% (Cotton-Myer grade 2). Four patients had grade 1 stenosis at presentation and 4 patients had grade 3 stenosis.

Median follow-up after surgery was 21.4 months (IQR, 5.1-43.1). For the 11 patients receiving at least 2 dilations, the median time between the first and second procedures was 23.7 months. Three patients ultimately proceeded to definitive cricotracheal resection.

Preoperative Versus Postoperative PFT Assessment

Seventeen of the 25 patients had a preoperative PFT in addition to at least 1 postoperative PFT performed within 8 weeks of surgery. Four parameters demonstrated a statistically significant improvement after intervention: (1) PEF (absolute change = 2.54 L/s), (2) PIF (absolute change = 1.57 L/s), (3) FEV1/PEF (absolute change = 0.44), and (4) FIF50% (absolute change = 1.71 L/s). FEV1, FVC, FEF25%-75%, and PEF/PIF did not change significantly (Table 2). Preoperative and postoperative PFT values were then examined in the context of balloon size. Improvement in the PIF was the only parameter that was affected by the size of dilation, with rank-order testing indicating a greater degree of improvement with use of a larger balloon (P = .047) (Table 3).

Changes in PFT Over Time

Seven patients of the 25 in this sample had at least 3 PFTs taken following the initial dilation. PEF, PIF, FEV1/PEF, and FIF50% from these samples were plotted as a function of time. There is a linear relationship between time and both PEF and FEV1/PEF (P = .0307 and P < .001, respectively). The slope of the line was unique to each patient (Figures 1 and 2). PIF and FIF50% generally decrease as the time from surgery increases, but a linear relationship could not be established (data not shown). Five of the 7 patients did have a second procedure but had not accumulated a sufficient number of subsequent PFTs during the study period for analysis. (The patients represented by a star and hexagon had only 1 procedure.)

Discussion

Not all patients are ideal candidates for endoscopic management of subglottic stenosis. Historically, previous failed dilations, stenosis length greater than 1 cm, circumferential stenosis, evidence of cartilage loss/damage, a history of severe bacterial infection with tracheotomy, posterior glottic stenosis with arytenoid fixation, and involvement of the inferior margin of the vocal folds were considered poor

Table 2. Comparison of Pre-dilation and Post-dilation Pulmonary Function Data.

PFT Parameter	Pre-dilation	Post-dilation	Median Change, Absolute	Median Change, %	*P* Value
FEV1, L/s	3.23	3.22	0.14	3.9	.177
FVC, L	2.4	2.57	0.25	10.0	.720
FEV1/FVC	0.74	0.77	0.5	5.8	.155
FEF25%-75%	2.05	2.88	0.41	21.2	.156
PEF, L/s	3.89	6.67	2.54	56.7	< .001
PIF, L/s	2.36	4.21	1.57	66.8	.001
PEF/PIF	1.66	1.67	0.13	8.4	.906
FEV1/PEF	0.82	0.39	0.443	56.0	.001
FIF50%, L/s	1.83	3.97	1.71	92.0	.001

Abbreviation: PFT, pulmonary function test.

Table 3. Influence of Dilation Size on Pulmonary Function Test (PFT) Parameters.

PFT Parameter	Dilation Size, mm	N	Median Change, Absolute	*P* Value
PEF, L/s	15.0	5	4.46	.497
	16.5	3	2.4	
	18.0	9	2.45	
PIF, L/s	15.0	5	2.47	.047
	16.5	3	1.35	
	18.0	9	1.51	
FEV1/PEF	15.0	5	–0.63	.441
	16.5	3	–0.62	
	18.0	9	–0.36	
FIF50%, L/s	15.0	4	2.13	.329
	16.5	3	1.38	
	18.0	9	1.58	

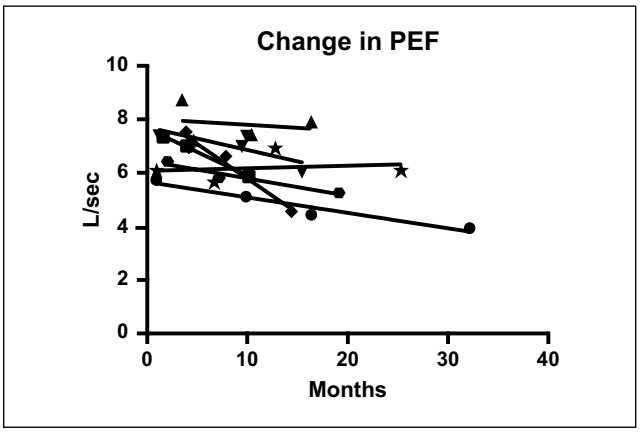

Figure 1. Change in PEF postoperatively. PEF declines in a linear fashion after intervention. The slope of the line is unique to each patient (*P* = .0307).

prognostic indicators.[23] Patients with isolated subglottic disease, however, can successfully be managed endoscopically in approximately 87% of cases.[24] The expectation is that more than 85% of these patients will have recurrence of their stenosis within 5 years, requiring a return to the operating suite.[25]

Clinically, the goal is to identify the techniques and adjuvant therapies that provide the greatest improvement in airflow and result in the slowest rate of restenosis. Not surprising, time between surgeries is a commonly reported metric used to demonstrate the efficacy of an intervention.[26,27] Although clinically relevant, confounding variables related to the surgical interval raise questions about its reliability for use in research. Surgeon availability certainly affects the time between procedures. Physically active patients are more likely to notice impairment of airflow than sedentary patients and may seek intervention sooner. Financial factors may affect a patient's decision to seek surgery, as well. It is clear that an objective measure is needed.

Using PFT data to evaluate airway stenosis is not a new concept. In the 1970s, a number of studies tried to identify which values or ratio of values could be used to diagnose UAO. Empey[18] reported on a series of 10 patients (most with bilateral vocal cord paralysis), noting that the FEV1/PEF ratio was greater than 10 in all cases and that the larger the ratio, the greater the degree of obstruction. In comparing UAO with chronic obstructive pulmonary disease, Rotman et al[19] identified 4 measurements that were found to differentiate upper from lower airway disease: (1) FEF50%/FIF50% > 1, (2) FEV1/PEFR > 10, (3) FIF50% < 100 L/min, and (4) FEV1/FEV0.5 > 1.5, with the latter 2 measures being less sensitive.

Pulmonary function tests have been used in a limited fashion to assess postoperative outcomes following tracheal resection[28] and endoscopic dilation,[29] but reports have been generally nonspecific about the degree of improvement and which measures are of greatest utility in quantifying results. To examine this issue, Wasserman et al[21] created a model of fixed obstruction using mouthpieces of decreasing inner

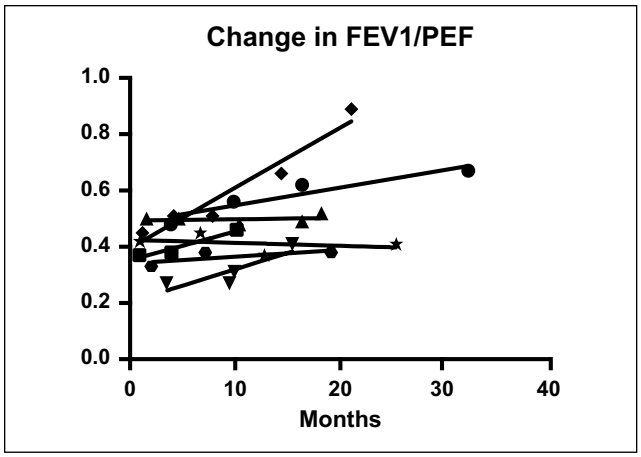

Figure 2. Change in FEV1/PEF postoperatively. FEV1/PEF similarly changes in a unique linear fashion after intervention ($P < .0001$).

diameters (15, 10, 8, and 6 mm) in line with a spirometer. PEF and PIF were found to be the most sensitive measures by which to assess changes in airway resistance. Nouraei et al[22] performed a similar experiment in 2007 in which the resistors were designed to generate a greater resistance to inspiratory flow than expiratory flow. The MEF_{50}/MIF_{50} and the ratio of the areas under the expiratory/inspiratory curves were the most sensitive and specific parameters.

Both the location of a stenosis and its characteristics (length, radius, boundary conditions) affect airflow, making every stenosis unique. In this study, we elected to specifically look at iSGS as a model of fixed obstruction. We chose to exclude glottic and tracheal stenosis from our analysis, both of which have some component of variability due to the contribution of Bernoulli forces and compression of membranous tracheal wall, respectively. Not surprising, we found that the PFT values that changed after intervention were the same as those reported by Wasserman et al. We suspect that an examination of PFT values affected by intervention for glottic or tracheal stenosis would yield results more closely resembling those recently reported by Nouraei and colleagues for a variable obstructive model.

In our evaluation, we examined the effect of radial incision and dilation on PFT values. Pulmonary function tests are an attractive means of evaluating stenosis patients because the study is inexpensive; widely available; and, unlike current techniques using CT for computational fluid dynamic studies, avoids the need for ionizing radiation. Whereas CT examines the structural component of stenosis, PFTs are a physiologic measure of the effect of stenosis on airflow and the patient's actual respiratory function. There are limitations to the PFT, the primary issue being that the study is dependent on patient effort. In addition, not all studies are conducted uniformly. Spirometry software is programmed to select the loop with the best expiratory effort for evaluation. Previous studies have demonstrated that the number reported by the computer for PIF and FIF50% did not represent the largest value in 50% and 69% of cases, respectively.[30,31] It is necessary to look at all loops or to establish a protocol for maximal inspiratory effort to mitigate this. Finally, some data suggest that the test may not be adequately sensitive for mild stenosis. Miller and Hyatt[32] reported that the trachea would have to be narrowed to 8 mm (~ 80%) prior to creating detectable changes in flow.

The retrospective nature of this study is an inherent limitation. As this is a single surgeon experience, the surgical technique was fairly consistent between patients. However, there was some variability in selecting the initial dilation size. The radius of the fluid conduit being one of the greatest influences on airflow has the potential to affect postoperative PFT values. Our data suggest that PIF may be influenced by dilation size. In addition, pulmonary function data were not collected at consistent time points during follow-up. Our data demonstrate that restenosis occurs in a linear manner that appears to be unique for each patient. It will be important in future studies to standardize the time and manner in which PFTs are administered in order to mitigate these confounding variables.

Conclusion

In the isolated subglottic stenosis model, PEF, PIF, FEV1/PEF, and FIF50% demonstrate a significant change after endoscopic intervention. The postoperative rate of change in the PEF and FEV1/PEF is linear but appears to be unique to each patient. PIF and FIF50% also change as restenosis occurs, but the lack of standardized protocols makes the rate of change more difficult to predict. These pilot data suggest that the change in PEF, FEV1/PEF, PIF, and FIF50% can be used to assess outcomes in the endoscopic management and treatment of iSGS. Furthermore, the rate of change in these values (PEF and FEV1/PEF, in particular) could potentially be used as an objective means of comparing the efficacy of different techniques or adjuvant therapies. Moving forward, prospective studies should focus on establishing uniform time points and protocols for collecting PFT data. Given the low incidence of the entity and the unique characteristics of each stenosis, garnering adequate numbers and long-term follow-up will likely require multi-institutional collaboration.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Brandenburg JH. Idiopathic subglottic stenosis. *Trans Am Acad Ophthalmol Otolaryngol*. 1972;76(5):1402-1406.
2. Mark EF, Meng F, Kradin RL. Idiopathic tracheal stenosis: a clinicopathologic study of 63 cases and comparison of the pathology with chondromalacia. *Am J Surg Path*. 2008;32(8):1138-1143.
3. Dedo HH, Catten MD. Idiopathic progressive subglottic stenosis: findings and treatment in 52 patients. *Ann Otol Rhinol Laryngol*. 2001;110(4):305-311.
4. Valdez TA, Shapsay SM. Idiopathic subglottic stenosis revisited. *Ann Otol Rhinol Laryngol*. 2008;111(8):690-695.
5. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope*. 1991;101(4, pt 2)(suppl 53):1-78.
6. Maldonado F, Loiselle A, DePew ZS, et al. Idiopathic subglottic stenosis: an evolving therapeutic algorithm. *Laryngoscope*. 2014;124(2):498-503.
7. Damrose EJ. On the development of idiopathic subglottic stenosis. *Med Hypoth*. 2008;71:122-125.
8. Dumoulin E, Stather DR, Gelfand G, Maranda B, Maceachern P, Tremblay A. Idiopathic subglottic stenosis: a familial predisposition. *Ann Thorac Surg*. 2013;95(3):1084-1086.
9. Ashiku SK, Kuzucu A, Grillo HC, et al. Idiopathic laryngotracheal stenosis: effective definitive treatment with laryngotracheal resection. *J Thorac Cardioavasc Surg*. 2004;127:99-107.
10. Shapshay SM, Beamis JF, Hybels RL, et al. Endoscopic treatment of subglottic and tracheal stenosis by radial laser incision and dilation. *Ann Otol Rhinol Laryngol*. 1987;96:661-664.
11. Parker NP, Bandyopadhyay D, Misono S, Goding GS Jr. Endoscopic cold incision, balloon dilation, mitomycin C application, and steroid injection for adult laryngotracheal stenosis. *Laryngoscope*. 2013;123(1):220-225.
12. Rutter MJ, Cohen AP, de Alarcon A. Endoscopic airway management in children. *Curr Opin Otolaryngol Head Neck Surg*. 2008;16(6):525-529.
13. Wicke C, Halliday B, Allen D, et al. Effects of steroids and retinoids on wound healing. *Arch Surg*. 2000;135(11):1265-1270.
14. Perepelitsyn I, Shapshay SM. Endoscopic treatment of laryngeal and tracheal stenosis—has mitomycin C improved the outcome? *Otolaryngol Head Neck Surg*. 2004;131(1):16-20.
15. Hata T, Hoshi T, Kanamori K, et al. Mitomycin, a new antibiotic from Streptomyces. I. *J Antibiot*. 1956;9(4):141-146.
16. McGaha TL, Phelps RG, Spiera H, Bona C. Halofuginone, an inhibitor of type-I collagen synthesis and skin sclerosis, blocks transforming-growth-factor-beta-mediated Smad3 activation in fibroblasts. *J Invest Dermatol*. 2002;118(3):461-470.
17. Cincik H, Gungor A, Cakmak A, et al. The effects of mitomycin C and 5-fluorouracil/triamcinolone on fibrosis/scar tissue formation secondary to subglottic trauma. *Am J Otolaryngol*. 2005;26(1):45-50.
18. Empey DW. Assessment of upper airways obstruction. *Br Med J*. 1972;3:503-505.
19. Rotman HH, Liss HP, Weg JG. Diagnosis of upper airway obstruction by pulmonary function testing. *Chest*. 1975;68(6):796-799.
20. Hyatt RE. Evaluation of major airway lesions using the flow volume loop. *Ann Otol Rhinol Laryngol*. 1975;84:635-642.
21. Wassermann K, Gitt A, Weyde J, Eckel HE. Lung function changes and exercise-induced ventilatory responses to external resistive loads in normal subjects. *Respiration*. 1995;62(4):177-184.
22. Nouraei SAR, Winterborne C, Nouraei M, et al. Quantifying the physiology of laryngotracheal stenosis: changes in pulmonary dynamics in response to graded extrathoracic resistive loading. *Laryngoscope*. 2007;117:581-589.
23. Simpson GT, Strong MS, Healy GB, et al. Predictive factors of success or failure in the endoscopic management of laryngeal and tracheal stenosis. *Ann Otol Rhinol Laryngol*. 1982;91:384-388.
24. Nouraei SAR, Sandhu GS. Outcome of a multimodality approach to the management of idiopathic subglottic stenosis. *Laryngoscope*. 2013;123:2474-2484.
25. Perotin JM, Jeanfaivre T, Thibout Y, et al. *Ann Thoracic Surg*. 2011;92:297-301.
26. Hseu AF, Benninger MS, Haffey TM, Lorenz R. Subglottic stenosis: a ten-year review of treatment outcomes. *Laryngoscope*. 2014;124(3):736-741.
27. Hirshoren N, Eliashar R. Wound-healing modulation in upper airway stenosis—myths and facts. *Head Neck*. 2009;31(1):111-126.
28. Farmer W, Littner MR, Gee BL. Assessment of therapy of upper airway obstruction. *Arch Intern Med*. 1977;137:309-312.
29. Roediger FC, Orloff LA, Courey MS. Adult subglottic stenosis: management with laser incisions and mitomycin-C. *Laryngoscope*. 2008;118(9):1542-1546.
30. Sterner JB, Morris MJ, Sill JM, Hayes JA. Inspiratory flow-volume curve evaluation for detecting upper airway disease. *Resp Care*. 2009;54(4):461-466.
31. Ewald FW, Tenholder MF, Waller R. Analysis of the inspiratory flow-volume curve: should it always precede the forced expiratory maneuver? *Chest*. 1994;106(3):814-818.
32. Miller RD, Hyatt RE. Obstructing lesions of the larynx and trachea: clinical and physiologic characteristics. *Mayo Clin Proc*. 1969;44(3):145-161.

Reprinted by permission of JAMA Otolaryngol Head Neck Surg. 2013; 139(1):76-81.

ORIGINAL ARTICLE

Clinical Manifestations and Treatment of Idiopathic and Wegener Granulomatosis–Associated Subglottic Stenosis

Stanford C. Taylor, BS; Daniel R. Clayburgh, MD, PhD; James T. Rosenbaum, MD; Joshua S. Schindler, MD

Objective: To compare and contrast the manifestations and surgical management of subglottic stenosis in patients with airway obstruction attributed to granulomatosis with polyangiitis (GPA), previously known as Wegener granulomatosis, and those with idiopathic subglottic stenosis (iSGS).

Design: Retrospective medical chart review. Review of subglottic stenosis cases seen in the otolaryngology department of an academic medical center from 2005 through 2010. Data were obtained on disease presentation, operative management. and findings.

Setting: Tertiary referral center.

Participants: A total of 24 patients with iSGS and 15 patients with GPA-associated subglottic stenosis (GPA-SGS).

Results: All individuals with iSGS were female, and 40% of patients with GPA-SGS were male ($P<.01$). Patients with iSGS tended to have a higher Myer-Cotton stenosis grade at the time of dilation than those with GPA-SGS ($P=.02$). Individuals with GPA-SGS were more likely to undergo tracheotomy as a result of disease-related complications than individuals with iSGS ($P<.01$). No patients with an open airway reconstruction in the iSGS group required follow-up mechanical dilation. In contrast, all patients with open airway reconstructions in the GPA-SGS group underwent more than 1 subsequent airway dilation ($P<.01$).

Conclusions: While surgical utilization is the mainstay of treatment in iSGS and GPA-SGS, iSGS occurs almost exclusively in females and presents with a greater degree of stenosis at the time of endoscopic dilation. In contrast, GPA-SGS is associated with greater rates of tracheotomy. Open airway reconstruction may be used in the treatment of iSGS and GPA-SGS but is much more effective in iSGS.

JAMA Otolaryngol Head Neck Surg. 2013;139(1):76-81

Author Affiliations: School of Medicine (Mr Taylor), Department of Otolaryngology–Head and Neck Surgery (Drs Clayburgh and Schindler), and Division of Arthritis and Rheumatic Diseases, Department of Medicine (Dr Rosenbaum), Oregon Health and Science University, Portland.

Acquired subglottic stenosis (SGS) describes airway narrowing in the area of the cricoid and is associated with prolonged intubation or external trauma. Other patients acquire SGS from a systemic autoimmune etiology, such as granulomatosis with polyangiitis (GPA), previously known as Wegener granulomatosis. However, in about 20% of cases there is no identifiable precipitant, and patients are therefore classified as having idiopathic subglottic stenosis (iSGS).[1] While posttraumatic SGS is often evident based on patient history, discriminating between iSGS and GPA-SGS can be diagnostically challenging, especially in cases in which SGS is the presenting symptom of GPA. Presenting symptoms of dyspnea, noisy breathing, and voice changes can occur in both groups.[2,3] Anatomically, the cohorts appear different, as scar formation in patients with GPA-SGS and iSGS typically involves soft tissue (membranous stenosis), but intubation and trauma-related SGS typically leads to firm, cartilaginous scar tissue.[4] GPA-SGS is diagnosed based on the presence of autoantibodies or characteristic findings on biopsy, but negative results cannot reliably rule out autoimmune-mediated SGS, especially when the disease has limited organ involvement.[5,6] Indeed, given the many similarities between patients with iSGS and those with GPA-SGS, some have suggested that iSGS may represent a spectrum of autoimmune-mediated SGS that is not yet fully understood and that GPA-SGS is but a part of this autoimmune SGS spectrum.

Despite the similarities noted between iSGS and GPA-SGS, there is a paucity of data in the literature directly comparing these entities, specifically with regard to therapeutic and surgical outcomes. Given

the systemic nature of GPA, it seems reasonable to assume that these patients would have more airway difficulty than those with idiopathic SGS. Furthermore, one would predict that patients with GPA would have a worse clinical response to surgical treatment, including a need for more procedures and more frequent failure of open procedure, although data to support this are lacking. To further clarify this, we undertook this study to compare the manifestations and surgical management for SGS in patients with airway obstruction attributed to GPA and those with iSGS.

METHODS

Following approval by the Oregon Health and Science University institutional review board, a retrospective review was performed of medical records of patients with GPA-SGS or iSGS seen in the otolaryngology department at our institution from 2005 through 2010. To identify patients, the department billing records and operative records were queried for all encounters associated with a diagnosis of GPA, laryngotracheal stenosis, or patients who had undergone endoscopic airway dilation or open airway reconstruction. Patients were classified as having GPA if (1) they had at least 1 clinical feature, such as SGS, consistent with the disease, and they were antineutrophil cytoplasmic antibody (ANCA) positive; (2) they had a biopsy finding consistent with GPA (granulomatous inflammation, vasculitis, and/or rapidly progressive glomerular nephritis); or (3) they manifested at least 2 signs of disease (laryngotracheal involvement, septal perforation, sinonasal involvement, nasolacrimal involvement, recurrent otitis media, or characteristic renal or pulmonary involvement). Patients were classified as having iSGS if they did not have a history of laryngotracheal trauma or tracheotomy and the airway narrowing could not be attributed to another cause, such as malignant disease or a systemic autoimmune condition.

Once patients were identified, all documentation, including pre-2005 encounters, was reviewed. Data were obtained on age at diagnosis, diagnostic procedures and laboratory tests, and therapeutic management, including immunosuppressive therapy and surgical procedures. SGS was diagnosed and evaluated by flexible fiber-optic examination or by intraoperative direct laryngoscopy. Extent of laryngotracheal involvement and gross characteristics of lesions were assessed. The Myer-Cotton staging system (MCS), which was originally developed as a pediatric SGS scale but has since been implemented in monitoring adult SGS, was used to describe the stenosis based on the percentage relative reduction in cross-sectional area of the subglottis. Four grades of stenosis are described: grade 1 lesions have less than 50% obstruction, grade 2 lesions have 51% to 70% obstruction, grade 3 lesions have 71% to 99% obstruction, and grade 4 lesions have no detectable lumen or complete stenosis.[7] Grade of stenosis was not documented in the operative record in 17 dilations. Airway dilations were performed with direct microlaryngoscopy using continuous radial expansion balloons (Boston Scientific), Jackson laryngeal dilators, or rigid bronchoscopic dilation. When comparing rates of surgical outcome, a minimum follow-up time of 6 months after the operative date was required for inclusion. Gastroesophageal reflux disease was diagnosed via esophagoscopy demonstrating esophagitis in 3 patients, pH probe testing in 5 patients, and clinical improvement of reflux symptoms with proton pump inhibitor therapy in 4 patients.

Descriptive statistics, *t* test, Fisher exact test, Mann-Whitney U test, and χ^2 analysis for categorical data were performed ($P < .05$ denoted significance).

Table 1. Background Information

Characteristic	GPA (n = 15)	Idiopathic (n = 24)	*P* Value
Sex			
Female	9	24	<.01
Male	6	0	
Ethnicity			
Hispanic	0	1	>.99
White, non-Hispanic	15	23	
Comorbidities			
GERD	4	12	.19
Asthma	1	2	>.99
Lifetime tracheotomy history	6	0	<.01
Age, median, y			
At GPA diagnosis	31.7	NA	
At SGS diagnosis	36.3	45.2	.24

Abbreviations: GERD, gastroesophageal reflux disease; GPA, granulomatosis with polyangiitis; NA, not applicable; SGS, subglottic stenosis.

RESULTS

A total of 39 patients were identified for the study, 24 with iSGS and 15 with GPA-associated subglottic airway obstruction (**Table 1**). Aside from the absence of males in the iSGS group, no significant differences in patient demographics were noted. Of note, 4 patients with GPA-SGS in our cohort (27%) were diagnosed as having GPA when younger than 20 years. At the date of last follow-up, 7 patients with GPA-SGS (47%) exhibited disease involvement restricted to the head and neck while 8 (53%) had systemic involvement, including renal and/or pulmonary manifestations. The cohort was followed for 177 patient-years. The mean and median periods of follow-up for the GPA-SGS group were 8.2 and 9.9 years, respectively. In comparison, the mean and median lengths of follow-up for the iSGS group were 2.8 and 1.8 years, respectively ($P < .01$).

Diagnosis of the 15 patients with GPA is illustrated as follows:

Patients who are ANCA positive, %	93
Patients with biopsy proven diagnosis, %	47
Patients who are ANCA positive and with diagnostic biopsy, No.	7
Patients with ANCA and nondiagnostic biopsy, No.	7
Patients diagnosed by clinical features alone, No.	1

The patient diagnosed by clinical features alone was male and, as is typical with GPA,[8] had additional disease involving the nose and sinuses. Furthermore, although his autoantibody ANCA titers were not positive, they were interpreted as having an atypical pattern. Given these considerations, his SGS was attributed to GPA rather than an idiopathic etiology.

The severity and location of stenosis observed during dilation was assessed. Patients with iSGS were found to have significantly worse stenosis based on MCS grading than patients with GPA-SGS (**Table 2**). There were no significant differences in the location of stenosis seen at initial dilation, although there was a trend toward more circumferential stenoses in the patients with GPA (**Table 3**).

©2013 American Medical Association. All rights reserved.

Table 2. Myer-Cotton Staging (MCS) at Time of Endoscopic Dilation

Characteristic	No. (%)		
	GPA	Idiopathic	Combined
Patients undergoing endoscopic dilation. No.	14	23	38
Dilations with known MCS, No.	36	50	86
MCS[a]			
1	8 (22.2)	2 (4)	10 (11.6)
2	16 (44.4)	14 (28)	30 (34.9)
3	12 (33.3)	34 (68)	46 (53.5)
4	0	0	0

Abbreviation: GPA, granulomatosis with polyangiitis.
[a] $P = .02$.

Table 3. Stenosis Location at Time of First Endoscopic Dilation

Characteristic	No. (%)		
	GPA	Idiopathic	Combined
Patients undergoing endoscopic dilation, No.	14	23	38
Dilations with known stenosis location at time of first endoscopic dilation, No.	11	19	30
Stenosis location[a]			
Anterior	1 (9.1)	8 (42.1)	9 (30)
Posterior	1 (9.1)	2 (10.5)	3 (10)
Circumferential	9 (81.2)	9 (47.4)	18 (60)

Abbreviation: GPA, granulomatosis with polyangiitis.
[a] $P = .42$.

Surgical treatment of SGS consisted of both endoscopic dilation and cricotracheal resection. Endoscopic dilation technique did not vary substantially between groups, aside from less frequent use of the carbon dioxide laser in the GPA group (**Table 4**). Patients with GPA underwent a mean of 3.53 surgical dilations per patient compared with 2.54 in those with iSGS ($P = .44$). Seven patients with GPA-SGS (47%) required fewer than 2 airway dilations compared with 11 of those with iSGS (46%) ($P < .99$).

Definitive operative resection or reconstruction was attempted in both groups; 5 of those with GPA-SGS (33%), and 6 of those with iSGS (25%). While no patients with an open airway reconstruction in the iSGS group required follow-up mechanical dilation, all patients with open airway reconstructions in the GPA-SGS group underwent more than 1 subsequent airway dilation ($P < .01$). Following open airway reconstruction, 1 patient with GPA-SGS underwent subsequent tracheotomy. Open airway reconstruction led to permanent decannulation of 2 previously tracheotomy-dependent patients with GPA-SGS. Six patients with GPA-SGS (40%) underwent tracheotomy as a result of disease-related complications and 2 (13%) remained tracheotomy dependent at the date of last follow-up. No patients with iSGS required tracheotomy as a result of a disease-related complications ($P < .01$).

We examined the impact of various factors on the success of airway procedures within each group (**Table 5** and **Table 6**). The presence of gastroesophageal reflux disease (GERD) and the operative use of carbon dioxide laser were not found to have an impact on the rate of surgical utilization. Male patients with GPA-SGS had a shorter time until additional procedures were needed than female patients with GPA-SGS, while the presence of a previous tracheostomy showed a nonsignificant trend toward worse outcomes. When patients from both cohorts (GPA-SGS and iSGS) were pooled into a single group, none of these factors (sex and history of tracheostomy or GERD) had any significant impact on time until additional procedures were needed.

We analyzed the utilization of systemic immunotherapy within both groups. All patients with GPA-SGS and 50% of iSGS individuals received systemic immunotherapy at some point throughout follow-up as part of disease management; all patients with iSGS received corticosteroids, while patients with GPA received a mix of corticosteroids (n = 13), methotrexate sodium (n = 11), and cyclophosphamide (n = 9). Use of immunosuppressive medication was not associated with longer procedure-free intervals.

COMMENT

When traumatic causes are not readily identifiable by patient history, determining the etiology of SGS can be diagnostically challenging. While many patients with nontraumatic SGS may have a systemic autoimmune condition such as GPA, many others will have an unrevealing autoimmune workup. In the absence of any identifiable cause, these patients are considered to have iSGS, although there is some speculation that this may be due to some unknown autoimmune mechanism. This study was conducted to better define the similarities and differences in presentation and therapeutic management of iSGS and GPA-SGS.

Previously, it has been demonstrated that GPA-SGS affects men and women equally.[2] In contrast, iSGS almost exclusively affects women[3,9,10]; it is thought to predominantly affect women owing to estrogen-mediated alterations to wound-healing responses in the subglottic airway.[11] Our study is consistent with these observations, further confirming the tendency for iSGS to disproportionately affect men and for GPA-SGS to affect both men and women. Interestingly, we observed that male GPA-SGS patients underwent more frequent subglottic airway surgical procedures than female patients with GPA-SGS. This may be due, in part, to previous trends noting that male patients with GPA tend to develop a more severe form of the disease than female patients with GPA.[12]

The median age of initial presentation in patients with GPA-SGS (36.3 years) was almost 9 years younger than that of the iSGS group (45.2 years). Although not statistically significant ($P = .24$), this finding is broadly in line with that of previous research, and we speculate that had our cohort been larger, the observations would have been significant. A study of 52 patients with iSGS found the average age of initial presentation to be 43.5 years,[9]

©2013 American Medical Association. All rights reserved.

Table 4. Operative Technique at Time of Endoscopic Dilation

Characteristic, No.	GPA	Idiopathic	Combined	*P* Value
Patients undergoing endoscopic dilation	14	23	38	NA
Total dilations	48	55	103	NA
Operative use of laser	15	35	50	<.01
Intraoperative injections				
No injection	10	5	15	.10
Triamcinolone	31	39	70	.53
Mitomycin C	1	1	2	>.99
Triamcinolone and mitomycin C	6	10	16	.59

Abbreviations: GPA, granulomatosis with polyangiitis; NA, not applicable.

Table 5. Therapeutic Airway Procedure Frequency[a]

	Granulomatosis With Polyangiitis, No.					Idiopathic, No.				
Characteristic	Patients	Procedures	*P* Value	DBP, Mean	*P* Value	Patients	Procedures	*P* Value	DBP, Mean	*P* Value
All patients	15	48		557		24	46		495	NA
GERD history										
Yes	4	9	.74	565	.96	12	32	.35	506	.60
No	11	39	.74	555	.96	12	14		470	
Lifetime tracheotomy history										
Yes	6	33	.24	367	.11	NA	NA	NA	NA	NA
No	9	15	.24	975	.11	NA	NA	NA	NA	NA
Sex										
Male	6	31	.24	318	.04	NA	NA	NA	NA	NA
Female	9	17	.24	994	.04	NA	NA	NA	NA	NA

Abbreviations: DBP, days between procedures; GERD, gastroesophageal reflux disease; NA, not applicable.
[a] Excludes surgical procedures with less than 6 months' postoperative follow-up.

whereas the median age at which SGS was diagnosed in a cohort of patients with GPA was 26 years.[2] Furthermore, patients with GPA and SGS are frequently diagnosed as having GPA at a very young age; in fact, up to 44% are diagnosed before the age of 20 years.[13] 27% of patients with GPA-SGS in our cohort were diagnosed as having GPA when younger than 20 years.

GERD has been implicated in the development of SGS and has been identified as a probable precipitant of iSGS.[14-17] However, some question the existence of a direct association.[9] GERD has also been explored as a possible cause of GPA-SGS, but evidence of a definitive link has yet to be identified.[18] The most compelling data to date come from a study by Blumin and Johnston[19] demonstrating pepsin in the larynx and trachea in 59% of patients with iSGS, but none in matched control patients. Half of our iSGS group either had a history of, or was empirically treated for, GERD, which was statistically no different from the comparison GPA-SGS group. Furthermore, the rate of surgical utilization between those with a diagnosis of GERD and those without was no different in both groups. While our results fail to demonstrate a difference in the rate of GERD and SGS in the iSGS and GPA-SGS groups, understanding the impact of GERD on the development of SGS will be best accomplished through continued prospective studies.

Operative management strategies for subglottic stenosis are focused on improving the airway, either via endoscopic dilation of the stenosis, excision of the stenosis with laryngotracheal reconstruction, or bypassing the stenosis with tracheostomy. Carbon dioxide laser resection and/or intralesional corticosteroid injection are common adjuvant treatments to endoscopic dilation. Interestingly, in our series we found that laser resection was utilized more frequently in patients with iSGS than in those with GPA-SGS. This may in part be explained by practices of the operating surgeon or a reluctance to use the carbon dioxide laser if there is a possibility of active

Table 6. Endoscopic Dilation Frequency[a] Based on Myer-Cotton Staging (MCS) at Time of Endoscopic Dilation

Characteristic	GPA	Idiopathic	*P* Value
Patients undergoing endoscopic dilation, No.	14	23	NA
Dilations with known MCS	36	50	NA
Days between procedures based on MCS, mean, No.			
1	829	358	NA[b]
2	562	602	.03
3	462	477	.23
4	NA	NA	NA

Abbreviations: GPA, granulomatosis with polyangiitis; NA, not applicable.
[a] Excludes surgical procedures with less than 6 months' postoperative follow-up.
[b] Only 1 dilation in the idiopathic MCS 1 group with more than 6 months' postoperative follow-up.

GPA within the stenosis. While previous studies have shown the successful use of the carbon dioxide laser for GPA-SGS,[20] a general principle in the treatment of GPA-SGS is to avoid as much airway manipulation as possible when active disease is present.

Intralesional corticosteroid injections at the time of manual dilation were documented in 85% of cases. Previous research of an intratracheal dilation-injection technique using glucocorticoids in GPA-SGS has shown this to be effective[21] and possibly a preferred method of immunosuppressive therapy in GPA isolated strictly to the subglottis.[2] While the role of intralesional corticosteroids in iSGS is less clear, it is often considered an adjunct to dilation to prolong the time between procedures. Definitive treatment of iSGS is thought to be most likely achieved with open airway reconstruction.[3,10] More recently, mitomycin C has also been used as an inhibitor of fibroblastic-mediated scar formation in laryngotracheal stenosis.[11,22,23] In our cohort, only 2 patients received mitomycin C; thus, it is impossible to derive any conclusions about this therapy. Future research will be needed to better define its role in the management of GPA-SGS and iSGS.

Our data demonstrated more severe stenosis as measured by MCS at the time of dilation in patients with iSGS than those with GPA-SGS, with 33% of dilations in patients with GPA-SGS and 68% of dilations in patients with iSGS classified as MCS 3. Although there is a perceived reluctance to operate on patients with GPA, our experience indicates they often undergo dilation for smaller degrees of stenosis. This may indicate that patients with GPA-SGS experience more clinically significant symptoms than those with iSGS for a given grade of stenosis. Alternatively, underlying sinonasal or pulmonary involvement with consequent increased work of breathing in individuals with GPA-SGS could explain the larger diameter airway at the time of dilation. Patients with GPA-SGS may also have longer or more irregular sections of stenosis that result in more turbulence and poorer airflow than those with iSGS with comparatively discrete and symmetric stenoses. Further research will be needed to explore the differences between the dyspnea in these 2 groups.

Although the MCS at the time of dilation was different between the 2 groups, we found identical percentages of patients with GPA-SGS and iSGS undergoing more than 1 endoscopic dilation and open airway reconstruction. Our data further indicate that iSGS cases classified as having MCS 2 have more days between mechanical airway dilations than those with GPA-SGS. A similar trend, albeit not significant, is noted in average days between dilations in patients classified as having MCS 3.

The percentage of patients with GPA-SGS undergoing open airway reconstruction in our group is similar to those of other studies,[13,20] and we recently reported on the efficacy of airway reconstruction in GPA-related laryngotracheal stenosis.[24] It is important to note that open airway reconstruction was much more effective for iSGS than for GPA-SGS. In patients with iSGS, open airway reconstruction could be considered definitive management, with no need for tracheostomy afterward and rare need for further airway interventions. However, in GPA-SGS, further dilation is the norm; as we have shown previously, the major benefit of open airway reconstruction for GPA-SGS is to effect decannulation.[24] Forty percent of patients with GPA-SGS in our cohort required tracheotomy as part of disease treatment. This is consistent with other research demonstrating that between 41% and 52% of patients with GPA-SGS require tracheotomy.[2,13,20] No patients with iSGS in our cohort required tracheotomy owing to disease-related complications, which is less than a previous study showing a 20% tracheotomy rate.[25]

Owing to the nature of retrospective medical chart reviews, our study has several inherent limitations. While specialists at tertiary referral centers follow patients with chronic medical conditions longitudinally for many years, the same is not true of some conditions, such as SGS, that may resolve after 1 or more treatments. The mean length of follow-up for patients with iSGS at our institution was 2.8 years compared with 8.2 years in those with GPA-SGS. This discrepancy, although informational for comparing disease chronicity between groups, does not allow for an accurate comparison of rate of surgical dilations over time. Thus, we viewed the observation that patients with GPA-SGS undergo less frequent surgical utilization as being due to loss of follow-up and attributable to the nature of retrospective reviews involving tertiary referral centers. It should be noted we also explored the possibility of systemic immunosuppressive therapy, which was used by all patients with GPA-SGS and half of those with iSGS, as an additional factor contributing to the decreased rate of surgical utilization in the GPA-SGS group. However, owing to a lack of numbers and prescribing variability between patients and procedures, the effect of immunosuppressive therapy on time between airway procedures could not be accurately assessed. The true rate of surgical utilization and systemic immunosuppressive therapy efficacy in these patients will be best determined by future prospective studies. In further considering the operative demand of both diseases, it is important to note that the median number of airway procedures in both the GPA-SGS and iSGS groups was 1 surgical intervention per patient. Thus, while both groups had individuals requiring chronic follow-up and multiple airway procedures, 1 operation provided definitive treatment for many of the patients.

When individually considering the duration of follow-up of the GPA-SGS group, the finding of an average of 0.47 surgical dilations per patient-year of follow-up could be viewed as representative for patients with GPA-SGS requiring referral to a head and neck surgeon. A similarly constructed previous study that followed patients for an average of 6.4 years demonstrated a comparable rate of surgical utilization in those with GPA-SGS: 0.36 surgical procedures per patient-year of follow-up.[20]

In conclusion, although several similarities exist between GPA-SGS and iSGS, iSGS occurs more often in women and presents with a greater degree of stenosis. GPA-SGS requires more long-term management and is associated with a higher rate of tracheotomy. While open airway reconstruction may be used in both iSGS and GPA-SGS, it is much more effective in iSGS. Within the GPA-SGS group, the rate of surgical utilization in individuals with GPA-SGS requiring tracheotomy was not significantly different from those whose disease did not re-

quire tracheotomy and male patients with GPA-SGS required more frequent subglottic airway procedures than female patients with GPA-SGS.

Submitted for Publication: July 13, 2012; final revision received September 27, 2012; accepted October 15, 2012.
Correspondence: Joshua S. Schindler, MD, Department of Otolaryngology–Head and Neck Surgery, Oregon Health and Science University, 3181 SW Sam Jackson Park Rd, Mail Code PV-01, Portland, OR 97239 (schindlj @ohsu.edu).
Author Contributions: All authors had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Taylor, Clayburgh, Rosenbaum, and Schindler. *Acquisition of data:* Taylor, Clayburgh, Rosenbaum, and Schindler. *Analysis and interpretation of data:* Taylor, Clayburgh, Rosenbaum, and Schindler. *Drafting of the manuscript:* Taylor and Clayburgh. *Critical revision of the manuscript for important intellectual content:* Clayburgh, Rosenbaum, and Schindler. *Statistical analysis:* Taylor and Clayburgh. *Study supervision:* Rosenbaum and Schindler.
Conflict of Interest Disclosures: Dr Rosenbaum is a paid consultant for Genentech.
Funding/Support: Research by Dr Rosenbaum is, in part, funded by Research to Prevent Blindness (New York), Genentech, the Stan and Madelle Rosenfeld Family Trust, the William and Mary Bauman Foundation, and the William C. Kuzell Foundation.
Previous Presentation: This study was presented at the Annual American Broncho-Esophagological Association Meeting; April 19, 2012; San Diego, California.
Additional Contributions: Jess Mace, MPH, provided assistance with the statistical analysis.

REFERENCES

1. Herrington HC, Weber SM, Andersen PE. Modern management of laryngotracheal stenosis. *Laryngoscope*. 2006;116(9):1553-1557.
2. Langford CA, Sneller MC, Hallahan CW, et al. Clinical features and therapeutic management of subglottic stenosis in patients with Wegener's granulomatosis. *Arthritis Rheum*. 1996;39(10):1754-1760.
3. Grillo HC, Mark EJ, Mathisen DJ, Wain JC. Idiopathic laryngotracheal stenosis and its management. *Ann Thorac Surg*. 1993;56(1):80-87.
4. Roediger FC, Orloff LA, Courey MS. Adult subglottic stenosis: management with laser incisions and mitomycin-C. *Laryngoscope*. 2008;118(9):1542-1546.
5. Finkielman JD, Lee AS, Hummel AM, et al; WGET Research Group. ANCA are detectable in nearly all patients with active severe Wegener's granulomatosis. *Am J Med*. 2007;120(7):643-, e9-e14.
6. Devaney KO, Travis WD, Hoffman G, Leavitt R, Lebovics R, Fauci AS. Interpretation of head and neck biopsies in Wegener's granulomatosis: a pathologic study of 126 biopsies in 70 patients. *Am J Surg Pathol*. 1990;14(6):555-564.
7. Myer CM III, O'Connor DM, Cotton RT. Proposed grading system for subglottic stenosis based on endotracheal tube sizes. *Ann Otol Rhinol Laryngol*. 1994; 103(4, pt 1):319-323.
8. Hoffman GS, Kerr GS, Leavitt RY, et al. Wegener granulomatosis: an analysis of 158 patients. *Ann Intern Med*. 1992;116(6):488-498.
9. Dedo HH, Catten MD. Idiopathic progressive subglottic stenosis: findings and treatment in 52 patients. *Ann Otol Rhinol Laryngol*. 2001;110(4):305-311.
10. Ashiku SK, Kuzucu A, Grillo HC, et al. Idiopathic laryngotracheal stenosis: effective definitive treatment with laryngotracheal resection. *J Thorac Cardiovasc Surg*. 2004;127(1):99-107.
11. Valdez TA, Shapshay SM. Idiopathic subglottic stenosis revisited. *Ann Otol Rhinol Laryngol*. 2002;111(8):690-695.
12. Stone JH; Wegener's Granulomatosis Etanercept Trial Research Group. Limited versus severe Wegener's granulomatosis: baseline data on patients in the Wegener's granulomatosis etanercept trial. *Arthritis Rheum*. 2003;48(8):2299-2309.
13. Lebovics RS, Hoffman GS, Leavitt RY, et al. The management of subglottic stenosis in patients with Wegener's granulomatosis. *Laryngoscope*. 1992;102 (12, pt 1):1341-1345.
14. Jindal JR, Milbrath MM, Shaker R, Hogan WJ, Toohill RJ. Gastroesophageal reflux disease as a likely cause of "idiopathic" subglottic stenosis. *Ann Otol Rhinol Laryngol*. 1994;103(3):186-191.
15. Maronian NC, Azadeh H, Waugh P, Hillel A. Association of laryngopharyngeal reflux disease and subglottic stenosis. *Ann Otol Rhinol Laryngol*. 2001;110 (7, pt 1):606-612.
16. Poetker DM, Ettema SL, Blumin JH, Toohill RJ, Merati AL. Association of airway abnormalities and risk factors in 37 subglottic stenosis patients. *Otolaryngol Head Neck Surg*. 2006;135(3):434-437.
17. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope*. 1991;101(4, pt 2)(suppl 53):1-78.
18. Church AC, Goldsmith K, Sivasothy P. Aspiration and development of subglottic stenosis in patients with Wegener's granulomatosis. *J Laryngol Otol*. 2010; 124(4):393-396.
19. Blumin JH, Johnston N. Evidence of extraesophageal reflux in idiopathic subglottic stenosis. *Laryngoscope*. 2011;121(6):1266-1273.
20. Gluth MB, Shinners PA, Kasperbauer JL. Subglottic stenosis associated with Wegener's granulomatosis. *Laryngoscope*. 2003;113(8):1304-1307.
21. Hoffman GS, Thomas-Golbanov CK, Chan J, Akst LM, Eliachar I. Treatment of subglottic stenosis, due to Wegener's granulomatosis, with intralesional corticosteroids and dilation. *J Rheumatol*. 2003;30(5):1017-1021.
22. Smith ME, Elstad M. Mitomycin C and the endoscopic treatment of laryngotracheal stenosis: are two applications better than one? *Laryngoscope*. 2009; 119(2):272-283.
23. Perepelitsyn I, Shapshay SM. Endoscopic treatment of laryngeal and tracheal stenosis: has mitomycin C improved the outcome? *Otolaryngol Head Neck Surg*. 2004;131(1):16-20.
24. Wester JL, Clayburgh DR, Stott WJ, Schindler JS, Andersen PE, Gross ND. Airway reconstruction in Wegener's granulomatosis-associated laryngotracheal stenosis. *Laryngoscope*. 2011;121(12):2566-2571.
25. Giudice M, Piazza C, Foccoli P, Toninelli C, Cavaliere S, Peretti G. Idiopathic subglottic stenosis: management by endoscopic and open-neck surgery in a series of 30 patients. *Eur Arch Otorhinolaryngol*. 2003;260(5):235-238.

Hsc

Home Study Course

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Empowering otolaryngologist—head and neck surgeons to deliver the best patient care

1650 Diagonal Road, Alexandria, Virginia 22314-2857 U.S.A.
www.entnet.org