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

Neoplastic and Inflammatory Diseases of the Head and Neck

Hsc
Home Study Course

Home Study Course

Section 7
February 2017

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

THE
HOME STUDY COURSE
IN
OTOLARYNGOLOGY -- HEAD AND NECK SURGERY

February 2017

SECTION 7

Neoplastic and Inflammatory Diseases of the Head and Neck

SECTION FACULTY:

Mark E. Prince, MD******
Bhuvanesh Singh, MD PhD **
Jeffrey S. Wolf, MD, FACS**
Salvatore M. Caruana, MD
Ted H. Leem, MD, MS, FACS
Matthew O. Old, MD
Bradley Schiff, MD

American Academy of Otolaryngology - Head and Neck Surgery Foundation

Section 7 exam deadline: March 13, 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 neoplastic and inflammatory diseases of the head and neck. 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 on Neoplastic and Inflammatory Diseases of the Head and Neck.

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
The participant who has successfully completed this section should be able to:

1. Evaluate the benefits and limitation of in office ultrasound use by otolaryngologists.
2. Define the role of ultrasound in assessing thyroid/parathyroid and neck disease.
3. Recognize role of PET and PET/CT in the diagnosis and management of head and neck cancer.
4. Define role of PET-CT in the management and outcome of oropharynx cancers.
5. Incorporate preventive exercises to maintenance of structure and swallowing in patients undergoing chemoradiation therapy for head and neck cancers.
6. Consider high risk for late effects of chemoradiation treatment for head neck cancer on swallowing function.
7. Explain impact of prophylactic central neck dissection on risk for locoregional failure in head and neck cancers.
8. Consider role of radioactive iodine in development of second primary malignancies.
9. Describe association between thyroid cancer incidence and ease of access to health care.
10. Define role of sentinel node biopsies in management of oral and cutaneous squamous cell carcinomas.
11. Recognize effects of head and neck cancer of Health-related quality of life.
12. Define role of surgery in the management of oropharyngeal squamous cell carcinomas.
13. Define role of novel systemic and injected therapies in the treatment of melanoma and thyroid cancer.
14. Explain effects of gastroesophageal reflux disease of the risk for head neck cancer.
15. Recognize effects of treatment with proton pump inhibitors and histamine 2 blockers on overall survival in patients with head and neck cancer.

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 Neoplastic and Inflammatory Diseases of the Head and Neck 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 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 completion 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 opportunity will be 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.**

2016-17 Section 7 NEOPLASTIC AND INFLAMMATORY DISEASES OF THE HEAD AND NECK FACULTY

Co-Chairs:

Mark E. Prince, MD, Professor and Interim Chair, Department of Otolaryngology-HNS, University of Michigan, Ann Arbor, Michigan
Disclosure: No relationships to disclose

Bhuvanesh Singh, MD PhD, Director, Laboratory of Epithelial Cancer Biology; Director, Speech and Hearing Center; Attending Surgeon, Head and Neck Service, Memorial Sloan-Kettering Cancer Center; Professor of Otolaryngology, Weill Medical College of Cornell University, New York, New York
Disclosure: No relationships to disclose

Jeffrey S. Wolf, MD, FACS Associate Professor and Medical Director; Associate Chair of Clinical Practice; Otolaryngology-Head and Neck Surgery Program in Oncology, University of Maryland School of Medicine, Baltimore, Maryland
Disclosure: Stock/Stock options: Maryland Development Center; Stock/Stock options: Tesserae Medical LLC; Stock/Stock Options: Aerea Medical LLC; Intellectual Property Rights: Tesserae Medical LLC.

Faculty:

Salvatore M. Caruana, MD, Associate Professor Department of Otolaryngology-Head and Neck Surgery, Columbia University College of Physicians and Surgeons; Director, Division of Head and Neck Surgery, New York, New York
Disclosure: Salary: Olympus

Ted H. Leem, MD, MS, FACS, Southern California Permanente Medical Group Downey, California
Disclosure: No relationships to disclose

Matthew O. Old, MD, FACS, Assistant Professor, Department of Otolaryngology-Head and Neck Surgery, The James Cancer Hospital and Solove Research Institute, Wexner Medical Center at The Ohio State University, Columbus, Ohio
Disclosure: No relationships to disclose

Bradley Schiff, MD, Associate Professor Department of Otorhinolaryngology-Head and Neck Surgery, Albert Einstein College of Medicine, Bronx, New York
Disclosure: No relationships to disclose.

Planner(s):

Linda Lee, AAO–HNSF Education Senior Manager	No relationships to disclose
Stephanie Wilson, Stephanie Wilson Consulting, LLC; Production Manager	No relationships to disclose
Alfred A. Simental, Jr, MD	No relationships to disclose
Richard V. Smith, MD, chair, Head and Neck Surgery Education	Expert Witness: Various legal firms

This 2017 Home Study Course Section does not include any discussion of drugs and devices that 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.

March 13, 2017: Suggested section 7 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

FEBRUARY 2017 SECTION 7
NEOPLASTIC AND INFLAMMATORY DISEASES OF THE HEAD AND NECK

I. HEAD AND NECK ULTRASOUND

II. APPLICATION OF PET/CT/MRI IN MANAGEMENT OF HEAD AND NECK CANCER

III. COMPLICATIONS: DYSPHAGIA PREVENTION AND MANAGEMENT AFTER THERAPY FOR HEAD AND NECK CANCER

IV. MANAGEMENT OF THYROID NODULES AND THYROID MALIGNANCY

V. LYMPHATIC SYSTEM: SENTINEL NODE BIOPSY

VI. QUALITY OF LIFE

VII. TREATMENT MODALITIES UPDATE: ROBOTIC SURGERY

VIII. TREATMENT MODALITIES UPDATE: IMMUNOTHERAPY

IX. INFLAMMATORY: GERD ROLE IN CANCER DEVELOPMENT AND PREVENTION

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

FEBRUARY 2017 SECTION 7
NEOPLASTIC AND INFLAMMATORY DISEASES OF THE HEAD AND NECK

II. Application of PET/CT/MRI in Management of Head and Neck Cancer

Summary: This paper looks at PET/CT for detecting residual/recurrent head and neck squamous cell carcinoma. The study is a meta-analysis that found that PET/CT has high sensitivity and specificity.

Summary: This paper prospectively examines how adding PET/CT to the work up of head and neck squamous cell carcinoma patients affects management and prognosis. The authors found that PET/CT changed the TNM stage in about one-third of patients, PET/CT work-up was more accurate than conventional work up, and patients upstaged by PET/CT work-up had a worse prognosis.

Summary: This paper presents a literature on the use of PET/CT in oropharyngeal squamous cell carcinoma.

III. Complications: Dysphagia Prevention and Management After Therapy for Head and Neck Cancer

Summary: This is a randomized controlled trial of 58 patients undergoing chemoradiation therapy (CRT) for head and neck cancer. Comparison groups were sham swallowing exercises, no exercises, and intense swallowing therapy during treatment. As measured by clinical swallowing outcome and by muscle bulk of the genioglossus and hyoglossus and mylohyoid muscles as determined by T2-weighted MRI, the intense therapy group did much better and had reduced muscle atrophy relative to the other groups. A short follow up, well-done study.

Summary: This is a retrospective study looking at all patients treated for head and neck cancer with chemoradiation therapy (CRT) and comparing the swallowing outcomes of patients who were compliant with swallowing therapy during treatment with those who were not compliant. Patients who were not compliant with speech and swallowing exercises during and after treatment did worse than patients who were compliant with regard to swallowing function as documented by FOSS scores in these groups. This study demonstrates that swallow rehabilitation and exercise can improve functional outcomes for patients receiving CRT or radiation therapy for head and neck cancer.

Summary: This paper reports on the results of a retrospective cohort study of all patients treated for larynx cancer at a single institution who would have met criteria for the 91-11 trial. Patients were carefully followed for the development of severe late dysphagia that developed after 5 years of follow up and therefore not reported in that trial. They identified that 26% of patients developed severe late dysphagia as a result of therapy after 5 years of follow up.

IV. Management of Thyroid Nodules and Thyroid Malignancy

Summary: This is a prospective case-controlled study comparing observation versus surgical intervention for patients with papillary thyroid microcarcinoma. The results show that observation is adequate for many, and that progression during observation does not adversely affect survival or salvage rates.

Summary: This is a meta-analysis of the locoregional recurrence and complications in patients who underwent prophylactic central neck dissection compared to those who did not. Evidence shows the benefit of the prophylactic central neck dissection in patients with N0 neck. Locoregional recurrence was reduced in patients undergoing central neck dissection.

Summary: Retrospective study of all patients treated with radioactive iodine (RAI) for differentiated thyroid cancer (DTC) within a single healthcare system in China. The 895 patients identified for study were followed for a minimum of 2 years; 645 patients received RAI as part of their treatment, while 249 patients did not. Controlling for other factors, RAI-positive and RAI-negative patients were compared the subsequent developments of second primary malignancies (SPMs). A statistically significant deference in the incidence of SPMs was noted in the RAI group, while the RAI-negative group had baseline levels of SPM development (13.5% vs 3.1%; $p = 0.015$). This study is one of several that strongly suggest that RAI therapy can have significant long-term effects on patients receiving this therapy and indirectly argues that RAI should be given selectively.

Summary: This study uses the SEER database and correlates the well-recognized increased incidence of papillary thyroid cancer (PTC) diagnosis in the U.S. to the availability of and access to healthcare among the more affluent population. The study, in conjunction with others, shows that the majority of the increased cases of PTC are from small, likely indolent, PTCs, and is driven by increased detection in an already existing pool of patients with subclinical disease.

V. Lymphatic System: Sentinel Node Biopsy

Summary: Sentinel lymph node biopsy using [99mTc]tilmanocept accurately predicted nodal status in oral cavity head and neck squamous cell carcinoma with a low false-negative rate, high negative predicative value, and high accuracy. This study demonstrates this may be a method used in conjunction with or in lieu of elective neck dissection, but future studies are warranted.

Summary: This study conducted a retrospective review of sentinel lymph node biopsy in cutaneous squamous cell carcinoma. Analysis by serial step sectioning and immunohistochemistry increased the sentinel lymph node biopsy positivity rate to 15.1%.

Summary: This paper presents a review of the role of sentinel lymph node biopsy in early-stage oral cavity carcinoma.

Summary: This is a prospective study of sentinel lymph node biopsy in oral cancer. The results show excellent sensitivity, positive predicative value, and survival when employed for oral cancer.

VI. Quality of Life

Summary: This study is a population-based longitudinal cohort study which attempts to identify sociodemographic, behavioral, and clinical factors associated with health-related quality of life (HRQOL) for head and neck cancer patients over time by administering a questionnaire at baseline, 22 months, and 42 months. Its strength is the largenumber of patients (587).

Summary: Quality of life (QOL) for older individuals with head and neck squamous cell carcinoma was examined using the SEER database. The records of 1653 patients were examined. The authors noted that QOL declines both before and after head and neck squamous cell carcinoma, and any observed posttreatment recovery is likely an artifact of shorter survival among individuals with the lowest QOL.

IX. Inflammatory: GERD Role in Cancer Development and Prevention

Summary: This study is a large population case-control study of head and neck cancer in North Carolina. The authors found no increased odds of head and neck cancer with self-reported heartburn symptoms or self-reported medical diagnosis of gastroesophageal reflux disease. These results held true for subgroup analysis for specific tumor sites as well.

Summary: This is a large prospective cohort of head and neck squamous cell carcinoma patients in which histamine receptor-2 antagonists (H2RAs) and proton pump inhibitor (PPI) use and treatment outcomes were examined. The findings demonstrated that both medications were significant prognostic factors for overall survival ,but that only H2RAs were associated with recurrence-free survival in HPV16-positive oropharyngeal squamous cell carcinoma patients.

2017 SECTION 7 ADDITIONAL REFERENCES

Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with the Afirma gene expression classifier. *J Clin Endocrinol Metab*. 2014; 99(1):119-125.

Ali S, Palmer FL, Yu C, et al. A predictive nomogram for recurrence of carcinoma of the major salivary glands. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(7):698-705.

Almeida JP, Sanabria AE, Lima EN, Kowalski LP. Late side effects of radioactive iodine on salivary gland function in patients with thyroid cancer. *Head Neck*. 2011; 33(5):686-690.

Annane D, Depondt J, Aubert P, et al. Hyperbaric oxygen therapy for radionecrosis of the jaw: a randomized, placebo-controlled, double-blind trial from the ORN96 study group. *J Clin Oncol*. 2004; 22(4):4893-4900.

Asher SA, White HN, Kejner AE, et al. Hemorrhage after transoral robotic-assisted surgery. *Otolaryngol Head Neck Surg*. 2013; 149(1):112-117.

Bhatti RM, Stelow EB. IgG4-related disease of the head and neck. *Adv Anat Pathol*. 2013; 20(1):10-16.

Caudell JJ, Schaner PE, Meredith RF, et al. Factors associated with long-term dysphagia after definitive radiotherapy for locally advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2009; 73(2):410-415.

Chang JS, Lo HI, Wong TY, et al. Investigating the association between oral hygiene and head and neck cancer. *Oral Oncol*. 2013; 49(10):1010-1017.

Chen AM, Chen LM, Vaughan A, et al. Head and neck cancer among lifelong never-smokers and ever-smokers: matched-pair analysis of outcomes after radiation therapy. *Am J Clin Oncol*. 2011; 34(3):270-275.

Chen AM, Daly ME, Farwell DG, et al. Quality of life among long-term survivors of head and neck cancer treated by intensity-modulated radiotherapy. *JAMA Otolaryngol Head Neck Surg*. 2014; 140(2):129-133.

Chia SH, Gross ND, Richmon JD. Surgeon experience and complications with Transoral Robotic Surgery (TORS). *Otolaryngol Head Neck Surg*. 2013; 149(6):885-892.

Cibas ES, Ali SZ. The Bethesda System for reporting thyroid cytopathology. *Thyroid*. 2009; 19(11):1159-1165.

Dziegielewski PT, Tekno, TN, Durmus K, et al. Transoral robotic surgery for oropharyngeal cancer: long-term quality of life and functional outcomes. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(11):1099-1108.

Erman AB, Collar RM, Griffith KA, et al. Sentinel lymph node biopsy is accurate and prognostic in head and neck melanoma. *Cancer*. 2012; 118(4):1040-1047.

Ettl T, Gosau M, Brockhoff G, et al. Predictors of cervical lymph node metastasis in salivary gland cancer. *Head Neck.* 2014; 36(4):517-523.

Giordano D, Valcavi R, Thompson GB, et al. Complications of central neck dissection in patients with papillary thyroid carcinoma: results of a study on 1087 patients and review of the literature. *Thyroid*. 2012; 22(9):911-917.

Hamdan AL, Sarieddine D. Laryngeal manifestations of rheumatoid arthritis. *Autoimmune Dis*. 2013; doi:10.1155/2013/103081. [Epub ahead of print].

Herman MP, Werning JW, Morris CG, et al. Elective neck management for high-grade salivary gland carcinoma. *Am J Otolaryngol*. 2013; 34(3):205-208.

King SN, Dunlap NE, Tennant PA, Pitts T. Pathophysiology of radiation-induced dysphagia in head and neck cancer. *Dysphagia*. 2016; 31(3):339-351.

Koss SL, Russell MD, Leem TH, et al. Occult nodal disease in patients with failed laryngeal preservation undergoing surgical salvage. *Laryngoscope*. 2014; 124(2):421-428.

Kupferman ME, Kubik MW, Bradford CR, et al. The role of sentinel lymph node biopsy for thin cutaneous melanomas of the head and neck. *Am J Otolaryngol*. 2014; 35(2):226-232.

Langevin SM, Michaud DS, Marsit CJ, et al. Gastric reflux is an independent risk factor for laryngopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2013; 22(6):1061-1068.

Manzoor NF, Russell JO, Bricker A, et al. Impact of surgical resection on survival in patients with advanced head and neck cancer involving the carotid artery. *JAMA Otolaryngol Head Neck Surg*. 2013; 139(11):1219-1225.

Nikiforov YE, Seethala RR, Tallini G, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol*. 2016; 2(8):1023-1029.

Nixon IJ, Wang LY, Ganly I, et al. Outcomes for patients with papillary thyroid cancer who do not undergo prophylactic central neck dissection. *Br J Surg*. 2016; 103(3):218-225.

Pisanu A, Porceddu G, Podda M, et al. Systematic review with meta-analysis of studies comparing intraoperative neuromonitoring of recurrent laryngeal nerves versus visualization alone during thyroidectomy. *J Surg Res*. 2014; 188(1):152-161.

Sawka AM, Thabane L, Parlea L, et al. A second primary malignancy risk after radioactive iodine treatment for thyroid cancer: a systematic review and meta-analysis. *Thyroid*. 2009; 19(5):451-457.

Sekulic A, Migden MR, Oro AE, et al. Efficacy and safety of vismodegib in advanced basal-cell carcinoma. *N Engl J Med*. 2012; 366(23):2171-2179.

Seup Kim B, Kang KH, Park SJ. Robotic modified radical neck dissection by bilateral axillary breast approach for papillary thyroid carcinoma with lateral neck metastasis. *Head Neck*. 2015; 37(1):37-45.

Sharma A, Patel S, Baik FM, et al. Survival and gastrostomy prevalence in patients with oropharyngeal cancer treated with transoral robotic surgery vs chemoradiotherapy. *JAMA Otolaryngol Head Neck Surg*. 2016; 142(7):691-697.

Strychowsky JE, Sommer DD, Gupta MK, et al. Sialendoscopy for the management of obstructive salivary gland disease: a systematic review and meta-analysis. *Arch Otolaryngol Head Neck Surg*. 2012; 138(6):541-547.

Sun GH, Peress L, Pynnonen MA. Systematic review and meta-analysis of robotic vs conventional thyroidectomy approaches for thyroid disease. *Otolaryngol Head Neck Surg*. 2014; 150(4):520-532.

Vashishta R, Gillespie MB. Salivary endoscopy for idiopathic chronic sialadenitis. *Laryngoscope*. 2013; 123(12):3016-3020.

VanderWalde NA, Meyer AM, Deal AM, et al. Effectiveness of chemoradiation for head and neck cancer in an older patient population. *Int J Radiat Oncol Biol Phys*. 2014; 89(1):30-37.

Wang TS, Cheung K, Farrokhyar F, et al. A meta-analysis of the effect of prophylactic central compartment neck dissection on locoregional recurrence rates in patients with papillary thyroid cancer. *Ann Surg Oncol*. 2013; 20(11):3477-3483.

Zhang Y, Dai J, Wu T, et al. The study of the coexistence of Hashimoto's thyroiditis with papillary thyroid carcinoma. *J Cancer Res Clin Oncol*. 2014; 140(6):1021-1026.

The Journal of Laryngology & Otology (2014), ***128***, *447–453*.

doi:10.1017/S0022215114000760

MAIN ARTICLE

Otolaryngologist-performed head and neck ultrasound: outcomes and challenges in learning the technique

K BADRAN[1], P JANI[1], L BERMAN[2]

Departments of [1]ENT and [2]Radiology, Addenbrooke's Hospital, Cambridge, UK

Abstract

Objective: To assess the feasibility and accuracy of otolaryngologist-performed ultrasound in evaluating head and neck pathology.

Method: An ENT trainee, who had undergone basic training in neck ultrasonography, performed this on patients referred with suspected neck pathology. The trainee recorded the presence and nature of any abnormality. Findings were compared with those from a repeated scan performed by an experienced head and neck radiologist.

Results: The study included 250 patients. The absence or presence of lesion as reported by the trainee correlated with the radiologist's findings in 207 cases (83 per cent). There were 144 true positives, 63 true negatives, 32 false negatives and 11 false positives, yielding a sensitivity of 82 per cent, specificity of 85 per cent and accuracy of 83 per cent. Of the 144 true positive lesions, 81 per cent were interpreted concordantly with the radiologist.

Conclusion: Neck ultrasonography performed by an otolaryngologist is less accurate than that performed by an experienced radiologist, but is still a useful adjunct to clinical assessment, facilitating assessment in a 'one-stop' clinical setting.

Key words: Ultrasonography; Neck; Abnormalities; Otolaryngology

Introduction

Ultrasound is a valuable diagnostic tool used in many areas of medicine. It has been described as quick, portable, non-invasive and cost effective, and does not involve ionising radiation.[1–3] In mainland Europe, it is almost the exception for the radiologist rather than the clinician to perform ultrasound in some specialties. However, in the UK, with the exception of obstetric ultrasound, radiologists and radiographically trained sonographers have traditionally provided a service from centralised departments of radiology, where equipment and manpower can be concentrated cost-effectively.

There are increasing demands for other medical specialists to utilise ultrasound as a direct adjunct to clinical examination, and in some specialties it is becoming an integral part of the physician's diagnostic armamentarium and training. This trend is likely to be exacerbated by the increase in referrals and shortage of radiologists.[4] A recent survey distributed by ENT UK discussed the prospects and usefulness of British otolaryngologists learning this skill. Additionally, there is a demand by some European training boards to incorporate ultrasound into clinical training and accreditation. The Royal College of Radiologists recognises that it is appropriate for medical practitioners other than clinical radiologists to develop skills in ultrasound.[5]

The role of head and neck ultrasound performed by the ENT clinician, and the ability of the clinician to carry out the ultrasound and accurately interpret the findings, have not been investigated. This prospective study essentially describes the learning process of an ENT trainee with no previous specialist imaging experience, in acquiring neck ultrasound skills.

Materials and methods

Training

An ENT trainee attended head and neck ultrasound sessions in the radiological ultrasound department of a large teaching hospital for 12 months. A well-established 2-day practical ultrasound course (The Head and Neck Ultrasound Workshop, Morriston Hospital, Swansea) provided a basic introduction. Thereafter, the trainee attended several sessions with one of the course faculty members, observing neck ultrasound examinations. Informal tutorials covered physics and

Accepted for publication 13 August 2013

instrumentation, and ultrasound anatomy of the neck. Early practical experience was gained by practising on normal volunteer colleagues.

Following this induction, the trainee worked alongside a consultant radiologist with over two decades of experience in head and neck ultrasound (LB). This consultant radiologist works closely with all clinical departments at our centre, including surgery, endocrinology and oncology, helping with the management of patients. Ultrasound sessions included a weekly dedicated 'head and neck lump' clinic. These sessions include patients with no palpable mass, which typically involves a search for an undiagnosed parathyroid lesion in a patient with hypercalcaemia. This arrangement afforded the trainee one-to-one mentorship.

Following the studies of normal volunteers, the second stage of the learning process involved 50 ultrasound examinations of clinical referrals observed by the radiologist. All examinations were repeated by the radiologist who provided immediate feedback to the trainee. These 50 examinations were excluded from the final analysis of the 250 cases that comprise the current study. If any aspect of the trainee's examination was considered technically suboptimal, and where time constraints permitted, the scan was repeated by the ENT trainee following the radiologist's study.

Learning objectives included the identification of variations in normal neck structures and anatomical relationships, the recognition of any deviation from normal, and correct interpretation of an abnormality. A systematic approach to examination was emphasised. This included comprehensive scanning of neck anatomical triangles, comparing both sides of the neck, and use of Doppler ultrasound where appropriate. Teaching included advanced use of the machine controls, to a much higher level than usually achieved by practitioners other than radiologists or sonographers.

Main study

After the induction and training period described above, the trainee undertook examinations on patients referred to the neck ultrasound clinic. The trainee's study and conclusion was compared with the examination and conclusion of the radiologist. The 'gold standard' was taken to be the radiologist's report rather than eventual surgical or histological diagnosis if biopsy or surgery was undertaken.

Examinations were performed with Toshiba Aplio XG ultrasound apparatus (Toshiba Medical Systems, Crawley, UK) using appropriate high-frequency linear array transducers. All patients referred with palpable neck masses were included. Scans were undertaken with the patient in a semi-recumbent position with neck extension.

Following the scan, the trainee completed a proforma, on which the trainee indicated the presence or absence of a lesion, and commented on its nature and significance. If the lesion was considered indeterminate, the most likely diagnosis was described. Minor findings (e.g. reactive lymph nodes) were considered as lesions and were included in our analysis. The radiologist repeated the study and completed a similar proforma. It was not possible to blind the radiologist to the ultrasound findings described by the trainee because of time constraints and the evaluation process: as part of the evaluation, the radiologist scrutinised, and, if necessary, criticised and corrected the trainee's scanning technique.

Anonymised data were entered into a database. Results were placed in one of five categories (Table I): true negative (normal study), true positive (abnormal study), false negative (missed abnormality), false positive (normal study misinterpreted as abnormal), and misinterpretation (abnormality detected, but the nature or significance misinterpreted). There were therefore two aspects to the trainee's assessment. Firstly, identifying whether an abnormality was present, and secondly correctly interpreting any abnormal findings.

TABLE I
OUTCOME CATEGORIES AND DEFINITIONS

Category	Definition
True negative	No lesion is detected by trainee or radiologist; patient is reassured on same visit
True positive	Lesion is detected by both trainee & radiologist; trainee is asked to interpret nature of lesion
False negative	Lesion is not detected (i.e. is missed) by trainee but is detected by radiologist
False positive	Lesion is 'detected' by trainee but not radiologist; typically a normal structure misinterpreted as pathological
Misinterpretation	Lesion is detected by both trainee & radiologist (i.e. true positive), but nature of lesion is misinterpreted by trainee

Results

A total of 250 consecutive patients with suspected head and neck masses who attended over a 12-month period were included in the study. The median patient age was 50 years, with a male to female ratio of 1:1.7. The range of clinically suspected pathologies at the time of referral is shown in Table II.

Scans performed by the trainee indicated a positive finding in 155 patients. The findings of radiological repeat examinations concurred with the trainee's study in 144 examinations (true positives). Eleven of

TABLE II
SUSPECTED PATHOLOGY

Diagnosis on referral	Patients (*n* (%))
Anterior triangle lump	72 (29)
Posterior triangle lump	26 (10)
Thyroid	60 (24)
Parotid	37 (15)
Submandibular or submental	37 (15)
Parathyroid	18 (7)
Total	250 (100)

TABLE III
TRAINEE FALSE POSITIVE RESULTS*

Pathology	Trainee's misinterpretation	Radiologist's correct impression	Normal structure misinterpreted as pathological
Thyroid	Thyroiditis	Normal	Normal thyroid gland but thickened isthmus
Thyroid	Thyroid nodule	Normal	Normal heterogeneous thyroid gland
Parathyroid	Adenoma	No adenoma	Normal section in lower thyroid lobe
Parathyroid	Adenoma	No adenoma	Normal section in oesophagus
Submandibular†	Stone	Normal	Normal section in hyoid bone
Submandibular	Dilated duct	Normal	Normal section in mylohyoid muscle
Submandibular	Dilated duct	Normal	Normal section in blood vessel
Submandibular	Impinging ranula (mylohyoid defect)	Normal	Normal section in blood vessel passing through mylohyoid
Anterior triangle†	LN	Normal	Normal section in SCM

*11 patients. †$n = 2$. LN = lymph node; SCM = sternocleidomastoid muscle

the trainee's 155 'positive' findings were considered normal by the radiologist and were therefore deemed to be false positives (Table III).

The trainee examination indicated a negative finding in 95 patients. The radiologist's repeat examination indicated normal findings in 63 patients (true negatives). Therefore, according to the radiologist gold standard, the trainee missed abnormalities in 32 (34 per cent) of the abnormal scans (false negatives). These abnormalities included palpable and impalpable neck masses (Tables IV and V).

Of the trainee's 144 true positives, the trainee's interpretation of the lesion was concordant with that of the radiologist in 117 (81 per cent) of the abnormal scans. The trainee's interpretation of detected pathology was considered a misinterpretation in 28 cases (19 per cent of all abnormal scans) (Table VI).

Using the radiological opinion as a gold standard, the overall figures for sensitivity, specificity, positive predictive value, negative predictive value and accuracy of the trainee examinations were: 82, 85, 93, 67 and 83 per cent, respectively.

Of all the 250 examinations, we were able to reassure 127 patients by excluding serious pathology (50 patients) or excluding any lesion (77 patients). Only 16 patients required biopsies, of which 10 proved to be malignant. Of the 16 patients that underwent biopsy, the trainee failed to detect 1 malignant lesion (false negative) and misinterpreted 4 malignant lesions as benign (interpretive error).

Although it was not the purpose of this study to evaluate the use of ultrasound in expert hands, with a minimum follow-up period of two years, none of the patients have re-attended with a significant lesion.

Discussion

This is the first study to describe the process of an ENT trainee undertaking structured training in neck ultrasound. Head and neck ultrasound is difficult, and fraught with pitfalls. Nevertheless, the radiologist in this study (LB) has trained numerous radiologists to a level consistent with non-specialist general radiology practice. The experience required to define or interpret some lesions may be measured in years rather than months, and this would apply equally to a radiologist or sonographer learning head and neck ultrasound.

Surgeon-performed neck ultrasound is infrequently discussed in the literature, with most reports describing the value of peri-operative localisation of parathyroid lesions in shortening operation time.[6–8] Other studies focused on the advantage of clinic-based ultrasound in changing decisions about operative management of thyroid disease when compared to scans performed by a conventional ultrasound practitioner before the clinic visit.[9] Spurious lesions are frequent in head and neck ultrasound (Table III), commonly the result

TABLE IV
TRAINEE FALSE NEGATIVES: PALPABLE LUMPS*

Pathology	Lesion missed by trainee	Source of error
Submandibular	Stone	Scanning too quick
Submandibular	Sublingual ranula herniate thought mylohyoid muscle	Trainee considered ranula a normal structure (muscle)
Parotid†	Lipoma	Controls set to a deeper level‡
Parotid	Sebaceous cyst	Controls set to a deeper level‡
Parotid	Duct stricture with sialectasis	No comparison made to contralateral side (wider lumen)
Thyroid	Solid colloid inside large thyroid cyst	Failure to scan entire cyst
Anterior neck	Level III LN	Distraction by incidental adjacent thyroid nodule
Anterior neck	Prominent transverse process of vertebrae	Inadequate knowledge of US features of a bony structure**
Posterior triangle	Thrombosed blood vessel	Doppler scan was not used

*10 patients. †$n = 2$. ‡Lesion was in superficial skin layers. **Appears as white line as it reflects sound. LN = lymph node; US = ultrasound

TABLE V
TRAINEE FALSE NEGATIVES: IMPALPABLE LUMPS*

Pathology	Lesion missed by trainee	Source of error
Parathyroid†	Parathyroid adenoma	Failure to adjust image to correct depth, or lesion considered a normal structure
Submandibular‡	Ranula	Lesion considered a normal structure (muscle)
Thyroid**	Thyroid nodule	Incomplete scanning
Parotid**	LN	Area scanned too quickly, or some areas missed
Anterior neck‡	LN (1 malignant)	Unaware of need to actively search around IJV (where LNs often exist)
Anterior neck	Thyroglossal cyst	Failure to adjust magnification (so cyst appeared too small)
Anterior neck	Normal thyroid tissue (laryngectomised)	Inadequate knowledge of US features of normal thyroid tissue
Anterior neck‡	Lipoma	Failure to adjust image to correct depth
Anterior neck	Calcified thyroid cartilage	Failure to apply sufficient coupling gel

*23 patients. †$n = 8$; ‡$n = 2$; **$n = 3$. LN = lymph node; IJV = internal jugular vein; US = ultrasound

of a misinterpretation of a normal neck structure. This more likely occurs at an early stage, before the trainee becomes familiar with the radiological anatomy of the neck. Bony structures such as the hyoid or prominent transverse processes of vertebrae can simulate macrocalcification in a lesion or a calculus in Wharton's duct. A blood vessel can be confused with a duct, but this distinction can usually be made by skilled Doppler ultrasound technique.

The process of palpation before the scan does not necessarily facilitate the ultrasound study. Table IV comprises 10 cases where the ENT trainee suspected a definite palpable abnormality prior to performing the ultrasound study, yet nevertheless went on to miss the abnormality on the scan. The ultrasound study may need to go beyond confirming the organ of origin of a positive palpation finding. An example of this is the quest for a calculus following the identification of a sialectatic salivary gland or duct. It may be important to further characterise a lesion; for example, defining a solid component that may require a biopsy within an otherwise cystic lesion. Extremely superficial lesions such as lipomas or sebaceous cysts may easily be overlooked if the focus of the ultrasound apparatus is suboptimal or too much pressure is applied to the ultrasound transducer.

It is notable that false negative results and misinterpretations on the part of the trainee were the most frequent types of errors (Tables V and VI). We regard this as a constructive rather than a discouraging learning outcome, as we will continue to develop this skill. It is likely that many of these errors would have been made by radiologically qualified practitioners less experienced than the gold standard radiologist of the current study. We analysed the trend of our false negative results by equally dividing the total number of examinations into five consecutive blocks. Interestingly, most errors occurred at the initial stages; the learning curve showed subsequent improvement (10 of the 32 missed lesions occurred in the first 50 examinations, and this figure was reduced to 8, 6, 6 and 2 in subsequent blocks). Individual readers of this study will decide

TABLE VI
TRAINEE MISINTERPRETATIONS*

Pathology	Trainee's misinterpretation	Radiologist's correct impression	Source of misinterpretation
Thyroid (7)	Malignant nodule (5), benign nodule (2)	Benign nodule (5), malignant nodule (2)	Inadequate knowledge of pathological features of thyroid nodules
Thyroid	Paratracheal LN	Thyroid nodule	Location of lesion close to trachea
Parathyroid	Parathyroid lesion	Paratracheal LN	Location of lesion deep to thyroid gland
Parotid (3)	Pleomorphic (3)	Metastasis (2), Warthin's tumour (1)	Inadequate knowledge of pathological features of parotid lesions
Submandibular (4)	Stone (2), LN (2)	LN (2), stone (2)	Whitish hilum (i.e. hyperechoic) of LN, so confused with stone
Submandibular (2)	Malignant	Sialectasis	Inadequate knowledge of pathological features of submandibular gland
Anterior triangle	Thyroid malignancy	Level IV LN malignancy	Loss of LN structure
Anterior triangle	Thyroglossal cyst	LN	Location of LN near hyoid bone
Anterior triangle (5)	Malignant LN (4), reactive LN (1)	Reactive LN (4), malignant LN (1)	Inadequate knowledge of pathological features of LN
Anterior triangle	LN	CBT	Failure to recognise lesion at bifurcation of carotid (typical of CBT)
Anterior triangle	Branchial cyst	Haematoma	Failure to recognise lesion is solid, not cystic (even when non-vascular)
Posterior triangle	Lipoma	Synovial cyst	Failure to recognise origin of lesion (sternoclavicular joint)

Numbers in parentheses represent number of lesions. *28 patients. LN = lymph node; CBT = carotid body tumour

whether this is acceptable following a regime of training that is unlikely to be equalled or surpassed in other centres. The subjective impression of the radiologist participating in this study is that the level of the ENT trainee's ability surpasses that of general radiology trainees.

The use of ultrasound is expanding rapidly in the emergency room, surgical ward and critical care unit, and more recently in office practice.[6,10–15] The impetus driving this trend may sometimes be suspect, and will vary between differing medical cultures such as private fee-for-item practice as opposed to a British model of salaried public health provision. A catalogue of objections to clinician-based ultrasound frequently raised by radiologists has included: access to an ultrasound machine, medicolegal liability, lack of specific training and fear of lost revenue.[11,16,17] The policy adopted by the Royal College of Radiologists is that it is appropriate for practitioners other than clinical radiologists to seek to develop skills in the performance of ultrasound.[5,18]

There is growing literature to suggest that clinicians with limited experience in radiology can perform niche ultrasound examinations at a level comparable to radiologists. Specific studies have included the gall bladder,[19] breast,[13] parathyroid gland,[6] joints,[10] emergency hepatobiliary pathology,[14] general trauma,[15] and chest in both critical care and trauma settings.[12,20] Similarly, radiographers performed well when they were adequately trained.[21] Ultrasound has been shown to be a more sensitive technique than clinical evaluation in certain conditions and has been recommended as an extension to physical examination.[9,12,22] A further advantage of office-based ultrasound is that it allows clinical and imaging assessment at a single visit.[23]

Ultrasound teaching programmes for surgeons have been established for decades in mainland Europe, as pioneered at the University of Göttingen in 1982. Subsequently, the German Association of Surgery began requiring experience and competence in ultrasound for certification in general surgery, orthopaedics and urology.[11] In 1996, the American College of Surgeons launched an educational programme to train surgeons on the use of this technology, supported by interested surgical societies and professional bodies.[17,24–26] The American Board of Surgery advocates that surgeons 'have the opportunity to gain a working knowledge of ultrasonography of the head and neck, breast, abdomen, and endorectal ultrasound'.[27] Residents in the US are expected to complete a basic ultrasound course.[2,15,28]

There are many specialties (obstetrics and gynaecology, cardiology, emergency medicine, urology, and family practice) where ultrasound skills are included in the training, and model curricula have been developed.[29] Similarly, a robust training model exists for radiographically qualified ultrasonographers, which is delivered in a relatively short timescale.[4] The Royal College of Radiologists stated that radiologists have the background to provide guidelines for the training of medical non-radiologists, which should be to the same standard as those for radiologists, albeit restricted to the relevant area of their clinical expertise.[5] They proposed three levels of minimum training requirement, ranging from the ability to recognise normal anatomy, to performing specialised examinations and interventions. This is consistent with the minimum requirements of the European Federation of Societies for Ultrasound in Medicine and Biology.[5]

Many criteria would need to be met before the experience of the current authors could be extrapolated. Ultrasound training requires a motivated ENT trainee, and a dedicated head and neck radiologist with relevant ultrasound expertise. Short courses are adequate as an introduction, but adequate one-to-one training more than doubles the time taken for each patient ultrasound study. Additionally, there may be competing radiology trainees in a teaching radiology department, and it would be impractical to train more than one person on each patient.

- **Ultrasound is a valuable diagnostic tool used in many areas of medicine including ENT**
- **Provision of ultrasound service by clinicians other than radiologists has gained wide acceptance in USA and Europe, but less in UK**
- **A recent survey published by ENT UK discussed the prospects of otolaryngologist-performed neck ultrasound as a diagnostic tool**
- **This study reports the unique experience, outcomes and lessons of an ENT trainee learning this technique**
- **Although trainee results were less favourable compared with an experienced head and neck radiologist, improvements were steady**
- **We regard this as a constructive learning outcome and will continue to develop this skill**

Specific training and a range of supervised examinations have been suggested before a non-radiologist can be considered competent and credentialled to perform ultrasound. The number of necessary examinations before applying for certification can be between 50 and 400. This wide range probably reflects the individual variation in aptitude and the varying complexity of different organ systems.[5,24,29–31] Some studies have been hyperbolically optimistic about the length of training. In a study evaluating surgeon-performed ultrasound in trauma patients, it was demonstrated that with only 8 hours of didactic and hands-on training, surgeons could acquire the necessary skills to obtain and interpret ultrasound images to accurately detect haemoperitoneum.[17] The radiologist author of the current study (LB) is sceptical about much of this literature

and would disregard studies where there has not been participation of a skilled radiologist.

The National Ultrasound Steering Group (a subgroup of the National Imaging Board in the UK) recommends the establishment of a Clinical Governance Board for all providers of ultrasound imaging services that includes a clinical lead for each department using ultrasound.[3] Quality assurance is emphasised with regard to maintaining professional standards equivalent to those issued by the General Medical Council, the latter of which recommends that doctors recognise and work within the limits of their competence. The Royal College of Radiologists states that National Health Service trusts in the UK are unlikely to be able to mount any defence to an action brought against an untrained practitioner.[5]

In this series, we describe a unique one-to-one training process in neck ultrasound. We consider this model the gold standard for any ENT trainee attempting to learn this technique, as it allows close supervision and input by the radiologist. Although it might look labour intensive to some readers, the process becomes less demanding as skills are learned. Following our study period, the department acquired an ultrasound machine and the radiologist joined our one-stop neck lump clinic, which improved our partnership and made the training more streamlined.

Conclusion

This study evaluated a one-to-one training model of neck ultrasound for an ENT trainee. We identified important learning outcomes and explored potential errors during the initial stages of training that we significantly improved. Neck ultrasound performed by an otolaryngologist, while less accurate than an experienced radiologist, is a useful adjunct to clinical assessment, and can facilitate assessment in a one-stop clinical setting. A close collaboration with the radiology department is a key element in learning this technique. This study can become a platform for the incorporation of ultrasound training in future ENT curricula. The authors consider that the overriding consideration for extending head and neck ultrasound skills beyond the radiology department should be the welfare and management of the patient, rather than the academic or financial competing interests of other professional groups.

References

1 Rozycki GS, Ochsner MG, Jaffin JH, Champion HR. Prospective evaluation of surgeons' use of ultrasound in the evaluation of trauma patients. *J Trauma* 1993;**34**:516–26
2 Rozycki GS, Feliciano DV, Schmidt JA, Cushman JG, Sisley AC, Ingram W *et al*. The role of surgeon-performed ultrasound in patients with possible cardiac wounds. *Ann Surg* 1996;**223**:737–46
3 Ultrasound clinical governance (National Ultrasound Steering Group, 2008). In: http://www.bmus.org/policies-guides/ClinicalGovernanceInUltrasound-061108.pdf [22 March 2014]
4 Developing and Growing the Sonographer Workforce: Education and Training Needs (The Society and College of Radiographers, 2009). In: http://www.sor.org/learning/document-library/developing-and-growing-sonographer-workforce-education-and-training-needs [22 March 2014]
5 Ultrasound Training Recommendations for Medical and Surgical Specialties, 2nd edn (The Royal College of Radiologists, 2005). In: http://www.rcr.ac.uk/docs/radiology/pdf/BFCR(12)17_ultrasound_training.pdf [22 March 2014]
6 Jabiev AA, Lew JI, Solorzano CC. Surgeon-performed ultrasound: a single institution experience in parathyroid localization. *Surgery* 2009;**146**:569–75
7 Van Ginhoven TM, Morks AN, Schepers T, de Graaf PW, Smit PC. Surgeon-performed ultrasound as preoperative localization study in patients with primary hyperparathyroidism. *Eur Surg Res* 2011;**47**:70–4
8 Solorzano CC, Carneiro-Pla DM, Irvin GL. Surgeon-performed ultrasonography as the initial and only localizing study in sporadic primary hyperparathyroidism. *J Am Coll Surg* 2006;**202**:18–24
9 Mazzaglia PJ. Surgeon-performed ultrasound in patients referred for thyroid disease improves patient care by minimizing performance of unnecessary procedures and optimizing surgical treatment. *World J Surg* 2010;**34**:1164–70
10 Iannotti JP, Ciccone J, Buss DD, Visotsky JL, Mascha E, Cotman K *et al*. Accuracy of office-based ultrasonography of the shoulder for the diagnosis of rotator cuff tears. *J Bone Joint Surg Am* 2005;**87**:1305–11
11 Yiengpruksawan A, Ganepola GP, Freeman HP. Extended applications of ultrasonography by the surgeon. A preliminary report. *Am J Surg* 1987;**153**:221–5
12 Rozycki GS, Pennington SD, Feliciano DV. Surgeon-performed ultrasound in the critical care setting: its use as an extension of the physical examination to detect pleural effusion. *J Trauma* 2001;**50**:636–42
13 Whitehouse PA, Baber Y, Brown G, Moskovic E, King DM, Gui GP. The use of ultrasound by breast surgeons in outpatients: an accurate extension of clinical diagnosis. *Eur J Surg Oncol* 2001;**27**:611–16
14 Kell MR, Aherne NJ, Coffey C, Power CP, Kirwan WO, Redmond HP. Emergency surgeon-performed hepatobiliary ultrasonography. *Br J Surg* 2002;**89**:1402–4
15 Buzzas GR, Kern SJ, Smith RS, Harrison PB, Helmer SD, Reed JA. A comparison of sonographic examinations for trauma performed by surgeons and radiologists. *J Trauma* 1998;**44**:604–8
16 Filly RA. Ultrasound: the stethoscope of the future, alas. *Radiology* 1998;**167**:400
17 Staren ED, Knudson MM, Rozycki GS, Harness JK, Wherry DC, Shackford SR. An evaluation of the American College of Surgeons' ultrasound education program. *Am J Surg* 2006;**191**:489–96
18 Ultrasound training by radiology departments for other medical specialties: resource implications and requirements (The Royal College of Radiologists, 2007). In: http://www.rcr.ac.uk/docs/radiology/pdf/ultrasoundtraining.pdf [22 March 2014]
19 Fang R, Pilcher JA, Putnam AT, Smith T, Smith DL. Accuracy of surgeon-performed gallbladder ultrasound. *Am J Surg* 1999;**178**:475–9
20 Knudtson JL, Dort JM, Helmer SD, Smith RS. Surgeon-performed ultrasound for pneumothorax in the trauma suite. *J Trauma* 2004;**56**:527–30
21 Leslie A, Lockyer H, Virjee JP. Who should be performing routine abdominal ultrasound? A prospective double-blind study comparing the accuracy of radiologist and radiographer. *Clin Radiol* 2000;**55**:606–9
22 Chen SC, Lin FY, Hsieh YS, Chen WJ. Accuracy of ultrasonography in the diagnosis of peritonitis compared with the clinical impression of the surgeon. *Arch Surg* 2000;**135**:170–4
23 Rahman RL, Crawford S, Hall T, Bavosiet D, Quinlan R. Surgical-office-based versus radiology-referral-based breast ultrasonography: a comparison of efficiency, cost, and patient satisfaction. *J Am Coll Surg* 2008;**207**:763–6
24 American Society of Breast Surgeons. Breast Ultrasound Certification. In: https://www.breastsurgeons.org/new_layout/programs/certification/breast_ultrasound_certification.php [22 March 2014]
25 Rozychi GS, Strauch GO. Ultrasound for the general surgeon: an ACS initiative. *Bull Am Coll Surg* 1998;**83**:37–9

26 Society of American Gastrointestinal and Endoscopic Surgeons. Guidelines for granting of ultrasonography privileges for surgeons. In: http://www.sages.org/publications/guidelines/guidelines-for-granting-of-ultrasonography-privileges-for-surgeons/ [22 March 2014]
27 Grosfield JL. American Board of Surgery. *Bull Am Coll Surg* 1996;**81**:57–60
28 Rozychi GS. Surgeon-performed ultrasound: its use in clinical practice. *Ann Surg* 1998;**228**:16–28
29 Mateer J, Plummer D, Heller M. Model curriculum for physician training in emergency ultrasound. *Ann Emerg Med* 1994;**23**:95–102
30 Shackford SR, Rogers FB, Osler TM, Trabulsy ME, Clauss DW, Vane DW. Focused abdominal sonogram for trauma: the learning curve of nonradiologist clinicians in detecting hemoperitoneum. *J Trauma* 1999;**46**:553–64
31 American Institute of Ultrasound in Medicine. Ultrasound practice accreditation. In: http://www.aium.org/accreditation/accreditation.aspx [22 March 2014]

Address for correspondence:
Mr K Badran,
66 Balforn Drive,
Coatbridge ML5 4FF, UK

E-mail: Badran99@hotmail.com

Mr K Badran takes responsibility for the integrity of the content of the paper
Competing interests: None declared

World J Surg (2010) 34:1164–1170
DOI 10.1007/s00268-010-0402-y

ORIGINAL SCIENTIFIC REPORTS

Surgeon-Performed Ultrasound in Patients Referred for Thyroid Disease Improves Patient Care by Minimizing Performance of Unnecessary Procedures and Optimizing Surgical Treatment

Peter J. Mazzaglia

Published online: 4 February 2010

Abstract
Background Ultrasonography has become an indispensable tool in the evaluation of thyroid nodular disease, and most patients will have had a thyroid ultrasound prior to initial surgical evaluation. This study examines the added benefit of office-based, surgeon-performed ultrasonography in patients referred for thyroid disease.
Methods All patients referred to a single endocrine surgeon for evaluation of thyroid disease over a 2-year period were reviewed. Outside ultrasonographic findings were compared to the surgeon-performed ultrasound that was used to formulate treatment decisions.
Results Of 286 consecutive patients referred for surgical evaluation of thyroid disease, 261 had an outside ultrasound available for comparison. There were 239 women and 47 men. Mean age was 54.7 ± 16.6. In 46 patients (17.6%), differences between the two ultrasounds were significant enough to alter treatment plans. For 18 patients no distinct nodule was identified and biopsy was avoided. Nine of these patients had ultrasound characteristics of Hashimoto's disease. In five patients the nodule was significantly smaller than reported and biopsy was not warranted. Twelve patients had nonpalpable, enlarged lymph nodes not previously identified; these were biopsied. Three were positive for metastatic thyroid cancer, which prompted the addition of neck dissection to the operative procedure. In 8 of 132 patients undergoing thyroidectomy, the surgical procedure was significantly altered by the ultrasound findings.
Conclusions This study demonstrates a clear advantage for patients who undergo a surgeon-performed ultrasound. For many, unnecessary procedures were prevented. For others, substantial modifications to the extent of surgery were made when new ultrasonographic findings were identified during the preoperative investigation.

P. J. Mazzaglia (✉)
Department of Bio Med Surgery, Warren Alpert School of Medicine at Brown University, Rhode Island Hospital, 154 Waterman St, Providence, RI 02906, USA
e-mail: pmazzaglia@lifespan.org

Introduction

In many areas of medicine and surgery, ultrasound is fast becoming an extension of the physical exam. Certainly this is proving true in the field of endocrine surgery, where the physical exam sometimes provides little insight into what lies just below the surface, and nearly all patient evaluations now involve a thyroid ultrasound. Since a growing proportion of thyroid disease is first identified incidentally during imaging studies of the neck performed for other indications, a large percentage of the thyroid nodules evaluated by surgeons are not palpable [1]. Traditionally, endocrine surgeons have relied on radiologists for ultrasonographic characterization of thyroid nodular disease and identification of possible lymph node metastases. Ultrasound-guided biopsy of thyroid nodules and suspicious lymph nodes has also been the purview of radiology.

Recently, with the wider availability of portable ultrasound units, surgeons have rapidly acquired the knowledge and skills to become excellent ultrasonographers in multiple disciplines, including head and neck, vascular, breast, and abdomen [2–6]. For multiple reasons, thyroid and parathyroid diseases lend themselves to the rapid development of expertise in the performance and interpretation of thyroid and parathyroid ultrasound, and many endocrine surgeons have adopted this as part of their routine practice. Many endocrine surgeons have published data supporting

the practice of surgeon-performed ultrasound (SPUS) not only for characterizing thyroid disease, but also for identifying suspicious lymph nodes preoperatively, following thyroid cancer patients for recurrence, and for preoperative localization of parathyroid adenomas in hyperparathyroid patients [4, 7, 8]. This study looks specifically at the role of surgeon-performed thyroid ultrasound and its impact on the evaluation and management of patients referred for surgical evaluation of thyroid disease. Particular focus is given to identifying ways in which the SPUS differed from the preconsultation study and in turn how treatment was modified.

Patients and methods

All patients referred to a single endocrine surgeon for evaluation of thyroid disorders from September 2006 until July 2009 were included. After completing the history and physical examination, all patients underwent a surgeon-performed thyroid ultrasound, including bilateral examination of the lateral cervical lymph node compartments. Ultrasound examination was performed with a Terason t3000 portable unit with a linear array transducer (Terason Ultrasound, Burlington, MA), set to a frequency of 12.5 kHz (Fig. 1). All thyroid lobes and nodules, suspicious lymph nodes, and any other abnormal findings were permanently imaged and measured. Both digital and hard copies were saved as part of the medical record.

If there was indication for biopsy of a thyroid nodule or cyst, as defined by the American Association of Clinical Endocrinologists guidelines or the American Thyroid Association guidelines, or if a suspicious lymph node was identified, an ultrasound-guided fine-needle aspiration (FNA) biopsy was performed [9, 10]. Biopsy was accomplished with a 22-gauge needle on a 20-cc syringe held with a Cameco syringe holder (Belpro Medical, Anjou, Quebec, Canada) and was performed during the same visit. All cytology was evaluated by the cytopathologists at Rhode Island Hospital. All pertinent history, ultrasound findings, biopsy results, and surgical pathology were entered into a prospective database, which was analyzed for the purposes of this study.

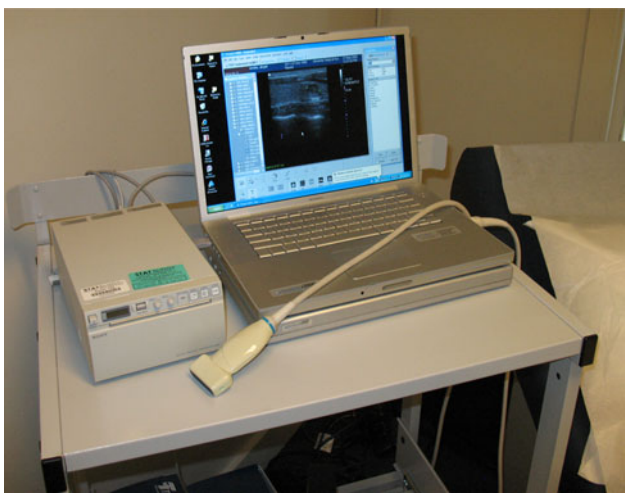


Fig. 1 Portable ultrasound unit and image printer

Results

There were 364 consecutive patients referred for endocrine surgery evaluation of thyroid disease. Three hundred thirty-four had an ultrasound exam performed prior to referral, and the report was available for comparison with the SPUS. There were 282 women and 52 men. Mean age was 54.7 ± 16.6 years. The referral diagnoses were 80.8% nodular thyroid disease, 9.6% thyroid cancer, 3.9% follicular neoplasm, 3.3% thyroiditis, and the remainder consisted of lymphadenopathy, non-nodular goiter, and cystic disease.

In 64 patients (19.2%) there were findings on the SPUS that significantly differed from those on the prereferral study. Those differences led to an alteration in management for 58 patients (17.4%) (Table 1). For 28 patients (8.4%) referred with the diagnosis of a new or growing thyroid nodule, the SPUS findings did not meet standard criteria for FNA biopsy as outlined by the American Thyroid Association [11]. Therefore, biopsy was not performed. In 16 of these 28 patients, no definite nodule could be identified in the location described by the outside study, or the nodule in question was significantly smaller than reported. In the remaining 12 patients, ultrasound findings were strongly characteristic of Hashimoto's thyroiditis, showing a diffusely hypoechoic gland and marked gland heterogeneity, without a definite nodule (Fig. 2).

Nineteen patients (5.7%) had nonpalpable enlarged cervical lymph nodes that were either 1 cm or larger or highly suspicious in appearance and were not reported by the outside ultrasound (Fig. 3). Thirteen of these patients then underwent ultrasound-guided FNA biopsy of the enlarged node. Three of the 13 were found to have metastatic papillary thyroid cancer and the rest were benign. In 6 of the 19 patients with cervical adenopathy, biopsy was not indicated given a benign ultrasound appearance. Additional nodules were identified in seven patients that had not been identified on the outside ultrasound. In seven patients the nodule described on the outside study as thyroid was suspicious in appearance for an enlarged parathyroid gland (Fig. 4), and FNA was sent for parathyroid hormone level (PTH).

Table 1 Nonoperative management changes made based on surgeon-performed ultrasound at the time of initial surgical consultation

Difference between outside and surgeon-performed ultrasound	Action taken	No. patients
Hashimoto's thyroiditis without distinct nodule	Biopsy deferred	12
Nodule <1 cm or not present	Biopsy deferred	9
Nodule had not enlarged as reported	Biopsy deferred	7
Nodule detected that was not reported on outside ultrasound	Biopsy performed	7
Nodule felt to represent parathyroid adenoma	Aspirate sent for PTH	7
Posterior thyroid cyst identified	Biopsy performed	1
Enlarged cervical nodes detected	Lymph node biopsy performed	13

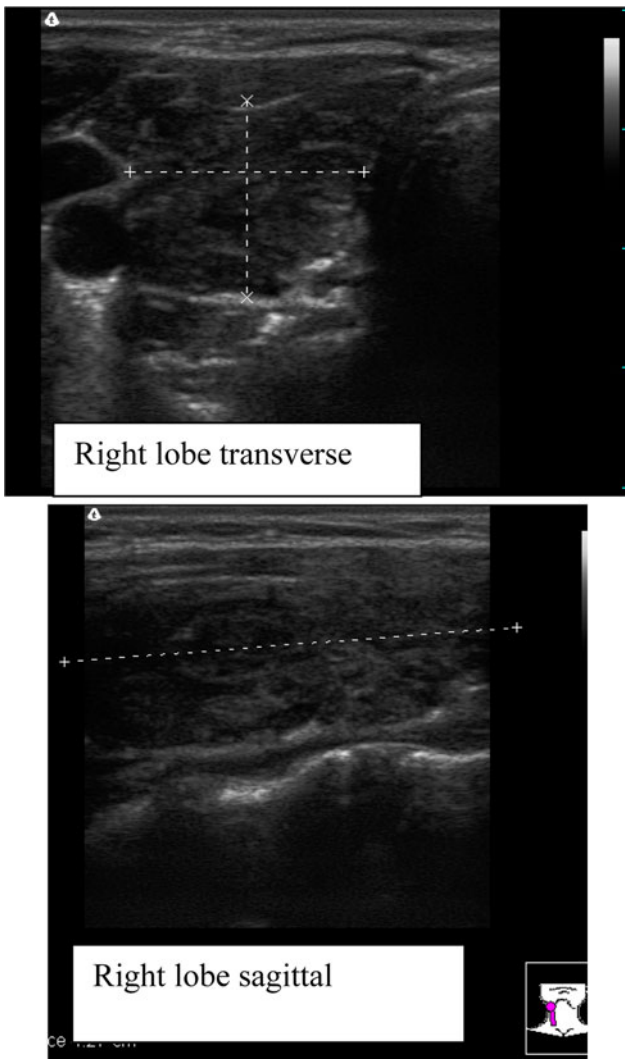


Fig. 2 Classic appearance of a thyroid lobe in a patient with Hashimoto's thyroiditis. The gland is diffusely hypoechoic and heterogeneous

The SPUS directly altered the operative plan for 12 patients (Table 2). For the three patients identified with metastatic thyroid cancer, a simultaneous lymph node dissection was planned preoperatively: two modified radical neck dissections and one central neck dissection.

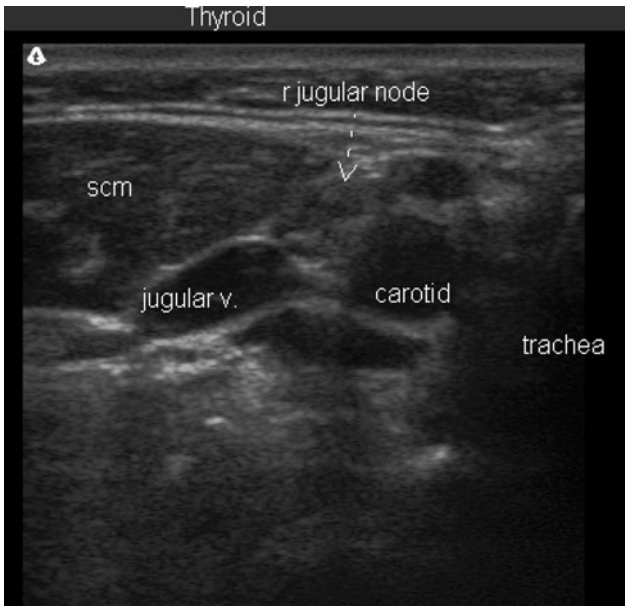


Fig. 3 Suspicious jugular lymph node that was not reported on an outside ultrasound. FNA confirmed metastatic papillary thyroid cancer

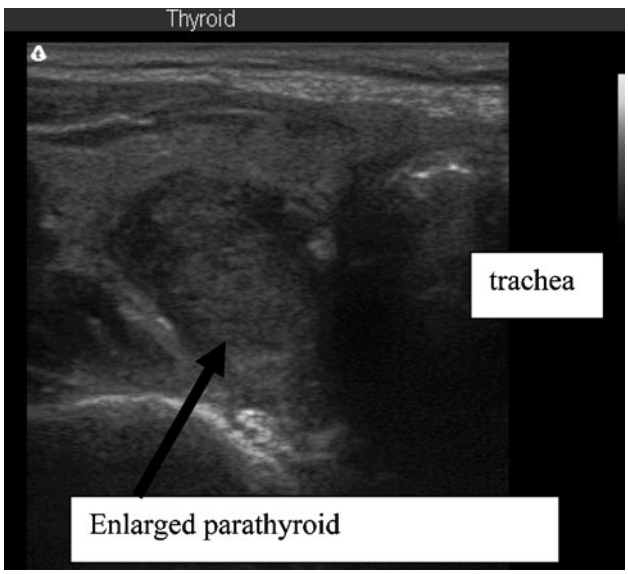


Fig. 4 Parathyroid adenoma mistaken for a thyroid nodule

Table 2 Operative management changes made based on surgeon-performed ultrasound

Difference between outside and surgeon-performed ultrasound	Action taken	No. patients
Differentiated exophytic thyroid nodules from incorrectly diagnosed metastatic lymph nodes	Decreased extent of surgery or eliminated need for surgery	2
Nodule strictly confined to isthmus	Isthmusectomy versus lobectomy	2
Nonpalpable contralateral nodules identified	Total thyroidectomy versus lobectomy	2
Nodule significantly smaller than reported	Surgery avoided	2
Stable 2-cm nodule identified as being posterior compressing esophagus	Thyroidectomy performed	1
Metastatic lymph nodes identified	Cervical lymph node dissection	3

Nonpalpable contralateral nodules were discovered in two patients, and the operation was upgraded from a lobectomy and isthmusectomy to a total thyroidectomy. In two patients the ultrasound demonstrated that the nodule was limited to the isthmus without abnormalities in either of the lobes, and thus the surgery was limited to an isthmusectomy.

Two patients were mistakenly diagnosed with metastatic thyroid cancer and both were referred for a total thyroidectomy with lymph node dissection. These diagnoses were made when exophytic thyroid nodules were interpreted to be abnormal lymph nodes and were biopsied (Fig. 5). The outside FNA biopsies in both patients showed Hürthle cells. Based on this finding of thyrocytes in what were misinterpreted to be level VI lymph nodes, the patients were both told that they had metastatic thyroid cancer. At the time of surgical evaluation, SPUS differentiated between the exophytic thyroid nodules and adjacent lymph nodes, allowing for ultrasound-guided biopsy of the lymph nodes and nodules in question. In both cases, the lymph node biopsies were benign. One of the patients had ultrasound findings of Hashimoto's thyroiditis and did not require surgery since her thyroid nodule had had a previous benign biopsy and was stable in size over a period of years. The other patient underwent a lobectomy and isthmusectomy for what turned out to be a benign thyroid nodule, as opposed to a total thyroidectomy and lymph node dissection.

In two patients surgery was avoided altogether because the nodule was either not present or was significantly smaller than reported. Finally, in one patient who complained of new onset dysphagia, despite the fact that her multinodular goiter had not changed in size over time, a total thyroidectomy was deemed appropriate, because the SPUS demonstrated a 2.1-cm nodule located extremely posteriorly, compressing the esophagus.

As noted above, seven patients had ultrasound-guided FNA biopsy to distinguish possible thyroid nodules from parathyroid glands. In these cases the outside ultrasound report identified a hypoechoic lesion as being consistent with a thyroid nodule. In three patients ultrasound-guided FNA biopsy had already been performed and the cytology results were suggestive of a follicular neoplasm. However, at the time of SPUS in these patients, the lesions in question were more suggestive of parathyroid adenomas than thyroid nodules. To determine whether such a questionable lesion represents a parathyroid adenoma, an ultrasound-guided FNA biopsy can be performed, as described by previous authors, and assessed for parathyroid hormone (PTH) content [12–14]. At Rhode Island Hospital the FNA biopsy is sent for both PTH level and cytological analysis. The cytopathologist is alerted to the question of possible parathyroid origin so appropriate testing can be performed. For PTH analysis, the aspirate is suspended in 10 cc of normal saline and sent for PTH assay, which is run on the

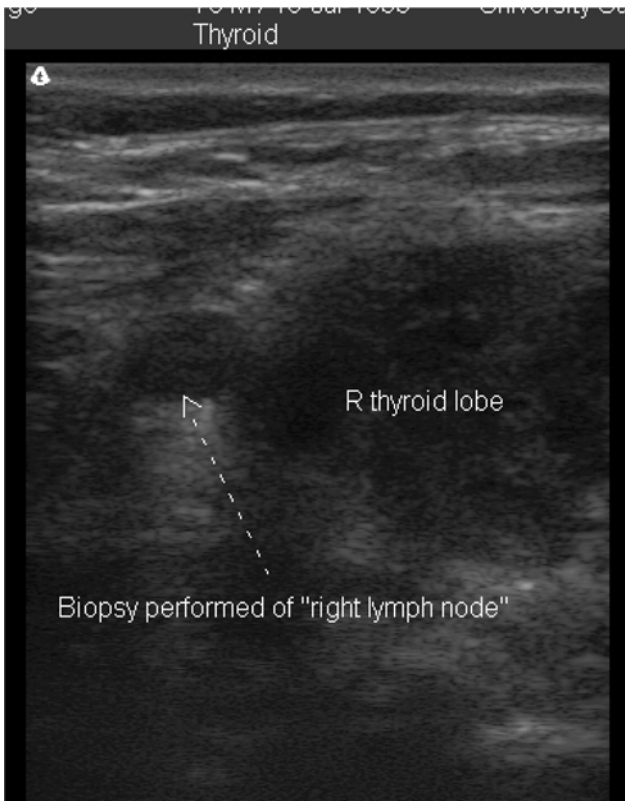


Fig. 5 Exophytic thyroid nodule that was misinterpreted as a lymph node, which led to the incorrect diagnosis of metastatic thyroid cancer based on FNA

same equipment as a blood sample. Any level greater than 40 pg/ml is considered diagnostic of parathyroid tissue [15]. In three of the seven patients tested, parathyroid hormone levels on the aspirates were positive, and two were operated on for hyperparathyroidism.

Discussion

Using ultrasound as an extension of the head and neck exam, the surgeon gains a wealth of information that previously only existed within the text of a radiology report, or on a monitor in a radiology suite far from the operating room or surgeon's office. It comes as no surprise that information gathered by a surgeon performing a thyroid ultrasound sometimes differs from that collected by an ultrasound technician or radiologist. Even among experienced ultrasonographers, concordance of ultrasound characteristics of thyroid nodules is far from 100% [16], so it is expected that similar if not greater discrepancies would exist between the interpretations of a radiologist and a surgeon.

As the surgeon is performing the study in preparation for a possible operation, attention to the contralateral lobe, location and overall appearance of the gland, and presence of central and lateral lymph nodes will naturally be higher. The scope of disease processes being examined is by nature more limited for the endocrine surgeon than for his radiology colleagues. Therefore, within a short time his experience and expertise with thyroid and parathyroid ultrasound rapidly accumulates [2]. Armed with the full understanding of thyroid and parathyroid pathophysiology, the endocrine surgeon can more aptly make decisions regarding which lesions should and should not be biopsied. The adequacy of SPUS-guided thyroid FNA biopsy is generally excellent. A recent review of 447 patients biopsied by surgeons revealed a 3.6% nondiagnostic rate, 3.8% suboptimal, and 92.6% adequate [17].

Specific findings in this study included the identification of 28 patients who were referred for biopsy of a thyroid nodule that did not exist or was significantly smaller than the outside ultrasound report. While prior studies have shown that 15% of "palpable thyroid nodules" are without abnormality on ultrasound evaluation [18], the findings of this study are consistent with previously documented interobserver variability in the interpretation of thyroid ultrasound characteristics [16]. Almost half of these 28 patients had marked gland heterogeneity characteristic of Hashimoto's thyroiditis, which when examined in isolation could be misinterpreted as nodular thyroid disease. Having the advantage of being able to question and examine the patients in real time, the surgeon can definitively make this fine distinction. All 28 patients were spared FNA biopsy of either nonexistent nodules, subcentimeter nodules, or nodules that were stable in size over a period of years. This is not insubstantial, as surely some of these biopsies would have shown follicular neoplasm or even false-positive papillary thyroid cancer, thereby relegating the patients to thyroidectomy.

In several patients the finding of additional nonpalpable contralateral nodules not noted on the outside ultrasound was important in planning surgical strategy. Making the diagnosis of multinodular goiter in contrast to a unilateral nodule is important since the patient needs to be counseled about the long-term risk of recurrence if disease is left behind, which can be upward of 40% [19]. Also, some patients with contralateral benign nodules prefer the option of total thyroidectomy and lifelong thyroid hormone supplementation instead of the possible need for a second operation or yearly ultrasound examinations.

Often abnormal cervical lymph nodes are nonpalpable [20], and in this study 19 patients were found to have abnormal cervical lymph nodes not reported on the outside ultrasound. Since the outside ultrasonographer is focusing on the thyroid gland, incidentally enlarged cervical nodes may not always be noticed, especially if the diagnosis of thyroid cancer has not previously been made. Also, the ability to distinguish lymphadenopathy from thyroid nodular disease is sometimes difficult, as illustrated by the two patients in this study who mistakenly had biopsies performed of thyroid nodules that were thought to be lymph nodes. For both of those patients, biopsies of exophytic thyroid nodules interpreted to be lymph nodes caused the cytologist interpreting the FNA biopsy to conclude that the patients likely had metastatic thyroid cancer. When the SPUS was performed, familiarity with the surgical anatomy of the thyroid and central compartment lymph nodes allowed for clarification of the ultrasound findings and a significantly altered treatment plan for both patients.

There is growing evidence that SPUS can improve the initial evaluation and surveillance of patients with thyroid cancer [2, 4, 8]. Many investigators have published studies examining ultrasound characteristics of thyroid nodules predictive of malignancy, such as hypoechogenicity, irregular borders, microcalcifications, and hypervascularity [21, 22]. Recent review of close to 500 SPUS exams of thyroid nodules showed a near 80% positive predictive value for malignancy if three of the following four characteristics were present: irregular borders, height greater than width on cross section, hypoechogenicity, and microcalcifications [23]. Unfortunately, the sensitivity and specificity for any of these characteristics are insufficient to allow for ultrasound to supplant the role of FNA biopsy. However, for the radiologist, endocrinologist, and endocrine surgeon alike, ultrasound is indispensable for the localization of nonpalpable nodules and for guidance

during biopsies. Ultrasound-guided FNA cytology is currently the best method of distinguishing between benign and malignant thyroid nodules prior to surgery with accuracy approaching 95% [24–26]. There is evidence that it also improves diagnostic yield for palpable lesions, allows for accurate sampling of complex nodules, and reduces rates of nondiagnostic FNA from 15–20% down to 5–10% [27–29]. In addition to these established advantages, the findings of this study show that the treatment algorithm was significantly modified in 17.4% of patients. Most of these patients were spared unnecessary biopsies and/or operations. For several patients, metastatic disease was detected preoperatively, allowing for all disease to be dealt with during a single operation.

For patients with differentiated thyroid cancer, cervical ultrasound has become the standard of care as part of the postoperative surveillance for disease recurrence. Ultrasound is more sensitive than radioactive iodine scanning and thyroglobulin measurements [30], and the majority of patients who recur will do so in the ipsilateral central or lateral neck [31]. While it has been shown that preoperative ultrasound in patients with thyroid cancer detects nonpalpable locoregional disease in close to 20% of primary operations and 30–60% of reoperations [2, 4, 8, 32], this study shows a specific advantage for SPUS over ultrasound exams performed elsewhere.

The increased accuracy of SPUS is predictable when one considers the relatively high volume of cervical ultrasounds being performed in an endocrine surgery practice. For the surgeon who will ultimately be performing the thyroidectomy with possible neck dissection, there is a strong motivation to map out all disease in the neck prior to surgery. The rewards of doing so for the patient and surgeon include decreased incidences of positive postoperative ultrasounds and whole-body radioiodine scans and a greater likelihood of having a normal postoperative thyroglobulin [33]. There is also an advantage to performing same-day ultrasound guidance in the operating room prior to incision for reoperative thyroid surgery [34].

In addition to improving preoperative planning and postoperative outcomes, another major advantage of office-based SPUS is that it streamlines patient care [35]. The ultrasound evaluation, ultrasound-guided FNA biopsy, and surgical consultation can all take place during a single visit, which not only saves the patient's time but also should decrease costs by eliminating additional ultrasounds and second visits. Currently, many patients proceed through the following sequence of visits prior to seeing an endocrine surgeon: The patient or a practitioner palpates a nodule, or a nodule is identified on an imaging study of the neck performed for other reasons, most commonly CT, MRI, or ultrasound. A dedicated thyroid ultrasound is recommended and ordered by the primary care physician. If a thyroid nodule greater than 1 cm is present, the proper next step would be ultrasound-guided FNA biopsy, which requires an additional visit to either radiology, surgery, or endocrinology. If the biopsy result returns as anything but benign or inadequate specimen, or if the nodule is symptomatic, the patient should be referred for surgical evaluation. Where surgeons with ultrasound expertise are available, several steps in this lengthy sequence could be eliminated, especially in cases where it is likely that the patient has a nodule 1 cm or larger.

In 1995 the American Board of Surgery issued a statement requiring exposure to surgical ultrasound as part of the residency training curriculum, and in 1996 the mission statement of the American College of Surgeons (ACS) advised that general surgeons obtain a "working knowledge" of head and neck, breast, abdomen, and endorectal ultrasound [36]. Currently, the ACS offers courses in basic ultrasound as well as in focused areas such as head and neck, breast, and abdominal imaging, with the goal of promulgating SPUS through surgeon training in its effective use. Becoming credentialed involves taking the online basic ultrasound course, followed by a specific training session and exam within one of the focus areas, if so desired. These courses are offered at ACS meetings and at the meetings of some of the subspecialty divisions of surgery; they are helping to increase the numbers of qualified surgeon ultrasonographers [37].

Conclusions

The results of this study echo the findings of the initial pioneers in SPUS. The evidence strongly suggests that SPUS improves and expedites the care of patients with thyroid nodular disease. In particular, it shows that in practices focused on endocrine surgery, surgeon familiarity with physiology and anatomy of thyroid diseases enables a more comprehensive interpretation of the ultrasound exam. These results should serve as a call for all endocrine surgeons to adopt SPUS as part of the routine evaluation of patients with thyroid diseases. Finally, there should be ongoing efforts to formally educate surgical residents and fellows in the use of ultrasound, which will only serve to enhance their diagnostic acumen and improve patient outcomes.

References

1. Pinchera A (2007) Thyroid incidentalomas. Horm Res 68(Suppl 5):199–201
2. Milas M, Stephen A, Berber E et al (2005) Ultrasonography for the endocrine surgeon: a valuable clinical tool that enhances

diagnostic and therapeutic outcomes. Surgery 138(6):1193–1200 discussion 1200-1201
3. Milas M, Mensah A, Alghoul M et al (2005) The impact of office neck ultrasonography on reducing unnecessary thyroid surgery in patients undergoing parathyroidectomy. Thyroid 15(9):1055–1059
4. Solorzano CC, Carneiro DM, Ramirez M et al (2004) Surgeon-performed ultrasound in the management of thyroid malignancy. Am Surg 70(7):576–580 discussion 580-582
5. Fine RE, Staren ED (2004) Updates in breast ultrasound. Surg Clin North Am 84(4):1001-1034, v-vi
6. Rozycki GS, Cava RA, Tchorz KM (2001) Surgeon-performed ultrasound imaging in acute surgical disorders. Curr Probl Surg 38(3):141–212
7. Solorzano CC, Lee TM, Ramirez MC et al (2005) Surgeon-performed ultrasound improves localization of abnormal parathyroid glands. Am Surg 71(7):557–562 discussion 562-563
8. Kouvaraki MA, Shapiro SE, Fornage BD et al (2003) Role of preoperative ultrasonography in the surgical management of patients with thyroid cancer. Surgery 134(6):946–954 discussion 954-955
9. Frates MC, Benson CB, Charboneau JW et al (2005) Management of thyroid nodules detected at US: Society of Radiologists in Ultrasound consensus conference statement. Radiology 237(3):794–800
10. Gharib H, Papini E, Valcavi R et al (2006) American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. Endocr Pract 12(1):63–102
11. Cooper DS, Doherty GM, Haugen BR et al (2006) Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 16(2):109–142
12. Clark OH, Gooding GA, Ljung BM (1981) Locating a parathyroid adenoma by ultrasonography and aspiration biopsy cytology. West J Med 135(2):154–158
13. Doppman JL, Krudy AG, Marx SJ et al (1983) Aspiration of enlarged parathyroid glands for parathyroid hormone assay. Radiology 148(1):31–35
14. Marcocci C, Mazzeo S, Bruno-Bossio G et al (1998) Preoperative localization of suspicious parathyroid adenomas by assay of parathyroid hormone in needle aspirates. Eur J Endocrinol 139(1):72–77
15. Stephen AE, Milas M, Garner CN et al (2005) Use of surgeon-performed office ultrasound and parathyroid fine needle aspiration for complex parathyroid localization. Surgery 138(6):1143–1150 discussion 1150-1151
16. Wienke JR, Chong WK, Fielding JR et al (2003) Sonographic features of benign thyroid nodules: interobserver reliability and overlap with malignancy. J Ultrasound Med 22(10):1027–1031
17. Bhatki AM, Brewer B, Robinson-Smith T et al (2008) Adequacy of surgeon-performed ultrasound-guided thyroid fine-needle aspiration biopsy. Otolaryngol Head Neck Surg 139(1):27–31
18. Brander A, Viikinkoski P, Tuuhea J et al (1992) Clinical versus ultrasound examination of the thyroid gland in common clinical practice. J Clin Ultrasound 20(1):37–42
19. Moalem J, Suh I, Duh QY (2000) Treatment and prevention of recurrence of multinodular goiter: an evidence-based review of the literature. World J Surg 32(7):1301–1312
20. Marqusee E, Benson CB, Frates MC et al (2000) Usefulness of ultrasonography in the management of nodular thyroid disease. Ann Intern Med 133(9):696–700
21. Papini E, Guglielmi R, Bianchini A et al (2002) Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. J Clin Endocrinol Metab 87(5):1941–1946
22. Iannuccilli JD, Cronan JJ, Monchik JM (2004) Risk for malignancy of thyroid nodules as assessed by sonographic criteria: the need for biopsy. J Ultrasound Med 23(11):1455–1464
23. Mendez W, Rodgers SE, Lew JI et al (2008) Role of surgeon-performed ultrasound in predicting malignancy in patients with indeterminate thyroid nodules. Ann Surg Oncol 15(9):2487–2492
24. Hegedus L (2004) Clinical practice. The thyroid nodule. N Engl J Med 351(17):1764–1771
25. Cappelli C, Pirola I, Castellano M et al (2007) Fine needle cytology of complex thyroid nodules. Eur J Endocrinol 157(4):529–532
26. Ko HM, Jhu IK, Yang SH et al. (2003) Clinicopathologic analysis of fine needle aspiration cytology of the thyroid. A review of 1,613 cases and correlation with histopathologic diagnoses. Acta Cytol 47(5):727-732
27. Bellantone R, Lombardi CP, Raffaelli M et al (2004) Management of cystic or predominantly cystic thyroid nodules: the role of ultrasound-guided fine-needle aspiration biopsy. Thyroid 14(1):43–47
28. Mittendorf EA, Tamarkin SW, McHenry CR (2002) The results of ultrasound-guided fine-needle aspiration biopsy for evaluation of nodular thyroid disease. Surgery 132(4):648–653 discussion 653-654
29. Yokozawa T, Miyauchi A, Kuma K et al (1995) Accurate and simple method of diagnosing thyroid nodules the modified technique of ultrasound-guided fine needle aspiration biopsy. Thyroid 5(2):141–145
30. Frasoldati A, Pesenti M, Gallo M et al (2003) Diagnosis of neck recurrences in patients with differentiated thyroid carcinoma. Cancer 97(1):90–96
31. Machens A, Hinze R, Thomusch O et al (2002) Pattern of nodal metastasis for primary and reoperative thyroid cancer. World J Surg 26(1):22–28
32. Stulak JM, Grant CS, Farley DR et al (2006) Value of preoperative ultrasonography in the surgical management of initial and reoperative papillary thyroid cancer. Arch Surg 141(5):489–494 discussion 494-496
33. Binyousef HM, Alzahrani AS, Al-Sobhi SS et al (2004) Preoperative neck ultrasonographic mapping for persistent/recurrent papillary thyroid cancer. World J Surg 28(11):1110–1114
34. McCoy KL, Yim JH, Tublin ME et al (2007) Same-day ultrasound guidance in reoperation for locally recurrent papillary thyroid cancer. Surgery 142(6):965–972
35. Layeequr Rahman R, Crawford S, Hall T et al (2008) Surgical-office-based versus radiology-referral-based breast ultrasonography: a comparison of efficiency, cost, and patient satisfaction. J Am Coll Surg 207(5):763–766
36. Rozycki GS, Strauch GO (1998) Ultrasound for the general surgeon: an ACS initiative. Bull Am Coll Surg 83(10):25–28
37. Staren ED, Knudson MM, Rozycki GS et al (2006) An evaluation of the American College of Surgeons' ultrasound education program. Am J Surg 191(4):489–496

Ann Surg Oncol (2015) 22:422–428
DOI 10.1245/s10434-014-4089-4

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – ENDOCRINE TUMORS

All Thyroid Ultrasound Evaluations are Not Equal: Sonographers Specialized in Thyroid Cancer Correctly Label Clinical N0 Disease in Well Differentiated Thyroid Cancer

Sarah C. Oltmann, MD, David F. Schneider, MD, MS, Herbert Chen, MD, and Rebecca S. Sippel, MD

Section of Endocrine Surgery, Department of Surgery, University of Wisconsin- Madison, Madison, WI

ABSTRACT

Background. Ultrasound (US) is a standard preoperative study in thyroid cancer. Accurate identification of lymph node (LN) disease in the central neck by US is debated, leading some surgeons to perform prophylactic central dissection. The purpose of this study was to evaluate if US performed by a surgeon with specialization in thyroid sonography correctly determined clinical N0 status.

Methods. Retrospective identification of cN0 thyroid cancer patients from a prospectively maintained database was performed. Exclusion criteria included LN dissection with thyroidectomy or missing preoperative US. Demographics and outcomes were reviewed. Patients were categorized by who performed the thyroid US (surgeon vs. non-surgeon). Additional radioactive iodine (RAI) treatments or subsequent positive pathology defined recurrence.

Results. From 2005 to 2012, 177 patients met criteria. Forty-eight patients had surgeon US versus 129 patients with non-surgeon US. Groups were equivalent in age, gender, and tumor size. Forty-six percent had a preoperative diagnosis of cancer, whereas 19 % had benign and 35 % had indeterminate diagnoses. Surgeon US documented LN status more frequently (69 vs. 20 %, $p < 0.01$). RAI treatment and dose were equivalent. RAI uptake was lower with surgeon US (0.06 % ± 0.02 vs. 0.20 % ± 0.03, $p < 0.01$). Recurrence rates were higher in non-surgeon US (12 vs. 0 %, $p = 0.01$). Median time to recurrence was 11 months.

Conclusions. Surgeons with thyroid US expertise correctly identify patients as N0, which may eliminate the need for prophylactic LN dissection without increasing risk of early recurrence. Because not all thyroid cancers are diagnosed preoperatively, US examination of the thyroid should include routine evaluation of the cervical LNs.

Cervical lymph node (LN) involvement in well-differentiated thyroid cancer (DTC) is common. For patients older than age 45 years, it also impacts staging.[1,2] Preoperative physical exam and ultrasound (US) are the mainstays for determining LN involvement prior surgery, although occasionally suspicious central LNs are encountered at time of operation prompting a therapeutic central lymph node dissection (LND).[1,3–8] Patients felt to be clinically node-negative (cN0) based on preoperative US do not need a therapeutic LND, although the use of prophylactic central LND in cN0 patients is hotly debated.[9–11]

Currently, preoperative assessment of the cervical LN in thyroid cancer patients is performed via US due to increased sensitivity to detect metastatic involvement of LN compared with manual palpation.[1,3–8,10] Traditionally, this assessment was performed by radiologists; however, in the recent decade, US has become a common tool for the surgeon and endocrinologist alike.[3–5,7,12–22] Use of US during surgical training has become integrated into multiple different specialties: trauma, breast, abdominal, vascular, critical care, and head and neck surgery.[23] Because interpretation of US images can vary greatly, expertise in thyroid imaging as well as consistency of whom is performing the study results in optimal outcomes.[11,15,16,24,25] Access to a specialized thyroid sonographer is not available at all institutions. In cases where the department of radiology does not have the

Poster Presentation at American Thyroid Association, San Juan, Puerto Rico, October 2013.

First Received: 21 March 2014;
Published Online: 19 September 2014

R. S. Sippel, MD
e-mail: sippel@surgery.wisc.edu

resources to dedicate a single individual or team with expertise in thyroid imaging, the surgeon sonographer with specialization in the care of thyroid cancer can provide consistency in interpretation and expertise in thyroid imaging.[3,11,12,15,18,20,26]

The purpose of this study was to assess recurrence rates in cN0, DTC patients and to determine if surgeon-performed US in contrast to non-surgeon-performed US resulted in differences in early disease recurrence.

METHODS

With institutional review board approval, a retrospective review of a prospectively collected thyroid database at a large tertiary referral center was performed. Patients with cN0, DTC with a minimum of 6 months of follow-up were included. The diagnosis of DTC was based on either fine-needle aspiration (FNA) cytology or final surgical pathology. In some instances, the diagnosis of cancer was not known at time of US examination or surgery. Because institutional practice involves compartment-based LND for clinically N1a or N1b disease, patients undergoing LND, either central or lateral, at the time of initial thyroidectomy were excluded. Prophylactic LND of the central or lateral compartment for well-differentiated thyroid cancer is not performed at our institution. Patients without documented preoperative US were excluded. Patients found to have micropapillary thyroid cancer (PTC, <1 cm) were only included if an additional worrisome feature was noted on final pathology (multifocality, extrathyroidal extension, lymphovascular invasion, or positive margins).

Patients were categorized by who performed the US: the operative surgeon or a non-surgeon. The surgeon performing thyroid ultrasound had successfully completed the American College of Surgeons Head and Neck US course and currently serves as a course instructor. Surgeon-performed US occurred during initial clinic visit; occasionally these were repeated in the operating room prior incision. The study institution does not have a dedicated individual or team of radiologists who specialize in thyroid cancer; thyroid US is done by a variety of different radiologists with expertise in US but not necessarily thyroid cancer. To determine if the central and lateral compartments were assessed during US, the provider needed to specifically comment on LN with an associated descriptor as well as which compartments were evaluated. If no comment was specifically found regarding LN in both the central and lateral neck, the patient was classified as no LN evaluation.

Some patients had multiple tumor histologies on final pathology (i.e., PTC and follicular, PTC and Hurthle cell, etc.). For this reason, the frequency of each tumor type was totaled. Administration and dosing of radioactive iodine ablation (RAI) was determined by the endocrinologists within the study institution. Patients were monitored for recurrence by endocrinology with suppressed thyroglobulin levels and an US examination at 6 months, followed by a stimulated thyroglobulin level and US examination at 1 year.[1] Diagnostic whole body scan was generally performed if US or thyroglobulin results were concerning for residual or recurrent disease. Follow-up after 1 year relied on annual suppressed thyroglobulin level and US evaluation of the neck. Disease recurrence was defined as the need for additional RAI treatment, as positive FNA or positive final pathology on operative reexploration. Staged lymph node dissections or staged completion thyroidectomies were not considered recurrences. Time to recurrence was calculated in months from time of initial operation to time of subsequent intervention (RAI or surgical resection).

Statistical analysis was performed using IBM SPSS Statistics, version 20.0. Pearson's χ^2, Fisher's exact, and unpaired t tests were performed as appropriate. Kaplan–Meier survival analysis was performed with outcome listed as time to recurrence or time to last disease-free follow-up. Comparison of estimated disease-free curves was performed using Mantel–Cox log-rank. Data are expressed as mean $\pm$ SE of the mean or as number (percentage) unless otherwise specified. A p value ≤ 0.05 was determined to be significant.

RESULTS

Between 2005 and 2012, 322 patients with cN0, DTC were identified. Seventy-three patients were excluded for less than 6 months follow-up available within the electronic medical record. An additional 59 patients with micro-PTC with low-risk features on histology (unifocal, intrathyroidal, no lymphovascular invasion, and negative margins) also were excluded. Finally, 13 (4 %) patients were noted to have no documented preoperative US, by either radiology report or by reference via clinician note and were excluded. The final study population was 177 patients.

The study population had an average age of 49 $\pm$ 1.1 years, and 73 % were female. Eighty-one patients (46 %) had a diagnosis of thyroid cancer before surgery, whereas 63 patients (35 %) had indeterminate biopsy results, and 34 patients (19 %) were undergoing surgery for a presumed benign condition (Graves', goiter, etc.). Surgeon-performed US occurred in 48 cases (27 %), whereas the remaining 129 patients (73 %) had a non-surgeon-performed US. Regardless of the sonographer (surgeon vs. non-surgeon), only 59 patients (33 %) had a full LN evaluation documented at time of US. However, the timing of the US may have occurred before the

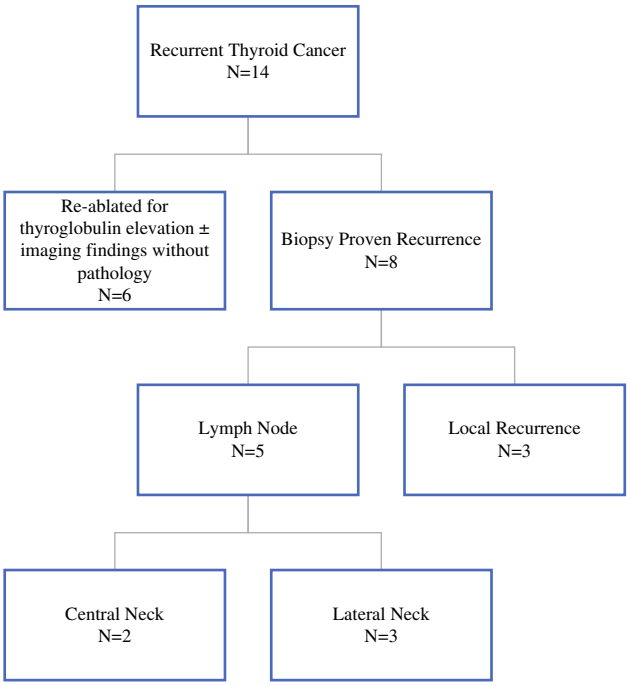


FIG. 1 Breakdown of patients considered to have persistent and/or recurrent disease based on treatment and/or location/type of disease

TABLE 1 Patient preoperative demographic information

	Non-surgeon sonographer	Surgeon sonographer	*p* value
N	129	48	
Age (yrs)	50 ± 1.4	49 ± 2.2	0.75
Female	92 (71 %)	37 (77 %)	0.57
Preop diagnosis			0.28
Benign	23 (18 %)	11 (23 %)	
Indeterminate	43 (33 %)	20 (42 %)	
Malignant	63 (49 %)	17 (35 %)	
Documented assessment of cervical lymph nodes with US	26 (20 %)	33 (69 %)	**<0.01**

Data expressed as mean ± SE of the mean or number (percentage) unless otherwise indicated, *p* values in bold denote statistical significance

diagnosis of cancer was established. Overall, 14 recurrences (8 %) were noted (Fig. 1).

Patients were grouped based on who performed their US evaluation: surgeon or non-surgeon. Patient age ($p = 0.77$) and gender ($p = 0.57$) were equivalent between groups (Table 1). The preoperative diagnosis based on FNA results and/or clinical diagnosis (i.e., Graves') were of similar distribution of benign, indeterminate, and malignant between groups ($p = 0.26$). Patients with a surgeon-performed US were much more likely to have evaluation of their cervical LN than those patients undergoing ultrasound evaluation by a non-surgeon provider (69 vs. 20 %, $p < 0.01$).

With the exception of the surgeon-performed US group having a higher incidence of follicular thyroid carcinoma (19 vs. 4 %, $p < 0.01$), the groups had equal rates of PTC ($p = 0.21$), Hürthle cell carcinoma ($p = 1$), and background thyroiditis ($p = 0.60$; Table 2). The surgeon-performed US group had a greater incidence of lymphovascular invasion noted on histology (13 vs. 3 %, $p = 0.03$), whereas the remaining histologic characteristics of the primary tumor were equivalent. On final pathology, tumor size ($p = 0.13$) and total gland weight ($p = 0.93$) did not differ.

RAI was used with equal frequency ($p = 0.41$) and equivalent doses ($p = 0.31$; Table 3). Median follow-up was shorter in the surgeon-performed US group (20 vs. 34 months, $p < 0.01$). However, median time to recurrence was 11 months, with first recurrence detected at 6 months and last recurrence detected at 6 years. Only two recurrences were diagnosed beyond 15 months and occurred between 4 and 6 years after initial surgery. Of the remaining patients, disease was detected within the first year from surgery in seven, and in five patients, shortly after the 1 year anniversary of their initial operation. No patient in the surgeon-performed US group had evidence of disease recurrence at time of last follow-up compared with 14 patients (12 %) in the non-surgeon-performed US group ($p = 0.01$).

Grouping patients based on if the operative surgeon performed an US evaluation of the neck, a Kaplan–Meier curve for disease-free interval was constructed (Fig. 2a). Patients having US exam performed only by a non-surgeon were disease-free 94 % at 1 year, 89 % at 2 years, and 87 % at 5 years. This was in marked contrast to the group with surgeon-performed US who were disease-free 100 % at 1, 2, and 5 years ($p = 0.04$). To ensure that the specialty of the individual performing the US evaluation was not a confounder for LN assessment, an additional analysis specific to documented LN assessment also was performed (Fig. 2b). Estimated disease-free status did not differ between these groups ($p = 0.66$).

DISCUSSION

In the hands of an experienced thyroid surgeon, trained in thyroid US, the classification of a patient as cN0 and forgoing prophylactic LND resulted in no recurrences to date, with actuarial follow-up to 5 years. In contrast, patients undergoing a non-surgeon US experienced a recurrence rate of 12 %, with 86 % of recurrences occurring within the first 15 months of diagnosis. This early time to recurrence is suggestive of unrecognized disease present

TABLE 2 Tumor type and pathologic characteristics

	Non-surgeon sonographer	Surgeon sonographer	*p* value
N	129	48	
Tumor type			
Papillary	121 (94 %)	42 (88 %)	0.21
Follicular	5 (4 %)	9 (19 %)	**<0.01**
Hürthle carcinoma	5 (4 %)	1 (2 %)	1.0
Micro PTC	26 (20 %)	8 (17 %)	0.39
Pathologic characteristics			
Multifocal	75 (58 %)	30 (63 %)	0.73
Extrathyroidal extension	13 (10 %)	4 (8 %)	1.0
Positive margin	8 (6 %)	3 (7 %)	1.0
Lymphovascular invasion	4 (3 %)	6 (13 %)	**0.03**
Lymphocytic thyroiditis	44 (34 %)	19 (40 %)	0.60
Tumor size (cm)	1.8 ± 0.1	2.2 ± 0.2	0.13
Size of micro PTC (cm)	0.6 ± 0.05	0.6 ± 0.07	0.71
Gland weight (g)	25 ± 3.9	26 ± 2.6	0.92

Data expressed as number (percentage) or as mean ± SE of the mean unless otherwise indicated, *p* values in bold denote statistical significance

TABLE 3 Postoperative management and disease specific outcomes

	Non-surgeon sonographer	Surgeon sonographer	*p* value
RAI	114 (88 %)	45 (94 %)	0.41
RAI Dose (mCi)	83 ± 5	93 ± 6	0.31
Remnant Uptake	0.2 ± 0.03	0.06 ± 0.02	**<0.01**
Follow Up (Months)	34 (16-64)	20 (10-34)	**<0.01**
Disease Recurrence	14 (12 %)	0	**0.01**
Time to Recurrence (Months)	11 (6.6)	0	**<0.01**

Data expressed as number (percentage), mean ± SE of the mean, or median (interquartile range) as appropriate, *p* values in bold denote statistical significance

at time of initial diagnosis, or persistent disease. These data also support previous reports that a negative US of the central neck by experienced sonographers predicts long-term regional control and that the microscopic disease found during prophylactic dissection may not impact short-term disease-free survival.[10,11,27,28]

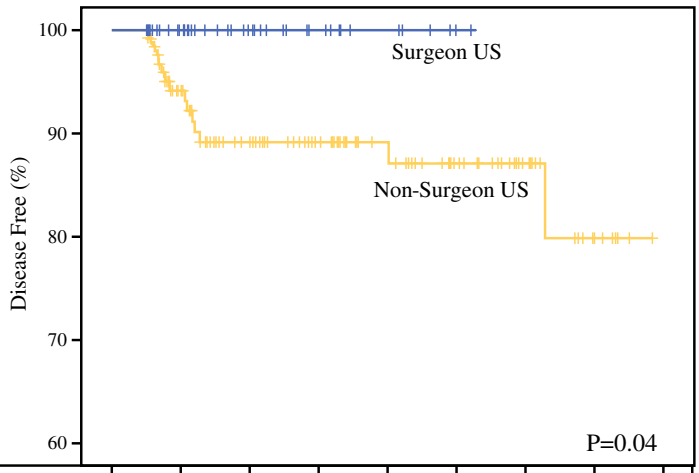
US is a highly operator-dependent modality and variability in image interpretation between sonographers is problematic.[15,16,24,25,29] Rosario evaluated US assessment of the cervical LN during surveillance in patients with known high-risk PTC.[29] Radiologists at a diagnostic imaging center, without specific specialization in thyroid imaging, missed half of the cervical metastasis caught 2 weeks later by a specialized thyroid sonographer. Previous work from this institution, as well as from other authors, has described the omission of LN commentary on thyroid US reports, even when the evaluation of the cervical LN were specifically requested.[3,26,29,30] For purposes of this study, patients with omitted LN commentary were classified as not having the assessment performed.

Surgeons have access to all pertinent clinical information at time of US, excellent understanding of the local anatomy, as well as feedback from final pathologic results to continue to learn the finer nuances of ultrasound findings within the neck.[3,4,15,18,20] Radiology educational literature emphasizes the importance of repetition, in addition to familiarity with the key imaging characteristics, for greater accuracy of thyroid US interpretation.[24,25] Thyroid surgeons, by using US weekly in both the clinic and operating room, can quickly develop the skills needed to perform thyroid US proficiently and accurately.

The timing of thyroid US during the course of patient workup also may influence image interpretation. During a thyroid nodule workup, US and thyroid function tests are initially ordered.[1] As the patient is deemed to need further evaluation and if necessary, referral for endocrinology or surgical consultation, the underlying index of suspicion for malignancy increases. At this time, a provider specialized in the care of thyroid cancer can scrutinize the US characteristics of the nodule, the remaining thyroid, as well as

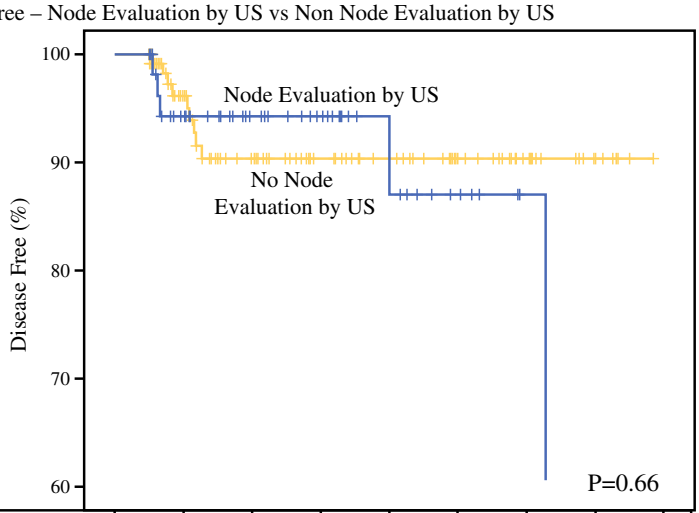
FIG. 2 Disease-free status. **a** Based on who performed the preoperative ultrasound evaluation. **b** Based on whether preoperative ultrasound evaluation included a lymph node assessment

A Patients Disease Free – Surgeon Performed US vs Non-Surgeon Performed US



Months		0	12	24	36	48	60	72	84	96
Patients at Risk	Surgeon US	48	34	19	11	6	0	0	0	0
	Non-Surgeon US	129	98	72	57	45	32	17	6	0

B Patients Disease Free – Node Evaluation by US vs Non Node Evaluation by US



Months		0	12	24	36	48	60	72	84	96
Patients at Risk	Node Eval	59	44	35	21	13	6	1	0	0
	No Node Eval	118	88	63	46	38	27	16	6	0

assess LN appearance. These variables can be placed within the context of the patient history, physical findings, and biopsy results to formulate an opinion regarding both the suspicion for malignancy and LN involvement.[3,7,20,31] The findings of improved short-term disease-free survival with surgeon US within this study are supportive of this as well.

Only 46 % of our study population had an established diagnosis of cancer before surgery, which likely influenced the extent of the preoperative ultrasound evaluation. The remainder of patients had indeterminate or benign pathologies, requiring operative intervention. Given these nonmalignant diagnoses, under current guidelines, LN assessment would not be indicated.[32] However, when patients undergo a diagnostic lobectomy for indeterminate cytology, and final pathology returns as malignant, LN assessment in a recently operated neck may be less reliable. Findings of suspicious cervical lymphadenopathy in the setting of suspicious or indeterminate cytology may prompt additional evaluation and confirm the diagnosis of malignancy in time to alter the operative plan.[3,30,33]

LN assessment is recommended to occur via physical exam at the initial stage of thyroid nodule workup; however, studies have shown that US is superior to physical

exam in detecting worrisome LN.[1,7,8,33,34] For these reasons, the authors advocate the routine cervical LN assessment with clear documentation of findings during initial thyroid US.[2–4,20,35] Evaluation of the LN at time of initial thyroid US would not add a substantial amount of time to the examination and would streamline care by avoiding additional appointments for dedicated LN assessment.[18] Results of this study have prompted ongoing quality improvement and continuing medical education within the study institution, as well as the surrounding medical community, emphasizing the importance of lymph node involvement at time of thyroid US.

Management of thyroid cancer requires a strong interdisciplinary team to facilitate the diagnosis, management, and long-term follow-up. Dedicated endocrinologists, surgeons, radiologist, and nuclear medicine physicians are critical to ensure a successful thyroid cancer program.[6] However, not every institution has all of these resources at their disposal, and overlapping skill sets between the providers may be necessary.[12,20,33] While these results are specific to surgeon performed US within the study institution, a dedicated thyroid sonographer of any specialty could achieve comparable outcomes.

Because this study is retrospective in nature, it is has its inherent flaws. The study population consists of only patients with negative findings on US, who did not undergo LN excision. Given the initial patient selection based on an absence of LND at time of initial surgery, as well as the presence of cancer on final pathology, it is unknown how many patients had negative US imaging, but during thyroidectomy suspicious LN were encountered prompting subsequent LND. Therefore, sensitivity, specificity, positive, or negative predictive value of US on the detection of LN metastases cannot be calculated. While clinically significant disease was not identified in follow-up, this does not equate to the absence of microscopic disease. The length of follow-up included can attest to early recurrence or persistence, but long-term (>5 year) outcomes cannot be assumed based on these data. Ongoing data collection for these cohorts of patients is being performed to see how long-term recurrence rates may differ between the cohorts. This also will determine the durability of the initial US evaluation.

While the study population does not differ in basic patient demographics, they are inherently different by the mere fact that a portion of the non-surgeon group includes patients erroneously categorized as cN0 who with follow-up have evidence of persistent disease. This disease was likely present at time of initial preoperative consultation but was missed. Patients undergoing surgeon US had this disease initially detected and were able to undergo therapeutic LND. However, this very fact drives home the point that surgeon US can correctly stratify patients before operative intervention.

CONCLUSIONS

We demonstrated that a surgeon sonographer with expertise in thyroid cancer can provide an accurate assessment of the LN status in both the central and lateral neck, as demonstrated by the 100 % disease-free status at time of last follow-up. This implies that a thorough US examination of the cervical LN can detect clinically relevant disease in DTC. A negative, high-quality US of the cervical LN may obviate the need for a prophylactic central LND. Because not all patients have an established diagnosis of cancer at time of thyroid US, additional information provided by a LN evaluation can lead to the correct diagnosis. Assessment of the cervical LN should be a standard part of any thyroid US. It is critical that an experienced sonographer provide this assessment to enable the proper extent of surgery and reduce early recurrence.

REFERENCES

1. American Thyroid Association Guidelines Taskforce on Thyroid N, Differentiated Thyroid C, Cooper DS, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2009;19(11):1167–214.
2. American Thyroid Association Surgery Working G, American Association of Endocrine S, American Academy of O-H, et al. Consensus statement on the terminology and classification of central neck dissection for thyroid cancer. *Thyroid.* 2009;19(11):1153–8.
3. Mazzaglia PJ. Surgeon-performed ultrasound in patients referred for thyroid disease improves patient care by minimizing performance of unnecessary procedures and optimizing surgical treatment. *World J Surg.* 2010;34(6):1164–70.
4. Milas M, Stephen A, Berber E, Wagner K, Miskulin J, Siperstein A. Ultrasonography for the endocrine surgeon: a valuable clinical tool that enhances diagnostic and therapeutic outcomes. *Surgery.* 2005;138(6):1193–1200; discussion 1200–1.
5. Miller BS, Gauger PG, Broome JT, Burney RE, Doherty GM. An international perspective on ultrasound training and use for thyroid and parathyroid disease. *World J Surg.* 2010;34(6):1157–63.
6. Schneider DF, Chen H. New developments in the diagnosis and treatment of thyroid cancer. *CA Cancer J Clin.* 2013;63(6):374-94.
7. Solorzano CC, Carneiro DM, Ramirez M, Lee TM, Irvin GL 3rd. Surgeon-performed ultrasound in the management of thyroid malignancy. *Am Surg.* 2004;70(7):576–80; discussion 580–2.
8. Kouvaraki MA, Shapiro SE, Fornage BD, et al. Role of preoperative ultrasonography in the surgical management of patients with thyroid cancer. *Surgery.* 2003;134(6):946–54; discussion 954–5.
9. Carling T, Carty SE, Ciarleglio MM, et al. American Thyroid Association design and feasibility of a prospective randomized controlled trial of prophylactic central lymph node dissection for papillary thyroid carcinoma. *Thyroid.* 2012;22(3):237–44.
10. Randolph GW, Duh Q-Y, Heller KS, et al. The prognostic significance of nodal metastases from papillary thyroid carcinoma can be stratified based on the size and number of metastatic lymph nodes, as well as the presence of extranodal extension. *Thyroid.* 2012;22(11):1144–52.
11. Moreno MA, Edeiken-Monroe BS, Siegel ER, Sherman SI, Clayman GL. In papillary thyroid cancer, preoperative central neck ultrasound detects only macroscopic surgical disease, but

negative findings predict excellent long-term regional control and survival. *Thyroid.* 2012;22(4):347–55.
12. Al-azawi D, Mann GB, Judson RT, Miller JA. Endocrine surgeon-performed US guided thyroid FNAC is accurate and efficient. *World J Surg.* 2012;36(8):1947–52.
13. Choi JS, Chung WY, Kwak JY, Moon HJ, Kim MJ, Kim EK. Staging of papillary thyroid carcinoma with ultrasonography: performance in a large series. *Ann Surg Oncol.* 2011;18(13):3572–8.
14. Goldfarb M, Gondek SS, Sanchez Y, Lew JI. Clinic-based ultrasound can predict malignancy in pediatric thyroid nodules. *Thyroid.* 2012;22(8):827–31.
15. Hwang HS, Orloff LA. Efficacy of preoperative neck ultrasound in the detection of cervical lymph node metastasis from thyroid cancer. *Laryngoscope.* 2011;121(3):487–91.
16. Kabaker AS, Tublin ME, Nikiforov YE, et al. Suspicious Ultrasound Characteristics PredictBRAFV600E-Positive Papillary Thyroid Carcinoma. *Thyroid.* 2012;22(6):585–9.
17. Kangelaris GT, Kim TB, Orloff LA. Role of ultrasound in thyroid disorders. *Otolaryngol Clin N Am.* 2010;43(6):1209–27, vi.
18. Lee CY, Snyder SK, Lairmore TC, Dupont SC, Jupiter DC. Utility of surgeon-performed ultrasound assessment of the lateral neck for metastatic papillary thyroid cancer. *J Oncol.* 2012;2012:973124.
19. The American Association of Clinical Endocrinologists and the American Association of Endocrine Surgeons position statement on the diagnosis and management of primary hyperparathyroidism. *Endocr Pract.* 2005;11(1):49–54.
20. Mendez W, Rodgers SE, Lew JI, Montano R, Solorzano CC. Role of surgeon-performed ultrasound in predicting malignancy in patients with indeterminate thyroid nodules. *Ann Surg Oncol.* 2008;15(9):2487–92.
21. Akinci B, Demir T, Yener S, et al. Beneficial effect of endocrinologist-performed ultrasonography on preoperative parathyroid adenoma localization. *Endocr Pract.* 2009;15(1):17-23.
22. Zangeneh F, Powell CC, Gharib H. A survey on the use of thyroid ultrasonography in clinical endocrinology training programs. *Endocr Pract.* 2003;9(2):162–3.
23. Rozycki GS. Surgeon-performed ultrasound: its use in clinical practice. *Ann Surg.* 1998;228(1):16–28.
24. Kim HG, Kwak JY, Kim EK, Choi SH, Moon HJ. Man to man training: can it help improve the diagnostic performances and interobserver variabilities of thyroid ultrasonography in residents? *Eur J Radiol.* 2012;81(3):e352–6.
25. Kim SH, Park CS, Jung SL, et al. Observer variability and the performance between faculties and residents: US criteria for benign and malignant thyroid nodules. *Korean J Radiol.* 2010;11(2):149–55.
26. Poehls JL, Chen H, Sippel RS. Preoperative ultrasonography findings predict the need for repeated surgery in papillary thyroid cancer. *Endocr Pract.* 2012;18(3):403–9.
27. Shen WT, Ogawa L, Ruan D, Suh I, Duh QY, Clark OH. Central neck lymph node dissection for papillary thyroid cancer: the reliability of surgeon judgment in predicting which patients will benefit. *Surgery.* 2010;148(2):398–403.
28. Raffaelli M, De Crea C, Sessa L, et al. Prospective evaluation of total thyroidectomy versus ipsilateral versus bilateral central neck dissection in patients with clinically node-negative papillary thyroid carcinoma. *Surgery.* 2012;152(6):957–64.
29. Rosario PW. Ultrasonography for the follow-up of patients with papillary thyroid carcinoma: how important is the operator? *Thyroid.* 2010;20(7):833–4.
30. Roy R, Kouniavsky G, Venkat R, et al. The role of preoperative neck ultrasounds to assess lymph nodes in patients with suspicious or indeterminate thyroid nodules. *J Surg Oncol.* 2012;105(6):601–5.
31. Sippel RS, Elaraj DM, Khanafshar E, Kebebew E, Duh QY, Clark OH. Does the presence of additional thyroid nodules on ultrasound alter the risk of malignancy in patients with a follicular neoplasm of the thyroid? *Surgery.* 2007;142(6):851–7; discussion 857 e851–2.
32. AIUM Practice Guideline for the performance of ultrasound examinations of the head and neck. *J Ultrasound Med.* 2014; 33(2):366–82.
33. Lew JI, Solorzano CC. Use of ultrasound in the management of thyroid cancer. *Oncologist.* 2010;15(3):253-8.
34. Giacomini CP, Jeffrey RB, Shin LK. Ultrasonographic evaluation of malignant and normal cervical lymph nodes. *Semin Ultrasound CT MR.* 2013;34(3):236–47.
35. Stack BC, Jr., Ferris RL, Goldenberg D, et al. American Thyroid Association consensus review and statement regarding the anatomy, terminology, and rationale for lateral neck dissection in differentiated thyroid cancer. *Thyroid.* 2012;22(5):501–8.

Reprinted by permission of Otolaryngol Head Neck Surg. 2016; 154(3):421-432.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Otolaryngology–Head and Neck Surgery
2016, Vol. 154(3) 421–432

DOI: 10.1177/0194599815621742
http://otojournal.org
SAGE

Detecting Residual/Recurrent Head Neck Squamous Cell Carcinomas Using PET or PET/CT: Systematic Review and Meta-analysis

Phylannie K. F. Cheung, BVSc, MBBS[1], Ronald Y. Chin, FRACS, MBBS[2], and Guy D. Eslick, DrPH, PhD[3]

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

Abstract

Objective. To evaluate the diagnostic accuracy of positron emission tomography (PET) and PET/computed tomography (CT) for detecting residual and/or recurrent local and regional disease and distant metastases in patients with head and neck squamous cell carcinomas (HNSCCs) following radiotherapy with or without chemotherapy.

Data Sources. A systematic review with no language restrictions was conducted using PREMEDLINE, MEDLINE, EMBASE, and Google Scholar.

Review Methods. Only prospective studies with histopathological and/or clinical follow-up that assessed the diagnostic accuracy of PET and PET/CT in detecting residual and/or recurrent disease following radiotherapy with or without chemotherapy in patients with HNSCCs were included.

Results. Twenty-seven studies were identified. The pooled sensitivity and specificity of PET and PET/CT for detecting residual or recurrent disease at the primary site was 86.2% and 82.3%, respectively. For residual and recurrent neck disease, the sensitivity and specificity were 72.3% and 88.3%, while for distant metastases, the values were 84.6% and 94.9%.

Conclusions. PET and PET/CT are highly accurate in detecting residual and/or recurrent HNSCC. PET/CT is more specific than PET alone. Specificity is also greater for scans performed more than 12 weeks after radiotherapy with or without chemotherapy. The authors support the use of PET/CT after 12 weeks posttreatment for the assessment of residual or recurrent disease.

Keywords

squamous cell carcinoma of the head and neck, locoregional neoplasm recurrence, systematic review, meta-analysis, positron emission tomography

Received September 2, 2015; revised October 29, 2015; accepted November 19, 2015.

Squamous cell carcinomas (SCCs) account for more than 90% of head and neck cancers.[1,2] Early-stage disease is typically treated with unimodality treatment (surgery or radiotherapy), while locally advanced tumors require a multimodality approach consisting of a combination of surgery and radiotherapy with or without chemotherapy.[3] Despite treatment, up to 40% patients with advanced tumors will have a locoregional recurrence,[4,5] and up to 25% will have distant metastases,[6,7] with most of these locoregional recurrences occurring in the first 2 years post-treatment.[8] Patients with early-stage recurrences have a better prognosis compared with those with advanced-stage disease.[9] Those with residual or recurrent disease confined to the head and neck may be candidates for salvage surgery and reirradiation. However, palliative measures may be more appropriate for those in whom distant metastases are present at the time of recurrence. Furthermore, the recognition of patients with a complete response postchemoradiotherapy can reduce the need for unnecessary tissue biopsies and neck dissections following treatment. Thus, from a clinical perspective, the ability to accurately detect residual or recurrent locoregional disease and exclude distant metastases is important as it can help guide ongoing management of patients after chemoradiotherapy.

Functional imaging using 18F-fluorodeoxyglucose positron emission tomography (FDG PET) and integrated FDG PET/computed tomography (CT) is now widely used in the assessment of residual or recurrent disease in patients with head and neck squamous cell carcinoma (HNSCC). As

[1]Sydney Medical School, The University of Sydney, New South Wales, Australia
[2]Department of Otolaryngology Head and Neck Surgery, The University of Sydney, Nepean Hospital, Penrith, New South Wales, Australia
[3]The Whiteley-Martin Research Centre, Discipline of Surgery, The University of Sydney, Nepean Hospital, Penrith, New South Wales, Australia

Corresponding Author:
Ronald Y. Chin, Department of Otolaryngology Head and Neck Surgery, The University of Sydney, Nepean Hospital,1 Hope St, Penrith, NSW 2750, Australia.
Email: drronaldchin@gmail.com

proliferating neoplastic cells consume glucose at a higher rate than normal cells do, the glucose analogue 18-F-fluorodeoxyglucose (18-FDG) accumulates at higher rates within malignant cells. However, nonspecific increases in the uptake of FDG within cells can also occur in normal salivary glands and lymphoid tissues and in the setting of infection and inflammation, such as that which occurs after radiotherapy.

The aim of our systematic review and meta-analysis was to evaluate the diagnostic accuracy of PET and PET/CT for detecting residual and/or recurrent local and regional disease and distant metastases in patients with HNSCCs following radiotherapy or chemoradiotherapy. Hereafter, PET and PET/CT will be collectively referred to as PET, with distinctions made where necessary.

Methods

Search Methodology

We searched EMBASE, PREMEDLINE, MEDLINE, and GoogleScholar for studies evaluating the diagnostic performance of FDG-PET in head and neck cancers. Additional relevant studies were identified by reviewing the reference list of articles retrieved and searching the Cochrane Database for Systematic Reviews. We used a search strategy based on a variety of keywords and Medical Subject Heading (MeSH) terms, with the search algorithm modified as necessary for each database Search terms included *positron emission tomography, head and neck neoplasm, squamous cell carcinoma, local neoplasm recurrence, residual neoplasm, squamous cell carcinoma, local neoplasm recurrence, sensitivity and specificity* (see Supplementary Appendix 1 and 2 at www.otojournal.org/supplemental). There were no language restrictions for our search, and we included all prospective studies published until February 28, 2015.

Study Selection and Eligibility Criteria

Two reviewers were involved in the selection of studies, data collection, and quality assessment process; any disagreements were resolved by consensus or by discussion with a third reviewer.

Citations were initially screened to determine whether they pertained to the use of imaging in head and neck cancers. The abstracts were then assessed for eligibility for inclusion based on the following criteria:

- FDG-PET or FDG-PET/CT for posttreatment response assessment or surveillance for residual or recurrent head and neck cancer after treatment with radiotherapy or chemoradiotherapy
- Histopathological analysis and/or close clinical and imaging follow-up was used as the reference standard
- Data on the number of true-positive, true-negative, false-positive, and false-negative results were available or could be extracted based on the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) provided
- Minimum of 10 patients

The full texts of these potentially eligible articles were retrieved and evaluated to ensure that all inclusion criteria were satisfied. Review articles, case reports, commentaries, conference proceedings, and letters to the editor were excluded. Retrospective studies were also excluded as these may potentially overestimate the diagnostic accuracy. Only patients with HNSCCs were included in this review; in studies in which the population had a mixture of histology, an attempt was made to extract just the data on patients with SCCs. Studies that did not specify SCCs or where it was not possible to separate out the data on SCCs from other nonepithelial tumors were excluded. Studies were also excluded if part of the study population received surgery alone as the treatment, if the primary treatment modality was not reported, or if dual head coincidence gamma cameras were used to capture the images.

Data Collection Process and Data Items

Data from each study were extracted onto a standardized data extraction form. One reviewer collected the data, and a second reviewer checked the extracted data. We recorded the author names; journal; year of publication; sample size; initial treatment modality; description of study population including age, gender, site, and stage of disease; time to initial PET imaging; definition of positive PET scan (visual, semiquantitative); location of recurrence (local, nodal, distant, all sites considered together); reference standard; and duration of follow-up. The number of true-positive, true-negative, false-positive, and false-negative results was recorded or extracted onto a 2×2 table based on the sensitivity, specificity, PPV, NPV, and sample size data provided. Based on the data in these tables, the sensitivity, specificity, PPV, NPV, and overall accuracy were calculated for each study.

Quality Assessment of Studies

The quality of each article was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool. This 14-item assessment tool was developed by the Centre for Reviews and Dissemination at the University of New York and the Academic Medical Centre at the University of Amsterdam to allow for the consistent and reliable assessment of the quality of diagnostic accuracy studies included a systematic review.[10] Specifically, the tool assists in assessing the risk of bias, sources of variation, and reporting quality of diagnostic accuracy studies. We weighed each of these items equally and attributed a summary score to each study based on the responses to each question, with 1 for ''yes,'' 0 for ''no,'' and 0.5 for ''unclear.''

Statistical Analysis

The weighted mean pooled sensitivity, specificity, PPV, NPV, diagnostic odds ratio (DOR), and their 95% confidence

intervals (CIs) were calculated using the DerSimonian and Laird random effects model because of anticipated heterogeneity. Random effects modeling takes into account both within-study and between-study variation. To correct for any continuity errors, 0.5 was added to all cells with a frequency of 0 in order to calculate the pooled estimates.

Summary receiver operating characteristics (SROC) curves were fitted using the Moses-Shapiro-Littenberg method, and the area under the curve (AUC), Q* index, and their respective standard errors were estimated.

The Spearman's correlation coefficient was calculated to assess for threshold effect, and variability between individual studies was evaluated by plotting the diagnostic accuracy estimates on a forest plot. Heterogeneity was quantified using the I^2 index. Potential heterogeneity between individual studies was explored using single-factor meta-regression with the following covariates: sample size, QUADAS score, site of initial tumor, imaging type, timing of posttreatment scan, method of image interpretation (visual vs semiquantitative or quantitative), and clinical presentation of recurrence (symptomatic vs asymptomatic or not reported). Covariates were considered to be explanatory for the heterogeneity if the regression coefficients were statistically significant ($P < .05$).

Publication bias was quantified using the Egger's regression model, with the effect of bias assessed using the fail-safe number and trim-and-fill method. The fail-safe number was the number of studies that we would need to have missed for our observed result to be nullified to statistical nonsignificance at the $P < .05$ level. Publication bias is generally regarded as a concern if the fail-safe number is less than $5n + 10$, with n being the number of studies included in the meta-analysis.

The impact of imaging modality, method of image interpretation, and timing of scan on sensitivity and specificity separately was also assessed using subgroup analysis, and a Z test was performed to determine the statistical differences between subgroups.

Statistical analyses were performed using Meta-Disc (version 1.4, Unit of Clinical Biostatics, Ramon y Cajal Hospital, Madrid, Spain), GraphPad Prism (version 6.0, GraphPad Software, San Diego, CA), and Microsoft Excel (version 14.2.0, Microsoft, 2011).

Results

Study Selection

The search strategy identified 3411 citations, of which 312 abstracts were considered relevant. Based on the predetermined selection criteria, 150 full-text articles were evaluated, and 27 studies met our inclusion criteria and provided test accuracy data (**Table 1**; **Figure 1**).

Study Characteristics

There were a total of 1195 patients in the 27 selected studies, with the number of patients in each study varying from 12 to 98. The time from treatment to imaging ranged from 2 to 260 weeks. The timing or duration of follow-up was noted in 23 studies and ranged from 6 to 86 months. Twenty-two studies reported on the diagnostic accuracy of FDG-PET, while 5 studies reported on the use of FDG-PET/CT. Scans were assessed qualitatively in 13 studies and semiquantitatively in 10 studies, with a specific cutoff value reported in 3 studies; 4 studies did not specify whether scans were interpreted visually or semiquantitatively.

The vast majority of studies included SCCs from a variety of locations on the head and neck; 1 study[23] reported specifically on oral cancer and 1 study[37] on nasopharyngeal cancers. Fourteen studies reported on the stage of the initial tumor, with 8 of these studies[17,20,26,27,30,32,34,35] specifically enrolling patients with stage III or IV head and neck cancers. Six studies[16,19,21,29,30,35] included only patients in whom there was no evidence of distant metastases at initial diagnosis, another 4 studies[13,15,20,37] did not have such an inclusion criterion, but the study population consisted only of patients in whom distant metastases were not present initially, and 4 studies[18,22,23,36] included at least 1 patient in whom distant metastases were detected at the initial diagnosis.

Treatment involved radiotherapy without chemotherapy in 3 studies,[13,19,34] radiotherapy with chemotherapy in 9 studies,* radiotherapy with and without chemotherapy in 8 studies,[11,12,18,21,27,32,33,35] and intra-arterial chemotherapy in 3 studies.[20,23,24] The remaining 4 studies[14-16,25] included at least 1 patient either who underwent radiotherapy postoperatively or in whom neck dissection was performed in addition to radiotherapy. We could not meaningfully compare the diagnostic accuracy of using PET to detect residual/recurrent disease after radiotherapy alone versus radiotherapy with chemotherapy, as there were insufficient studies once we considered primary site and neck recurrences separately.

Only 1 study[19] specified that the study population was clinically asymptomatic for disease. Four studies[24,25,31,32] recruited clinically symptomatic patients or patients with suspected recurrence, 1 study[13] noted that at least some of the patients in the study population were symptomatic, while 21 studies did not report on the patient's clinical presentation at recurrence.

Publication Bias

The primary and nodal groups were assessed for publication bias using an Egger's regression model; no publication was observed for primary sites ($P = .48$). However, publication bias was detected for nodal sites ($P = .006$), with the fail-safe number being 1445 studies. Given the comprehensive literature search strategy used, we feel it is extremely unlikely that this large number of studies was missed.

Quality Assessment of Studies

The QUADAS score ranged from 10 to 13 out of a maximum of 14, with a median of 11.5. Most papers scored well on the items relating to variability and reporting. However, the scores for presence of bias were more variable. Only 4 studies[23,26,34,35] reported that all patients received the same reference test

*References 17, 19, 22, 26, 28, 30, 31, 36, 37.

Table 1. Characteristics of Studies

Study	Patients, n	Age, y	Male, %	Patient Characteristics	Scan Type	Interpretation of Scan	Time from Treatment to Scan	Reference Standard	Duration of Follow-up	QUADAS
Brkovich et al[11] 2006	19	Mean, 58.6 (range, 47-74)	100	Various HNSCC; RT/CRT; primary site confirmed disease free; neck disease only	PET	NR	Mean, 8.95 wk (range, 7-12 wk)	Planned neck dissection	Mean, 11.5 mo (range, 8-24 mo)	12.5
Ceulemans et al[12] 2001	40	Range, 34-80	83	Various HNSCC; RT ± chemo	PET	Visual	4 mo	Biopsy or clinical follow-up >4 mo	Median, 26 mo (range, 7-50 mo)	12
Chaiken et al[13] 1993	14	Range, 23-84	64	Various HNSCC; RT, some symptomatic at recurrence	PET	Semiquantitative	2-120 wk	Pathology or clinical and imaging follow-up	NR	10.5
Conessa et al[14] 2004	42	Mean, 60 (range, 43-78)	88	Various HNSCC, stage II, III, IV; RT postchemo/sx	PET	Visual	3-6 mo	Biopsy or clinical and imaging follow-up	Mean, 17 mo (range, 8-36 mo)	11
Fakhry et al[15] 2006	61	Mean, 57.5 (range, 41-82)	85	Various HNSCC; Sx + RT/CRT/RT/CCRT	PET	Semiquantitative	Mean, 3.5 mo (range, 2-5 mo)	Biopsy or clinical follow-up >6 mo	Mean, 7.4 mo (range, 6-14 mo)	11.5
Goerres et al[16] 2000	28	Range, 28-82	75	Various HNSCC; M0 only; RT ± chemo ± neck dissection	PET	NR	Mean, 13 mo (range, 6-35 mo)	Biopsy or clinical follow-up 6 mo	NR	11.5
Goerres et al[17] 2004	26	Mean, 56 (range, 35-76)	96	Various HNSCC; stage III/IV; CCRT	PET	Visual	Mean, 45 d (range, 40-59 d)	Histopathology or clinical follow-up 6 mo	NR	11
Gupta et al[18] 2010	57	Median, 54 (range, 33-65)	89	OPX/HPX/LX HNSCC; RT ± chemo	PET/CT	Semiquantitative	Median, 9 wk (range, 5-39 wk)	Local: biopsy or clinical follow-up 12 mo Nodal: FNAC ± neck dissection	Median, 36 mo (range, 7-45 mo)	12.5
Inohara et al[19] 2009	48	Mean, 43 (range, 41-48)	85	OPX/HPX/LX HNSCC; M0 only; CRT; Asymptomatic at recurrence; neck disease only	PET	Visual	7 wk	Histopathology or clinical follow-up 10 mo	Median, 20 mo (range, 10-41 mo)	12.5
Ito et al[20] 2010	53	Mean, 59.4	68	Various HNSCC, stage III/IV; intra-arterial CRT	PET	Visual	8-12 wk	Biopsy or clinical and imaging follow-up 9 mo	NR	12.5
Kim et al[21] 2007	97	Median, 57 (range, 17-84)	88	Various HNSCC; M0 only; RT ± chemo	PET	Semiquantitative (cutoff: 3)	1 mo	Biopsy or clinical follow-up 6 mo	Median, 20 mo (range, 6-30 mo)	12
Kishino et al[22] 2012	28	Range, 40-83	89	Various HNSCC; CCRT	PET	Visual	Median, 5 wk (range, 3-11 wk)	Biopsy/neck dissection or clinical and imaging follow-up	Median, 39 mo (range, 12-65 mo)	11
Kitagawa et al[23] 2003	23	Mean, 63.8 (range, 47-85)	78	Oral HNSCC; intra-arterial CRT	PET	Semiquantitative	Mean, 38 d	Histopathology or clinical follow-up 12 mo	NR	13
Kubota et al[24] 2004	36	Range, 19-82	86	Various HNSCC; intra-arterial CRT; symptomatic at recurrence	PET	Visual	Median, 4 mo	Histopathology or clinical follow-up 12 mo	NR	11.5
Li et al[25] 2001	43	NR	NR	Various HNSCC; RT ± chemo ± sx; symptomatic at recurrence	PET	Visual	Mean, 11 mo (range, 4-36 mo)	Biopsy or clinical follow-up 6 mo	NR	11
McCollum et al[26] 2004	40	Median, 54 (range, 29-78)	78	Various HNSCC; stage III/IV; induction chemo + CRT	PET	Visual	4-12 wk	Histopathology	Median, 20.5 mo (range, 9.7-30 mo)	11
Moeller et al[27] 2009	98	Mean, 58 (range, 36-79)	85	OPX/HPX/LX HNSCC; stage III/IV; RT ± chemo	PET/CT	Semiquantitative (cutoff: local 6.5, nodal 2.8)	5-12 wk	Histopathology or clinical and imaging follow-up 6 mo	Range, 37-122 wk	11.5
Mori et al[28] 2011	65	Mean, 65 (range, 36-85)	88	Various HNSCC, CCRT	PET	Semiquantitative	4-6 wk	Local: Hispathology Nodal: FNAC	NR	12.5
Nam et al[29] 2005	24	Median, 59 (range, 17-78)	96	Various HNSCC; M0 only; RT	PET	Semiquantitative (cutoff: 3)	4-6 wk	Histopathology or clinical follow-up 6 mo	Median, 12 mo (range, 7-16 mo)	10
Nayak et al[30] 2007	43	NR	NR	Various HNSCC; stage IV; M0 only; CRT; neck disease only	PET/CT	NR	2-5 mo	Histopathology/neck dissection or clinical follow-up 5 mo	Median, 18.1 mo	11
Ng et al[31] 2011	79	Mean, 52.4 (range, 33-74)	89	OPX/HPX HNSCC; stage II/III/IV; CCRT; suspected recurrence	PET/CT	Visual	Mean, 6.5 mo (range, 2.8-35.6 mo)	Biopsy or clinical and imaging follow-up 12 mo	NR	11.5
Porceddu et al[32] 2005	39	Median, 55 (range, 37-89)	74	Various HNSCC; stage III/IV; RT ± chemo; symptomatic at recurrence	PET	NR	Median, 12 wk (range, 8-32 wk)	Histopathology or clinical follow-up 12 mo	Median, 34 mo (range, 16-86 mo)	10
Rege et al[33] 1994	15	Range, 23-80	80	Various HNSCC; RT ± chemo	PET	Visual	2.5-192 wk	HIstopathology OR clinical follow-up	NR	12
Rogers et al[34] 2004	12	NR	NR	Various HNSCC; stage III/IV; RT; neck disease only	PET	Visual	1 mo	Planned neck dissection	NR	12
Sjovall et al[35] 2014	82	Median, 62 (range, 34-89)	76	OPX/HPX/LX HNSCC; stage III/IV; M0 only; RT ± chemo	PET/CT	Visual	Mean, 44 d	Endoscopy ± biopsy or clinical follow-up 9 mo	NR	11.5
Wang et al[36] 2009	44	Median, 50 (range, 34-78)	95	Various HNSCC; induction chemo + CCRT	PET	Semiquantitative	Median, 16 wk (range, 12-17 wk)	Biopsy or clinical follow-up 12 mo	NR	12.5
Yen et al[37] 2006	39	Median, 51 (range, 15-79)	69	NPC; CCRT	PET	Semiquantitative	3 mo	Histopathology or clinical and imaging follow-up 12 mo	Mean, 24.2 mo (range, 11.1-40.6 mo)	11.5

Abbreviations: chemo, chemotherapy; CCRT, concurrent chemoradiotherapy; CRT, chemoradiotherapy; CT, computed tomography; FNAC, fine-needle aspiration cytology; HNSCC, head and neck squamous cell carcinoma; HPX, hypopharynx; LX, larynx; M0, no metastasis; OPX, oropharynx; QUADAS, Quality Assessment of Diagnostic Accuracy Studies; RT, radiotherapy; NPC, nasopharyngeal carcinoma; NR, not reported; PET, positron emission tomography; Sx, surgery.

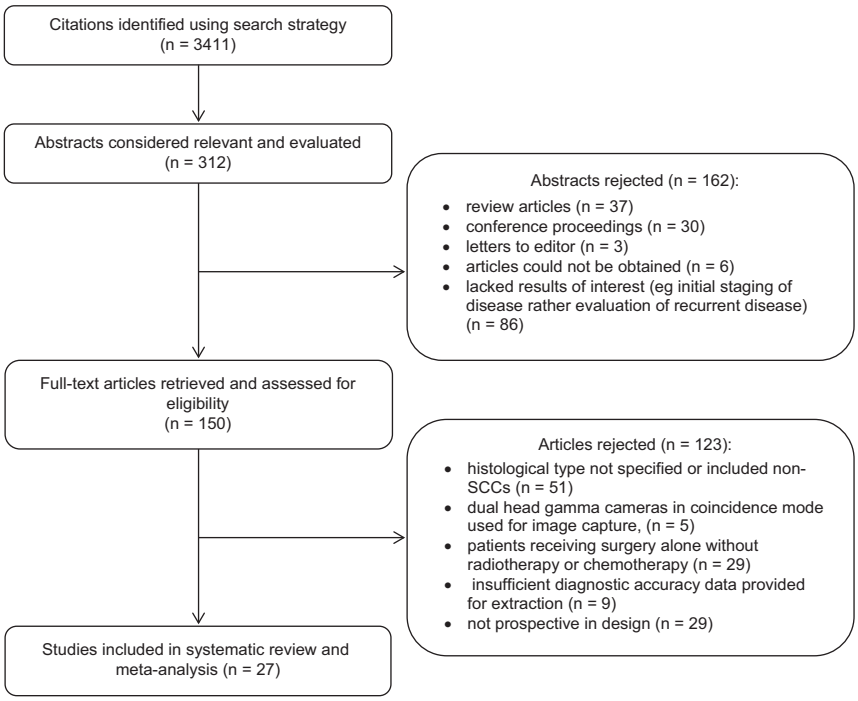


Figure 1. Literature flow diagram.

regardless of the index test result (item 6). Only 6 studies[15,20,24,26,33,36] made it clear whether or not clinical information was available at the time of image interpretation (item 12). A graphical summary of the QUADAS assessment is provided (Supplementary Appendix 3 at www.otojournal.org/supplemental).

Diagnostic Accuracy of PET and PET/CT

Primary site. For PET imaging of the primary site, there were 18 studies with a total of 854 patients. The sensitivity and specificity of FDG-PET or FDG-PET/CT in detecting local residual disease or recurrences ranged from 50.0% to 100.0% and from 31.8% to 100.0%, respectively. The Spearman correlation coefficient between the logit of sensitivity and logit of 1-specificity was 0.298 (P = .230), which suggested that a threshold effect was not present. The pooled sensitivity, specificity, PPV, NPV, and DOR were 86.2% (95% CI: 79.8%-91.1%), 82.3% (95% CI: 79.3%-85.1%), 52.7% (95% CI: 46.4%-58.9%), 96.3% (95% CI: 94.4%-97.7%), and 32.93 (95% CI: 19.17-56.56), respectively (**Figure 2**). The area under the SROC curve was 0.91 (standard error 0.02) with a Q* index of 0.85 (standard error 0.02), suggesting good diagnostic accuracy (**Figure 3**).

Neck nodes. For PET imaging of residual disease or recurrences in neck nodes, there were 15 studies involving 725 patients. The sensitivity ranged from 45.5% to 100.0%, while the specificity ranged from 53.3% to 100.0%. The pooled sensitivity, specificity, PPV, NPV, and DOR were 72.3% (95% CI: 63.1%-80.4%), 88.3% (95% CI: 85.4%-90.8%), 72.3% (95% CI: 63.1%-80.4%), 88.3% (95% CI: 85.4%-90.8%), and 22.84 (95% CI: 9.42-55.38; **Figure 4**), with no correlation between sensitivity and 1-specificy (Spearman correlation coefficient –0.359, P = .189). The area under the SROC curve was 0.86 (standard error 0.04) with a Q* index of 0.80 (standard error 0.03; **Figure 5**).

Distant metastases. For PET imaging for the detection of distant metastases at recurrence, there were 3 studies involving 182 patients. The sensitivity and specificity ranged from 69.2% to 100.0% and from 92.2% to 97.4%, respectively. The pooled sensitivity, specificity, PPV, NPV, and DOR were 84.6% (95% CI: 65.1%-95.6%), 94.9% (95% CI: 90.1%-97.8%), 84.6% (95% CI: 65.1%-95.6%), 94.9% (95% CI: 90.1%-97.8%), and 81.47 (95% CI: 21.6-307.31; **Figure 6**). The Spearman correlation coefficient between logit of sensitivity and logit of 1-specificity was 0.5 (P = .67), suggesting an absence of threshold effect. The area under the curve was 0.978 (standard error 0.02), with a Q* index of 0.93 (standard error 0.04).

Multiple sites combined. Some studies reported on the detection of residual/recurrent disease at all sites combined (ie, local, nodal, and distant residual/recurrences considered together). This included 9 studies with a total of 662 patients, of which 4 studies reported on data for local, nodal, and/or distant recurrent disease separately as well as together. For the detection of recurrent or residual disease at any site, the sensitivity ranged from 53.8% to 100%, while the specificity ranged from 47.7% to 95.5%. The pooled sensitivity, specificity, PPV, NPV, and DOR were 81.6% (95% CI: 75.1%-87%), 86.3% (95% CI: 82.9%-89.3%), 81.6% (95% CI: 75.0%-87.0%), and 33.60 (95% CI: 16.16-

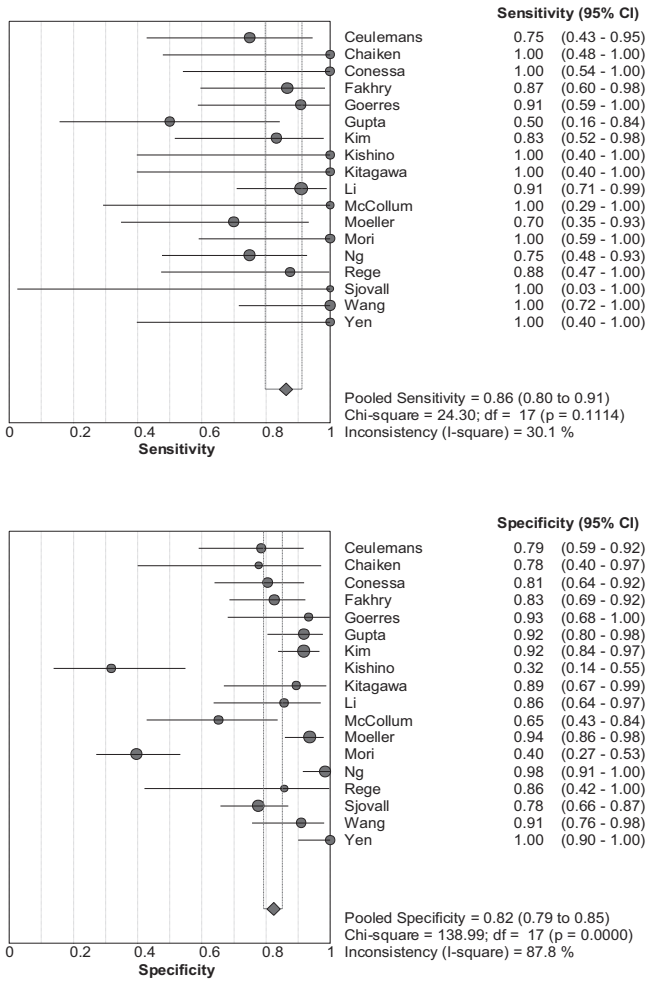


Figure 2. Forest plots of sensitivity and specificity for positron emission tomography and positron emission tomography/computed tomography in the diagnosis of residual and/or recurrent disease at the primary site.

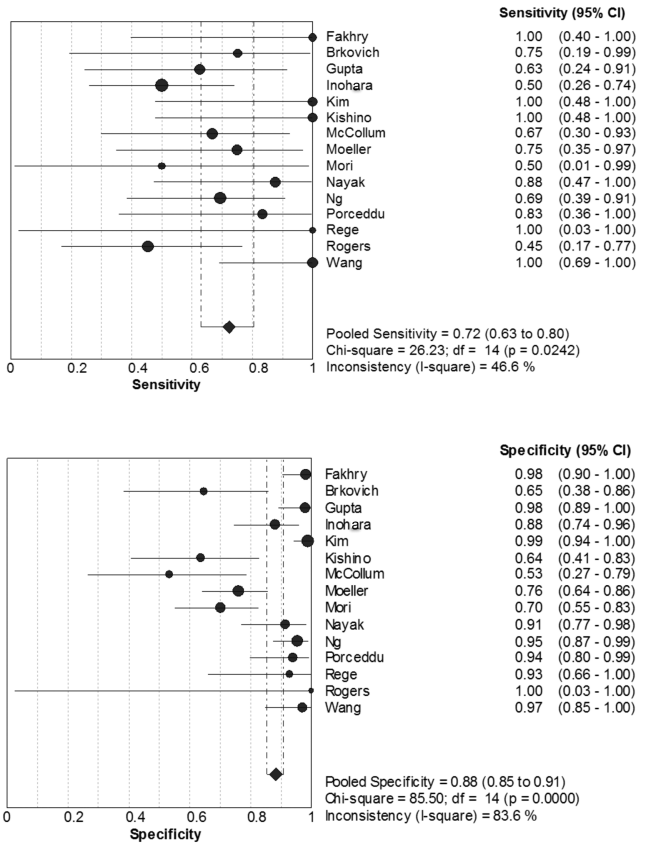


Figure 4. Forest plots of sensitivity and specificity for positron emission tomography and positron emission tomography/computed tomography in the diagnosis of residual and/or recurrent disease at the neck nodes.

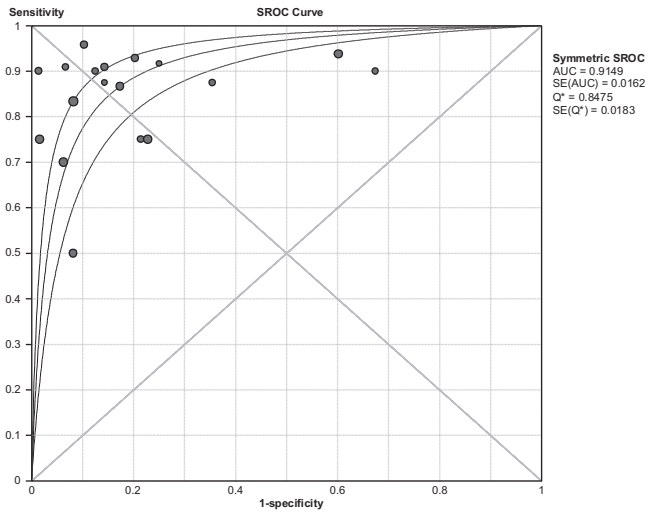


Figure 3. Summary receiver-operating characteristic curves for positron emission tomography and positron emission tomography/computed tomography in the diagnosis of residual and/or recurrent disease at the primary site.

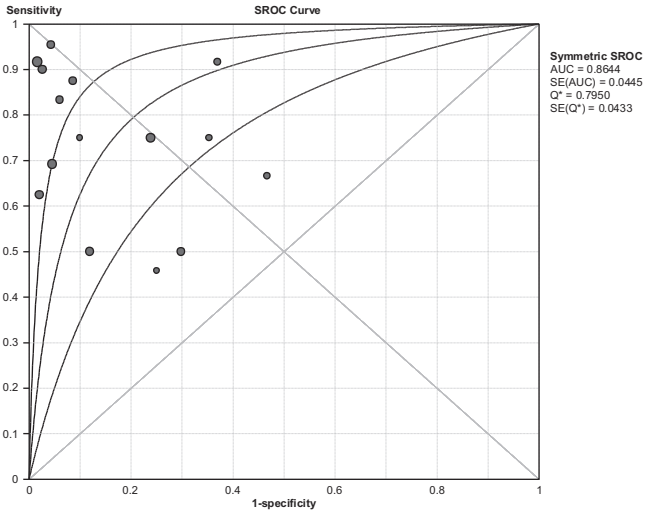


Figure 5. Summary receiver-operating characteristic curves for positron emission tomography and positron emission tomography/computed tomography in the diagnosis of residual and/or recurrent disease at the neck nodes.

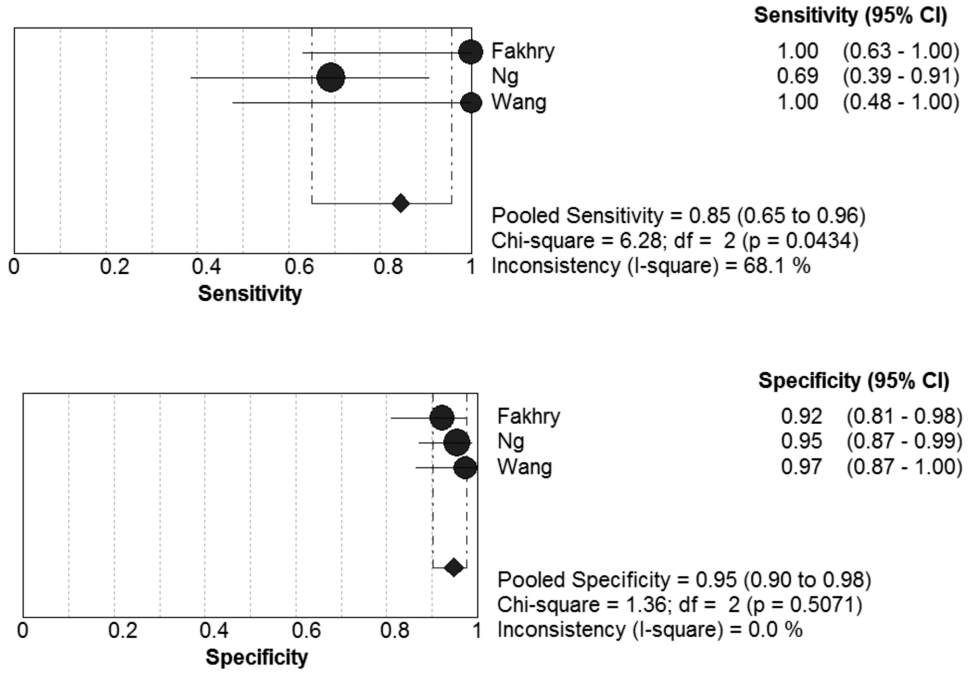


Figure 6. Forest plots of sensitivity and specificity for positron emission tomography and positron emission tomography/computed tomography in the diagnosis of distant metastases in recurrent head and neck squamous cell carcinomas.

69.82; **Figure 7**). The AUC was 0.93 (standard error 0.02) with a Q* index of 0.86 (standard error 0.02).

Meta-regression Analysis

On univariate meta-regression analysis, sample size, QUADAS score, imaging type, timing of posttreatment scan, and method of image interpretation did not affect the diagnostic odds ratio for detection of local, nodal, or overall tumor recurrence (all P values $>.05$). There were insufficient data to assess the impact of the initial tumor site and clinical presentation at recurrence on test accuracy. Because of the limited number of studies reporting on the detection of distant recurrences, the effect of these factors on test accuracy could also not be assessed.

Subgroup Analysis

Subgroup analysis could not be performed for studies evaluating the detection of distant metastases as there were insufficient studies.

PET versus PET/CT. PET/CT was found to be more specific than PET alone in the detection of residual/recurrent disease at the primary site ($P < .001$). No significant difference in sensitivity was noted between the 2 modalities for local recurrence ($P = .07$). There was no statistical difference in the sensitivity or specificity between studies using PET to detect residual/recurrent disease at neck sites compared with those using PET/CT (**Table 2**).

Visual vs semiquantitative analysis. No statistical difference was found in the sensitivity or specificity between visual and semiquantitative analysis of scans for the detection of disease at the primary site or in the neck (**Table 2**).

Timing before 12 weeks versus after 12 weeks. Studies that had scans performed both before and after the 12-week cutoff were excluded from the subgroup analysis. The specificity was significantly higher for scans performed more than 12 weeks after treatment for both local and nodal recurrence ($P = .009$ and $P = .0043$, respectively). There was no significant difference found in the sensitivity of scans (**Table 2**).

Discussion

The aim of our meta-analysis was to evaluate the diagnostic accuracy of PET and PET/CT for the detection of residual and/or recurrent disease in the post-(chemo)radiotherapy setting. We found that PET and PET/CT were highly accurate in the detection of residual and/or recurrent disease at local, nodal, and distant sites, although the timing of the scan did have an impact on the accuracy of such scans. PET/CT was more specific than PET alone in the detection of disease at the primary site. However, no difference was found between scans that were interpreted visually compared with those analyzed semiquantitatively using standard uptake values.

While there have been previous meta-analyses summarizing the diagnostic accuracy of PET and PET/CT in the detection of recurrences at locoregional[38,39] and distant[40] sites, these reviews have included retrospective as well as prospective studies, and this may overestimate the diagnostic test accuracy by introducing bias. Moreover, the meta-analysis by Gao et al[40] included patients with head and neck cancers, not specifically SCCs, while the study by Isles et al[39] included data on dual-head gamma detection systems that have inferior resolution compared with dedicated full-ring PET scanners.

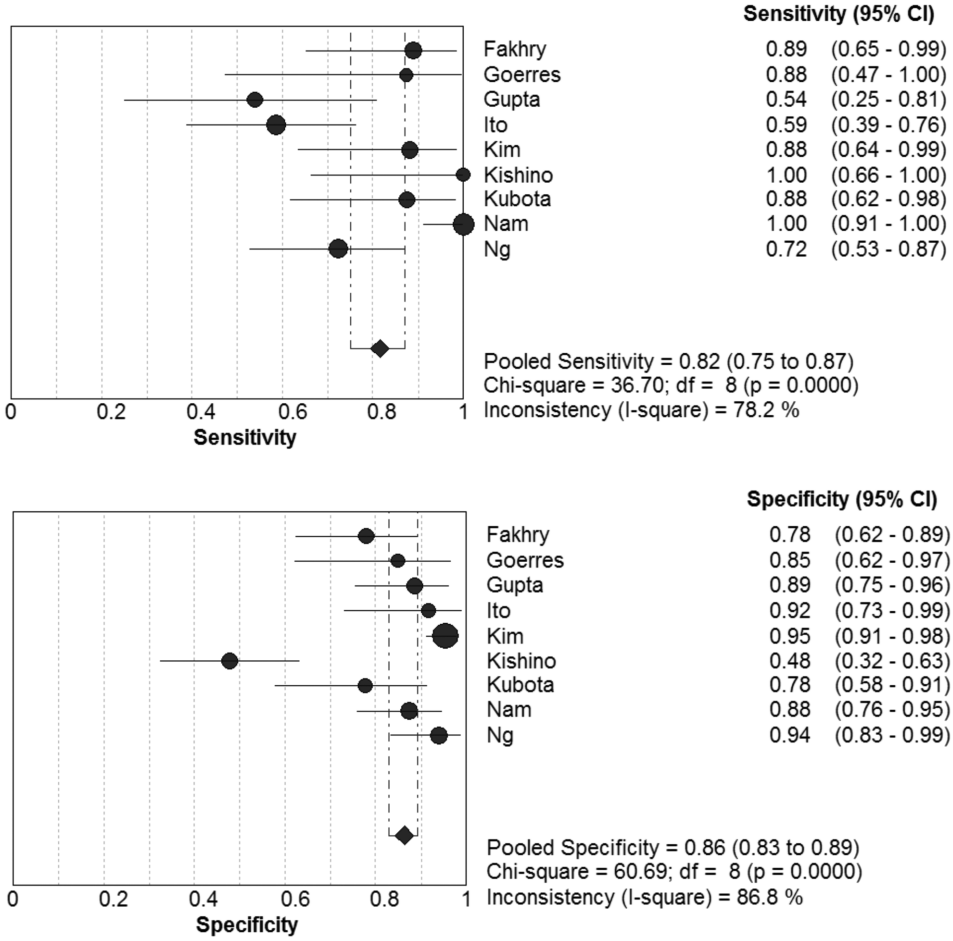


Figure 7. Forest plots of sensitivity and specificity for positron emission tomography and positron emission tomography/computed tomography in the diagnosis of residual and/or recurrent disease at multiple sites combined.

Table 2. Summary Estimates of Sensitivity and Specificity and the Results of Subgroup Analysis for Imaging Modality, Image Analysis, and Timing of Scan

	Local Residual/Recurrent Disease		Nodal Residual/Recurrent Disease	
Subgroup	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Imaging modality				
PET	90.7 (82.5-95.9)	76.5 (72.0-80.6)[a]	72.0 (60.4-81.8)	87.7 (83.9-90.8)
PET/CT	80.8 (69.9-89.1)	89.8 (85.8-92.9)[a]	73.0 (55.9-86.2)	89.4 (84.5-93.2)
Image analysis				
Visual	86.7 (78.4-92.7)	80.2 (75.5-84.4)	63.9 (50.6-75.8)	88.4 (83.4-92.4)
Semiquantitative	85.2 (73.8-93.0)	84.2 (80.1-87.8)	81.8 (64.5-93.0)	88.6 (84.3-92.0)
Timing of scan				
<12 wk	84.8 (75.0-91.9)	79.9 (76.1-83.4)[b]	67.5 (56.1-77.6)	86.2 (82.6-89.3)[c]
>12 wk	87.5 (77.6-94.1)	88.4 (83.0-92.6)[b]	82.6 (61.2-95.0)	96.0 (90.1-98.9)[c]

Abbreviations: CI, confidence interval; CT, computed tomography; PET, positron emission tomography.
[a]Statistically significant difference ($P < .001$).
[b]Statistically significant difference ($P = .009$).
[c]Statistically significant difference ($P = .004$).

In our systematic review and meta-analysis, we included only prospective studies that used dedicated PET or PET/CT scans for the detection of residual/recurrent disease in patients with SCCs of the head and neck. Our review included 27 prospective studies, comprising 1195 patients, published between 1993 and 2014. There were 6 studies

that reported exclusively on residual/recurrent disease at the primary site and 5 studies on residual/recurrent disease at the neck only. There were no prospective studies that evaluated the detection of distant metastases exclusively; all 3 studies included in our analysis for distant disease also reported on residual/recurrent disease at the primary site and in the neck.

We found that FDG-PET and FDG-PET/CT had a high overall accuracy in detecting local, nodal, and distant residual/recurrent disease after (chemo)radiotherapy in patients initially diagnosed with HNSCCs. The pooled sensitivity, specificity, PPV, and NPV for local and regional residual/recurrent disease were similar to the findings of previous meta-analyses. Gupta et al[38] reported a sensitivity of 79.9% and 72.7%, specificity of 87.5% and 87.6%, PPV of 58.6% and 52.1%, and NPV of 95.1% and 94.5% for primate site and nodal recurrences, respectively. Similarly, a systematic review by Isles et al[39] that included data from dual-head gamma cameras reported a sensitivity of 94% and 74%, specificity of 82% and 88%, PPV of 75% and 49%, and NPV of 95% and 96% for local and regional recurrent disease, respectively. For the detection of distant metastases in recurrent head and neck cancers, not specifically SCCs, Gao et al[40] reported a sensitivity of 92% and a specificity of 95%.

Our results indicated that the overall diagnostic accuracy was slightly lower for the detection of residual and recurrent disease in the neck, when compared with detection at primary, distant, or multiple sites considered together. The lower sensitivity of PET for nodal disease compared with other sites may be related to the spectrum of disease in the studies that reported on regional recurrence; nearly a third of these studies included only patients in whom nodal disease was present at initial diagnosis. Micrometastases in the lymph nodes may not be detected by imaging, leading to a higher false-negative rate and lower sensitivity.

The pooled NPVs for residual and recurrent disease at local, nodal, distant, and all sites combined were nevertheless quite high, suggesting that PET and PET/CT scans can reliably exclude residual/recurrent locoregional disease and distant metastases. A negative posttreatment scan can therefore guide the ongoing management of patients with HNSCCs and potentially reduce the need for more invasive diagnostic procedures.

PET/CT has largely superseded the use of PET alone in clinical practice, and we wanted to explore whether this newer technology would have an impact on the test diagnostic accuracy. We found that there was a small benefit of PET/CT over PET alone for the detection of residual/recurrent disease but only at the primary site ($P < .001$). The results of our subgroup analyses suggest that PET/CT has greater specificity, but no difference in sensitivity, when compared with PET alone for the detection of local recurrences. We found no significant difference between the imaging modalities in terms of sensitivity or specificity in the detection of residual/recurrent nodal disease.

The increased specificity with PET/CT for the identification of recurrent disease at the primary site may be related to the improvement in anatomical localization possible with the co-registration of anatomical and functional information. Our results differ from the findings of a previous meta-analysis by Gupta et al,[38] which showed no difference between PET and PET/CT in terms of diagnostic performance. While there were no prospective studies directly comparing the use of PET/CT and PET in head and neck cancers, the few retrospective studies directly comparing the use of the 2 modalities generally reflect the findings of our study. Fakhry et al[41] compared the use of PET and PET/CT in 32 patients who presented with a suspicion of recurrent HNSCC. They found no difference in sensitivity (94% for both modalities) and a nonsignificantly higher specificity for PET/CT (57% vs 36%-50%), and they concluded that PET/CT was more accurate than PET alone. Likewise, a study by Ishitaka et al[42] involving 129 patients with suspected head and neck (including thyroid) cancer recurrence demonstrated no significant sensitivity benefit of PET/CT over PET (sensitivity 93.9% vs 91.4%, respectively) but a significant improvement in specificity when integrated PET/CT is used (specificity 97.2% vs 74.4%). Similarly, when Chan et al[43] compared the use of the 2 modalities in 67 patients with papillomavirus-associated oropharyngeal SCCs, the findings showed that PET/CT had a better NPV compared with PET alone (98.2%-95% vs 95.7%-100%) for the detection of nodal recurrence. On the other hand, a study by Halpern et al[44] in patients with suspected local recurrence found that integrated PET/CT did not significantly improve the detection of recurrence compared with PET alone.

FDG uptake by tissues can be assessed qualitatively using visual comparison of the abnormal and normal tissue or semiquantitatively through the calculation of standardized uptake values (SUVs). While results indicated that there was a trend toward a greater sensitivity or specificity with the use of semiquantitative methods for image assessment, the difference was not statistically significant at the primary site or in the neck, suggesting that either method can be used to interpret PET scans with a reliable degree of accuracy.

This is consistent with a previous study[45] that suggested that the accuracy of visual interpretation by an experienced nuclear physician is comparable to SUV-based assessments. While the calculation of SUV may be viewed as a more objective index in assessing the uptake of FDG, it is nevertheless affected by technical aspects such as the uptake time and the selection of the region of interest.[46] Moreover, despite 3 of the included studies[21,27,29] nominating a specific cutoff value for diagnostic purposes, it has been argued that such thresholds are somewhat arbitrary because of the considerable overlap in SUVs between patients with benign and malignant lesions.[47] While many institutions report SUVs as part of their protocol, based on the results of our study, visual assessment alone is sufficient for characterizing residual/recurrent HNSCCs.

Retrospective studies directly comparing the diagnostic accuracy of PET and PET/CT at different time points after

treatment have consistently found that the accuracy of the scans varies with timing, with a longer time interval associated with a greater accuracy.[48-51] Our study confirms these findings, with scans for local or regional disease performed after 12 weeks having a greater specificity (P = .009 and P = .004, respectively), but no difference in sensitivity, when compared with those performed before 12 weeks. These results are slightly different from 2 previous meta-analyses that demonstrated an improvement in sensitivity, rather than specificity, with a delay of 10 to 12 weeks after the completion of treatment before scanning. However, both of these meta-analyses included retrospective as well as prospective studies in their analysis.

The lower specificity in the immediate posttreatment period found in our review is likely related to the increased vascularity, edema, and inflammatory changes at the primary site and in the neck after (chemo)radiation, which results in an increased physiological uptake of FDG[52] and hence more false-positive readings. However, a deliberate delay before the first posttherapy scan is not without consequences; a prolonged period between the completion of chemotherapy and salvage surgery allows time for extensive postradiation fibrosis to develop, leading to an increased frequency and severity of surgical complications.[53] In determining the optimal time for the initial scan, we must therefore balance the need for prompt diagnosis and management of disease against the risk of misleading results if scans are performed too early. However, the timing of the first posttreatment scan remains somewhat controversial despite numerous diagnostic accuracy studies. Based on the results of our review, we would support a delay of 12 weeks after (chemo)radiotherapy before imaging because of the improvement in diagnostic accuracy seen with a later scan.

Strengths and Limitations

There are several strengths of this meta-analysis. We included only prospective studies in our review, thus reducing the number of articles included in our study compared with previous meta-analyses. However, this inclusion may help reduce the risk of bias that may be found with retrospective studies. Because studies with positive results are more likely to be published, there is always the risk of publication bias with systematic reviews. We attempted to minimize the potential for such bias by using a comprehensive search strategy with no language restrictions. Our exclusion of conference abstracts, letters, editorials, and gray literature may affect the results; however, we believe that this would have minimal impact overall. Publication bias was detected for the nodal sites only, and based on the large fail-safe number ($>$1000), we believe it is highly unlikely that these studies would have not been found using our comprehensive search strategy.

The studies identified in our review had some limitations. Most notably, the reference standard was not consistent across all studies; histopathology was performed in every patient in only 4 of the 27 included studies. In most cases, histopathological confirmation was used only in patients with a positive PET or PET/CT because of the invasive nature of biopsies and neck dissections. Clinical follow-up, with and without conventional imaging, formed the basis of the reference standard in those with negative PET scans. This may potentially result in the overestimation of test sensitivity and underestimation of test specificity.[54]

There was also substantial variability in the sensitivity and specificity estimates among studies. Although the difference in imaging modality and timing explained this to some extent, some heterogeneity remained despite subgroup analysis. Other variables such as the stage and location of the tumor at initial diagnosis, the reference standard used, and the clinical presentation at recurrence may have contributed to the heterogeneity among studies. We could not assess the impact of these factors on test accuracy because of inconsistent reporting of data.

Conclusion

This is a meta-analysis focused on the diagnostic accuracy of PET and PET/CT for the detection of residual and/or recurrent local and regional disease and distant metastases in patients with HNSCCs using only prospective data. We found that both modalities had a good overall diagnostic accuracy for detection of residual and/or recurrent disease at local, nodal, and distant sites, with PET/CT being more specific than PET alone for the detection of disease at the primary site. The accuracy of visual assessment and semiquantitative analysis of images were comparable at local, nodal, and distant sites. The timing of the scan had an impact on accuracy, with later scans being more specific than earlier scans. This study has determined that the most ideal strategy for follow-up scans is after 12 weeks posttreatment with the use of combined PET and CT.

Author Contributions

Phylannie K. F. Cheung, study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, and statistical analysis; **Ronald Y. Chin**, study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, and study supervision; **Guy D. Eslick**, study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, and study supervision.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

Supplemental Material

Additional supporting information may be found at http://otojournal.org/supplemental.

References

1. Curado MP, Hashibe M. Recent changes in the epidemiology of head and neck cancer. *Curr Opin Oncol*. 2009;21:194-200.
2. Marur S, Forastiere AA. Head and neck cancer: changing epidemiology, diagnosis, and treatment. *Mayo Clinic Proc*. 2008; 83:489-501.

3. Gregoire V, Lefebvre JL, Licitra L, Felip E. Squamous cell carcinoma of the head and neck: EHNS-ESMO-ESTRO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2010;21(suppl 5):v184-v186.
4. Ang KK, Trotti A, Brown BW, et al. Randomized trial addressing risk features and time factors of surgery plus radiotherapy in advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2001;51:571-578.
5. Koness RJ, Glicksman A, Liu L, et al. Recurrence patterns with concurrent platinum-based chemotherapy and accelerated hyperfractionated radiotherapy in stage III and IV head and neck cancer patients. *Am J Surg*. 1997;174:532-535.
6. Gourin CG, Watts TL, Williams HT, Patel VS, Bilodeau PA, Coleman TA. Identification of distant metastases with positron–emission tomography-computed tomography in patients with previously untreated head and neck cancer. *Laryngoscope*. 2008;118:671-675.
7. Perlow A, Bui C, Shreve P, Sundgren PC, Teknos TN, Mukherji SK. High incidence of chest malignancy detected by FDG PET in patients suspected of recurrent squamous cell carcinoma of the upper aerodigestive tract. *J Comput Assist Tomogr*. 2004;28:704-709.
8. Leemans CR, Tiwari R, Nauta JJ, van der Waal I, Snow GB. Recurrence at the primary site in head and neck cancer and the significance of neck lymph node metastases as a prognostic factor. *Cancer*. 1994;73:187-190.
9. Salaun PY, Abgral R, Querellou S, et al. Does 18fluoro-fluorodeoxyglucose positron emission tomography improve recurrence detection in patients treated for head and neck squamous cell carcinoma with negative clinical follow-up? *Head Neck*. 2007;29:1115-1120.
10. Whiting P, Rutjes AW, Dinnes J, Reitsma J, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess*. 2004;8:iii, 1-234.
11. Brkovich VS, Miller FR, Karnad AB, Hussey DH, McGuff HS, Otto RA. The role of positron emission tomography scans in the management of the N-positive neck in head and neck squamous cell carcinoma after chemoradiotherapy. *Laryngoscope*. 2006; 116:855-858.
12. Ceulemans G, Voordeckers M, Farrag A, Verdries D, Storme G, Everaert H. Can 18-FDG-PET during radiotherapy replace post-therapy scanning for detection/demonstration of tumor response in head-and-neck cancer? *Int J Radiat Oncol Biol Phys*. 2011;81:938-942.
13. Chaiken L, Rege S, Hoh C, et al. Positron emission tomography with fluorodeoxyglucose to evaluate tumor response and control after radiation therapy. *Int J Radiat Oncol Biol Phys*. 1993;27:455-464.
14. Conessa C, Foehrenbach H, Hervé S, Poncet J-L. FDG-PET scan in local follow-up of irradiated head and neck squamous cell carcinomas. *Ann Otol Rhinol Laryngol*. 2004;113:628-635.
15. Fakhry N, Jacob T, Paris J, et al. Contribution of 18-F-FDG PET for detection of head and neck carcinomas with an unknown primary tumor. *Ann Otolaryngol Chir Cervicofac*. 2006;123:17-25.
16. Goerres GW, Haenggeli CA, Allaoua M, et al. Direct comparison of F-18-FDG PET and ultrasound in the follow-up of patients with squamous cell cancer of the head and neck. *Nuklearmedizin*. 2000;39:246-250.
17. Goerres GW, Schmid DT, Bandhauer F, et al. Positron emission tomography in the early follow-up of advanced head and neck cancer. *Arch Otolaryngol Head Neck Surg*. 2004;130:105-109.
18. Gupta T, Jain S, Agarwal JP, et al. Diagnostic performance of response assessment FDG-PET/CT in patients with head and neck squamous cell carcinoma treated with high-precision definitive (chemo) radiation. *Radiother Oncol*. 2010;97:194-199.
19. Inohara H, Enomoto K, Tomiyama Y, et al. The role of CT and 18F-FDG PET in managing the neck in node-positive head and neck cancer after chemoradiotherapy. *Acta Otolaryngol*. 2009; 129:893-899.
20. Ito K, Yokoyama J, Kubota K, Morooka M, Shiibashi M, Matsuda H. 18F-FDG versus 11C-choline PET/CT for the imaging of advanced head and neck cancer after combined intra-arterial chemotherapy and radiotherapy: the time period during which PET/CT can reliably detect non-recurrence. *Eur J Nucl Med Mol Imaging*. 2010;37:1318-1327.
21. Kim SY, Lee S-W, Nam SY, et al. The feasibility of 18F-FDG PET scans 1 month after completing radiotherapy of squamous cell carcinoma of the head and neck. *J Nucl Med*. 2007;48:373-378.
22. Kishino T, Hoshikawa H, Nishiyama Y, Yamamoto Y, Mori N. Usefulness of 3′-deoxy-3′-18F-fluorothymidine PET for predicting early response to chemoradiotherapy in head and neck cancer. *J Nucl Med*. 2012;53:1521-1527.
23. Kitagawa Y, Nishizawa S, Sano K, et al. Prospective comparison of 18F-FDG PET with conventional imaging modalities (MRI, CT, and 67Ga scintigraphy) in assessment of combined intraarterial chemotherapy and radiotherapy for head and neck carcinoma. *J Nucl Med*. 2003;44:198-206.
24. Kubota K, Yokoyama J, Yamaguchi K, et al. FDG-PET delayed imaging for the detection of head and neck cancer recurrence after radio-chemotherapy: comparison with MRI/CT. *Eur J Nucl Med Mol Imaging*. 2004;31:590-595.
25. Li P, Zhuang H, Mozley PD, et al. Evaluation of recurrent squamous cell carcinoma of the head and neck with FDG positron emission tomography. *Clin Nucl Med*. 2001;26:131-135.
26. McCollum AD, Burrell SC, Haddad RI, et al. Positron emission tomography with 18F-fluorodeoxyglucose to predict pathologic response after induction chemotherapy and definitive chemoradiotherapy in head and neck cancer. *Head Neck*. 2004;26:890-896.
27. Moeller BJ, Rana V, Cannon BA, et al. Prospective risk-adjusted [18F] fluorodeoxyglucose positron emission tomography and computed tomography assessment of radiation response in head and neck cancer. *J Clin Oncol*. 2009;27:2509-2515.
28. Mori M, Tsukuda M, Horiuchi C, et al. Efficacy of fluoro-2-deoxy-D-glucose positron emission tomography to evaluate responses to concurrent chemoradiotherapy for head and neck squamous cell carcinoma. *Auris Nasus Larynx*. 2011;38:724-729.
29. Nam SY, Lee S-W, Im KC, et al. Early evaluation of the response to radiotherapy of patients with squamous cell carcinoma of the head and neck using 18 FDG-PET. *Oral Oncol* 2005;41:390-395.
30. Nayak JV, Walvekar RR, Andrade RS, et al. Deferring planned neck dissection following chemoradiation for stage IV

head and neck cancer: the utility of PET-CT. *Laryngoscope*. 2007;117:2129-2134.
31. Ng S-H, Chan S-C, Yen T-C, et al. PET/CT and 3-T whole-body MRI in the detection of malignancy in treated oropharyngeal and hypopharyngeal carcinoma. *Eur J Nucl Med Mol Imaging*. 2011;38:996-1008.
32. Porceddu SV, Jarmolowski E, Hicks RJ, et al. Utility of positron emission tomography for the detection of disease in residual neck nodes after (chemo) radiotherapy in head and neck cancer. *Head Neck*. 2005;27:175-181.
33. Rege S, Maass A, Chaiken L, et al. Use of positron emission tomography with fluorodeoxyglucose in patients with extracranial head and neck cancers. *Cancer*. 1994;73:3047-3058.
34. Rogers JW, Greven KM, McGuirt WF, et al. Can post–rt neck dissection be omitted for patients with head-and-neck cancer who have a negative pet scan after definitive radiation therapy? *Int J Radiat Oncol Biol Phys*. 2004;58:694-697.
35. Sjovall J, Brun E, Almquist H, Kjellen E, Wahlberg P. Radiotherapy response in head and neck cancer: evaluation of the primary tumour site. *Acta Otolaryngol*. 2014;134:646-651.
36. Wang YF, Liu RS, Chu PY, et al. Positron emission tomography in surveillance of head and neck squamous cell carcinoma after definitive chemoradiotherapy. *Head Neck*. 2009;31:442-451.
37. Yen TC, Lin CY, Wang HM, et al. 18F-FDG-PET for evaluation of the response to concurrent chemoradiation therapy with intensity-modulated radiation technique for Stage T4 nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys*. 2006;65:1307-1314.
38. Gupta T, Master Z, Kannan S, et al. Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging*. 2011;38:2083-2095.
39. Isles MG, McConkey C, Mehanna HM. A systematic review and meta-analysis of the role of positron emission tomography in the follow up of head and neck squamous cell carcinoma following radiotherapy or chemoradiotherapy. *Clin Otolaryngol*. 2008;33:210-222.
40. Gao S, Li S, Yang X, Tang Q. 18FDG PET-CT for distant metastases in patients with recurrent head and neck cancer after definitive treatment: a meta-analysis. *Oral Oncol*. 2014;50:163-167.
41. Fakhry N, Lussato D, Jacob T, Giorgi R, Giovanni A, Zanaret M. Comparison between PET and PET/CT in recurrent head and neck cancer and clinical implications. *Eur Arch Otorhinolaryngol*. 2007;264:531-538.
42. Ishikita T, Oriuchi N, Higuchi T, et al. Additional value of integrated PET/CT over PET alone in the initial staging and follow up of head and neck malignancy. *Ann Nucl Med*. 2010;24:77-82.
43. Chan JY, Sanguineti G, Richmon JD, et al. Retrospective review of positron emission tomography with contrast-enhanced computed tomography in the posttreatment setting in human papillomavirus–associated oropharyngeal carcinoma. *Arch Otolaryngol Head Neck Surg*. 2012;138:1040-1046.
44. Halpern BS, Yeom K, Fueger BJ, Lufkin RB, Czernin J, Allen-Auerbach M. Evaluation of suspected local recurrence in head and neck cancer: a comparison between PET and PET/CT for biopsy proven lesions. *Eur J Radiol*. 2007;62:199-204.
45. Wong RJ, Lin DT, Schoder H, et al. Diagnostic and prognostic value of [(18)F]fluorodeoxyglucose positron emission tomography for recurrent head and neck squamous cell carcinoma. *J Clin Oncol*. 2002;20:4199-4208.
46. Keyes JW Jr.SUV: standard uptake or silly useless value? *J Nucl Med*. 1995;36:1836-1839.
47. Lapela M, Eigtved A, Jyrkkiö S, et al. Experience in qualitative and quantitative FDG PET in follow-up of patients with suspected recurrence from head and neck cancer. *Eur J Cancer*. 2000;36:858-867.
48. Ghanooni R, Delpierre I, Magremanne M, et al. 18F-FDG PET/CT and MRI in the follow-up of head and neck squamous cell carcinoma. *Contrast Media Mol Imaging*. 2011;6:260-266.
49. Greven KM, Williams DW, Keyes JW, et al. Positron emission tomography of patients with head and neck carcinoma before and after high dose irradiation. *Cancer*. 1994;74:1355-1359.
50. Kim JW, Roh JL, Kim JS, et al. 18F-FDG PET/CT surveillance at 3–6 and 12 months for detection of recurrence and second primary cancer in patients with head and neck squamous cell carcinoma. *Br J Cancer*. 2013;109:2973-2979.
51. Lonneux M, Lawson G, Ide C, Bausart R, Remacle M, Pauwels S. Positron emission tomography with fluorodeoxyglucose for suspected head and neck tumor recurrence in the symptomatic patient. *Laryngoscope*. 2000;110:1493-1497.
52. Purohit BS, Ailianou A, Dulguerov N, Becker CD, Ratib O, Becker M. FDG-PET/CT pitfalls in oncological head and neck imaging. *Insights Imaging*. 2014;5:585-602.
53. Lango MN, Myers JN, Garden AS. Controversies in surgical management of the node-positive neck after chemoradiation. *Semin Radiat Oncol*. 2009;19:24-28.
54. Zhou XH. Correcting for verification bias in studies of a diagnostic test's accuracy. *Stat Methods Med Res*. 1998;7:337-353.

European Journal of Cancer 63 (2016) 88–96

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Impact of ^{18}F-FDG PET/CT staging on management and prognostic stratification in head and neck squamous cell carcinoma: A prospective observational study

In Sun Ryu [a], Jong-Lyel Roh [b,*], Jae Seung Kim [c], Jeong Hyun Lee [d], Kyung-Ja Cho [e], Seung-Ho Choi [b], Soon Yuhl Nam [b], Sang Yoon Kim [b]

[a] *Department of Otolaryngology, Gangneung Asan Hospital, University of Ulsan College of Medicine, Gangneung, Republic of Korea*
[b] *Department of Otolaryngology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea*
[c] *Department of Nuclear Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea*
[d] *Department of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea*
[e] *Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea*

Received 8 February 2016; received in revised form 17 April 2016; accepted 2 May 2016
Available online 9 June 2016

KEYWORDS
^{18}F-FDG PET/CT;
Head and neck squamous cell carcinoma;
Staging;
Management;
Prognosis

Abstract ***Background:*** Accurate assessment of the extent of cancer is essential for appropriate treatment planning and outcome prediction. This study prospectively evaluated whether adding ^{18}F-fluorodeoxyglucose (^{18}F-FDG) positron emission tomography (PET)/computed tomography (CT) to the routine initial staging practice in head and neck squamous cell carcinoma (HNSCC) improved management and prognosis.

Methods: All consecutive patients with newly diagnosed HNSCC who presented in October 2010 – December 2012 underwent conventional workups (CWU) followed by PET/CT. The clinical stage and management plans before and after PET/CT were compared. PET/CT was deemed to have no/low, moderate, and high impact on management planning depending on whether PET/CT changed the treatment modality or goal. The appropriateness of PET/CT staging and management impact was confirmed by histopathology and clinical follow-up, and its association with survival was analysed.

Findings: Of the 248 patients, PET/CT changed the Tumour Node Metastasis (TNM) classification in 79 (31.9%). In the patients with discordant staging, PET/CT staging was significantly more sensitive and accurate than CWU staging (both $P < 0.001$). PET/CT had high or moderate impact on management in 39 (15.7%) patients. Patients with PET/CT upstaged disease had significantly worse progression-free survival (PFS) and overall survival (OS) than patients with no CWU-stage changes (3-year PFS = 56.8% versus 74.5%, $P = 0.043$; 3-year

* *Corresponding author*: Department of Otolaryngology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Seoul 05505, Republic of Korea. Tel.: +82 2 3010 3965; fax: +82 2 489 2773.
E-mail address: rohjl@amc.seoul.kr (J.-L. Roh).

http://dx.doi.org/10.1016/j.ejca.2016.05.002

OS = 61.3% versus 85.3%, $P = 0.006$). Multivariate analyses revealed that PET/CT staging and second primary cancer were independent predictive factors for both PFS and OS ($P < 0.05$, each).

Interpretations: ^{18}F-FDG PET/CT added important staging information that improved management and prognostic stratification in HNSCC.

1. Introduction

Head and neck squamous cell carcinoma (HNSCC) accounts for 90% of head and neck cancers and 3–5% of all human malignancies [1,2]. A recent analysis of the Surveillance Epidemiology and End Results database indicated that, in 2005–2011, the overall 5-year survival rate for all HNSCC stages was ~60% [3,4]. The 5-year relative survival rate for patients with localised disease was ~80.0%. However, approximately two third of HNSCC patients are initially diagnosed with advanced stage disease, including regional lymph node metastasis [5]. In cases of nodal and distant metastasis, the 5-year relative survival decreases to 44.5% and 35.2%, respectively [3,4]. Although various clinicopathological factors correlate with HNSCC prognosis, the most significant factor is cancer stage at diagnosis [6]. Thus, precise cancer staging is essential as it allows clinicians to select the appropriate treatment strategies and predict the prognosis of the patients.

The conventional workups (CWU) for initial HNSCC staging include physical examination, endoscopy, computed tomography (CT), and/or magnetic resonance imaging (MRI) of the head and neck to evaluate the extent of the primary tumour and whether cervical lymph nodes are involved. CT scans of the chest are also usually included because the lung is the most common site of second or metastatic HNSCC cancer [7]. However, a more sensitive method that screens the whole body may be more accurate and less time consuming [8].

^{18}F-fluorodeoxyglucose (^{18}F-FDG) positron emission tomography (PET) was rapidly adopted in oncological practice over the past decade because it is an effective imaging modality that provides both functional and anatomical information [9]. Previous reports have demonstrated that adding ^{18}F-FDG PET/CT to CWU stages HNSCC more accurately than CWU alone and may alter the clinical management [10–13]. Recent studies also suggest that PET/CT detects regional or distant metastases and second primary cancers (SPCs) better than PET alone and CWU alone [14–16]. Nevertheless, the potential role of PET/CT in primary HNSCC staging has yet to be defined, and the clinical guideline only recommend PET/CT as an option for stage III–IV HNSCC [17]. The impact of the additional information provided by PET/CT on HNSCC management and prognosis also remains poorly understood [18,19]. We, therefore, evaluated whether ^{18}F-FDG PET/CT staging affects the management plan and prognostic stratification of patients with newly diagnosed HNSCC.

2. Materials and methods

2.1. Patients

This prospective study was approved by the institutional review board of our institution. Informed consent was obtained from all enrolled patients. The primary end-point was the clinical impact of PET/CT-induced change in CWU-determined stage on the CWU-based treatment plan. The secondary end-point was the prognostic value of incorporating PET/CT in the initial staging process.

All consecutive patients (≥18 years old) with pathologically confirmed untreated HNSCC of the oral cavity, oropharynx, larynx, or hypopharynx who underwent CWU for primary cancer staging within 3 weeks of the initial treatment between October 2010 and December 2012 were enrolled. All surviving patients were followed for at least 12 months. The exclusion criteria were patients with no available data of either pre-treatment CT/MRI or ^{18}F-FDG PET/CT ($n = 33$) and with no adequate follow-up information ($n = 21$). During the study period, a total of 248 eligible patients were included in this study.

2.2. Study design

CWU stage was determined on the basis of CWU before PET/CT. According to the protocol of our institution, CWU includes physical and endoscopic examinations, contrast-enhanced CT and/or MRI of the head and neck, CT of the chest, and flexible oesophagogastroduodenoscopy because synchronous cancers in HNSCC are predominantly located in the upper aerodigestive tract [20]. The CWU results are then reviewed for diagnostic quality during our institutional multidisciplinary head and neck oncology team meetings. The team consists of experienced surgical, medical, and radiation oncologists. The tumours are staged according to the *American Joint Committee on Cancer Staging Manual* (7th ed., 2010) [21].

The team also determines which treatment is appropriate on the basis of the CWU stage, and then the decision was written in medical records. Depending on the location of the primary tumour and clinical stage, the treatment options are definitive definite radiotherapy (RT), chemoradiotherapy (CRT), induction chemotherapy (ICT), and/or surgery.

During the study period, all patients underwent ^{18}F-FDG PET/CT using a Biograph Sensation 16 or True Point 40 System (Siemens Medical Systems, Knoxville, TN) after CWU. Patients were required to fast for an average time of 13.6 h (standard deviation [SD], 16.0; range, 7–20). The average patient blood glucose level was 102 (SD, 16.1; range, 67–149) mg/dL. Patients were injected with an average of 398.6 (SD, 216.5; range, 372–555) MBq of ^{18}F-FDG and incubated for an average period of 60.2 min (SD, 6.2; range, 51–70). Before acquiring the PET emission data, spiral CT scanning was performed in spiral mode from the skull base to the proximal thigh at 100 mAs and 120 kV, with a section width of 5 mm and collimation of 0.75 mm. No oral or intravenous contrast medium was used. The PET results were reconstructed using CT attenuation correction, an attenuation-weighted algorithm (2 iterations and 16 subsets), and a post-reconstruction smoothing Gaussian filter (full width at half maximum = 6 mm). Images were reconstructed using a 168 × 168 matrix (pixel size = 5.3 mm).

The PET/CT findings were then reviewed on the workstation by an experienced nuclear medicine physician (J.S.K.) who was blinded to the CWU findings. Increased focal ^{18}F-FDG uptake in the tumour and metastatic nodes were graded from 1 to 4, where grades 3 and 4 were regarded as evidence of tumour involvement. Visual and semiquantitative analyses were used to determine abnormally increased focal ^{18}F-FDG uptake in comparison with the background and blood-pool activity in the mediastinum. But strict standardised uptake value cutoffs were not used. The CT signs for assessing nodal metastases are based on nodal size (shortest axial diameter > 11 mm in the jugulodigastric regions or > 10 mm in other cervical regions) and shape, the presence of central necrosis, and the presence of a localised group of nodes in an expected node-draining area for a specific primary tumour. The cartilage or bone destruction by tumour was also used for image interpretation.

The PET/CT results were added to the CWU findings during the separate decision-making meeting. Whether this changed the TNM classification (i.e. the T, N, and/or M stage was altered) and management plan was then recorded prospectively. The impact of PET/CT on the management plan was classified as follows [10]: high (change in planned treatment modality or purpose, e.g. surgery to CRT, curative to palliative), moderate (change in delivery within the same treatment modality, e.g. a change in the RT target volume or a change in extent of surgical resection), low (no change in proposed management), or no (PET/CT result ignored).

Table 1
Patient characteristics (N = 248).

Characteristics	N (%)
Gender	
Male/female	208 (83.9)/40 (16.1)
Age, years	
Median (IQR)	61 (54–69)
Smoking, >20 pack-year	144 (58.1)
Alcohol drinking, ≥1 drink per day	173 (69.8)
Site of primary tumour	
Oral cavity	62 (25.0)
Oropharynx	56 (22.5)
Larynx	99 (40.0)
Hypopharynx	31 (12.5)
Histological grade	
WD/MD/PD/NA	67 (27.0)/137 (55.2)/35 (14.1)/9 (3.7)
Treatment	
Surgery alone	70 (28.2)
Surgery + RT/CRT	77 (31.0)
IC + surgery ± RT/CRT	6 (2.4)
IC + CRT ± surgery	17 (6.9)
RT/CRT/CT alone	37 (14.9)/40 (16.2)/1 (0.4)
Treatment intention	
Curative	241 (97.2)
Palliative	7 (2.8)
Follow-up	
Follow-up period, median (range), months	38.0 (12.3–55.3)
Disease progression	68 (27.4)
Last status, NED/AWD/DOD/DOC	191 (77.0)/11 (4.4)/37 (15.0)/9 (3.6)
Synchronous SPC found at initial staging	18 (7.3)

The data are shown as number (%) unless otherwise indicated.
Abbreviations: AWD, alive with disease; CRT, chemoradiotherapy; CT, chemotherapy; DOC, died of other cause; DOD, died of disease (index cancer); ICT, induction chemotherapy; IQR, interquartile range; MD, moderated differentiated; NA, not available; NED, no evidence of disease; PD, poorly differentiated; RT, radiotherapy; SPC, second primary cancer; WD, well differentiated.

The validation was determined by assessing the histopathology for the only cases in which there was the discrepant staging and/or management change between CWU and CWU + PET/CT results. For some patients who underwent nonsurgical treatment, subsequent serial imaging and clinical follow-up were also considered when histopathologic diagnosis was not obtained because of difficulty in approaching the suspicious malignant lesions. Of these, the validation by clinical follow-up was regarded as 'not assessable' in some cases of the use of treatment intervention (e.g. RT/CRT was applied both neck side in case of advanced T stage) that could alter disease extent. The latter cases were not included in the analysis.

After the initial treatment, all patients underwent physical and endoscopic examinations at each clinic visit, and serial imaging workups were performed regularly.

2.3. Statistical analyses

Groups were compared in terms of categorical variables by using the Fisher's exact test. PET/CT and CWU were compared in terms of the sensitivity and accuracy with which they staged HNSCC by using the McNemar test. The Kaplan–Meier method was used to estimate overall survival (OS) and index cancer progression (progression-free survival [PFS]). OS and PFS were defined as the time between the first day of treatment to the date of death or progression, respectively, or to the last clinical follow-up. Disease progression was defined as the appearance of new lesions or enlargement of the initial primary tumour and/or development of metastatic disease [22]. The log-rank test was used to compare survival rates according to stage and management impact. The Cox proportional hazards model was used to identify the prognostic variables for univariate and multivariate predictions of PFS and OS. The tests were based on the likelihood ratio statistic, and the estimated hazard ratio (HR) and 95% confidence intervals (CIs) were calculated. All statistical analyses were two sided and were performed by using SPSS software, version 22.0. $P < 0.05$ was deemed to indicate statistical significance.

3. Results

The patient clinical characteristics are presented in Table 1. PET/CT changed the CWU-based TNM classification of 83 lesions in 79 (31.9%) patients; in two patients, both T and N changed. In the remaining two patients, both N and M changed. In the remaining 169 (68.1%) patients, PET/CT and CWU findings presented identical TNM classification. Of the 79 patients with discordant TNM classification, histopathology was available in 68 (86.0%). In another two and one patient, the TNM classification was confirmed by subsequent imaging and clinical follow-up, respectively. It was not possible to definitively confirm the stage in the remaining eight patients. These cases were not included when comparing the PET/CT stage and CWU stage in terms of diagnostic accuracy.

The discrepancies of T classification were identified in 24 (30.3%) patients. The extent of the primary tumour was not confirmed in two patients because ICT was performed. There were no false-positive PET/CT results for the detection of primary tumour. However, PET/CT failed to detect 15 of the CWU-staged T1 tumours (15 of 101, 14.8%). The discrepancies of N classification were identified in 55 (69.6%) patients (Supplementary Table S1). The discordant nodal stage was confirmed by histopathology in 46 and serial imaging in 1. PET/CT classified the N classification more accurately and sensitively than CWU (both $P < 0.05$, Supplementary Table S2). The discrepancies of M classification were identified in four (5.0%) patients. PET/CT downstaged two of four patients, correctly in all cases, and upstaged one patient correctly. The remaining one was not assessable because of therapeutic interventions for unconfirmed site of disease visualised only on imaging.

In terms of detecting synchronous SPCs, CWU and PET/CT differed in 21 (11%) patients. CWU detected SPCs in 12 of these 21 patients: these SPCs were in the oesophagus ($n = 4$), stomach ($n = 2$), thyroid ($n = 2$), or lung ($n = 4$). PET/CT accurately excluded the SPC in the lung in three of the latter four cases. However, PET/CT failed to detect the four cases of CWU-detected SPC in the oesophagus: this was inaccurate. PET/CT accurately detected the remaining one in the lung, the two in the stomach, and the two in the thyroid. Additionally, PET/CT, but not CWU, detected SPCs in the remaining nine patients: the SPCs were in the thyroid ($n = 3$), lung ($n = 2$), colon ($n = 1$), breast ($n = 1$), palatine tonsil ($n = 1$), and epiglottis ($n = 1$).

Overall, PET/CT stage and CWU stage were discordant in 79 patients (31.9%), for whom a validation was available in 71 patients. PET/CT staging was significantly more sensitive and accurate than CWU staging (both $P < 0.001$; Table 2). Considering the whole population of the study, we hypothesise that the stages would be correct for patients with identical PET/CT and CWU–TNM stages because changes of management were not expected in these patients. The overall accuracy of PET/CT staging was significantly higher than those of CWU staging (87.1% versus 82.0%; $P < 0.001$).

Patients whose CWU-determined stage was upstaged by PET/CT staging had a significantly worse PFS and OS than those whose clinical stage did not change (both $P < 0.05$) (Fig. 1).

3.1. Primary outcome

Overall, the PET/CT staging led to management changes in 39 (15.7%) of the 248 patients. PET/CT had a significantly higher impact on the CWU-determined management plan in patients who were CWU staged as III–IV than in those who were CWU staged as I–II (21.4% versus 9.8%, $P = 0.014$).

In 12 patients (4.8%), PET-CT had a high impact on the CWU-determined management plan. In most cases, this was because PET/CT detected distant metastasis and SPCs (details are given in Supplementary Table S3). All these lesions were histologically confirmed. PET/CT correctly changed the management of these patients. In 27 (10.9%) patients, PET/CT had a moderate impact on the management plan. This was mainly because PET-CT upstaged the nodal stage. This led to modification of the radiation field and/or dose ($n = 9$) and surgical extent ($n = 18$). The actual disease stage could be validated in 24 patients: PET/CT correctly changed the management in 19 of these patients. PET/CT had a low impact in 206 (83%) patients, predominantly because it concurred with

Table 2
Comparison of diagnostic accuracy of staging by conventional workups with and without ^{18}F-FDG PET/CT in discordant cases.

Imaging	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
TNM staging[a] (n = 71)									
CWU	16	12	31	12	34.0 (22.9–45.0)	50.0 (38.3–61.6)	39.4 (28.0–50.7)	57.1 (45.5–68.6)	27.9 (17.4–38.3)
CWU + PET/CT	29	14	17	11	63.0 (51.7–74.2)	44.0 (32.4–55.5)	56.3 (44.7–67.8)	67.4 (56.5–78.3)	39.2 (27.8–50.5)
P value[b]					**< 0.001**	0.500	**< 0.001**		

Abbreviations: CWU, conventional work-ups; ^{18}F-FDG, fluorine 18-fluorodeoxyglucose; FN, false negative; FP, false positive; NPV, negative predictive value; PET, positron emission tomography; PPV, positive predictive value; TN, true negative; TP, true positive.
[a] TNM staging represented as T + N + M, according to discrepancies.
[b] P values were determined by McNemar's test. Values in bold indicate $P < 0.05$.

the CWU staging. Three (1.2%) patients were classified as no impact because the PET/CT results were ignored in the management decision making.

Patients in whom PET/CT had a high impact on the management plan had significantly worse PFS and OS than those in whom PET/CT had low/no impact (both $P < 0.001$; Fig. 2).

3.2. Secondary outcome

The Kaplan–Meier estimates of 3-year PFS and OS rates in all patients were 72.8% and 83.0%, respectively. Univariate analyses for PFS showed that CWU stage, PET/CT stage, and SPC associated significantly with lower PFS (all $P < 0.05$). Multivariate analysis revealed that PET/CT stage III–IV (HR = 2.05, 95% CI = 1.25–3.44; $P = 0.007$) and SPCs (HR = 2.30, 95% CI = 1.16–4.54; $P = 0.016$) independently predicted PFS. Univariate analyses for OS showed that CWU stage, PET/CT stage, and SPC associated significantly with lower OS (all $P < 0.05$). Multivariate analysis demonstrated that PET/CT stage III–IV (HR = 4.70, 95% CI = 2.08–10.60; $P < 0.001$) and SPCs (HR = 3.07, 95% CI = 1.51–6.23; $P = 0.002$) independently predicted reduced OS (Table 3).

Subset analyses showed that the 3-year OS of the 122 patients with CWU stages I–II disease was 92.3% and the patients with CWU stages I and II did not differ in terms of OS ($P = 0.317$, Fig. 3A). However, after PET/CT, 98, 13, 7, and 4 of these 122 patients were re-staged as stage I, II, III, and IV, respectively. This PET/CT re-staging was of prognostic significance as the 3-year OS rates of these four groups were 94.8%, 92.3%, 85.7%, and 50.0%, respectively ($P = 0.002$, Fig. 3B).

4. Discussion

The current study evaluated the impact of incorporating PET/CT findings into the initial staging process on the management and prognostic stratification of patients with newly diagnosed HNSCC. A few prospective studies have investigated the management impact of PET alone or PET/CT in a subset of patients (17–40%) and assessed only specific focus [10,11,18,19]. Moreover, there was no research to present relevant follow-up data for patients with HNSCC in regard to the prognostic stratification of PET/CT staging. To our knowledge, the current study is the first to prospectively evaluate the incremental clinical impact of PET/CT for the above topic.

PET/CT altered the management of 15.7% of our patients. This is slightly higher than the rate reported previously: Lonneux et al. showed that PET altered the management of 13.7% of 233 patients [10]. This difference may be associated with superiority of PET/CT in detecting regional or distant metastases and SPCs that are critical for selecting treatment [14–16]. Further, in terms of the primary tumour, hybrid imaging using PET/CT has been shown to improve both anatomic localisation and extent of ^{18}F-FDG-avid lesions compared with PET alone [23,24]. All our patients underwent PET/CT scan, while 83% of patients were assessed by PET in the latter report [10]. However, two other prospective studies found that PET/CT had a much greater impact on the management of HNSCC: PET/CT changed the management of 26–33.8% of the patients, including those with only stage III–IV or considerable proportion of cervical metastasis of an unknown primary (CUP). And they regarded major impact on management in whom PET/CT detected the primary tumour in CUP [11,18]. The possible explanation for different impact rates is that we included the patients at all stage of HNSCC with known primary site. Further, the result of Cacicedo et al. [18] are similar with those of our patients staged as III–IV by CWU. Our recent study already demonstrated that PET/CT lead to improved therapeutic planning in CUP [25].

The main finding of the present study was that PET/CT staging has a prognostic role in HNSCC. PET/CT staging separated the HRs for both PFS and OS better than CWU staging, regardless of which treatment was employed. Multivariate analysis revealed that PET/CT staging independently predicted worse PFS and OS. In addition, patients whose CWU-determined disease was upstaged by PET/CT had significantly poorer PFS and OS than those whose clinical stage was unchanged. Thus, PET/CT staging can reveal disease extent better than CWU staging. Our results are in line with findings

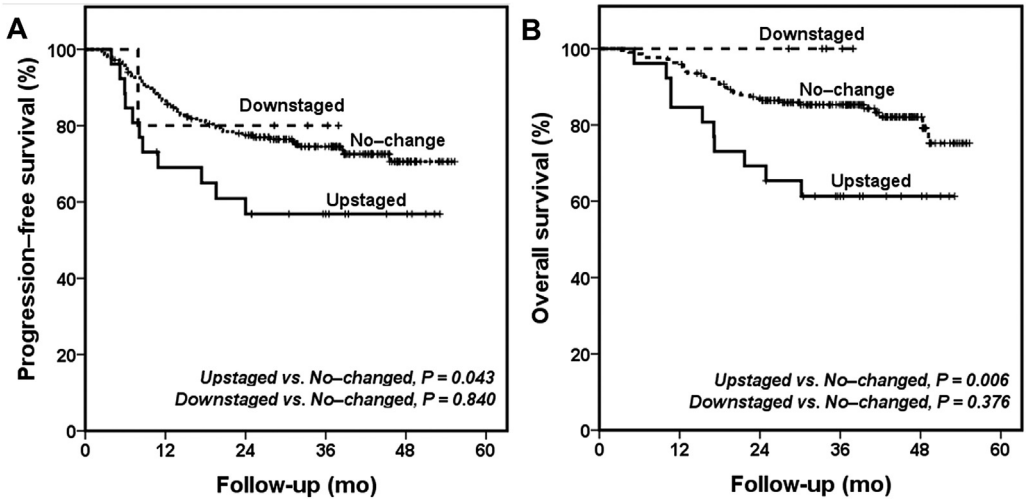


Fig. 1. Kaplan–Meier curves of progression-free survival (PFS) (A) and overall survival (OS) (B) according to the effect of PET/CT-induced changes in conventional workup (CWU) stage of the 248 eligible patients. Patients with upstaged diseases had worse PFS and OS than those whose CWU stage did not change after PET/CT (3-year PFS = 56.8% versus 74.5%, $P = 0.043$; 3-year OS = 61.3% versus 85.3%, $P = 0.006$).

in other human malignancies [26–28]. Interestingly, we found that patients whose management plan was highly impacted by PET/CT had significantly worse PFS and OS than the patients with low/no impact; in contrast, patients with moderate impact did not differ from patients with no/low impact in terms of PFS or OS. Notably, the survival curves of the patients with moderate and no/low impact converged toward the end of follow-up. This suggests that despite nodal upstaging, PET/CT appropriately modified the treatment plan in the patients with moderate impact. However, the patients whose management was highly impacted by PET/CT failed to obtain a survival benefit because PET/CT detected distant metastasis and SPCs. Detection of these lesions in patients with HNSCC generally leads to palliative treatment and death [20].

The current clinical guidelines recommend performing PET/CT as an optional imaging modality in advanced stage disease because of a low diagnostic yield of PET/CT in patients with stage I or II disease [17,29]. Indeed, we observed that PET/CT failed to detect the lesion in 14.8% of patients with T1 tumours, due to limitations of PET/CT in assessment of the early T1 stage [30]. However, we also found that PET/CT could upstage several of our patients with CWU stage I–II disease into stages III and IV; after these restaging, significant differences in stage-related prognosis were found, whereas the CWU stage I and II patients did not differ in terms of prognosis. These PET/CT-induced changes also changed the management of 9.8% of the CWU stage I and II patients. The number of these patients is relatively small and results should be interpreted carefully. Nevertheless, our findings suggest that PET/CT may need to be implemented in the routine imaging workup for initial staging in all patients with HNSCC, not only those with advanced stage disease.

In this study, PET/CT staging was significantly more accurate than CWU staging: it improved the staging accuracy in 16.1% (40/248) of the patients. Our data are similar with previous studies, which show that most PET/CT-induced stage migration is the result of nodal upstaging [10,11,18]. However, contrary to a previous meta-analysis [31], we did not find that CWU and PET/CT differ significantly in terms of the accuracy with which they detect SPCs. This may reflect the failure of PET/CT to detect synchronous oesophageal cancer. This was also observed by two recent studies [29,32]. Thus, PET/CT may detect early oesophageal cancer in HNSCC with particularly low sensitivity.

Our study had several limitations. First, the CWU and PET/CT results were only validated in the patients with discordant PET/CT and CWU stages and in those in whom PET/CT changed the management plan. Because changes of management were not expected in patients with identical PET/CT and CWU stages, we did not evaluate these patients further. We believe that this reflects the clinical practice of HNSCC staging workup and does not alter our findings. Second, our study cohort included heterogeneous tumour sites, although they were histologically identical (HNSCC). The disparities in T staging, natural courses, and clinical behaviour in HNSCC patients may exist according to the primary tumour location. Third, histopathological confirmation was unavailable in some of the patients. Fourth, this study was not designed to assess the cost-effectiveness of PET/CT in initial routine imaging workup. A randomised trial that addresses this issue is warranted. Nevertheless, our results are valuable as they

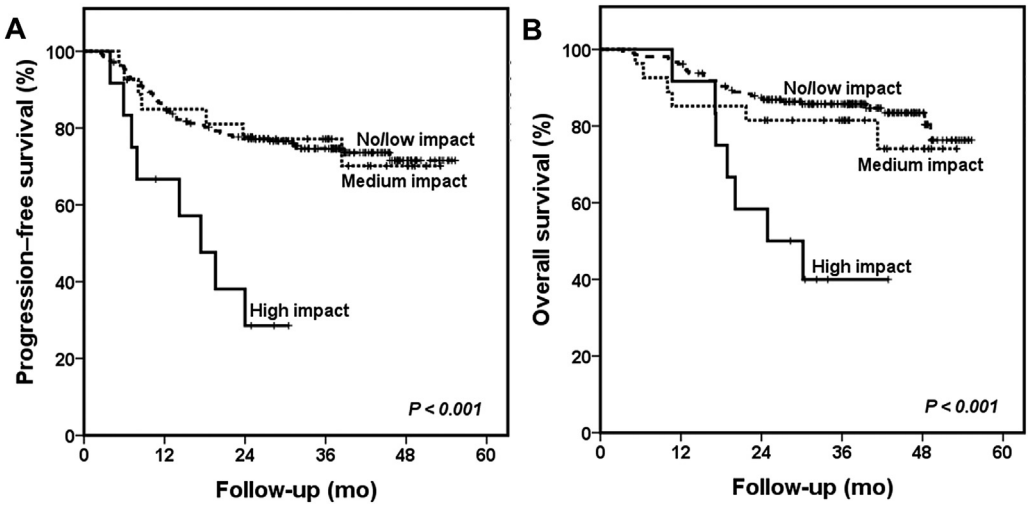


Fig. 2. Kaplan–Meier curves of progression-free survival (PFS) (A) and overall survival (OS) (B) according to the impact of PET/CT on the conventional workup-determined management plan of all patients. Patients with high impact had significantly worse PFS and OS than those with no/low impact (3-year PFS = 28.6% versus 74.6%, P 0.001; 3-year OS = 40% versus 85.7%, $P < 0.001$). However, patients with moderate impact did not significantly differ in terms of PFS or OS from patients with no/low impact.

Table 3
Factors affecting progression-free and overall survival outcomes in the study patients ($N = 248$).

Variable	Progression-free survival						Overall survival					
	Univariate			Multivariate			Univariate			Multivariate		
	HR	95% CI	P	HR	95% CI	P[a]	HR	95% CI	P	HR	95% CI	P[a]
Age >60 years	1.48	0.91–2.40	0.111				1.43	0.80–2.58	0.224			
Sex, female	1.68	0.94–2.98	0.075				1.142	0.53–2.45	0.733			
Smoking >20 pack-year	1.20	0.43–13.32	0.718				1.04	0.58–1.87	0.878			
Alcohol ≥1 drink per day	1.05	0.62–1.78	0.831				1.428	0.72–2.81	0.304			
Tumour site												
Non-oropharynx	1.87	0.95–3.67	0.066				1.44	0.67–3.10	0.343			
Tumour differentiation, poor	1.30	0.73–2.32	0.366				1.56	0.34–7.13	0.565			
Primary treatment[b]												
Nonsurgical treatment	0.73	0.43–1.21	0.226				0.93	0.51–1.70	0.826			
Conventional workup staging												
Nodal classification, N2–3	1.76	1.09–2.85	**0.020**				2.67	1.49–4.77	**0.001**			
TNM stage, III–IV	1.71	1.04–2.59	**0.031**				3.83	1.90–7.73	**<0.001**			
PET/CT staging												
Nodal classification, N2–3	1.67	1.02–2.76	**0.009**				3.11	1.69–5.72	**<0.001**			
TNM stage, III–IV	2.10	1.26–3.52	**0.005**	2.05	1.25–3.44	**0.007**	5.21	2.33–11.67	**<0.001**	4.70	2.08–10.60	**< 0.001**
Second primary cancer	2.63	1.34–5.17	**0.005**	2.30	1.16–4.54	**0.016**	4.01	1.99–8.10	**<0.001**	3.07	1.51–6.23	**0.002**

Abbreviations: CI, confidence interval; CT, computed tomography CRT, concurrent chemoradiation therapy; HR, hazard ratio; ICT, induction chemotherapy; PET, positron emission tomography; RT, radiotherapy.

[a] In multivariate analysis, Cox proportional hazard regression analyses were performed with backward elimination using variables with P values < 0.05 on univariate analyses. Values in bold indicate $P < 0.05$.

[b] The treatment modalities were divided into two major categories as follows: surgical treatment included surgery alone, surgery plus adjuvant RT or CRT, and ICT followed by definite surgery. Non-surgical treatment included definite CRT or RT and ICT followed by definite CRT with or without salvage surgery.

showed that initial PET/CT is important not only for staging and management planning but also for prognostic stratification.

In conclusion, this large prospective study demonstrated that incorporating ^{18}F-FDG PET/CT in CWU staging provided valuable additional information that altered the management plan in 15.7% of patients, largely because this modality detected metastatic disease or SPCs. PET/CT staging was significantly more predictive of OS and PFS outcomes than CWU staging. Our findings suggest that the incorporation of PET/CT into routine clinical practice for the primary staging of HNSCC could aid the planning of treatment and the prediction of survival outcomes.

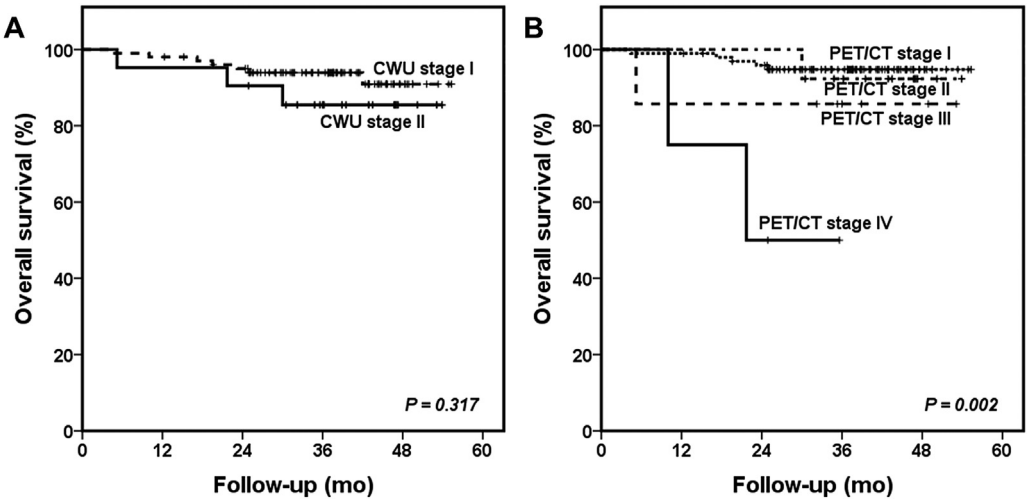


Fig. 3. Overall survival (OS) of the 122 patients with conventional work-ups (CWU)-determined stage I–II. (A) The patients with CWU stage I and II disease did not differ in OS ($P = 0.317$). (B) However, PET/CT re-staging stratified this population into groups with stage I–IV disease. The OS declined significantly as the PET/CT stage increased ($P = 0.002$).

Conflict of interest statement

None declared.

Acknowledgements

This study was supported by a grant (2015R1A2A1A15054540) from Basic Science Research Program through the National Research Foundation of Korea (NRF), Ministry of Science, ICT, and Future Planning and a grant (no. HI14C23050000) from the Korean Health Technology R&D Project, Ministry of Health & Welfare, Seoul, Republic of Korea (J.L. Roh).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejca.2016.05.002.

References

[1] A snapshot of head and neck cancer. Incidence and mortality. http://www.cancer.gov/researchandfunding/snapshots/headandneck [accessed 16.02.01].
[2] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics 2012. CA Cancer J Clin 2015;65(2):87–108.
[3] SEER stat fact sheets: oral cavity and pharynx cancer. http://seer.cancer.gov/statfacts/html/oralcav.html [accessed 16.02.01].
[4] SEER stat fact sheets: larynx cancer. http://seer.cancer.gov/statfacts/html/laryn.html [accessed 16.02.01].
[5] Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. Lancet 2008;37(9625):1695–709.
[6] Lefebvre JL. Current clinical outcomes demand new treatment options for SCCHN. Ann Oncol 2005;16(Suppl. 6):vi7–12.
[7] de Bree R, Deurloo EE, Snow GB, Leemans CR. Screening for distant metastases in patients with head and neck cancer. Laryngoscope 2000;110(3 Pt 1):397–401.
[8] Brouwer J, de Bree R, Hoekstra OS, Golding RP, Langendijk JA, Castelijns JA, et al. Screening for distant metastases in patients with head and neck cancer: is chest computed tomography sufficient? The Laryngoscope 2005;115(10):1813–7.
[9] Beyer T, Townsend DW, Brun T, Kinahan PE, Charron M, Roddy R, et al. A combined PET/CT scanner for clinical oncology. J Nucl Med 2000;41(8):1369–79.
[10] Lonneux M, Hamoir M, Reychler H, Maingon P, Duvillard C, Calais G, et al. Positron emission tomography with [18F]fluorodeoxyglucose improves staging and patient management in patients with head and neck squamous cell carcinoma: a multicenter prospective study. J Clin Oncol 2010;28(7):1190–5.
[11] Scott AM, Gunawardana DH, Bartholomeusz D, Ramshaw JE, Lin P. PET changes management and improves prognostic stratification in patients with head and neck cancer: results of a multicenter prospective study. J Nucl Med 2008;49(10):1593–600.
[12] Pohar S, Brown R, Newman N, Koniarczyk M, Hsu J, Feiglin D. What does PET imaging add to conventional staging of head and neck cancer patients? Int J Rad Oncol Biol Phys 2007;68(2):383–7.
[13] VanderWalde NA, Salloum RG, Liu TL, Hornbrook MC, O'Keeffe Rosetti MC, Ritzwoller DP, et al. Positron emission tomography and stage migration in head and neck cancer. JAMA Otolaryngol Head Neck Surg 2014;140(7):654–61.
[14] Rohde M, Dyrvig AK, Johansen J, Sorensen JA, Gerke O, Nielsen AL, et al. 18F-fluoro-deoxy-glucose-positron emission tomography/computed tomography in diagnosis of head and neck squamous cell carcinoma: a systematic review and meta-analysis. Eur J Cancer Oxf Eng 1990 2014;50(13):2271–9.
[15] Kim SY, Roh JL, Yeo NK, Kim JS, Lee JH, Choi SH, et al. Combined 18F-fluorodeoxyglucose-positron emission tomography and computed tomography as a primary screening method for detecting secondprimary cancers and distant metastases in patients with head and neck cancer. Ann Oncol 2007;18(10):1698–703.
[16] Kyzas PA, Evangelou E, Denaxa-Kyza D, Ioannidis JP. 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. J Natl Cancer Inst 2008;100(10):712–20.
[17] NCCN Clinical Practice Guidelines in Oncology. Head and neck cancer. Version 2. 2014. http://www.nccn.org/professionals/physician gls/f_guidelines.asp#head-and-neck [accessed 16.02.01].

[18] Cacicedo J, Fernandez I, Del Hoyo O, Dolado A, Gomez-Suarez J, Hortelano E, et al. Should PET/CT be implemented in the routine imaging work-up of locally advanced head and neck squamous cell carcinoma? A prospective analysis. Eur J Nucl Med Mol Imaging 2015;42(9):1378–89.

[19] Connell CA, Corry J, Milner AD, Hogg A, Hicks RJ, Rischin D, et al. Clinical impact of, and prognostic stratification by, F-18 FDG PET/CT in head and neck mucosal squamous cell carcinoma. Head Neck 2007;29(11):986–95.

[20] Strobel K, Haerle SK, Stoeckli SJ, Schrank M, Soyka JD, Veit-Haibach P, et al. Head and neck squamous cell carcinoma (HNSCC)–detection of synchronous primaries with (18)F-FDG-PET/CT. Eur J Nucl Med Mol Imaging 2009;36(6):919–27.

[21] Edge SB, Compton CC, Fritz AG, Greene FL, Trotti A. AJCC cancer staging manual. 7th ed. New York, NY: Springer; 2010. p. 21–67.

[22] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer Oxf Eng 1990 2009;45(2):228–47.

[23] Lee JR, Kim JS, Roh JL, Lee JH, Baek JH, Cho KJ, et al. Detection of occult primary tumors in patients with cervical metastases of unknown primary tumors: comparison of (18)F FDG PET/CT with contrast-enhanced CT or CT/MR imaging-prospective study. Radiology 2015;274(3):764–71.

[24] Barber TW, Duong CP, Leong T, Bressel M, Drummond EG, Hicks RJ. 18F-FDG PET/CT has a high impact on patient management and provides powerful prognostic stratification in the primary staging of esophageal cancer: a prospective study with mature survival data. J Nucl Med 2012;53(6):864–71.

[25] Syed R, Bomanji JB, Nagabhushan N, Hughes S, Kayani I, Groves A, et al. Impact of combined (18)F-FDG PET/CT in head and neck tumours. Br J Cancer 2005;92(6):1046–50.

[26] Branstetter BFt, Blodgett TM, Zimmer LA, Snyderman CH, Johnson JT, Raman S, et al. Head and neck malignancy: is PET/CT more accurate than PET or CT alone? Radiology 2005;235(2):580–6.

[27] Takeuchi S, Khiewvan B, Fox PS, Swisher SG, Rohren EM, Bassett Jr RL, et al. Impact of initial PET/CT staging in terms of clinical stage, management plan, and prognosis in 592 patients with non-small-cell lung cancer. Eur J Nucl Med Mol Imaging 2014;41(5):906–14.

[28] Krammer J, Schnitzer A, Kaiser CG, Buesing KA, Sperk E, Brade J, et al. (18) F-FDG PET/CT for initial staging in breast cancer patients – is there a relevant impact on treatment planning compared to conventional staging modalities? Eur Radiol 2015;25(8):2460–9.

[29] Hanamoto A, Takenaka Y, Shimosegawa E, Ymamamoto Y, Yoshii T, Nakahara S, et al. Limitation of 2-deoxy-2-[F-18]fluoro-D-glucose positron emission tomography (FDG-PET) to detect early synchronous primary cancers in patients with untreated head and neck squamous cell cancer. Ann Nucl Med 2013;27(10):880–5.

[30] Quon A, Fischbein NJ, McDougall IR, Le QT, Loo Jr BW, Pinto H, et al. Clinical role of 18F-FDG PET/CT in the management of squamous cell carcinoma of the head and neck and thyroid carcinoma. J Nucl Med 2007;48(Suppl. 1):58S–67S.

[31] Xu GZ, Guan DJ, He ZY. (18)FDG-PET/CT for detecting distant metastases and second primary cancers in patients with head and neck cancer. A meta-analysis. Oral Oncol 2011;47(7):560–5.

[32] Yabuki K, Kubota A, Horiuchi C, Taguchi T, Nishimura G, Inamori M. Limitations of PET and PET/CT in detecting upper gastrointestinal synchronous cancer in patients with head and neck carcinoma. Eur Arch Otorhinolaryngol 2013;270(2):727–33.

Review

Use of ^{18}F-Fludeoxyglucose–Positron Emission Tomography/Computed Tomography for Patient Management and Outcome in Oropharyngeal Squamous Cell Carcinoma
A Review

Mehdi Taghipour, MD; Sara Sheikhbahaei, MD, MPH; Wael Marashdeh, MD; Lilja Solnes, MD; Anna Kiess, MD, PhD; Rathan M. Subramaniam, MD, PhD, MPH

^{18}F-fludeoxyglucose–positron emission tomography/computed tomography (FDG-PET/CT) has been performed widely in diagnosis and management of patients with oropharyngeal squamous cell carcinoma (OPSCC). This review summarizes the literature on this tool in the management of these patients. The use of FDG-PET/CT helps in accurate staging of primary tumor, nodal involvement, and distant metastasis of patients with OPSCC. Contrast-enhanced FDG-PET/CT combines high-resolution CT and functional FDG-PET, providing the optimum imaging information for patient management. Using contrast-enhanced PET/CT leads to a combined anatomic and metabolic approach to radiation therapy planning in OPSCC. Moreover, PET/CT not only is a good modality for therapy assessment but also is a powerful tool in early recurrence detection of OPSCC. Finally, the PET/CT parameters provide survival information in patients with OPSCC; however, further studies are needed to introduce a scoring system to use clinically for prognosis prediction.

JAMA Otolaryngol Head Neck Surg. 2016;142(1):79-85. doi:10.1001/jamaoto.2015.2607
Published online November 19, 2015.

Supplemental content at jamaotolaryngology.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Rathan M. Subramaniam, MD, PhD, MPH, Russell H. Morgan Department of Radiology and Radiology Science, Johns Hopkins Medical Institutions, 601 N Caroline St, JHOC 3235, Baltimore, MD 21287 (rsubram4@jhmi.edu).

Head and neck cancers are the sixth most common malignant neoplasms globally. Despite an improvement in cancer detection and treatment methods, they are still a substantial health care problem[1] and approximately 650 000 new cases are detected annually, with 350 000 deaths yearly worldwide.[2] Oropharyngeal cancer is the most common subtype of head and neck cancers, and oropharyngeal cancers account for approximately 25% of head and neck cancers in the United States.[3] Almost all oropharyngeal cancers are squamous cell carcinomas (SCCs).[2]

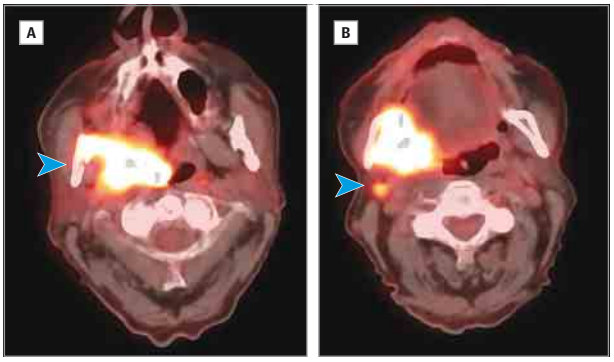
Smoking and alcohol use are the most important risk factors for oropharyngeal squamous cell carcinoma (OPSCC). Alcohol synergistically increases the carcinogenic effect of smoking. This carcinogenic effect is dose dependent for both alcohol and smoking. Human papillomavirus (HPV) infection, mainly HPV type 16, plays an important role in the etiology of head and neck cancers, especially OPSCC.[4] The incidence of HPV-related OPSCC has increased, and the palatine tonsils and base of the tongue are the most common sites of HPV-related head and neck cancers. Some studies have reported that as many as 72% of OPSCCs are positive for HPV.[5] Studies have shown that HPV-related squamous cell carcinoma (SCC) is epidemiologically, clinically, and biologically different from HPV-negative SCC. The former occurs more frequently in younger patients, predominantly in males and whites, and is associated with sexual behavior.[6] Human papillomavirus–related OPSCC is histologically different from HPV-negative disease, as it presents with higher mitotic rate, poor differentiation, no keratinization, and distinct basaloid appearance. Both prospective and retrospective studies, as well as meta-analysis, have confirmed better survival rates in patients with HPV-related OPSCC compared with HPV-negative disease, either with surgical or nonsurgical management.[7]

^{18}F-fludeoxyglucose–positron emission tomography/computed tomography (FDG-PET/CT) has been proven to be valuable in the management of many human solid tumors.[8,9] The purpose of this article was to review the literature to assess the value of FDG-PET/CT in the management of patients with OPSCC. A PubMed search was performed using the search terms "positron emission tomography" or "PET" or "PET/CT," "oropharyngeal neoplasm" or "oropharyngeal cancer," "oropharyngeal squamous cell carcinoma," "head and neck neoplasm," "staging," "therapy planing," "therapy assessment," "survival," or "prognosis" without restriction or filter for all relevant articles published through December 31, 2014.

Value of PET/CT in Primary Tumor Detection

Uptake of FDG in primary tumors depends on variable factors including tumor type, size, proliferation rate, and ratio of viable vs necrotic cells.[10] The overall sensitivity and specificity of PET scans in detecting primary tumor is greater than 90%, especially if a contrast CT scan is added to PET.[4] For primary tumor detection, CT and magnetic resonance imaging (MRI) effectively localize large tumors, but for small tumors PET is more effective[11] because of in-

Figure 1. Nodal Staging in Patient With Clinically N0 Disease



Axial fused positron emission tomography/computed tomography images of initial scan of a man in his 60s with a diagnosis of squamous cell carcinoma of the base of the tongue, extending to the soft palate and retromolar trigone on the right side. The ^{18}F-fludeoxyglucose–positron emission tomography/computed tomography scan was performed to stage the disease. A, Intense fludeoxyglucose uptake in the oropharyngeal tumor (standard uptake value, 12.83) (arrowhead). B, Intense fludeoxyglucose uptake in 1 ipsilateral hypermetabolic level II (standard uptake value, 5.26) lymph node (arrowhead). The patient was treated with radiation therapy (6000 cGy) and chemotherapy (7 weeks cetuximab), but because of persistent local disease, weekly palliative chemotherapy (weekly docetaxel) was continued. His last follow-up in our center was 7 months after completion of primary treatment.

tense FDG uptake in OPSCC. However, because the spatial resolution of PET is limited to 5 mm,[12] superficial mucosal lesions may be missed. Also, physiologic FDG uptake of normal oral cavity tissues and oropharyngeal lymphoid tissues (Waldeyer ring) may obscure small tumors.[12] Therefore, some authors claim that PET is inaccurate for T1 OPSCC diagnosis but performs well enough in diagnosis of the other stages of OPSCC.[4]

Standard uptake value (SUV) is a quantitative parameter of PET that has been performed in evaluation of OPSCC lesions. Metabolic tumor volume (MTV) and total lesion glycolysis (TLG) are other PET parameters that represent volume of metabolically active tumor and total activity of the tumor, respectively[13] (eTable 1 in the Supplement). Although multiple studies have assessed the role of these 2 parameters in prognosis of OPSCC, to our knowledge there is no published study evaluating the accuracy of these parameters in discriminating benign from malignant oropharyngeal lesions (eFigure 1 in the Supplement).

Value of PET/CT in Cervical Lymph Node Metastasis (N stage)

Because of high lymphatic drainage, OPSCC has a tendency to metastasize early to the cervical lymph nodes (LNs).[14] Accurate staging of OPSCC is the initial step in the management of this cancer, and detection of cervical LN involvement plays a major role in staging.[7] Positron emission tomography has a higher accuracy of detecting cervical LN involvement than CT or MRI, with a sensitivity and specificity of 89% and 81%, respectively.[4] This was verified by a meta-analysis that analyzed data from 32 previous studies that evaluated the accuracy of PET in cervical LN detection of head and neck cancers and reported 79% sensitivity and 86% specificity, compared with lower values for CT, MRI, or CT plus MRI (eTable 2 in the Supplement).[15]

The limitation of PET is in the evaluation of the clinically N0 neck. Despite PET performing better than conventional imaging in detection of involved LNs in a clinically N0 neck, the results are not optimal.[16] Although different studies applied different qualitative features, different SUV_{max} cutoffs, and different timing for detecting LN involvement, in most studies, the specificity of PET was reported between 85% and 90%, whereas the sensitivity was low (50%-70%).[17] A previous study reported that when PET/CT was applied to distinguish N0 neck from N-positive neck, it had limitations due to the high number of false-positive findings.[18] On the other hand, Kovacs et al[19] reported that only 0% to 20% of positive sentinel LNs had been diagnosed by PET and this was due to technical limitations, given that micrometastases smaller than 5 mm are beyond PET's resolution. The ongoing ACRIN 6685 multicenter study is investigating the value of PET/CT in clinical evaluation of the N0 neck, and the results of this study will aid in determining the best role for PET/CT in these patients (**Figure 1**).

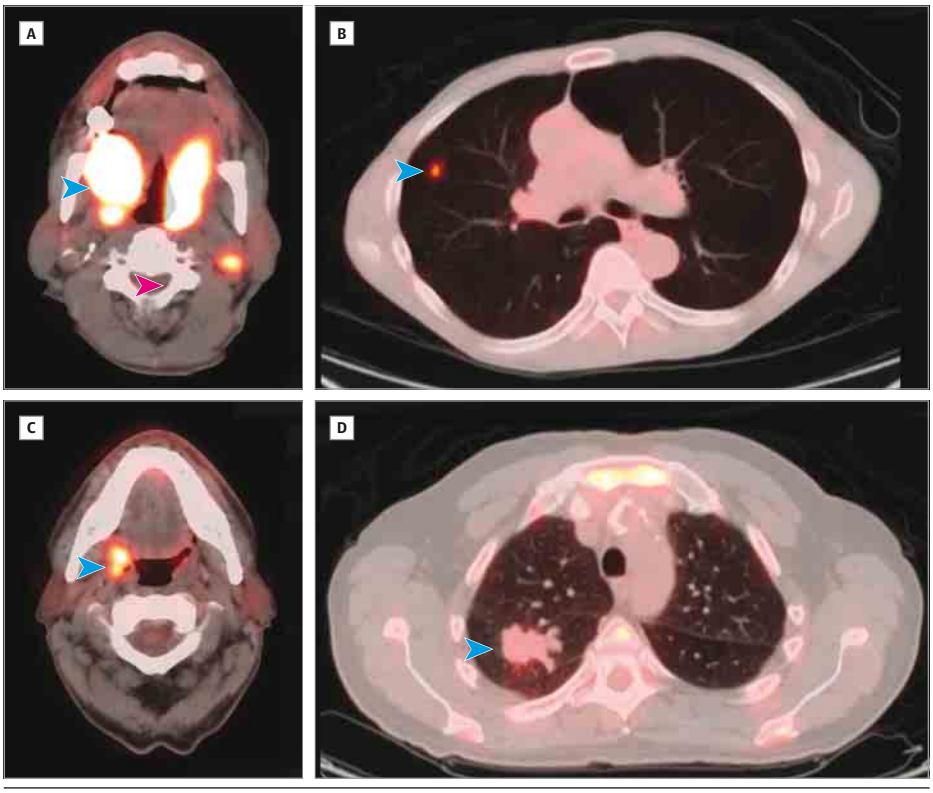
Value of PET/CT in Detecting Distant Metastasis (M stage)

The most common cause of failure of therapy in patients with OPSCC is undetected distant metastasis. The incidence of distant metastasis in OPSCC is 15% to 20%. Patients with distant metastasis from OPSCC usually receive palliative chemotherapy.[20] Therefore, it is important to detect patients with distant metastasis in the primary staging evaluation to determine the best strategy of management for them. In multiple previous studies, it has been shown that PET is a good imaging modality to find distant metastasis in patients with OPSCC.[20] In 1 study, the diagnostic capability of PET was shown to be higher than 3-T whole-body MRI for detecting distant metastasis in OPSCC.[21] The sensitivity and specificity were 92% and 93%, respectively. Therefore, PET scan is recommended as the first imaging modality for detecting distant metastasis in OPSCC[20] (**Figure 2**A and B).

Value of PET/CT in Detecting Second Primary Lesions

Common risk factors such as smoking, as well as genetic predisposition, may place patients with head and neck cancer at risk of developing second primary cancers.[22] The chance of developing second primary cancers in patients with primary head and neck cancer is 10- to 30-fold higher than in the general population.[21] Krabbe et al[23] reported that 10% of patients with oral and oropharyngeal SCC developed second primary cancer in approximately 7 months following initial diagnosis. In a larger study of oral and oropharyngeal SCC (917 cases), 16% of patients developed second primary cancers with 2.6 years median time of diagnosis. The risk of developing a second primary tumor increases with longer follow-up regardless of the patient's age and sex.[24] Although the sensitivity and specificity of PET/CT in detecting second primary tumors is 100% and 95.7%, respectively, and considerably higher than that of panendoscopy, PET/CT is often recommended for diagnosis of second primary cancer only in advanced stages.[25]

Figure 2. Value of Positron Emission Tomography/Computed Tomography (PET/CT) in Detecting Distant Metastasis and Second Primary Cancer in Patients With Oropharyngeal Squamous Cell Carcinoma



A and B, Axial fused PET/CT) images of initial scan of a man in his 60s with a diagnosis of poorly differentiated squamous cell carcinoma of the oropharynx. A, The PET/CT scan demonstrated an intensely fludeoxyglucose (FDG)-avid primary lesion (standard uptake value [SUV], 36.5) (blue arrowhead); fludeoxyglucose-avid metastasis to regional neck nodes (SUV, 11.8) (red arrowhead). B, Fludeoxyglucose-avid lung metastasis (SUV, 4.4) (arrowhead). C and D, Axial fused PET/CT image of initial PET/CT scan of a man in his 60s who presented with a mass in the right tonsil; PET/CT scan was performed for staging and evaluation for metastatic disease. C, The FDG PET/CT scan demonstrated an intensely FDG-avid lesion within the region of the right palatine tonsil (SUV, 9.3) (arrowhead). D, In addition, it revealed another moderately FDG-avid lung lesion (SUV, 7.45) (arrowhead) in the right upper lobe compatible with synchronous primary lung carcinoma, which biopsy proved to be an adenocarcinoma of the lung.

A population-based cohort study of 75 087 patients with head and neck SCC reported that the rate of developing second primary cancer in patients with OPSCC has decreased to the lowest levels of any subtype and claimed that this is because of the increasing rate of HPV-related OPSCC.[26] Gan et al[27] reported that smoking status (former, current, or nonsmoker) was an important factor in developing a second primary cancer in patients with OPSCC, and Rodriguez-Bruno et al[28] recommended not performing routine panendoscopy for detection of second primary tumor in patients who never have smoked. In general, it seems that because of the increasing incidence of HPV-related OPSCC and the low rate of a second primary cancer in this subgroup of patients, future studies are needed to provide a guideline for second primary cancer screening in OPSCC according to HPV and smoking status (Figure 2C and D).

Value of Contrast-Enhanced Head and Neck PET/CT in Staging

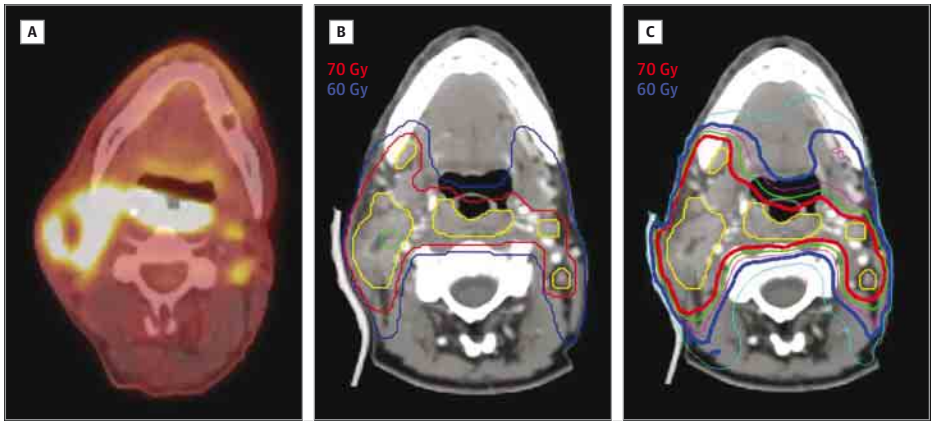
The TNM staging of OPSCC is an essential factor for determining surgical and radiation treatment strategies. Aside from anatomic evaluation of the tumor with CT and MRI, assessment of metabolic features seems to be important for accurate staging of head and neck cancers. As noted, PET scan provides advantages over anatomic imaging in the assessment of the primary tumor and cervical metastasis; in addition, PET is capable of detecting distant metastasis or a second primary tumor, if present. Altogether, PET increases the accuracy of pretreatment staging of OPSCC.[4]

It is known that PET alone or PET/CT without contrast does not provide sufficient anatomic detail for surgical planning. Krabbe et al[4] reported that performing a single PET/CT with contrast instead of a separate PET/CT without contrast and neck CT with contrast in initial staging of OPSCC has several advantages, such as providing fully optimized head and neck CT, reduced radiation dose, decreased false-positive results, and increased confidence for staging. Other advantages are improved lesion conspicuity, precise tumor delineation, evaluation of resectability of primary lesions, and detection of distant metastasis such as liver metastases. These advantages of PET/CT with contrast have been shown to change initial oncological management compared with PET/CT without contrast.[29] In addition, it is less efficient to prepare 2 images separately and write 2 separate reports, and if 2 different radiologists report separately, the other physicians may see conflicting interpretations that could lead to treatment delay.[29] It has been shown that baseline contrast-enhanced PET/CT decreases the need for a supplementary contrast CT scan and can provide both high-quality anatomic and functional information in a single study. Today, it is recommended that a single contrast-enhanced PET/CT study be performed for initial assessment and staging of patients with OPSCC.[29]

Value of PET/CT in Therapy Planning for OPSCC

Radiotherapy with or without chemotherapy is the mainstay of therapy for advanced OPSCC. To decrease the adverse effects of ra-

Figure 3. Value of Positron Emission Tomography/Computed Tomography (PET/CT) in Radiation Therapy Planning for Oropharyngeal Squamous Cell Carcinoma (OPSCC)



Imaging of a man in his 50s who received a diagnosis of stage cT3N2cM0 human papillomavirus–negative OPSCC. A, Pretreatment PET/CT. B, Simulation CT. These were fused for radiation therapy planning, which aided in identifying gross tumor volumes (yellow), including posterior oropharyngeal primary, bulky adenopathy of the right side of the neck with central necrosis, and small left neck nodes that did not meet CT size criteria but were FDG avid. Planning target volumes were prescribed 70 Gy (red) and 60 Gy (blue). C, The resulting intensity-modulated radiotherapy plan was conformal to the target contours, and the patient had a complete response to treatment.

diation therapy on the normal surrounding tissues and achieve the best treatment response, radiation dose and volume should be adjusted according to the extent of primary and nodal disease and the risk of subclinical disease in each area (**Figure 3**).[30] Modern radiotherapy technology and intensity-modulated radiotherapy allow for such dose adjustment.[31] Positron emission tomography in combination with CT or MRI can provide biological tumor information plus anatomic features, which can change radiotherapy target planning.[32] Chatterjee et al[31] reported that PET/CT is much better than contrast-enhanced CT in radiotherapy planning for OPSCC because it provides more clinically relevant information and decreases the chance of a geographical miss.

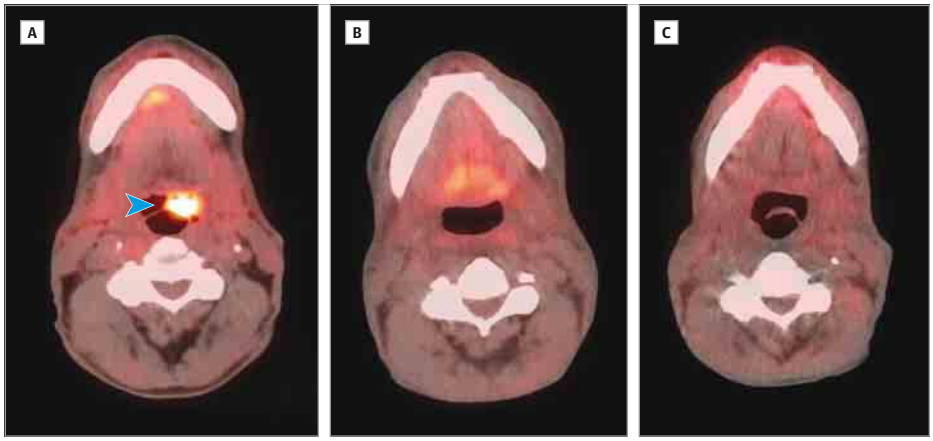
Gross tumor volume (GTV) is an important factor for intensity-modulated radiotherapy planning. Computed tomography, with or without contrast, often cannot assess GTV precisely in the oropharynx and neck as a result of lack of clear demarcation between tumor and normal surrounding tissues.[33] In addition, when using PET in preradiotherapy planning, the radiation field size and dose may increase as a result of identification of normal-size LNs with FDG uptake (Figure 3).[34] Gross tumor volume attained from PET/CT imaging was reported to be different from that of contrast CT in advanced stages of OPSCC, which changes the area and dosage of radiotherapy.[31] Paulino et al[35] showed that approximately 25% of patients did not receive the optimal radiation dosage by using CT GTV, and PET/CT not only changes the GTV for radiotherapy planning but also defines the initial extent of disease more precisely.[36] On the other hand, PET may underestimate tumor volume definition because of tumor necrosis or lack of metabolic activity in an area of tumor, especially in HPV-related nodal metastases (Figure 3); so alongside PET, other modalities such as clinical examination, contrast-enhanced CT, or MRI should be considered.[33] Not only is PET/CT highly accurate in detecting extent of tumor and nodal involvement for pretreatment radiotherapy planning, but it also provides associated staging and prognostic information for determination of whether to use concurrent chemotherapy with radiation therapy.[34] Using contrast-enhanced PET/CT leads to a combined anatomic, metabolic, and biologic approach to radiation therapy planning of head and neck cancers.[36]

Value of PET in Evaluating Therapy Response of OPSCC

Treatment of OPSCC with surgical or radiation strategies leads to tissue changes such as edema, hyperemia, and fibrosis, and these changes affect the accuracy of CT and MRI in detecting residual and/or recurrent lesions. Because fibrosis and/or scar tissue has no associated metabolic activity on PET, PET can better discriminate between fibrosis and/or scarring vs recurrent and/or residual disease than the other imaging modalities. Some studies evaluated the ability of PET to detect residual tumor and therapy response.[33] Most of these studies reported that PET is more effective than CT and MRI for estimating the response to chemoradiotherapy of OPSCC.[10] One of the most important factors in therapy assessment is using reliable, accurate, and valid interpretation criteria. Recently, Marcus et al[37] presented a standardized interpretation criterion for head and neck PET/CT (Hopkins criteria), which demonstrated excellent interreader reliability, accuracy, and survival prediction. According to the Hopkins criteria, the activity in the internal jugular vein and liver were taken as a reference and the FDG uptake of the suspicious area was compared with the uptake of the internal jugular vein and liver (eTable 3 in the Supplement).

Studies showed that FDG-PET can assess volume changes during radiotherapy and is a good imaging modality to evaluate early therapy response.[32] This could be used to help physicians assess the behavior of tumor during therapy and change the treatment strategies if needed. Other reports showed a significant association between pretreatment value of SUV and response to chemotherapy. There was a reverse association between SUV and therapy re-

Figure 4. Therapy Assessment



A, Axial fused positron emission tomography/computed tomography (PET/CT) image of initial scan of a man in his 60s with a history of left tongue base squamous cell carcinoma, which presented as a fludeoxyglucose-avid lesion (standard uptake value, 9.99) (arrowhead) in PET/CT. The patient was treated with chemoradiation (9 weeks cetuximab, 7000 cGy). B, Three months after treatment, PET/CT scan showed good response with diffuse uptake suggestive of postradiotherapy inflammation (diffuse uptake, standard uptake value, 6.74). C, The 9-month follow-up PET/CT showed complete response without any interval treatment.

sponse. The lower the pretreatment SUV value, the more response to chemotherapy was seen.[1] A meta-analysis demonstrated that PET is highly accurate for response monitoring or relapse detection after radiotherapy with or without chemotherapy in advanced head and neck cancers. This study showed 94% sensitivity, 82% specificity, 75% positive predictive value, and 95% negative predictive value for PET in detecting residual lesions.[38] Sjovall et al[39] evaluated 82 patients with head and neck cancer (85% OPSCC) and reported that PET/CT has a sensitivity and specificity of 100% and 78%, respectively, in detecting residual tumor after radiotherapy.

The other issue is the optimal timing to evaluate the treatment response with PET. Although on one hand we should wait for radiation therapy and chemotherapy to show their full effect and treatment inflammation to decrease (8-10 weeks), on the other, we do not want to wait too long to prevent tumor progression or missed treatment opportunities. Sensitivity of PET in therapy assessment is low when performed less than 10 weeks after treatment, whereas specificity seems not to be related to timing.[38] Most reports studied the capability of PET in treatment evaluation within 2 months after therapy.[23] Kim et al[40] reported high negative predictive value of PET on evaluation of radiotherapy after 1 month. However, PET used later than 12 weeks after treatment can decrease the number of diagnostic neck dissections needed.[41] In general, it seems that PET is a good modality for therapy assessment of OPSCC approximately 3 months after radiotherapy and/or chemotherapy (**Figure 4**).

Role of PET in Recurrence Detection of OPSCC

Despite all advances in treatment methods (surgery, radiotherapy, chemotherapy), the locoregional recurrence rate of HPV-negative OPSCC remains high. Some articles claim that up to 24% of patients with advanced head and neck cancer developed locoregional recurrence.[23] Although most of the time when the recurrence is detected, it is advanced and has aggressive tumor characteristics (eFigures 2 and 3 in the Supplement), a recent study by Fakhry et al[42] showed improved survival in patients with recurrent HPV-related OPSCC treated with salvage surgery. In recurrence detection, PET has a lower false-negative and false-positive rate when compared with conventional imaging.[43] A meta-analysis by Isles et al[38] showed that PET has a sensitivity and specificity of 94% and 82%, respectively, for detecting residual or recurrent head and neck cancer. In a study that compared PET/CT vs whole-body MRI in detecting recurrence in oropharyngeal and hypopharyngeal SCC, PET/CT had higher accuracy.[44] Positron emission tomography was reported to be a powerful tool in early recurrence detection of OPSCC[23] (eFigure 4 in the Supplement).

Role of PET in Prognosis of OPSCC

Some studies evaluated the role of PET in predicting prognosis of OPSCC and showed that a negative PET/CT result after therapy is associated with lower chance of recurrence and better prognosis. Therefore, this subgroup of patients needs less frequent radiologic surveillance. The effect of negative PET result is more prominent in HPV-positive patients and is a better predictor of survival than OPSCC stage for these patients.[45] Different studies evaluated the relationship between PET parameters (SUV_{max}, SUV_{mean}, SUV_{peak}, TLG, MTV) and prognostic factors such as disease-free survival (DFS), progression-free survival (PFS), overall survival (OS), and disease-specific survival (DSS). Because of its observer-independent measurement, SUV is a popular PET parameter used by multiple scientists.[46] Some previous studies claim that SUV_{max} predicts head and neck cancer outcome and higher SUV_{max} is associated with worse DFS.[13] Most of these studies evaluated tumors from various head and neck sites, with different tumor characteristics, risk factors, and different overall prognosis. However, several recent studies reported that SUV had poor predictive performance for treatment outcome and had no independent relation to OS or DFS, especially when corrected for stage.[13,46] Overall, it should not be a valuable predictive factor for prognosis.[3,46]

Other PET parameters evaluated to predict prognosis are volume-based measurements such as TLG and MTV. These factors can indicate total activity and volume of metabolically active tumor cells. Multiple studies have shown that MTV and TLG are important prognostic factors in OPSCC, independent of the stage of disease.[47] Garsa et al[13] reported that primary tumor MTV was a significant predictor of OS and DFS whereas primary tumor TLG was related only to OS. Although total MTV (including LN MTV) and total TLG were significant predictors of DFS and OS, there was no difference between PET parameters in HPV-positive and HPV-negative patients. Cheng et al[48] reported that TLG and uniformity (also called *angular second moment*, a measure of image homogeneity that is extracted from the normalized gray-level co-occurrence matrix) were independently associated with PFS and DSS, whereas MTV and uniformity were associated just with OS. In another study, Cheng et al[49] showed that primary tumor TLG in both HPV-positive and HPV-negative patients with OPSCC is a prognostic factor for OS, PFS, and DFS but nodal TLG is significant just for DFS. Kikuchi et al,[50] who evaluated OPSCC, claimed that although MTV and TLG of primary lesions, LNs, and total tumor lesions were significant prognostic factors for DFS, DSS, or OS, in multivariate Cox regression analysis only MTV for total tumor lesions remained an independent prognostic factor of DFS, DSS, and OS. On the other hand, Moon et al[46] reported that TLG is the only independent prognostic factor in tonsil SCC. Total lesion glycolysis is a combination of SUV and MTV and represents the metabolically active tumor uptake and size; therefore, it is theoretically reasonable that TLG is an ideal parameter of tumor burden[46] (eTable 4 in the Supplement). Today, it seems that a scoring system is needed to use clinically for prognosis prediction for each kind of head and neck cancer given that they behave differently.

Conclusions

^{18}F-fludeoxyglucose–positron emission tomography/computed tomography is a vital tool in the management of patients with oropharyngeal squamous cell carcinoma and is helpful in staging, therapy planning, evaluating therapy response, detecting recurrence, and predicting prognosis of these patients.

ARTICLE INFORMATION

Submitted for Publication: July 13, 2015; accepted July 30, 2015.

Published Online: November 19, 2015. doi:10.1001/jamaoto.2015.2607.

Author Affiliations: Russell H. Morgan Department of Radiology and Radiological Sciences, Johns Hopkins School of Medicine, Baltimore, Maryland (Taghipour, Sheikhbahaei, Marashdeh, Solnes, Subramaniam); Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins School of Medicine, Baltimore, Maryland (Kiess); Department of Otolaryngology and Head and Neck Surgery, Johns Hopkins School of Medicine, Baltimore, Maryland (Kiess, Subramaniam); Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, Maryland (Subramaniam); Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Subramaniam).

Author Contributions: Dr Subramaniam 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: Taghipour, Sheikhbahaei, Solnes, Subramaniam.
Acquisition, analysis, or interpretation of data: Marashdeh, Kiess.
Drafting of the manuscript: Taghipour, Sheikhbahaei, Marashdeh, Subramaniam.
Critical revision of the manuscript for important intellectual content: Sheikhbahaei, Marashdeh, Solnes, Kiess, Subramaniam.
Administrative, technical, or material support: Taghipour, Sheikhbahaei.
Study supervision: Marashdeh, Kiess, Subramaniam.

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Kawakita D, Masui T, Hanai N, et al. Impact of positron emission tomography with the use of fluorodeoxyglucose on response to induction chemotherapy in patients with oropharyngeal and hypopharyngeal squamous cell carcinoma. *Acta Otolaryngol*. 2013;133(5):523-530.

2. Paidpally V, Chirindel A, Lam S, Agrawal N, Quon H, Subramaniam RM. FDG-PET/CT imaging biomarkers in head and neck squamous cell carcinoma. *Imaging Med*. 2012;4(6):633-647.

3. Dibble EH, Alvarez AC, Truong MT, Mercier G, Cook EF, Subramaniam RM. 18F-FDG metabolic tumor volume and total glycolytic activity of oral cavity and oropharyngeal squamous cell cancer: adding value to clinical staging. *J Nucl Med*. 2012;53(5):709-715.

4. Krabbe CA, Balink H, Roodenburg JL, Dol J, de Visscher JG. Performance of 18F-FDG PET/contrast-enhanced CT in the staging of squamous cell carcinoma of the oral cavity and oropharynx. *Int J Oral Maxillofac Surg*. 2011;40(11):1263-1270.

5. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol*. 2010;11(8):781-789.

6. Tahari AK, Alluri KC, Quon H, Koch W, Wahl RL, Subramaniam RM. FDG PET/CT imaging of oropharyngeal squamous cell carcinoma: characteristics of human papillomavirus-positive and -negative tumors. *Clin Nucl Med*. 2014;39(3):225-231.

7. O'Rorke MA, Ellison MV, Murray LJ, Moran M, James J, Anderson LA. Human papillomavirus related head and neck cancer survival: a systematic review and meta-analysis. *Oral Oncol*. 2012;48(12):1191-1201.

8. Parikh U, Marcus C, Sarangi R, Taghipour M, Subramaniam RM. FDG PET/CT in pancreatic and hepatobiliary carcinomas: value to patient management and patient outcomes. *PET Clin*. 2015;10(3):327-343.

9. Marcus C, Marashdeh W, Ahn SJ, Taghipour M, Subramaniam RM. 18F-FDG PET/CT and colorectal cancer: value of fourth and subsequent posttherapy follow-up scans for patient management. *J Nucl Med*. 2015;56(7):989-994.

10. Chepeha DB, Sacco AG, Oxford LE, et al. Advanced squamous cell carcinoma of the oropharynx: efficacy of positron emission tomography and computed tomography for determining primary tumor response during induction chemotherapy. *Head Neck*. 2009;31(4):452-460.

11. Dammann F, Horger M, Mueller-Berg M, et al. Rational diagnosis of squamous cell carcinoma of the head and neck region: comparative evaluation of CT, MRI, and 18FDG PET [published correction appears in *AJR Am J Roentgenol*. 2005;184(6):1968]. *AJR Am J Roentgenol*. 2005;184(4):1326-1331.

12. Zafereo ME. Evaluation and staging of squamous cell carcinoma of the oral cavity and oropharynx: limitations despite technological breakthroughs. *Otolaryngol Clin North Am*. 2013;46(4):599-613.

13. Garsa AA, Chang AJ, Dewees T, et al. Prognostic value of 18F-FDG PET metabolic parameters in oropharyngeal squamous cell carcinoma. *J Radiat Oncol*. 2013;2(1):27-34.

14. Joo YH, Yoo IeR, Cho KJ, et al. Relationship between extracapsular spread and FDG PET/CT in oropharyngeal squamous cell carcinoma. *Acta Otolaryngol*. 2013;133(10):1073-1079.

15. Kyzas PA, Evangelou E, Denaxa-Kyza D, Ioannidis JP. 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. *J Natl Cancer Inst*. 2008;100(10):712-720.

16. Krabbe CA, Dijkstra PU, Pruim J, et al. FDG PET in oral and oropharyngeal cancer: value for confirmation of NO neck and detection of occult metastases. *Oral Oncol*. 2008;44(1):31-36.

17. Murakami R, Uozumi H, Hirai T, et al. Impact of FDG-PET/CT imaging on nodal staging for head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys*. 2007;68(2):377-382.

18. Piao Y, Bold B, Tayier A, et al. Evaluation of 18F-FDG PET/CT for diagnosing cervical nodal metastases in patients with oral cavity or

oropharynx carcinoma. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2009;108(6):933-938.

19. Kovács AF, Döbert N, Gaa J, Menzel C, Bitter K. Positron emission tomography in combination with sentinel node biopsy reduces the rate of elective neck dissections in the treatment of oral and oropharyngeal cancer. *J Clin Oncol*. 2004;22(19):3973-3980.

20. Wallowy P, Diener J, Grünwald F, Kovács AF. 18F-FDG PET for detecting metastases and synchronous primary malignancies in patients with oral and oropharyngeal cancer. *Nuklearmedizin*. 2009;48(5):192-199.

21. Chan SC, Wang HM, Yen TC, et al. ^{18}F-FDG PET/CT and 3.0-T whole-body MRI for the detection of distant metastases and second primary tumours in patients with untreated oropharyngeal/hypopharyngeal carcinoma: a comparative study. *Eur J Nucl Med Mol Imaging*. 2011;38(9):1607-1619.

22. Hujala K, Sipilä J, Grenman R. Panendoscopy and synchronous second primary tumors in head and neck cancer patients. *Eur Arch Otorhinolaryngol*. 2005;262(1):17-20.

23. Krabbe CA, Pruim J, Dijkstra PU, et al. 18F-FDG PET as a routine posttreatment surveillance tool in oral and oropharyngeal squamous cell carcinoma: a prospective study. *J Nucl Med*. 2009;50(12):1940-1947.

24. van der Haring IS, Schaapveld MS, Roodenburg JL, de Bock GH. Second primary tumours after a squamous cell carcinoma of the oral cavity or oropharynx using the cumulative incidence method. *Int J Oral Maxillofac Surg*. 2009;38(4):332-338.

25. Haerle SK, Strobel K, Hany TF, Sidler D, Stoeckli SJ. ^{18}F-FDG-PET/CT versus panendoscopy for the detection of synchronous second primary tumors in patients with head and neck squamous cell carcinoma. *Head Neck*. 2010;32(3):319-325.

26. Morris LG, Sikora AG, Patel SG, Hayes RB, Ganly I. Second primary cancers after an index head and neck cancer: subsite-specific trends in the era of human papillomavirus-associated oropharyngeal cancer. *J Clin Oncol*. 2011;29(6):739-746.

27. Gan SJ, Dahlstrom KR, Peck BW, et al. Incidence and pattern of second primary malignancies in patients with index oropharyngeal cancers versus index nonoropharyngeal head and neck cancers. *Cancer*. 2013;119(14):2593-2601.

28. Rodriguez-Bruno K, Ali MJ, Wang SJ. Role of panendoscopy to identify synchronous second primary malignancies in patients with oral cavity and oropharyngeal squamous cell carcinoma. *Head Neck*. 2011;33(7):949-953.

29. Subramaniam RM, Agarwal A, Colucci A, Ferraro R, Paidpally V, Mercier G. Impact of concurrent diagnostic level CT with PET/CT on the utilization of stand-alone CT and MRI in the management of head and neck cancer patients. *Clin Nucl Med*. 2013;38(10):790-794.

30. Kovalchuk N, Jalisi S, Subramaniam RM, Truong MT. Deformable registration of preoperative PET/CT with postoperative radiation therapy planning CT in head and neck cancer. *Radiographics*. 2012;32(5):1329-1341.

31. Chatterjee S, Frew J, Mott J, et al. Variation in radiotherapy target volume definition, dose to organs at risk and clinical target volumes using anatomic (computed tomography) versus combined anatomic and molecular imaging (positron emission tomography/computed tomography): intensity-modulated radiotherapy delivered using a tomotherapy Hi Art machine: final results of the VortigERN study. *Clin Oncol (R Coll Radiol)*. 2012;24(10):e173-e179.

32. Troost EG, Bussink J, Hoffmann AL, Boerman OC, Oyen WJ, Kaanders JH. 18F-FLT PET/CT for early response monitoring and dose escalation in oropharyngeal tumors. *J Nucl Med*. 2010;51(6):866-874.

33. Subramaniam RM, Truong M, Peller P, Sakai O, Mercier G. Fluorodeoxyglucose-positron-emission tomography imaging of head and neck squamous cell cancer. *AJNR Am J Neuroradiol*. 2010;31(4):598-604.

34. Menda Y, Graham MM. Update on 18F-fluorodeoxyglucose/positron emission tomography and positron emission tomography/computed tomography imaging of squamous head and neck cancers. *Semin Nucl Med*. 2005;35(4):214-219.

35. Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2005;61(5):1385-1392.

36. Ciernik IF, Dizendorf E, Baumert BG, et al. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys*. 2003;57(3):853-863.

37. Marcus C, Ciarallo A, Tahari AK, et al. Head and neck PET/CT: therapy response interpretation criteria (Hopkins criteria)—interreader reliability, accuracy, and survival outcomes. *J Nucl Med*. 2014;55(9):1411-1416.

38. Isles MG, McConkey C, Mehanna HM. A systematic review and meta-analysis of the role of positron emission tomography in the follow up of head and neck squamous cell carcinoma following radiotherapy or chemoradiotherapy. *Clin Otolaryngol*. 2008;33(3):210-222.

39. Sjövall J, Brun E, Almquist H, Kjellén E, Wahlberg P. Radiotherapy response in head and neck cancer—evaluation of the primary tumour site. *Acta Otolaryngol*. 2014;134(6):646-651.

40. Kim SY, Lee SW, Nam SY, et al. The feasibility of 18F-FDG PET scans 1 month after completing radiotherapy of squamous cell carcinoma of the head and neck. *J Nucl Med*. 2007;48(3):373-378.

41. Porceddu SV, Pryor DI, Burmeister E, et al. Results of a prospective study of positron emission tomography-directed management of residual nodal abnormalities in node-positive head and neck cancer after definitive radiotherapy with or without systemic therapy. *Head Neck*. 2011;33(12):1675-1682.

42. Fakhry C, Zhang Q, Nguyen-Tan PF, et al. Human papillomavirus and overall survival after progression of oropharyngeal squamous cell carcinoma. *J Clin Oncol*. 2014;32(30):3365-3373.

43. Chen AY, Vilaseca I, Hudgins PA, Schuster D, Halkar R. PET-CT vs contrast-enhanced CT: what is the role for each after chemoradiation for advanced oropharyngeal cancer? *Head Neck*. 2006;28(6):487-495.

44. Ng SH, Chan SC, Yen TC, et al. PET/CT and 3-T whole-body MRI in the detection of malignancy in treated oropharyngeal and hypopharyngeal carcinoma. *Eur J Nucl Med Mol Imaging*. 2011;38(6):996-1008.

45. Koshkareva Y, Branstetter BF IV, Gaughan JP, Ferris RL. Predictive accuracy of first posttreatment PET/CT in HPV-related oropharyngeal squamous cell carcinoma. *Laryngoscope*. 2014;124(8):1843-1847.

46. Moon SH, Choi JY, Lee HJ, et al. Prognostic value of 18F-FDG PET/CT in patients with squamous cell carcinoma of the tonsil: comparisons of volume-based metabolic parameters. *Head Neck*. 2013;35(1):15-22.

47. Alluri KC, Tahari AK, Wahl RL, Koch W, Chung CH, Subramaniam RM. Prognostic value of FDG PET metabolic tumor volume in human papillomavirus-positive stage III and IV oropharyngeal squamous cell carcinoma. *AJR Am J Roentgenol*. 2014;203(4):897-903.

48. Cheng NM, Fang YH, Chang JT, et al. Textural features of pretreatment 18F-FDG PET/CT images: prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. *J Nucl Med*. 2013;54(10):1703-1709.

49. Cheng NM, Chang JT, Huang CG, et al. Prognostic value of pretreatment ^{18}F-FDG PET/CT and human papillomavirus type 16 testing in locally advanced oropharyngeal squamous cell carcinoma. *Eur J Nucl Med Mol Imaging*. 2012;39(11):1673-1684.

50. Kikuchi M, Koyasu S, Shinohara S, et al. Prognostic value of pretreatment ^{18}F-fluorodeoxyglucose positron emission tomography/CT volume-based parameters in patients with oropharyngeal squamous cell carcinoma with known p16 and p53 status. *Head Neck*. 2015;37(10):1524-1531.

International Journal of
Radiation Oncology
biology • physics

www.redjournal.org

Clinical Investigation: Head and Neck Cancer

"Pharyngocise": Randomized Controlled Trial of Preventative Exercises to Maintain Muscle Structure and Swallowing Function During Head-and-Neck Chemoradiotherapy

Giselle Carnaby-Mann, M.P.H., Ph.D.,* Michael A. Crary, Ph.D.,† Ilona Schmalfuss, M.D.,‡ and Robert Amdur, M.D.§

*Departments of *Behavioral Science and Community Health, †Speech Language and Hearing Sciences, and §Radiation Oncology, University of Florida, Gainesville, FL; ‡Department of Radiology, North Florida/South Georgia Veterans Health System, Gainesville, FL*

Received Feb 18, 2011, and in revised form Jun 3, 2011. Accepted for publication Jun 3, 2011

Summary

Phayngo-esophageal dysfunction is common after chemo-radiation for HNC. A program of preventative exercise for swallowing was tested in a randomized phase II study. Subjects receiving the swallowing program demonstrated significant benefit over the comparator arms (usual care and placebo) in maintenance of swallow muscle composition and preservation of swallowing function, salivation and chemosensation. Thus simple swallowing exercises administered daily throughout chemo-radiation treatment may offer a cost effective way to prevent swallowing related morbidity

Purpose: Dysphagia after chemoradiotherapy is common. The present randomized clinical trial studied the effectiveness of preventative behavioral intervention for dysphagia compared with the "usual care."

Methods and Materials: A total of 58 head-and-neck cancer patients treated with chemoradiotherapy were randomly assigned to usual care, sham swallowing intervention, or active swallowing exercises (pharyngocise). The intervention arms were treated daily during chemoradiotherapy. The primary outcome measure was muscle size and composition (determined by T_2-weighted magnetic resonance imaging). The secondary outcomes included functional swallowing ability, dietary intake, chemosensory function, salivation, nutritional status, and the occurrence of dysphagia-related complications.

Results: The swallowing musculature (genioglossus, hyoglossuss, and mylohyoid) demonstrated less structural deterioration in the active treatment arm. The functional swallowing, mouth opening, chemosensory acuity, and salivation rate deteriorated less in the pharyngocise group.

Conclusion: Patients completing a program of swallowing exercises during cancer treatment demonstrated superior muscle maintenance and functional swallowing ability.

Keywords: Swallowing dysfunction, Chemoradiotherapy, Swallowing therapy, Randomized controlled trial

Reprint requests to: Giselle Carnaby-Mann, M.P.H., Ph.D., Tel: (352) 273-6164; Fax: (352) 392-7018; E-mail: gmann@phhp.ufl.edu

Conflict of interest: none.

Int J Radiation Oncol Biol Phys, Vol. 83, No. 1, pp. 210–219, 2012
0360-3016/$ - see front matter
doi:10.1016/j.ijrobp.2011.06.1954

Introduction

The swallowing deficits from oropharyngeal cancer and the therapies used to control the disease are devastating to functional feeding outcome (1, 2). Specifically, the swallowing outcomes of patients treated with external beam radiotherapy are poorer than those of patients treated by surgical intervention alone (3–5). One reason for the effect of external beam radiotherapy on swallowing is the development of deep tissue fibrosis (6, 7). The formation of radiation-induced fibrotic tissue and the acute radiation effects (*i.e.,* edema, mucositis, xerostomia) can act collectively to promote muscular disuse or atrophy, contributing to the decline in swallowing function (8, 9).

Skeletal muscle demonstrates remarkable plasticity in response to functional demand (8). Muscles atrophy rapidly after immobilization or disuse (9–11). In contrast, aggressive treatment with weight-loaded exercises results in the recovery of strength and work capacity in previously weakened muscles (12, 13). Head-and-neck cancer (HNC) patients undergoing chemoradiotherapy (CRT) frequently demonstrate muscle changes as a result of fibrosis, muscle edema, and fatty infiltration. Moreover, they have a reduced swallowing frequency owing to the discomfort resulting from the acute radiation effects (14). In essence, they demonstrate constraint-induced muscular weakness from swallowing avoidance. We postulated that swallowing exercises would facilitate maintenance of oropharyngeal muscle function. The present study evaluated the benefit of a battery of exercises on the maintenance of muscle composition and function for swallowing in HNC patients undergoing CRT. Specifically, the maintenance of oropharyngeal muscle size and composition as identified by T_2-weighted magnetic resonance imaging (MRI), level of functional swallowing ability, maintenance of nutritional and chemosensory indexes, and the occurrence of dysphagia-related complications.

Methods and Materials

The present study was undertaken at a university hospital cancer center. The local institutional review board approved the study protocol. All participants signed an approved consent form.

Patients

Patients presenting to the Cancer Center from 2001 to 2004 were screened for inclusion. The patients were included if they presented with *(1)* HNC of the oropharyngeal regions, confirmed by the clinical history and examination findings, with positive cross-sectional imaging studies and histopathologic biopsy, excluding other pathologic factors; *(2)* external beam radiotherapy was planned; *(3)* and they had no history of nonoral feeding for cancer-related illness and were able to undergo MRI procedures.

Study design

The present study followed a randomized, controlled trial design. The treatment allocation used a computer-generated blocked random numbers list. The randomization schedule was held in the trial office, remote from the study environment. After review by the study radiation oncologist, the eligible patients were informed about the trial and, after consenting, randomly assigned to one of three treatment options.

Interventions

The three treatment groups included usual care, standardized sham treatment, and high-intensity behavioral treatment (pharyngocise).

The usual care (control) group included patient management by the attending radiation oncologist "as usual." Treatment, if offered, consisted of supervision for feeding and precautions for safe swallowing (*e.g.,* positioning, slowed rate of feeding) by the hospital speech pathology service. The patients in this group received focused attention sessions during the course of CRT from a research assistant, consisting of weekly telephone calls to monitor the swallowing outcome.

Standardized sham therapy included a buccal extension maneuver ("valchuff") and appropriate dietary modification, under the direction of the study speech pathologist, twice daily for the duration of the CRT. The patients assigned to this group completed the exercise for 10 repetitions over 4 cycles, each of 10 minutes' duration. The treatment sessions were 45 minutes in duration.

Standardized high-intensity swallowing therapy ("pharyngocise") included a battery of exercises (*e.g.,* falsetto, tongue press, hard swallow, and jaw resistance/strengthening using the Therabite Jaw Motion Rehabilitation System [15]) and dietary modification, under the direction of the study speech pathologist, twice daily for the duration of the CRT (up to a maximum of 6 weeks). The patients assigned to this condition completed the four swallowing exercises in 10 repetitions over 4 cycles, each of 10 minutes' duration. The treatment sessions were 45 minutes in duration.

Masking/blinding

Only the treating speech pathologist and patients were aware of the intervention assignment. The study staff worked independently of the hospital service and did not share trial information. The speech pathologists in the hospital service continued to receive sporadic referrals from the radiation oncology staff. The attending radiation oncologists were unaware of the randomization assignment of their patients.

Outcome events

Before CRT, all subjects received a standard clinical and instrumental swallowing assessment, nutritional examination, quality-of-life questionnaires, and T_2-weighted MRI. All baseline measures were repeated at CRT completion and at 6 months after CRT.

The outcome was assessed by 2 independent speech pathologists (M.C., G.C.), who were unaware of the treatment allocation. The swallowing progress and occurrence of possible complications were sought from multiple overlapping sources. Information about the specific swallow treatment was not requested, and the direct treatment records were not reviewed to maintain the blinding. Additionally, patients in both the sham and the pharyngocise arms completed a daily home record of the exercise conducted between treatment sessions. The outcomes after discharge was recorded by the patient or caregiver in a diary and reviewed at monthly telephone interviews.

The primary outcome measure was the change in muscle size and composition identified by T_2-weighted MRI from before to after treatment and at 6 months after randomization.

T_2-weighted MRI

Magnetic resonance imaging was conducted to quantify the baseline muscle parameters in the oral cavity and pharynx. The muscle size, composition, and T_2 signal intensity was documented. The patients were scanned using a Siemens 1.5 T Vision MRI scanner and a phase array neck coil. Multiplaner localizer and subsequent T_1-weighted sagittal images were acquired through the face and upper neck for localization. Subsequently, a T_2 relaxation mapping sequence (Carr-Purcell-Meiboom-Gill sequence) was performed in the axial plane. This T_2 mapping sequence was performed with a repetition time of 2,000 ms and 16 different excitation times (23, 45, 68, 90, 113, 135, 158, 180, 203, 225, 248, 270, 293, 315, 338, and 360 ms) to allow objective calculation of the T_2 value of the different anatomic structures of interest. The T_2 relaxation images were performed in two separate sets of five images of 5-mm slice thickness and an interslice gap of 2.5 mm using a 180-mm field of view through the oral cavity and glottic region. The T_2 relaxation images in the coronal plane through the oral cavity were done using the same imaging parameters. The axial T_1-weighted images aligned parallel to the true vocal folds were done from the hard palate to the upper trachea with a repetition time of 700 ms, excitation time of 15 ms, and flip angle of 90° using the same field of view and slice thickness as used for the T_2 relaxation images.

Axial T_2 relaxation images through the oral cavity were used to measure the length, width, and T_2 relaxation time of the genioglossus muscle and the thickness and T_2 relaxation time of the mylohyoid, hyoglossus, and middle pharyngeal constrictor muscles. In addition, the thickness and T_2 relaxation time of the mylohyoid muscles were measured on the coronal T_2 relaxation images. Images through the glottic level were used to measure the thickness and T_2 relaxation times of the inferior pharyngeal constrictor and cricopharyngeus muscle, as well as of the cervical esophagus. For the measurement of the T_2 relaxation time, the regions of interest were placed into the widest portions of the visible muscle at the level of best differentiation of the muscle to the adjacent tissue planes. For patients with significant muscle wastage, the size of the regions of interest was adjusted to the size of the wasted muscles to avoid a skew of the readings by capturing the relaxation time of the adjacent tissue planes. The distance measurements were performed by a board-certified radiologist (I.S.), with qualification in neuroradiology, who was unaware of the clinical and disease status of the patients. The measurements were recorded for each side separately.

The secondary outcomes included the following:

Changes in the Functional Oral Intake Scale score (FOIS) (16). An abnormal diet was defined as nonoral feeding or oral intake requiring a restricted consistency or special preparation (*i.e.,* FOIS level of $\leq$5). Functional swallowing was defined as a return to the pre-CRT diet without swallowing-related complications.

Swallowing function measured using the Mann Assessment of Swallowing Ability (MASA) (17), confirmed by the videoendoscopic and videofluoroscopic evaluation findings. A significant change was defined as $\pm$10 points on the MASA. The videofluoroscopic assessment included a standard protocol of thin liquid, nectar-thick liquid, and pudding (Varibar, EZ-Em, Westbury, NY) in 5- and10-mL amounts. If appropriate (*i.e.,* did not place the patient at risk of airway compromise), the patients were offered a cup to drink self-selected volumes of liquids and a cracker coated with barium pudding to masticate and swallow. The videofluoroscopic assessment was conducted by a radiologist, who was unaware of the results of the clinical assessment. Scoring followed a published median-weighted scoring system (18, 19).

Change in mouth opening during the study period.

Change in nutritional status, reflected by patient weight during the study period.

Favorable outcome (*i.e.,* composite variable of weight loss $<$10%, maintenance of oral feeding, and change in MASA of $\leq$5 points).

Occurrence of dysphagia-related complications (*e.g.,* pneumonia, dehydration).

Change from baseline to 6 week assessment in unstimulated whole saliva production measured using standard saliometric techniques (20).

Change in smell and taste perception evaluated using the University of Pennsylvania Smell Identification Test (21) (Sensonics, Haddon Heights, NJ) and Accusens T Taste function kit (22) (Westport Pharmaceuticals, Westport, CT).

Statistical analysis

Sample size calculations were determined from previous reports that 30% of HNC patients with dysphagia returned to a pretreatment diet by 6 months. Because previous studies had not used concomitant swallowing therapy, we hypothesized that the patients assigned usual care would have greater muscle decline and that concomitant swallowing therapy would improve that rate by 20% in absolute terms to 50% at 6 months. Therefore, we estimated that 60 patients would provide 80% power at the 5% (two-tailed) significance level to identify this treatment effect.

Repeated measures analysis of variance were used to evaluate the primary MRI outcome. Post hoc testing used Dunnett's and Bonferroni's corrections. The risk ratios and 95% confidence intervals were derived for the functional outcomes. Chi-square tests were used for the discrete counts of patients with adverse events. The three treatment groups were directly compared as the numbers permitted. Subsequently, the primary comparison of interest was between the pharyngocise and usual care groups. A trend analysis was conducted using the chi-square test for linear trend in proportion for all three groups. Exploratory logistic regression analysis was conducted for a favorable outcome at the CRT endpoint.

Results

A total of 703 HNC patients were reviewed between November 2001 and April 2004. Of these 703 patients, 92 (13%) were eligible for inclusion (Fig. 1). Of the 92 eligible patients, 58 (70%) provided written informed consent and were randomized to the usual care (n = 20), sham (n = 18), and pharyngocise (n = 20) groups. The reasons for nonenrollment are provided in Fig. 1. The ineligible patients did not differ significantly from the enrolled subjects in tumor type ($p \leq .95$), location ($p \leq .81$), or size ($p \leq .57$). All randomized patients were included in the intent-to-treat analysis.

The three study arms were characterized by a similar proportion of baseline factors (Table 1). The mean interval to recruitment was 35.1 $\pm$ 28.6 days after diagnosis, and the mean interval to

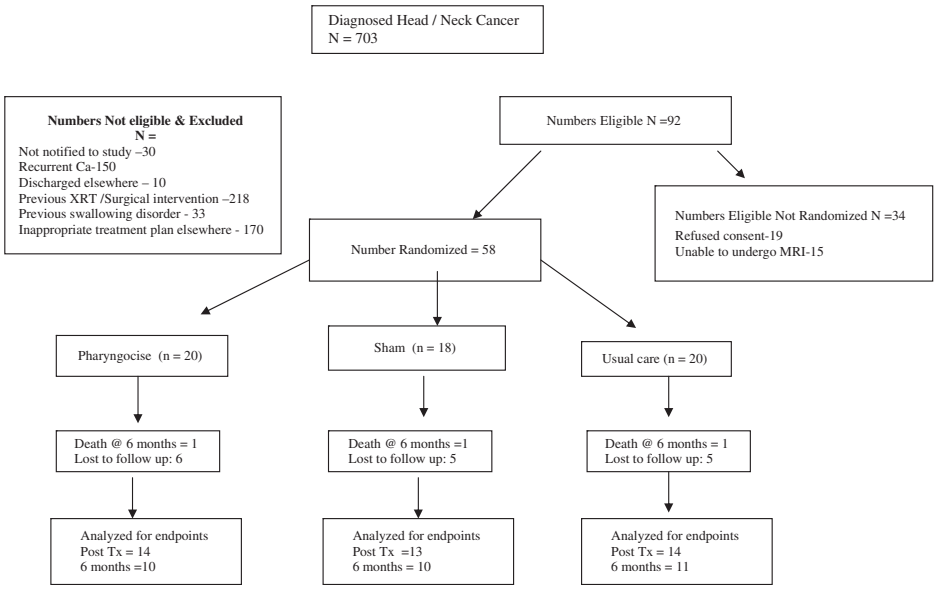


Fig. 1. Trial profile.

randomization in the study was 2.8 ± 8.2 days after the radiation oncology assessment. Of the 58 patients, 36 underwent radiotherapy and 22 underwent concurrent chemotherapy. The mean duration of CRT was not different among the three groups. No significant difference was found among the groups in age, gender, tumor size, tumor site, tumor location side, radiation dose administered, or provision of concurrent chemotherapy (Table 1). During the treatment course, 3 patients died of complications associated with their primary diagnosis or treatment.

Swallowing intervention

The number and duration of swallowing therapy sessions for the patients assigned to the treatment arms (pharyngocise and sham) were significantly greater than those for the usual care group [$F(2,81) = 4.8$, $p < .0001$]. No differences emerged between the treatment arms in the intervention length ($p \leq .58$), total work/exercise performed (cycles) ($p \leq .42$), or duration of sessions (minutes) ($p \leq .016$). The number of sessions received differed significantly between the groups (pharyngocise, 19.9; sham, 25.8; $t = -2.194$; $p \leq .03$).

Home practice

On average, 68% of the subjects complied with the home practice activities. Significantly more subjects in the sham group (28.3) than in the pharyngocise group (20.4; $t = -3.096$; $p < .007$) complied with home practice.

Follow-up

The follow-up data to 6 months were complete for 31 (56%) of the 55 survivors. The data from the 3 patients who died and the 24 patients lost to follow-up (16 at 6 weeks and 8 at 6 months) were censored for the time spent in the study and included in the analysis (Fig. 1).

Primary outcome

Maintenance of muscle composition

All groups demonstrated deterioration in muscle composition during CRT (Fig. 2). Our primary focus was to prevent the deterioration in muscle and swallowing characteristics. The MRI data calculated for the primary side of radiation exposure are presented in Table 2. The data for three muscle groups (*i.e.,* middle pharyngeal constrictor, inferior pharyngeal constrictor, and cervical esophageal wall) demonstrated movement and image artifact in the follow-up examinations and are not presented. From the remaining muscles groups, the muscle size and T_2 relaxation time were significantly different among the study arms (Table 2). Specifically, three muscles related to swallowing function demonstrated greater preservation in the pharyngocise group. The genioglossus showed more deterioration in the usual care group (length, $p \leq .03$; T_2 value, $p \leq .01$). Similar findings were obtained for the mylohyoid (thickness, $p \leq .02$; T_2 value, $p \leq .017$) and the hyoglossus (length, $p \leq .01$; T2 value, $p \leq .037$; Table 2). The T_2 relaxation time demonstrated a significant reduction in all three muscle groups for the pharyngocise group compared with the other study groups.

Secondary outcomes

Functional swallowing ability

Thirty-one percent of the patients demonstrated a significant reduction in the MASA score (defined as ≥10 points) during the CRT period. The functional swallowing ability deteriorated less (chi-square = 3.28, $p \leq .03$) in the pharyngocise group than in the usual care (Table 4) or sham (p for trend < .06; Table 5) groups. The absolute risk difference for achieving functional swallowing after treatment in the pharyngocise group was 36% compared with the usual care group.

Oral feeding

All patients consumed a normal oral diet at baseline. Only 9 patients (23%) were able to maintain a normal oral diet throughout

Table 1 Demographic characteristics

Characteristic	Usual care group	Sham group	Pharyngocise group
Age (y)	54 ± 11.3	60 ± 12.2	59 ± 10.4
Gender			
Male	15	11	18
Female	5	7	2
Interval after diagnosis (d)	33.4 ± 34.3	38.9 ± 32	33 ± 25.3
Interval to randomization (d)	2.5 ± 3.15	2.7 ± 2.5	2.8 ± 4
Tumor size (T grade)			
Median	2	2	2
Range	0–4	1–4	1–4
Tumor site (mode)			
Base of tongue	3	3	5
Tonsil	9	4	3
Tumor side			
Left	6	7	9
Right	5	5	6
Bilateral	9	6	5
Radiotherapy			
Conventional	9	6	9
IMRT	11	12	11
Plus chemotherapy *(n)*	10	6	6
Mean dose (cycles)	3.5 ± 5	2.72 ± 4.2	3.1 ± 3.9
Cisplatin *(n)*	8	2	4
Carboplatin *(n)*	3	4	2
Taxol *(n)*	4	4	3
Combined agents *(n)*	4	4	3
Radiotherapy dose (Gy)	67.5 ± 2.5	69.2 ± 1.4	72.5 ± 1.18
Neck dissection *(n)*	8	6	8
Left	3	1	4
Right	5	5	4
Baseline BMI (kg/m^2)	28.6 ± 1.3	26.9 ± 1.3	26.8 ± 1.0

Abbreviations: IMRT = intensity-modulated radiotherapy; BMI = body mass index.
Data presented as mean ± standard deviation, unless otherwise noted.

the CRT period. The patients in the pharyngocise group maintained oral feeding more often than those in the usual care group (42% vs. 14%, respectively). During CRT, 12 patients (31%) began nonoral (gastrostomy tube) feeding, including 10% with prophylactic tube placement. Fewer subjects received gastrostomy tube feeding in the pharyngocise group (20%) than in the usual care group (30%). At 6 months, 6 patients (21%) were not oral feeding, with most ($n = 4$) in the usual care arm.

Functional oral intake scale

All groups demonstrated diet alteration (reduction in the FOIS score) during CRT. Although the pharyngocise group demonstrated a greater median FOIS score after treatment. However, this change was not significantly different statistically among the groups after treatment (Table 3).

Video endoscopic and videofluoroscopic

The video endoscopic review demonstrated significant changes in pharyngeal structure across all groups during the study period (Fig. 3). Similarly, videofluoroscopic evaluation (Table 3) demonstrated an alteration in swallowing ability within all arms. The common changes included reduced tongue base retraction, hyolaryngeal elevation, and pharyngeal clearance. The weighted scores were not significantly different among the groups. The prevalence of aspiration was low (14%, $n = 8$), with no statistically significant differences among the groups.

Mouth opening

During the CRT period, the mouth opening reduced by a mean of 3.8 ± 5.08 mm. A greater declination in opening was noted in patients receiving radiotherapy (4.8 mm) than in those receiving CRT (2.7 mm). However, this difference was not statistically significant. The pharyngocise group demonstrated significantly less decline in mouth opening (1.6 mm) than did the sham and usual care groups [5.1 mm and 4.3 mm, respectively; $F(2,43) = 3.28$, $p \leq .47$]. The post hoc analysis identified a significantly superior outcome for the pharyngocise group (6.38, $p \leq .046$) compared with the usual care (Table 3).

Nutrition

The mean weight loss per patient during the study period was 6.69 kg (mean ± standard deviation, 14.75 ± 4.9 lb). A total of 23 patients (40%) lost >10% of their baseline body weight by the 6-week point. A greater number of subjects receiving CRT (61%) lost >10% of their body weight than those receiving RT alone (38%). The average weight loss was not significantly different among the groups after treatment.

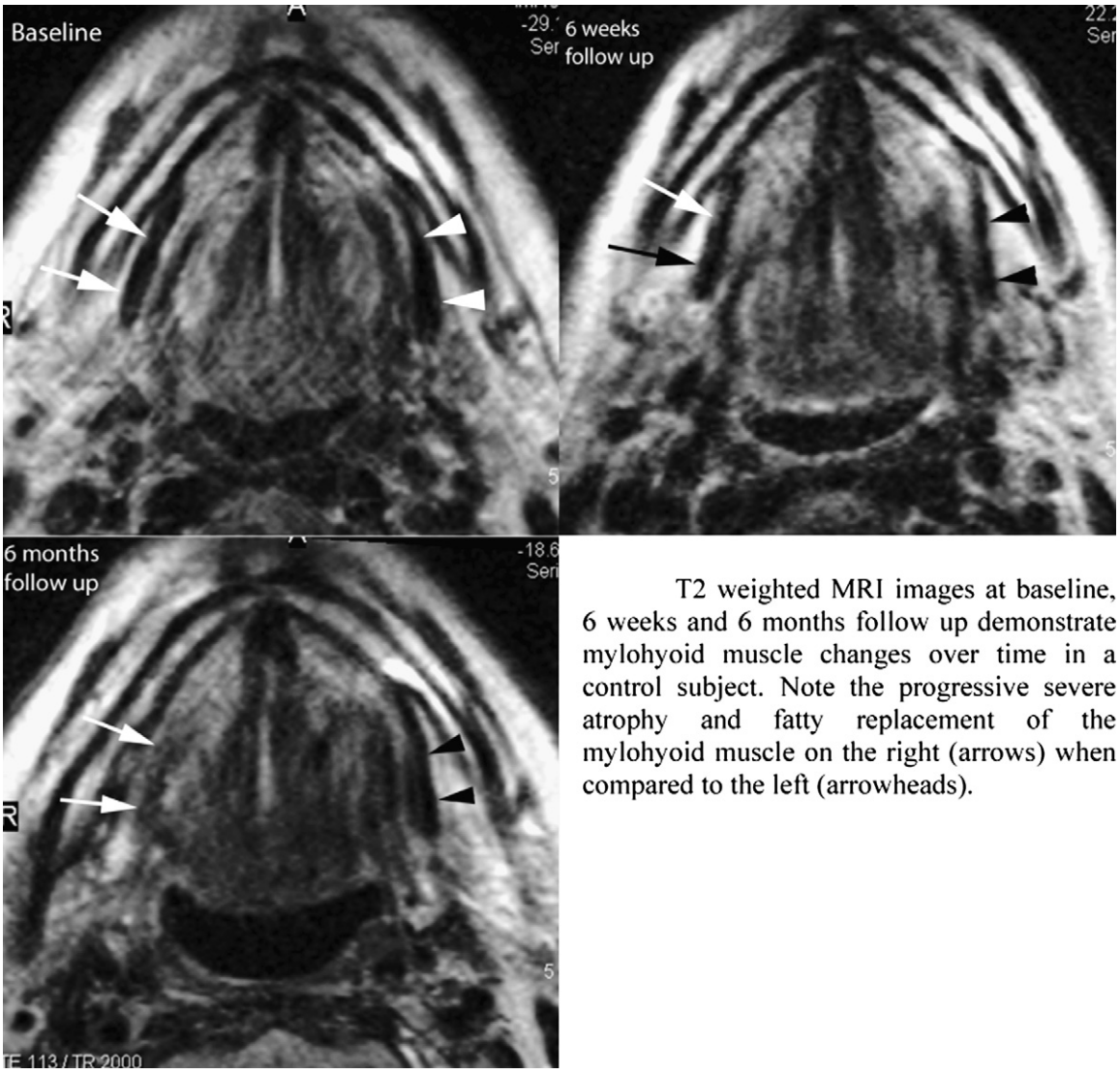


Fig. 2. Example of T_2-weighted muscle change in control arm subject.

Favorable outcome after CRT

The *a priori* composite for a favorable outcome (weight loss <10%, maintenance of oral feeding and minimal change in MASA score [≤5 points]) was reached by 57% ($n = 33$) of the sample at the post-treatment evaluation point. A greater proportion of patients in the intervention arms (86% in the pharyngocise and 82% in the sham groups) reached this endpoint than in the control arm (47%). Participation in the pharyngocise arm was associated with a more favorable outcome ($p \leq .009$). Exploratory logistic regression analysis ($n = 58$; 5 fitted variables) revealed that participation in the pharyngocise arm produced a superior benefit (odds ratio, 6; 95% confidence interval, 1−37.2). The final model indicated significant predictive power for the variables pharyngocise ($p = .05$) and sham ($p = .06$). The odds that a patient receiving pharyngocise treatment for swallowing would have a favorable outcome after CRT were six times greater than the corresponding odds for a patient who did not receive preventative exercise during CRT. In addition, the post hoc Homer-Lemeshow test from this model yielded a p value of .987, suggesting a model with adequate predictive value.

Salivation

Reduced salivary flow was identified in >80% of the patients by the end of CRT. The mean reduction in salivary flow was 0.182 ± 0.21 mL/min. Repeated measures analysis of variance demonstrated a significant difference in salivation decline [$F(1,36) = 30$, $p \leq .0001$] with the post hoc comparison [$F(1,36) = .238$, $p \leq .020$], demonstrating significant preservation of the salivary flow in the pharyngocise group. The absolute risk reduction for salivation decline in the pharyngocise group was 35% compared with the usual care group (Table 4).

Taste

Taste reduction was noted in 32 patients (82%) during the CRT period. The taste decline demonstrated a significant difference among the groups [chi-square (trend) = 5.8, $p \leq .053$]; with fewer patients in the pharyngocise group demonstrating a decline in taste acuity (Table 5). The absolute risk reduction for the taste decline in the pharyngocise group compared with the usual care group was 19% (Table 4).

Table 2 Muscle composition at 6 weeks

Muscle	Study arm			*p*
	Usual care	Sham	Pharyngocise	
Genioglossus*				
Length				<.03
Before	37.08 ± 6.4	34 ± 4.7	34.9 ± 4.8	
After	33.6 ± 5.7	32.5 ± 3.9	34.4 ± 2.7	
Change	3.67	1.5	0.5	
Thickness				NS
Before	7.31 ± 1.9	7.41 ± 0.7	7.54 ± 1.8	
After	6.89 ± 0.7	6.97 ± 0.6	7.11 ± 1.8	
Change	0.42	0.43	0.44	
T_2				<.01
Before	108.1 ± 5.2	107 ± 6.6	111.2 ± 3.8	
After	108.05 ± 2.1	104.9 ± 4.1	101.6 ± 5	
Change	0.05	2.1	9.6	
Hyoglossus*				
Length				<.018
Before	21.04 ± 4.1	17.9 ± 4.1	17.4 ± 3.9	
After	17.2 ± 3.6	16.9 ± 3.4	17.9 ± 3.07	
Change	3.84	1	−0.05	
Thickness				NS
Before	4.11 ± 0.88	3.1 ± 0.73	2.9 ± 0.95	
After	3.06 ± 0.86	3.2 ± 0.9	2.5 ± 0.6	
Change	1.05	−0.1	0.4	
T_2				<.037
Before	104.2 ± 4.1	106.8 ± 6.2	114.7 ± 8.8	
After	104.9 ± 3.7	105.1 ± 2.6	105.1 ± 2.6	
Change	−0.07	1.7	9.6	
Mylohyoid*				
Thickness				<.021
Before	4.4 ± 1.1	2.86 ± 0.7	3.86 ± 0.96	
After	2.8 ± 0.78	3.01 ± 1.0	3.8 ± 1.2	
Change	1.6	−0.15	0.06	
T_2				<.017
Before	104.1 ± 4.6	103.7 ± 4.4	111.8 ± 11.3	
After	106.3 ± 6.5	104.1 ± 5.6	103.8 ± 3.4	
Change	−2.2	−0.4	8	

Data presented as mean ± standard deviation, unless otherwise noted.
data displayed is from primary field of irradiation.
* Repeated measures analysis of variance within measures − time group.

Smell

Overall, 12 (32%) patients demonstrated a decline in olfactory acuity by the end of CRT. A significant difference between the pharyngocise and usual care groups was identified in olfactory decline (chi-square = 4.1, $p \leq .03$), with a superior outcome identified in the pharyngocise group. The absolute risk reduction for olfactory decline in the pharyngocise group was 39% compared with the usual care group (Table 4).

Dysphagia-related complications

No significant associations were noted between the treatment group and dysphagia-related complications. Pneumonia was uncommon, occurring in 3.4% ($n = 2$) of the patients. Dehydration was identified in 17.2% ($n = 10$) of the group and was significantly associated with concurrent chemotherapy (chi-square = 5.97, $p \leq .015$). Mucositis occurred in 35 patients (65%) during CRT and oral yeast infections in 8 patients (14%). No association was identified between the occurrence of mucositis and oral yeast infection and the treatment group.

Discussion

The results from the present study have demonstrated that swallowing exercises administered during CRT results in the maintenance of head-and-neck musculature and improved swallowing indexes. Furthermore, mobilizing swallowing muscles at any level could affect the feeding and chemosensory outcomes in this population.

The present study identified maintenance of muscle characteristics from swallow exercise during CRT. Previous research has reported muscle thickening and T_2 elongation associated

Table 3 Swallowing outcomes

Variable	Study arm			*p*
	Usual care	Sham	Pharyngocise	
MASA				
Baseline	195.5 ± 4	194.7 ± 3.5	195.1 ± 5.9	NS
6-wk Outcome	171.5 ± 14.2	173.6 ± 11.8	177.14 ± 12.5	≤.006
Change	24.16 ± 13.4	20.8 ± 12.9	17.7 ± 10.1	
FOIS				
Baseline				NS
Median	7	7	7	
Range	5–7	5–7	5–7	
6-wk Outcome				
Median	4	4	5	
Range	1–6	1–7	2–7	
VFE score				
Baseline	0.186 ± 0.09	0.272 ± 0.15	0.214 ± 0.02	NS
At 6 wk	0.214 ± 0.09	0.343 ± 0.16	0.200 ± 0.16	
Mouth opening				
Baseline	36.6 ± 8.05	39.2 ± 6.4	41.6 ± 8.4	NS
At 6 wk	32.3 ± 5.9	34.07 ± 7.3	40.05 ± 8.3	<.047*
Change	4.3	5.1	1.6	

Data presented as mean ± standard deviation, unless otherwise noted.
* Dunnett's post hoc comparison.

with edema in head-and-neck muscles receiving doses of >50 Gy (23). Accordingly, the reduction in T_2 relaxation time and maintenance of muscle size associated with the pharyngocise protocol might reflect a deterrent to inflammatory changes noted with CRT. Although T_2 declination could be influenced by multiple factors, the reduction in muscle edema or fatty infiltration is likely to be a contributing factor. The combination of T_2 declination with maintenance in the muscle structure and preservation of swallowing function in the pharyngocise group supports this conclusion. The MRI results for the sham group were between those of the pharyngocise and usual care groups, suggesting that patients might receive a benefit from lower intensity exercise regimens.

The subjects in all three groups were treated by the same team of radiation oncologists, received comparable CRT regimens, and did not differ in tumor site or disease extent. Specific swallowing muscle dosimetry was not available for all subjects to confirm the balanced exposure to the muscles of interest. Notwithstanding, we believe the application and exposure to medical intervention did not differ by group.

The present study is the first truly randomized trial to evaluate a systematic program of swallowing exercises completed during CRT. Two previously published studies suggested that pretreatment swallowing therapy improved the post-treatment quality of life and limited swallowing variables (epiglottic inversion and tongue base position) in HNC patients (24, 25). These studies, conducted by the same center, provided swallowing intervention for 2 weeks before CRT not concomitantly. Furthermore, the design of those studies (unmatched case control and cross sectional) was not as rigorous as the design of the present trial and the total number of patients was smaller ($n = 9$ and $n = 37$, respectively). Similar to our study, Van Der Molen *et al.* (26) described the application of swallowing exercises concurrent with CRT in 49 patients treated for HNC. That study did not

Table 4 Comparison of pharyngocise vs. usual care at 6 weeks

Outcome (at 6 wk)	Intervention		Analyses		
	Usual care	Pharyngocise	RR	95% CI	ARR (%)
Normal diet	2/14	5/12*	2.91	0.68–12.4	27
Nonoral feeding	6/14	3/12*	0.58	0.18–1.84	18
Functional swallowing	2/14	6/12*	3.5	0.86–14.2	36†
Weight loss (>10%)	6/13*	4/14	0.62	0.22–1.7	18
Salivation decline	12/13*	8/14	0.62	0.38–1.02	35†
Taste decline	10/12*	9/14	0.77	0.48–1.23	19
Smell decline	6/11*	2/13*	0.28	0.07–1.13	39†
Any complication	7/14	5/12*	0.71	0.31–1.6	17

Abbreviations: RR = relative risk; CI = confidence interval; ARR = absolute risk reduction (risk difference).
* Chi-square significance.
† Missing data points.

Table 5 Comparison of pharyngocise vs. sham vs. usual care at 6 weeks

Outcome (at 6 wk)	Intervention			Trend analyses, p for trend
	Usual care (n = 14)	Sham (n = 13)	Pharyngocise (n = 14)	
Normal diet	2	2	5	.185
Nonoral feeding	6	3	3	.295
Functional swallowing	2	2	6	.067*
Weight loss (>10%)	6	6	4	.604
Salivation loss	12	12	8	.061*
Taste decline	10	13	9	.053*
Smell decline	6	4	2	.123
Any complication	7	4	5	.597

* Trend toward significance from chi-square trend analysis.

include a control group but compared two forms of swallowing therapy. Both swallowing therapies involved patient-controlled and clinician-directed exercises. The results indicated significant decreases in oral intake, mouth opening, and weight at 10 weeks after CRT. However, the patients in both treatment groups demonstrated reduced feeding tube dependency. Thus, although their results did not address the efficacy of active exercise on the outcome, they did address the potential benefit from any exercise and the acceptability and feasibility of swallowing therapy for this population. In this respect, although limited, the results from previous studies support our results.

Our sample included both RT and CRT patients, providing greater generalization to the HNC treatment population. The exercise protocol used was significantly different between the groups and used validated muscle and swallowing outcome measures. Although the number of patients and outcome events at the 6-month follow-up period were small (because of morbidity and measurement artifact), we were able to demonstrate the consistency of results across several outcome events (all favoring the pharyngocise group), strongly suggesting a positive treatment effect.

Although our study results suggest benefit (physiologically and functionally) from swallowing exercises, the dose–response curve

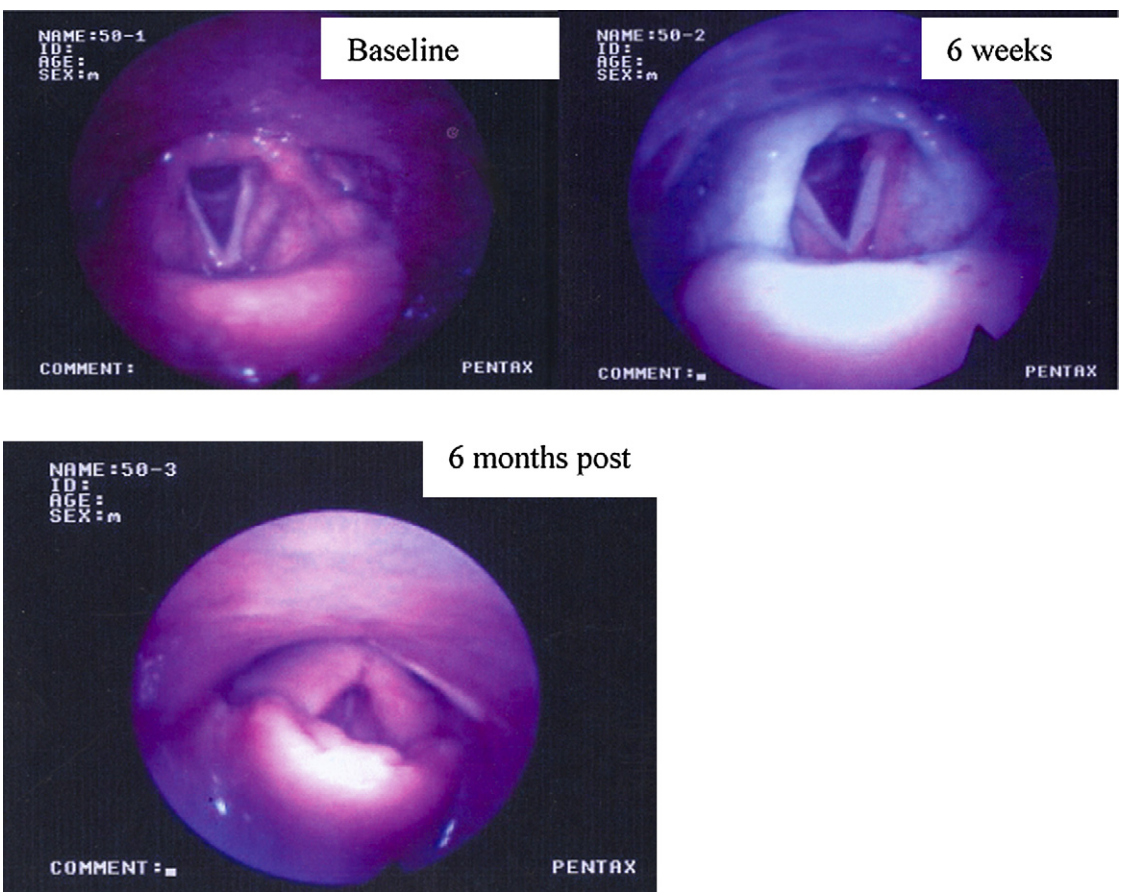


Fig. 3. Endoscopic image showing change in anatomy of oropharynx in control arm subject.

for this form of behavioral treatment remains unclear. Benefit was derived not only from the intensive intervention arm but also from sham intervention. These arms did not differ in length or duration of intervention or total work performed. Whether the benefits obtained by the sham group can be ascribed to a placebo effect of behavioral attention or to the affect of attenuated movement is unclear. A larger study is underway to review the dose−response effect of low- and high-intensity pharyngocise intervention.

The data from our study were most complete up to the 6-week post-treatment point. We experienced a withdrawal rate at 6 months that precluded the meaningful analysis of many outcomes to that point. This is not an unusual finding in the HNC population, for whom the high morbidity levels and associations with negative lifestyle factors elevate the lost-to-follow-up rates. A comparison between the enrolled patients with and without complete data in the present study did not reveal significant differences in age, cancer stage, or swallowing comorbidity, suggesting that our results are representative.

Conclusion

The results of the present study demonstrated a benefit from a program of simple swallowing exercises administered during CRT. This approach is novel in timing of delivery and preventative design. Given the health costs of dysphagia from HNC and positive outcomes reported from the present study, it is imperative that additional research be undertaken to refine the swallowing treatments and their delivery for this population. Preventative swallowing programs can offer a cost-effective alternative to prevent medically related complications and optimize functional outcome for HNC patients.

References

1. Colangelo LA, Logemann JA, Pauloski BR, *et al*. T stage and functional outcome in oral and oropharyngeal cancer patients. *Head Neck* 1996;18:259−268.
2. Chasen MR, Bhargava R. A descriptive review of the factors contributing to nutritional compromise in patients with head and neck cancer. *Support Care Cancer* 2009;17:1345−1351.
3. Pauloski B, Logemann JA, Rademaker AW, *et al*. Speech and swallowing function after oral and oropharyngeal resections: One year follow-up. *Head Neck* 1994;16:313−322.
4. Lazarus C. Effects of radiation therapy and voluntary maneuvers on swallow functioning in head and neck cancer patients. *Clin Commun Disord* 1993;3:11−20.
5. Lazarus C, Logemann JA, Pauloski BR, *et al*. Swallowing disorders in head and neck cancer patients treated with radiotherapy and adjuvant chemotherapy. *Laryngoscope* 1996;106:1157−1166.
6. Riekki R, Jukkola A, Sassi M, *et al*. Modulation of skin collagen metabolism by irradiation: Collagen synthesis is increased in irradiated human skin. *Br J Dermatol* 2000;142:874−880.
7. Kendall K, McKenzie M, Leonard R, Jones C. Structural mobility in deglutition after single modality treatment of head and neck carcinomas with radiotherapy. *Head Neck* 1998;20:720−725.
8. Tyml K, Mathieu-Costello O. Structural and functional changes in the microvasculature of disused skeletal muscle. *Front Biosci* 2001;6: D45−D52.
9. Berg HE. *Effects of unloading on skeletal muscle mass and function in man*. Stockholm: Physiology and Pharmacology, Karolinska Institutet; 1996.
10. Berg HE, Tesch PA. Changes in muscle function in response to 10 days of lower limb unloading in humans. *Acta Physiol Scand* 1996;157:63−70.
11. Piquet F, Stevens L, Butler-Browne G, Mounier Y. Differential effects of a six-day immobilization on newborn rat soleus muscles at two developmental stages. *J Muscle Res Cell Motil* 1998;19:743−755.
12. Warfield SK, Mulkern RV, Winalski CS, *et al*. An image processing strategy for the quantification and visualization of exercise-induced muscle MRI signal enhancement. *J Magn Reson Imaging* 2000;11:525−531.
13. Foley JM, Jayaraman RC, Prior BM, *et al*. MR measurements of muscle damage and adaptation after eccentric exercise. *J Appl Physiol* 1999;87:2311−2318.
14. Zuydam AC, Rogers SN, Brown JS, *et al*. Swallowing rehabilitation after oro-pharyngeal resection for squamous cell carcinoma. *Br J Oral Maxillofac Surg* 2000;38:513−518.
15. Buchbinder D, Currivan RB, Kaplan AJ, Urken ML. Mobilization regimens for the prevention of jaw hypomobility in the radiated patient: A comparison of three techniques. *J Oral Maxillofac Surg* 1993;51:863−867.
16. Crary MA, Mann GD, Groher ME. Initial psychometric assessment of a functional oral intake scale for dysphagia in stroke patients. *Arch Phys Med Rehabil* 2005;86:1516−1520.
17. Mann G. *The Mann assessment of swallowing ability*. Clifton Park, NJ: Singular, Thompson, Delmar Learning; 2002.
18. Mann G, Hankey G, Cameron D. Swallowing disorders following acute stroke: Prevalence and diagnostic accuracy. *Cerebrovasc Dis* 2000;10:380−386.
19. Carnaby-Mann G, Crary M. McNeill Dysphagia Therapy Program: A case-control study. *Arch Phys Med Rehabil* 2010;91:743−749.
20. Ben-Aryeh H, Miron D, Szargel R, Gutman D. Whole-saliva secretion rates in old and young healthy subjects. *J Dental Res* 1984;63:1147−1148.
21. Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: A standardized microencapsulated test of olfactory function. *Physiol Behav* 1984;32:489−502.
22. Henkin R. Human taste and smell disorders. In: Adelman G, editor. Encyclopedia of neuroscience. 2nd ed. Boston: Birkhauser; 1999. p. 2010−2013.
23. Popovtzer A, Cao Y, Feng F, Eisbruch A. Anatomical changes in the pharyngeal constrictors after chemoirradiation of the head and neck cancer and their dose−effect relationships: MRI-based study. *Radiother Oncol* 2009;93:510−515.
24. Carroll WR, Locher JL, Canon CL, *et al*. Pretreatment swallowing exercises improve swallow function after chemoradiation. *Laryngoscope* 2008;118:39−43.
25. Kulbersh BD, Rosenthal EL, McGrew BM, *et al*. Pretreatment, preoperative swallowing exercises may improve dysphagia quality of life. *Laryngoscope* 2006;116:883−886.
26. van der Molen L, van Rossum MA, Burkhead LM, *et al*. A randomized preventive rehabilitation trial in advanced head and neck cancer patients treated with chemoradiotherapy: Feasibility, compliance, and short-term effects. *Dysphagia* 2011,26(2):155-70.

Reprinted by permission of Otolaryngol Head Neck Surg. 2015; 152(5): 863-867.

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Otolaryngology–Head and Neck Surgery
2015, Vol. 152(5) 863–867

DOI: 10.1177/0194599815575508
http://otojournal.org
SAGE

A Swallow Preservation Protocol Improves Function for Veterans Receiving Chemoradiation for Head and Neck Cancer

Kevin A. Peng, MD[1], Edward C. Kuan, MD, MBA[1], Lindsey Unger, MS[2], William C. Lorentz, MD[3], Marilene B. Wang, MD[1,4], and Jennifer L. Long, MD, PhD[1,4,5]

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

Received October 24, 2014; revised January 2, 2015; accepted February 10, 2015.

Abstract

Objective. Determine the efficacy of a swallow preservation protocol (SPP) on maintaining swallow function in patients undergoing chemoradiation (CRT) or radiation therapy alone (RT) for head and neck squamous cell carcinoma (HNSCC).

Study design. Retrospective case series.

Setting. Veterans Affairs medical center.

Subjects and Methods. Patients treated with CRT or RT for HNSCC between February 2006 and November 2013 were studied. Those enrolled in the SPP participated in swallowing, jaw, and tongue exercises during cancer therapy. The comparator group received no swallowing intervention during CRT. A previously described functional outcome swallowing scale (FOSS; 0 = no symptoms and 5 = nonoral feeding for all nutrition) was used to quantify dysphagia prior to and at the completion of cancer therapy, and an analysis was performed to compare swallowing function.

Results. Forty-one (all male; mean age, 66 years) and 66 patients (all male; mean age, 61 years) were included in the SPP and comparator groups, respectively. In the SPP group, mean pre- and posttreatment FOSS scores were 2.2 and 2.2, respectively, while the corresponding scores in the comparator group were 1.8 and 2.7, respectively, with posttreatment FOSS scores being significantly worse than pretreatment FOSS scores in the comparator group only.

Conclusion. Patients enrolled in the SPP demonstrated preserved swallowing function over the course of cancer treatment compared with a comparator group. This confirms the importance of early evaluation and intervention for dysphagia prior to and during CRT or RT alone.

Keywords

dysphagia, head and neck cancer, chemoradiation, swallow preservation

Dysphagia is a debilitating side effect of organ-sparing treatment for head and neck squamous cell carcinoma (HNSCC).[1-3] Risk factors for development of dysphagia after combined chemotherapy and radiation therapy (CRT) for HNSCC include an oropharyngeal primary site, cessation of per os (PO) intake during treatment, and conventional 2D or 3D-conformal radiation therapy.[4-7] Manifestations of dysphagia include prolongation of mealtime, aspiration, weight loss, dietary limitations, and the need for nonoral nutrition.[8,9] Acute dysphagia during the course of cancer treatment may pose life-threatening challenges, particularly with regard to inadequate hydration and nutrition.[10] Late dysphagia may manifest as a pharyngoesophageal stricture or aspiration, and intensive therapy may be required to reverse gastrostomy tube dependence.[11-13]

Intensity-modulated radiation therapy with optimization of radiation dose to avoid constrictor musculature has emerged as an important technique to avoid both early and late dysphagia, in part by ameliorating inflammation, fibrosis, and eventual diminished mobility of pharyngeal structures.[7,14-17]

[1]Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA, Los Angeles, California, USA
[2]Department of Audiology and Speech Pathology, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California, USA
[3]Department of Radiation Oncology, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California, USA
[4]Department of Surgery, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California, USA
[5]Research Service, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California, USA

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

Corresponding Author:
Kevin A. Peng, MD, University of California, Los Angeles, Head and Neck Surgery, 10833 Le Conte Ave, CHS 62-132, Los Angeles, CA 90095, USA.
Email: kap@ucla.edu

Table 1. Functional Outcome Swallowing Scale (FOSS).[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, 1999. (© 2000 Karger Publishers, Basel, Switzerland.)

Notwithstanding this, dysphagia remains a common complaint following cancer therapy.

Subjective measures of dysphagia include the Performance Status Scale for Head and Neck Cancer patients (PSS-H&N) and the M.D. Anderson Dysphagia Inventory (MDADI).[6,18] However, disadvantages of these subjective scales include patient bias and inconsistency among subjects. Objective measures, including the modified barium swallow study (MBSS) and the videofluoroscopic swallow study (VFSS), offer detailed information about swallowing anatomy and physiology.[5,19] However, time and resource constraints may preclude the clinician from performing these studies at each consecutive visit, and the complexity of findings may make numerical grading and subsequent statistical analyses difficult. Therefore, to stage dysphagia, we elected to use an objective, clinician-determined scale of oropharyngeal dysphagia, the functional outcome swallowing scale (FOSS), first proposed by Salassa (**Table 1**).[20]

Strategies for rehabilitation of swallowing during and following CRT include postural adjustments, diet modification, range-of-motion exercises, and the strengthening of pharyngeal and suprahyoid musculature.[21,22] At our institution, we have implemented a swallow preservation protocol (SPP) comprising swallowing, jaw, and tongue exercises presented to patients prior to or within 2 weeks of beginning CRT. Exercises are performed for 10 repetitions 3 times daily for a total of 30 repetitions per exercise per day. A jaw motion rehabilitation system is used as necessary in patients who demonstrate trismus prior to or at any point during CRT, and patients are asked to self-report compliance with the SPP using a diary. Patients are seen by speech pathology practitioners every 1 to 2 weeks during CRT; following completion of CRT, patients are seen on a variable basis ranging from once every few weeks to once every several months.

In this study, we sought to investigate the effect of our SPP on dysphagia following CRT or radiation therapy (RT) alone for HNSCC and hypothesized that veterans participating in a SPP during CRT or RT would demonstrate better posttreatment swallowing outcomes compared with a comparator population.

Methods

Subjects

The Institutional Review Board of the Greater Los Angeles Veterans Affairs Health System approved this study. A retrospective chart review was conducted of all patients treated with CRT or RT alone for HNSCC at a Veterans Affairs Medical Center between February 2006 and November 2013, including both patients who did and did not participate in the SPP. Demographic and clinical information was gathered. Using clinical documentation by speech pathology and head and neck surgery, swallowing function was assessed using the FOSS within 1 to 2 weeks prior to the beginning of CRT/RT (''pretreatment'') and within 2 to 4 weeks after the termination of CRT/RT (''posttreatment''). Compliance to the SPP was also noted.

Swallow Preservation Protocol

Patients were enrolled in the SPP beginning in September 2010; by July 2013, nearly all veterans undergoing CRT or RT for HNSCC were enrolled in the SPP. No specific clinical factors influenced the decision to enroll a patient in the SPP.

The SPP consists of 2 jaw exercises, 2 tongue exercises, and 4 swallowing exercises. Jaw exercises include the jaw stretch and the lateral jaw stretch, comprising jaw opening and lateral jaw displacement in both directions 10 times in a row, 3 times daily. Tongue exercises include the tongue press (forced contraction of the tongue against the anterior hard palate) and anterior and lateral tongue stretch (forced contraction of the tongue anteriorly and to the left and right), also 10 times in a row, 3 times daily.

The 4 swallowing exercises, which compose the majority of the SPP, are the Shaker exercise, the Mendelsohn maneuver, the Masako tongue-hold, and the effortful swallow. The Shaker exercise, designed to strengthen the suprahyoid musculature and enhance opening of the upper esophageal sphincter (UES), consists of prolonged, forced flexion of the neck in a supine position followed by 3 fast repetitions of the same.[23] The Mendelsohn maneuver, designed to prolong

hyolaryngeal elevation at the peak of the swallow, also facilitates UES opening. Patients are instructed to palpate the cartilaginous laryngeal framework as they swallow without food (''dry swallow'') and develop voluntary motor control of hyolaryngeal elevation.[24]

In the Masako tongue-hold, the patient bites firmly but comfortably on the anterior oral tongue using the upper and lower incisors, thus rendering it immobile, and then performs dry swallows.[25] This procedure augments the anterior excursion of the posterior pharyngeal wall. Finally, in the effortful swallow, the patient imagines swallowing a large object (''swallow a large vitamin,''''swallow a ping-pong ball''), theoretically strengthening all muscle groups involved in swallowing.

With the exception of the Shaker exercise, which is performed 3 times each in prolonged and repetitive fashion, the swallowing exercises are performed 10 times in a row, 3 times daily, for a total of 30 repetitions daily. Patients were asked to log performance of jaw, tongue, and swallowing exercises in a provided diary.

Patients displaying trismus prior to, during, or following cancer therapy were also provided with and instructed in the use of a TheraBite Jaw Motion Rehabilitation System (Atos Medical AB, Hörby, Sweden) to maximize jaw opening.

Objective Assessment of Swallow Function

As described above, the FOSS, yielding ordinal scores, was used to quantify swallowing function prior to and following CRT or RT for HNSCC. The MBSS was variably performed on patients in the SPP and the comparator group, and these data were therefore excluded from analysis.

Statistical Analysis

Initially, subjects were analyzed in an intention-to-treat manner, and all patients enrolled in the SPP were included in the treatment cohort regardless of compliance. Student *t* tests and the *z* test were used to compare differences between the SPP and comparator groups. The FOSS scores were compared using Mann-Whitney *U* and Wilcoxon signed-rank tests. Statistical analysis was performed with SPSS 20 (SPSS, Inc, an IBM Company, Chicago, Illinois). Thereafter, patients who were compliant and noncompliant with the SPP were analyzed separately.

Results

The SPP and comparator groups comprised 41 and 66 patients, respectively. All subjects were male; there were no significant differences between the 2 groups with respect to mean age, mean TNM stage group at time of cancer diagnosis, and distribution of treatment modality (CRT vs RT; $P = .26$). Similarly, no significant difference was seen when comparing pretreatment FOSS scores between the SPP and comparator group (2.15 and 1.78, respectively; $P = .068$, Mann-Whitney *U;* **Table 2**). In the SPP group, compliance with treatment was 71%.

Pretreatment and posttreatment FOSS scores were compared pairwise for each subject within the SPP and comparator groups. In the SPP group, there was no significant difference between pre- and posttreatment FOSS (2.15 and 2.23, respectively; Wilcoxon signed-rank, $P = .343$). In the comparator group, a significant difference was observed between pre- and posttreatment FOSS (1.78 and 2.73, respectively; $P = .000$), consistent with worse swallow function posttreatment (**Table 3**).

Compliant and noncompliant patients in the SPP group were then analyzed separately. The compliant cohort had no statistically significant difference in swallowing function when comparing pretreatment with posttreatment FOSS score ($P = .887$, Wilcoxon signed-rank), while the noncompliant cohort demonstrated a trend toward worse swallowing function that did not reach significance ($P = .102$, Wilcoxon signed-rank).

As increasing age has previously been implicated in worse swallowing function after CRT, we stratified patients by age, considering patients 55 years and younger separately from those older than 55 years. In the SPP group, both age groups revealed no significant difference when comparing pre- and posttreatment FOSS ($P = .435$ and .655 for the younger and older age groups, respectively). In the comparator group, both age groups revealed statistically significantly worse swallowing function after treatment ($P = .000$

Table 2. Patient Characteristics.[a]

	SPP (n = 41)	Comparator (n = 66)
Age, y		
Mean (range)	66 (48-88)	61 (27-80)
≤55	3 (7)	10 (15)
>55	38 (93)	56 (85)
Cancer treatment received		
CRT	32 (78)	57 (86)
RT	9 (22)	9 (14)
Compliant with SPP	29 (71)	NA

Abbreviations: CRT, chemoradiation; NA, not applicable; SPP, swallow preservation protocol; RT, radiation therapy.
[a]Values are presented as number (%) unless otherwise indicated.

Table 3. Functional Outcome Swallowing Scale (FOSS) Scores Prior to ("Pretreatment") and following ("Posttreatment") Therapy for Head and Neck Cancer.

	SPP[a]	Comparator[b]
Pretreatment,[c] mean (SD)	2.15 (1.24)	1.78 (1.55)
Posttreatment, mean (SD)	2.23 (1.37)	2.73 (1.59)

Abbreviation: SPP, swallow preservation protocol.
[a]No statistically significant difference between pretreatment and posttreatment FOSS in the SPP group ($P = .343$, Wilcoxon signed-rank).
[b]Posttreatment FOSS was statistically significantly worse than pretreatment FOSS in the comparator group ($P = .000$, Wilcoxon signed-rank).
[c]No statistically significant difference between pretreatment FOSS in the SPP and comparator groups ($P = .068$, Mann-Whitney *U*).

and .017 for the younger and older age groups, respectively). Thus, no notable difference was seen when stratifying patients by age.

Discussion

Dysphagia following chemoradiation or radiation therapy alone for head and neck cancer is a significant detriment to quality of life following curative therapy.[2] Rehabilitation of swallowing after prolonged disuse is difficult, and recent strategies focus on early intervention to ameliorate acute symptoms as well as prevent the late sequelae of fibrosis and atrophy of involved musculature.[21]

At our institution, we have implemented an SPP for veterans undergoing CRT or RT for HNSCC. This protocol includes swallowing exercises, jaw exercises, and tongue exercises that are performed 3 times daily. The 4 swallowing exercises—the Shaker maneuver, the Mendelsohn maneuver, the Masako tongue-hold, and the effortful swallow—are the core of the protocol. Together, the swallowing exercises augment and prolong UES opening, enhance posterior pharyngeal wall excursion, and globally strengthen the pharyngeal musculature. When necessary, a jaw motion rehabilitation device is provided to treat trismus. Patients were prospectively enrolled in this SPP beginning in September 2010; by July 2013, nearly all veterans undergoing CRT or RT for HNSCC were enrolled in this protocol and underwent weekly to biweekly follow-up with speech pathology providers during the course of cancer therapy.

On intention-to-treat analysis, veterans enrolled without randomization in the SPP demonstrated no significant difference compared with a comparator group with respect to demographic parameters, cancer treatment, cancer stage, and pretreatment swallowing function as quantified by FOSS score. In contrast, following CRT or RT, the comparator group demonstrated statistically worse swallowing function compared with the beginning of cancer treatment; in the SPP group, there was no significant difference between pretreatment and posttreatment swallowing function. Overall, compliance in the SPP was 71%. When analyzing patients compliant with and not compliant with the SPP separately, compliant patients demonstrated no significant difference between pre- and posttreatment swallowing function. Noncompliant patients, however, demonstrated a trend toward worse swallowing function, approaching statistical significance. Taken together, these data suggest that participation in the SPP maintained swallowing function during CRT or RT.

Limitations of the current work include lack of randomization to the SPP. The comparator group did receive cancer therapy chronologically earlier, on average, than did the SPP group, and advances in CRT or even changes in oncologic protocols may have had an unidentified influence in producing the observed differences between the SPP and comparator groups. Furthermore, patients were not stratified by primary site, and future research must probe the efficacy of the SPP, and specifically the swallowing exercises, in patients with primary tumors involving sites other than the oropharynx and hypopharynx. Finally, posttreatment follow-up in our study was 2 to 4 weeks following completion of cancer therapy; long-term swallowing function must be assessed and compared.

Conclusion

Compared with a comparator group, participants in a swallow preservation protocol during chemoradiation or radiation therapy alone for head and neck squamous cell carcinoma demonstrated preservation of swallow function during and shortly following cancer treatment.

Author Contributions

Kevin A. Peng, data acquisition, drafting manuscript, approval of manuscript, accountability to accuracy and integrity; **Edward C. Kuan**, data acquisition, drafting manuscript, approval of manuscript, accountability to accuracy and integrity; **Lindsey Unger**, data acquisition, manuscript revision, approval of manuscript, accountability to accuracy and integrity; **William C. Lorentz**, data acquisition, manuscript revision, approval of manuscript, accountability to accuracy and integrity; **Marilene B. Wang**, conception and design of work, manuscript revision, approval of manuscript, accountability to accuracy and integrity; **Jennifer L. Long**, conception and design of work, manuscript revision, approval of manuscript, accountability to accuracy and integrity.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: This material is based upon work supported in part by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Biomedical Laboratory Research and Development, Career Development Award IK2BX001944 (Dr Jennifer L. Long). This work was supported with resources and facilities at the Greater Los Angeles VA Healthcare System.

References

1. 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;61:772-778.
2. Nguyen NP, Moltz CC, Frank C, et al. Dysphagia following chemoradiation for locally advanced head and neck cancer. *Ann Oncol*. 2004;15:383-388.
3. Gillespie MB, Brodsky MB, Day TA, et al. Swallowing-related quality of life after head and neck cancer treatment. *Laryngoscope*. 2004;114:1362-1367.
4. Caudell JJ, Schaner PE, Meredith RF, et al. Factors associated with long-term dysphagia after definitive radiotherapy for locally advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2009;73:410-415.
5. Eisbruch A, Schwartz M, Rasch C, et al. Dysphagia and aspiration after chemoradiotherapy for head-and-neck cancer: which anatomic structures are affected and can they be spared by IMRT?*Int J Radiat Oncol Biol Phys*. 2004;60:1425-1439.
6. Chen AY, Frankowski R, Bishop-Leone J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M. D. Anderson dysphagia inventory. *Arch Otolaryngol Head Neck Surg*. 2001;127:870-876.

7. Eisbruch A, Levendag PC, Feng FY, et al. Can IMRT or brachytherapy reduce dysphagia associated with chemoradiotherapy of head and neck cancer? The Michigan and Rotterdam experiences. *Int J Radiat Oncol Biol Phys*. 2007;69:S40-S42.
8. Garcia-Peris P, Paron L, Velasco C, et al. Long-term prevalence of oropharyngeal dysphagia in head and neck cancer patients: impact on quality of life. *Clin Nutr*. 2007;26:710-717.
9. Pauloski BR, Rademaker AW, Logemann JA, et al. Swallow function and perception of dysphagia in patients with head and neck cancer. *Head Neck*. 2002;24:555-565.
10. Murphy BA, Gilbert J. Dysphagia in head and neck cancer patients treated with radiation: assessment, sequelae, and rehabilitation. *Semin Radiat Oncol*. 2009;19:35-42.
11. Goguen LA, Posner MR, Norris CM, et al. Dysphagia after sequential chemoradiation therapy for advanced head and neck cancer. *Otolaryngol Head Neck Surg*. 2006;134:916-922.
12. Nguyen NP, Frank C, Moltz CC, et al. Aspiration rate following chemoradiation for head and neck cancer: an underreported occurrence. *Radiother Oncol*. 2006;80:302-306.
13. Nguyen NP, Moltz CC, Frank C, et al. Evolution of chronic dysphagia following treatment for head and neck cancer. *Oral Oncol*. 2006;42:374-380.
14. Caudell JJ, Schaner PE, Desmond RA, et al. Dosimetric factors associated with long-term dysphagia after definitive radiotherapy for squamous cell carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys*. 2010;76:403-409.
15. Dirix P, Abbeel S, Vanstraelen B, et al. Dysphagia after chemoradiotherapy for head-and-neck squamous cell carcinoma: dose-effect relationships for the swallowing structures. *Int J Radiat Oncol Biol Phys*. 2009;75:385-392.
16. Kendall KA, McKenzie SW, Leonard RJ, et al. Structural mobility in deglutition after single modality treatment of head and neck carcinomas with radiotherapy. *Head Neck*. 1998;20:720-725.
17. Li B, Li D, Lau DH, et al. Clinical-dosimetric analysis of measures of dysphagia including gastrostomy-tube dependence among head and neck cancer patients treated definitively by intensity-modulated radiotherapy with concurrent chemotherapy. *Radiat Oncol*. 2009;4:52.
18. List MA, D'Antonio LL, Cella DF, et al. The Performance Status Scale for Head and Neck Cancer Patients and the Functional Assessment of Cancer Therapy–Head and Neck Scale: A study of utility and validity. *Cancer*. 1996;77:2294-2301.
19. Humphreys B, Mathog R, Rosen R, et al. Videofluoroscopic and scintigraphic analysis of dysphagia in the head and neck cancer patient. *Laryngoscope*. 1987;97:25-32.
20. Salassa JR. A functional outcome swallowing scale for staging oropharyngeal dysphagia. *Dig Dis*. 1999;17:230-234.
21. Rosenthal DI, Lewin JS, Eisbruch A. Prevention and treatment of dysphagia and aspiration after chemoradiation for head and neck cancer. *J Clin Oncol*. 2006;24:2636-2643.
22. Logemann JA, Rademaker AW, Pauloski BR, et al. Effects of postural change on aspiration in head and neck surgical patients. *Otolaryngol Head Neck Surg*. 1994;110:222-227.
23. Shaker R, Easterling C, Kern M, et al. Rehabilitation of swallowing by exercise in tube-fed patients with pharyngeal dysphagia secondary to abnormal UES opening. *Gastroenterology*. 2002;122:1314-1321.
24. Mendelsohn MS, McConnel FM. Function in the pharyngoesophageal segment. *Laryngoscope*. 1987;97:483-489.
25. Fujiu M, Logemann JA. Effect of a tongue-holding maneuver on posterior pharyngeal wall movement during deglutition. *Am J Speech Lang Pathol*. 1996;5:23-30.

Oral Oncology 57 (2016) 21–26

Contents lists available at ScienceDirect

Oral Oncology

journal homepage: www.elsevier.com/locate/oraloncology

Severe late dysphagia and cause of death after concurrent chemoradiation for larynx cancer in patients eligible for RTOG 91-11 ☆

Matthew C. Ward [a,*], David J. Adelstein [b], Priyanka Bhateja [c], Tobenna I. Nwizu [b], Joseph Scharpf [d], Narcissa Houston [a], Eric D. Lamarre [d], Robert Lorenz [d], Brian B. Burkey [d], John F. Greskovich [a], Shlomo A. Koyfman [a]

[a] Department of Radiation Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, United States
[b] Department of Hematology and Medical Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, United States
[c] Department of Hematology & Oncology, University Hospitals Seidman Cancer Center, Case Western Reserve University, Cleveland, OH, United States
[d] Department of Otolaryngology, Head and Neck Institute, Cleveland Clinic, Cleveland, OH, United States

ARTICLE INFO

Article history:
Received 16 February 2016
Received in revised form 14 March 2016
Accepted 15 March 2016
Available online 1 April 2016

Keywords:
Late toxicity
Larynx cancer
Chemoradiation
Dysphagia
Larynx preservation

SUMMARY

Purpose: The long-term results of RTOG 91-11 suggested increased deaths not attributed to larynx cancer after concomitant chemoradiotherapy (CRT) despite no apparent increase in late effects. Because the timing of events was not reported by RTOG 91-11, one possibility is that severe late dysphagia (SLD) develops beyond five years and leads to unreported treatment-related deaths. Here we explore the timing of SLD after CRT.

Methods: Patients who would have met eligibility criteria for RTOG 91-11 and were treated with CRT between 1993 and 2013 were identified. Events occurring beyond 3 months after treatment and suggestive of SLD were recorded including esophageal stricture dilations, hospital admissions for aspiration pneumonia or feeding-tube insertion. Feeding-tube dependence beyond one year was also considered SLD. The cumulative incidence of SLD and its components was quantified using Gray's competing risk analysis with recurrence or death considered competing risks.

Results: Eighty-four patients were included with a median follow-up of 43 months. The 5-year overall survival was 70% (95% CI 58–80%). No death was directly a result of treatment-induced late dysphagia. The 5-year incidence of SLD was 26.5%. While 15 of 18 (83%) first stricture dilations occurred within 5 years after CRT, 3 of 5 (60%) aspiration admissions and 5 of 8 late feeding tube insertions occurred beyond five years from CRT.

Conclusions: SLD is common after CRT for larynx cancer and can occur beyond 5 years from the end of treatment, emphasizing the importance of survivorship follow-up. Despite the incidence of SLD, death related to dysphagia is uncommon.

Introduction

After the landmark Veterans Affairs Laryngeal Cancer Study Group (VALSG) larynx-preservation trial reported in 1991, induction chemotherapy and radiation became a viable organ-preservation strategy for the treatment of locoregionally-advanced larynx cancer [1]. Building on this trial, Radiation Therapy Oncology Group (RTOG) 91-11 compared three strategies for non-operative organ-preservation treatment: radiotherapy alone, radiotherapy with concurrent cisplatin and the VALSCG regimen of induction cisplatin and 5-fluorouracil followed by radiotherapy. The long-term report of this trial confirmed the continued efficacy of the two combined-modality arms and encouraged the adoption of larynx-preservation strategies in daily practice [2].

Although the concurrent cisplatin arm of RTOG 91-11 demonstrated a clear advantage with regards to locoregional control, this did not translate to an overall survival benefit when compared to the other arms. In the final update, there was a trend toward inferior overall survival in the concurrent arm when compared to induction chemotherapy (HR 1.25, $p = 0.08$) with an increase in non-cancer related deaths with concurrent chemotherapy but no

☆ Presented at the 2015 ASTRO Annual Meeting, San Antonio, TX, United States.

* Corresponding author at: Department of Radiation Oncology, Taussig Cancer Institute, Cleveland Clinic, 9500 Euclid Avenue, T28, Cleveland, OH 44195, United States. Tel.: +1 (216) 444 5571; fax: +1 (216) 445 1068.
E-mail address: wardm3@ccf.org (M.C. Ward).

http://dx.doi.org/10.1016/j.oraloncology.2016.03.014
1368-8375/

difference in late toxicity [2]. This result, combined with data from the National Cancer Database suggesting a downward trend in overall survival as chemoradiation was adopted has led some to consider the possibility that long-term toxicity unaccounted for on trial led late toxic deaths [3]. Late toxicity is difficult to measure and was reported as a maximum grade in RTOG 91-11 without comment on the timing or precise nature of the toxic events. The purpose of this report is to perform a detailed analysis of the cumulative incidence of late dysphagia observed after concurrent chemoradiation for larynx cancer in an attempt to identify its contribution to late mortality.

Methods

Patients

From an IRB-approved head and neck database all patients treated with definitive concurrent chemoradiotherapy at our institution between 1993 and 2013 who met eligibility criteria for RTOG 91-11 were retrospectively identified. The criteria included medically-fit patients with previously untreated squamous cell carcinoma of the glottis or supraglottis, stage III-IVB by AJCC version 3 criteria [4], the version in place during 91-11 accrual. Patients with disease classified as T4 due to soft-tissue extension beyond the larynx were excluded from RTOG 91-11 and hence the current study as well. Patients with T4 disease with $\leqslant$1 cm invasion into the base of tongue were included.

Treatment

All patients were treated with definitive concurrent chemoradiotherapy. Chemotherapy consisted of multi-agent cisplatin and 5-fluorouracil (5-FU) in the earlier years and single-agent cisplatin in the later years. Combination chemotherapy was given in two cycles of a 4-day continuous infusion of cisplatin 20 mg/m^2/day and fluorouracil 1000 mg/m^2/day. Single-agent chemotherapy was given as a bolus of cisplatin 100 mg/m^2 on weeks 1, 4 and 7 of radiotherapy. All patients were programmed to receive 70 Gy or greater of radiotherapy. In the early years twice-daily (BID) radiotherapy was often delivered to doses of 72–74.4 Gy for tumors of T3 or greater extent along with concurrent cisplatin and fluorouracil [5,6]. Prior to 2009, radiotherapy was delivered using standard conventional techniques including opposed lateral fields matched to an anteroposterior supraclavicular field. In 2009, intensity-modulated radiotherapy (IMRT) became the radiotherapy technique of choice at our institution. After completion of therapy an adjuvant neck dissection was performed for select patients within three months for patients with residual adenopathy at the discretion of the treating surgeon. Patients were subsequently followed by our multidisciplinary team as per National Comprehensive Cancer Network (NCCN) guidelines [7].

Severe late dysphagia, other toxicity and the cause of death

The goal of this study was to provide a detailed description of the incidence and timing of severe late dysphagia (SLD) in this patient cohort. Therefore, individual multidisciplinary follow-up encounters documented in the electronic medical record were investigated for signs of severe dysphagia and other late toxicity. Ninety days after the completion of radiotherapy was used as the definition of "late" events as per RTOG 91-11. Severe late dysphagia was defined as the occurrence of one or more of the following events 90 days after the completion of treatment or beyond: pharyngeal stricture dilation, admission to the hospital with a diagnosis of aspiration pneumonia or placement of a new feeding tube. Feeding-tube dependence greater than one year after treatment was also included in the definition of severe late dysphagia. This composite endpoint is similar to a previous cooperative group time-to-event analysis of physician-reported severe late dysphagia [8]. Furthermore, these events all correlate with the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 and RTOG scoring criteria used by RTOG 91-11 but offer a more specific definition.

In order to evaluate the incidence of death associated with severe late dysphagia and subsequent aspiration pneumonia, special attention was paid to the cause of death. Patients who expired after metastatic or locoregional failure that could not be salvaged were counted as a primary cancer death regardless of any dysphagia events occurring after failure. For patients who expired outside of the hospital system, the death certificate was obtained if possible through the Ohio death registry.

Statistical analysis

This was a single-cohort retrospective study. Overall survival and the cumulative incidence of salvage laryngectomy, locoregional failure and distant metastatic failure were assessed using the Kaplan-Meier method measuring time from the date of diagnosis to the date of event or last oncologic follow-up. The cumulative incidence of severe late dysphagia and its four components (esophageal stricture requiring dilation, placement of a new feeding tube, hospital admission for aspiration pneumonia or feeding tube dependence beyond one year) was calculated using Gray's competing risk analysis treating death or disease recurrence as competing events [9]. Time was measured from the end of radiotherapy to the first SLD event, with patients censored from the analysis at the time of the last multidisciplinary follow-up. Although late toxicity was the primary outcome of interest in this study, patients with short follow-up or who did not complete therapy were not excluded in order to capture deaths that may have occurred from acute toxicity. The cause of death was compared with the data reported by RTOG 91-11 on a per-patient basis using the Pearson Chi-square test. To investigate for factors associated with severe late dysphagia, relevant patient, tumor and treatment variables were entered into a competing risk regression as described by Fine and Gray with severe late dysphagia as the endpoint of interest [10]. All hypothesis testing was performed with significance assumed at the 0.05 level. Statistical analyses were performed using JMP Version 10 software (SAS Institute, Cary, NC) and competing risk analyses were performed using R statistical software (R Foundation, Vienna, Austria).

Results

Patient characteristics

Between 1993 and 2013, 84 patients were identified who met the inclusion criteria. Patient, disease and treatment characteristics are represented in Table 1. The majority of tumors were located in the supraglottis (71%), classified as T3 (78%) and were node-positive (63%). Combination cisplatin with 5-FU was the most common concurrent chemotherapy regimen (76%).

The median follow-up for survivors was 54 months (range 8.8–180). Twenty-eight patients (33%) were followed beyond 5 years and 13 (15%) were followed beyond 9 years. A median of 14 individual oncologic follow-up visits were recorded (range 0–44), totaling 1240 individual oncologic visits for the entire cohort. Fifty patients (60%) have experienced either severe late dysphagia or a competing event (failure or death) and further follow-up of these patients will not alter the incidence of severe

late dysphagia. Sixty-two patients (74%) have been followed beyond 5 years, or have experienced severe late dysphagia, death or disease recurrence.

The actuarial 5-year rate of locoregional failure was 20% (95% CI 12–31%), distant metastases 16% (95% CI 9–27%) and overall survival 70% (95% CI 58–80%). Fifty-three of 84 patients (63%) have either experienced toxicity, recurrence, death or have been followed beyond 9 years. Ten patients underwent a salvage laryngectomy after recurrence for a 5-year cumulative incidence of 15% (95% CI 8–25%). Eight of the ten salvage laryngectomies were successful and the patients were alive and without evidence of disease at the time of their last follow-up.

Late toxicity

Twenty-two patients experienced severe late dysphagia. The cumulative incidence of severe late dysphagia and its components are presented in Table 2 and Fig. 1. The overall cumulative incidence of severe late dysphagia at 5 years was 26.5% (95% CI 15.2–37.8%). Sixty-eight percent of patients required feeding tube support on-treatment (Fig. 2) but at one year after radiotherapy only 1.8% remained feeding tube dependent (95% CI 0.2–11.2%). No patient required a laryngectomy for toxicity. Stricture dilation within the first year after radiotherapy was the most common severe late dysphagia event. Of the 18 patients who required stricture dilation, 12 required multiple dilations. The median number of dilations for those who underwent dilation was 2 (range 1–8). Of the 22 patients who experienced severe late dysphagia, four (18%) experienced their first event beyond 5 years.

Cause of death

At the time of last follow-up, 31 patients had died (37%). Table 3 presents the distributions of deaths observed in comparison to the distribution seen on RTOG 91-11. The index head and neck cancer was the most common cause of death in both cohorts. Two patients in our series expired as a result of neutropenic fever while on-treatment. The cause of death could not be determined in 6 patients (19% deaths) who expired out of state or out of the country. The distribution is comparable to RTOG 91-11 (Pearson Chi-square $p = 0.454$). Of note, the specific number of deaths observed on RTOG 91-11 due to aspiration pneumonia is not specified, but three late RTOG grade 5 toxicities were observed within the concurrent chemotherapy arm: one pharynx/esophagus, one larynx and one "other" event, leading to a maximum crude rate of 2% or less.

Factors associated with severe late dysphagia

To investigate clinical factors which may be associated with severe late dysphagia, a Fine-Gray competing risk regression was performed. Results are presented in Table 4. Among all patient, tumor and treatment factors entered into the univariate regression, twice-daily radiotherapy fractionation was the only statistically-significant association with increased severe late dysphagia (HR 2.51, 95% CI 1.10–5.72, $p = 0.028$). The use of single agent cisplatin rather than multiagent chemotherapy, or IMRT as opposed to 3D planning were not associated with a reduction in severe late dysphagia on univariate analysis. A multivariate analysis was not performed given the univariate results.

Table 1
Study demographics ($n = 84$).

Age at first diagnosis	Median (Range)	60 (43–76)
Race	Caucasian	72 (87%)
	African American	10 (12%)
	Hispanic	1 (1%)
Gender	Male	60 (71%)
	Female	24 (29%)
Smoking history	Never smoker	3 (4%)
	Former smoker (Quit >3 months)	38 (45%)
	Current smoker	33 (39%)
	Use during or after radiation	9 (11%)
	Unknown smoking history	1 (1%)
Karnofsky score	Median (Range)	90 (80–90)
Tobacco pack-years	Median (Range)	40 (0–200)
Heavy alcohol consumption?	No	68 (81%)
	Yes	16 (19%)
Larynx subsite	Supraglottic	60 (71%)
	Glottic	24 (29%)
T Classification	2	13 (16%)
	3	65 (77%)
	4	6 (7%)
N Classification	0	31 (37%)
	1	18 (21%)
	2a	3 (4%)
	2b	12 (14%)
	2c	19 (23%)
	3	1 (1%)
Grouped stage	III	45 (54%)
	IV	39 (46%)
Lymph node dissection	No	74 (88%)
	Yes	10 (12%)
Chemotherapy	Cisplatin (CP)	17 (20%)
	CP/5FU	63 (75%)
	CP/5FU with Gefitinib	1 (1%)
	Other multiagent	3 (4%)
Radiation type	3D-RT	63 (75%)
	IMRT	21 (25%)
Altered fractionation	Daily	46 (55%)
	BID	31 (37%)
	6-Fractions per week	7 (8%)
Feeding tube placed during treatment	No	27 (33%)
	Yes	56 (68%)
Dose of RT	Median (Range)	72 Gy (62.4–74.4 Gy)
Number of fractions	Median (Range)	36 (32–62)
Duration of RT (Days)	Median (Range)	46 (29–64)
Months of follow-up (Survivors)	Median (Range)	53 (8.8–180)
Number of follow-up visits	Median (Range)	14 (0–44)

Table 2
Cumulative incidence (CI) of severe late dysphagia and its components (cumulative incidence and 95% confidence intervals are listed). Twenty-two patients experienced severe late dysphagia and 4 experienced the first event beyond 5 years.

CI of stricture dilation at 5 years	17.2% (8.9–25.6%)
CI of late feeding tube placed at 5 years	3.8% (0–8.0%)
CI of aspiration admission at 5 years	2.8% (0–6.9%)
Feeding tube dependent at 1 year	1.8% (0.2–11.2%)
CI of severe late dysphagia at 5 years	26.5% (15.2–37.8%)

Discussion

In this study patients with larynx cancer who met the inclusion criteria of RTOG 91-11 were retrospectively identified and a detailed time-to-event analysis of severe late dysphagia was performed while accounting for the competing risks of recurrence or death. This is the first analysis to apply a competing risk analysis

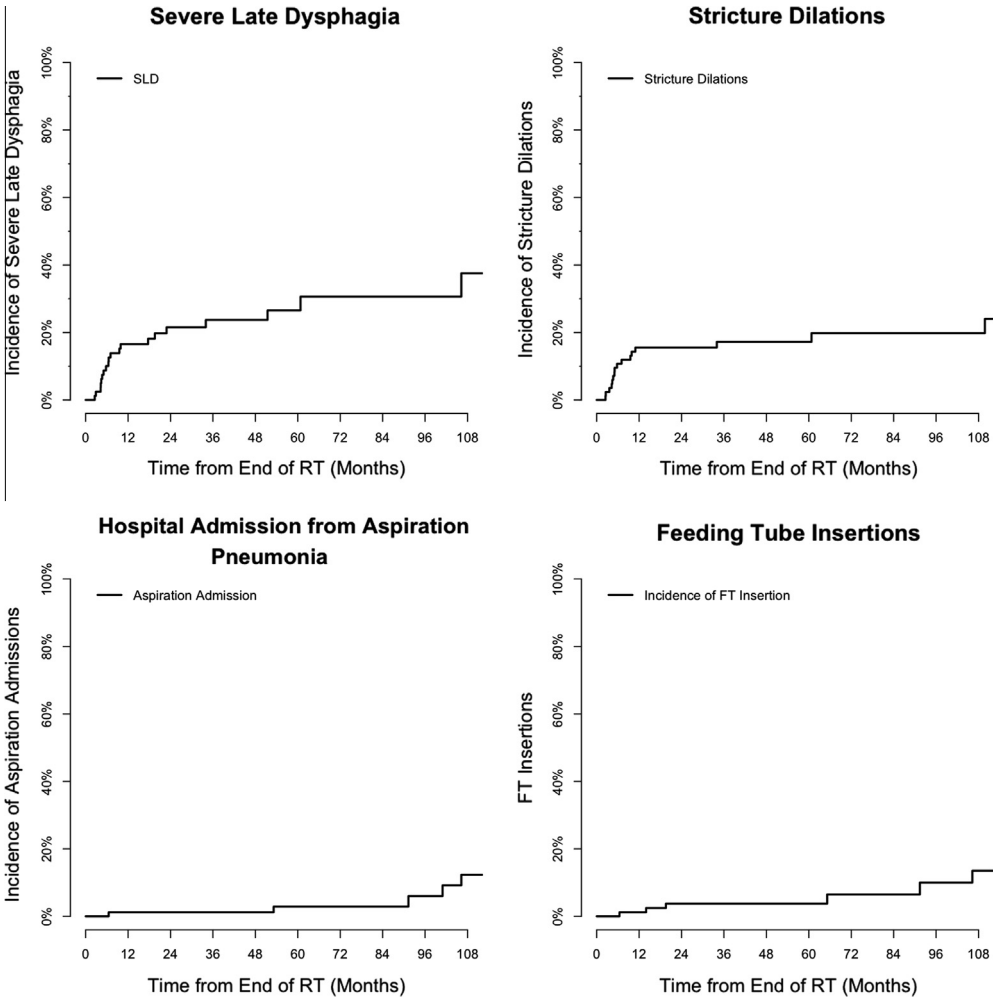


Fig. 1. Cumulative incidence of severe late dysphagia and its components.

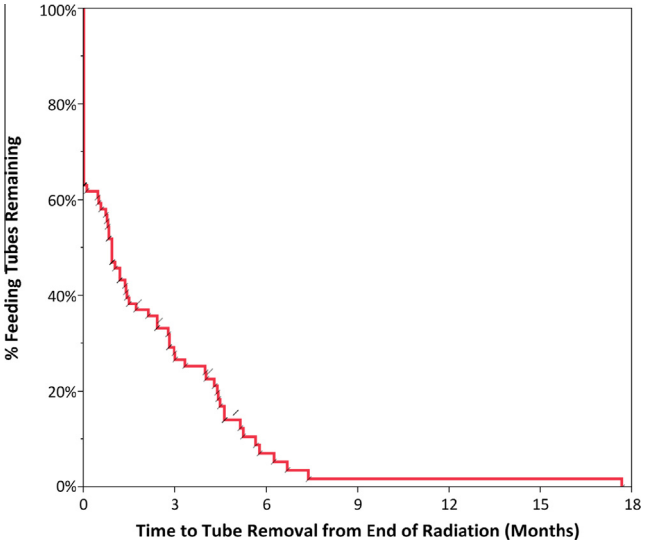


Fig. 2. Percent of feeding tubes remaining after the end of radiotherapy.

to this population. The results suggest that although the incidence of severe late dysphagia was significant (26.5% at 5 years), no deaths directly related to severe late dysphagia were observed.

Table 3
Cause of death.

Cause of death	Current study	RTOG 91-11
H&N cancer	13 (42%)	38 (29%)
Unknown	6 (19%)	23 (17%)
Co-morbid illness	5 (16%)	42 (32%)
Other non-H&N cancer	5 (16%)	18 (14%)
Acute toxicity	2 (7%)	9 (7%)
Aspiration pneumonia	0 (0%)	<2% (Not specified)

Chi-square $p = 0.454$.

Factors driving the trending decrease in overall survival seen in larynx cancer since the adoption of chemoradiation remain unclear, but in our experience deaths seem not to be clearly related to late toxicity. Other hypotheses including the treatment of patients with T4 disease with significant soft-tissue invasion or poor compliance with post-treatment follow-up protocols seem possible. The possibility of very late severe dysphagia beginning beyond five years is not excluded by this analysis as multiple patients experienced their first severe late dysphagia event beyond five years. Although the risk of severe late dysphagia was highest within the first two years, this risk did remain for years to come. This entity of very late dysphagia has been previously described by Hutcheson et al, who also described a component of dysphagia originating from cranial nerve dysfunction [11]. The current analysis may not have captured cranial nerve dysfunction if it did not require feeding tube placement or hospital admission. Regardless,

Table 4
Univariate Fine-Gray competing risk regression for factors associated with severe late dysphagia.

		HR	95% CI	*p*
Age		1.03	0.984–1.09	0.18
Smoking history	Current/After RT vs Never/Former	0.814	0.233–2.84	0.75
Pack-years smoking		0.995	0.985–1.00	0.32
Disease subsite	Supraglottic vs Glottic or NOS	1.35	0.526–3.46	0.53
T stage	T3–4 vs T2	1.82	0.417–7.93	0.43
N stage	N2a-3 vs N0-1	1.35	0.598–3.03	0.47
Grouped stage	Stage IV vs Stage III	1.71	0.745–3.92	0.21
Year treated		0.99	0.929–1.05	0.76
Neck dissection	Yes vs No	1.32	0.525–3.30	0.56
Chemo type	Multiagent vs Single agent	5.78	0.805–41.5	0.08
RT type	IMRT vs 3D	0.353	0.084–1.47	0.15
RT dose		0.937	0.743–1.18	0.58
Altered fractionation	BID vs QD	2.51	1.10–5.72	**0.028**
Feeding tube type	NG vs None	2.00	0.623–6.45	0.24
	PEG vs None	1.94	0.606–6.21	0.26
	NG vs PEG	1.03	0.407–2.62	0.94

the two reports together emphasize the point that patients should be closely followed by dedicated head and neck caregivers for the duration of their lifetime to screen for new onset severe late dysphagia.

Our study differs from RTOG 91-11 not only in its retrospective nature but also in the radiotherapy fractionation and the intensity of chemotherapy. Most patients in the current study received combination cisplatin and 5-FU concurrent and conventional radiotherapy and 37% received BID fractionation. Regardless, oncologic outcomes in the current study appear similar or improved to those described in RTOG 91-11 in terms of 5-year overall survival (58% 91–11 vs. 70% current study), locoregional control (54.8% 91–11 vs. 80% current study) and distant control (85.3% 91–11 vs. 84% current study). A low rate of competing events combined with a more intense chemotherapy regimen may have increased the incidence of severe late dysphagia in the current study compared to the current era of treatment with IMRT and single-agent chemotherapy although power to detect this on univariate analysis is limited.

This study highlights the challenge of measurement of late toxicity following radiotherapy in the management of head and neck cancer. The insights RTOG 91-11 provided into the incidence, timing and nature of late toxic events were limited by the method of recording a maximum grade late toxicity leading to the speculation that the increased late mortality in the concurrent chemoradiotherapy group reflected an increase in unmeasured or unrecorded severe late dysphagia. Our study provides a more detailed analysis accounting for competing risks of recurrence or death and suggests that this explanation is incorrect. The analysis method we used is therefore a strength of the current report in comparison to the methods used by most studies investigating physician-reported severe toxicity.

The univariate regression demonstrating an association between twice-daily (BID) fractionation and severe late dysphagia (Table 4) is likely related to practice patterns at our institution rather than an independent contribution of fractionation. Twice-daily radiotherapy with dose-escalation to 74.4 Gy was a regimen used during the earliest years of this study for patients with locoregionally advanced disease in attempt to improve oncologic outcomes [5]. Nearly all of these patients received combination cisplatin and 5-FU along with conventional radiotherapy. Therefore, the result on univariate analysis may be a surrogate for older treatment regimens rather than a true independent contribution of fractionation on severe late dysphagia. This is generally supported by the long-term results of RTOG 90-03 which did not demonstrate a clear difference in severe late toxicity within the hyperfractionated arm compared to standard fractionation [12].

The limitations of this study include its retrospective nature, a modest sample-size, patients with an unknown cause of death and the heterogeneity in radiotherapy and chemotherapy regimens used. While the study-size is modest, in comparison the concurrent chemotherapy arm of RTOG 91-11 analyzed 174 patients and experienced more attrition due to locoregional recurrence or death. The follow-up in the current study, although less than RTOG 91-11, appears sufficient to capture the majority of severe late dysphagia events according to the cumulative incidence curves presented (Fig. 2) but may be insufficient to capture very late dysphagia originating beyond 5 years and further study is required. Additionally, six patients in our report died but the cause of death could not be determined. This is a common challenge with elucidating a cause of death and a similar rate was observed on RTOG 91-11 despite the prospective nature of the trial. Although it is impossible to say for sure, there is no reason to suggest that these patients succumbed to severe late dysphagia at a higher rate than the rest of the cohort. In addition, the heterogeneity in chemotherapy and radiotherapy techniques was inherent to changes in practice patterns over the years. Although including older techniques limits the application to modern patients, the low incidence of deaths related to severe late dysphagia remains relevant in the modern era and enforces that good candidates for larynx-preservation should continue to be offered chemoradiotherapy. With modern IMRT techniques and chemotherapy regimens, the incidence of SLD is likely to continue to decrease.

Conclusion

In conclusion, after a detailed time-to-event analysis accounting for the competing risks of recurrence or death in patients otherwise eligible for larynx-preservation strategies, we did not identify a contribution of severe late dysphagia to late mortality. Larynx preservation should continue to be offered to patients who meet criteria for RTOG 91-11. Because severe late dysphagia can occur beyond five years, patients should be followed closely by a dedicated head and neck caregiver to monitor for recurrence, second primary or toxicity for the remainder of their lifetime. Future clinical trials should carefully track the incidence of late toxicity to allow for clarity in elucidating the timing and incidence of radiotherapy-induced dysphagia.

Conflicts of interest/Financial disclosures

None declared.

References

[1] Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. The department of veterans affairs laryngeal cancer study group. N Engl J Med. 1991;324(24):1685–90.

[2] Forastiere AA, Zhang Q, Weber RS, et al. Long-term results of RTOG 91-11: a comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer. J Clin Oncol 2013;31(7):845–52.

[3] Hoffman HT, Porter K, Karnell LH, et al. Laryngeal cancer in the United States: changes in demographics, patterns of care, and survival. Laryngoscope 2006;116(9 Pt 2 Suppl 111):1–13.

[4] Manual for Staging of Cancer. 3rd ed. Philadelphia: J.B. Lippincott Company; 1988.

[5] Horiot JC, Le Fur R, N'Guyen T, et al. Hyperfractionation versus conventional fractionation in oropharyngeal carcinoma: final analysis of a randomized trial of the EORTC cooperative group of radiotherapy. Radiother Oncol 1992;25(4):231–41.

[6] Fu KK, Pajak TF, Trotti A, et al. A radiation therapy oncology group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of RTOG 9003. Int J Radiat Oncol Biol Phys 2000;48(1):7–16.

[7] Pfister DG, Spencer S, Brizel DM, et al. Head and neck cancers, Version 2.2014. Clinical practice guidelines in oncology. J Natl Compr Canc Netw 2014;12(10):1454–87.

[8] Machtay M, Moughan J, Trotti A, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. J Clin Oncol 2008;26(21):3582–9.

[9] Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 1988;16(3):1141–54.

[10] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;1999(94):496–509.

[11] Hutcheson KA, Lewin JS, Barringer DA, et al. Late dysphagia after radiotherapy-based treatment of head and neck cancer. Cancer 2012;118(23):5793–9.

[12] Beitler JJ, Zhang Q, Fu KK, et al. Final results of local-regional control and late toxicity of RTOG 9003: a randomized trial of altered fractionation radiation for locally advanced head and neck cancer. Int J Radiat Oncol Biol Phys 2014;89(1):13–20.

World J Surg (2010) 34:28–35
DOI 10.1007/s00268-009-0303-0

An Observational Trial for Papillary Thyroid Microcarcinoma in Japanese Patients

Yasuhiro Ito · Akira Miyauchi · Hiroyuki Inoue · Mitsuhiro Fukushima · Minoru Kihara · Takuya Higashiyama · Chisato Tomoda · Yuuki Takamura · Kaoru Kobayashi · Akihiro Miya

Published online: 30 November 2009

Abstract
Background The recent development and spread of ultrasonography and ultrasonography-guided fine needle aspiration biopsy (FNAB) has facilitated the detection of small papillary microcarcinomas of the thyroid measuring 1 cm or less (PMC). The marked difference in prevalence between clinical thyroid carcinoma and PMC detected on mass screening prompted us to observe PMC unless the lesion shows unfavorable features, such as location adjacent to the trachea or on the dorsal surface of the thyroid possibly invading the recurrent laryngeal nerve, clinically apparent nodal metastasis, or high-grade malignancy on FNAB findings. In the present study we report comparison of the outcomes of 340 patients with PMC who underwent observation and the prognosis of 1,055 patients who underwent immediate surgery without observation.
Methods Between 1993 and 2004, 340 patients underwent observation and 1,055 underwent surgical treatment without observation. These 1,395 patients were enrolled in the present study. Observation periods ranged from 18 to 187 months (average 74 months).
Results The proportions of patients whose PMC showed enlargement by 3 mm or more were 6.4 and 15.9% on 5-year and 10-year follow-up, respectively. Novel nodal metastasis was detected in 1.4% at 5 years and 3.4% at 10 years. There were no factors related to patient background or clinical features linked to either tumor enlargement or the novel appearance of nodal metastasis. After observation 109 of the 340 patients underwent surgical treatment for various reasons, and none of those patients showed carcinoma recurrence. In patients who underwent immediate surgical treatment, clinically apparent lateral node metastasis (N1b) and male gender were recognized as independent prognostic factors of disease-free survival.
Conclusions Papillary microcarcinomas that are not associated with unfavorable features can be candidates for observation regardless of patient background and clinical features. If there are subsequent signs of progression, such as tumor enlargement and novel nodal metastasis, it would not be too late to perform surgical treatment. Even though the primary tumor is small, careful surgical treatment including therapeutic modified neck dissection is necessary for N1b PMC patients.

Introduction

Papillary carcinoma is the most common malignancy originating from the thyroid. Usually, papillary carcinoma is indolent and grows slowly, although cases having certain biological characteristics, such as clinically apparent node metastasis in the lateral compartment (N1b in the International Union Against Cancer [UICC] tumor node metastasis [TNM] classification [1]) and massive extrathyroid extension (pT4[1]) are progressive [2–4]. Papillary carcinoma measuring 1.0 cm or less is defined as papillary microcarcinoma (PMC) by the World Health Organisation (WHO) classification. Formerly, detection of

Y. Ito (✉) · A. Miyauchi · H. Inoue · M. Fukushima · M. Kihara · T. Higashiyama · C. Tomoda · Y. Takamura · K. Kobayashi · A. Miya
Department of Surgery, Kuma Hospital, 8-2-35 Shimoyamate-dori, Chuo-ku, Kobe 650-0011, Japan
e-mail: ito01@kuma-h.or.jp

PMC without clinically apparent lymph nodal and/or distant metastasis was exceedingly difficult because PMC is rarely palpable. Therefore, in the past, PMC could be classified into three categories based on the circumstances of detection: (1) latent PMC, which are detected in autopsy specimens; (2) occult PMC, which are discovered as the origin of lymph node and/or distant metastasis; and (3) incidental PMC, which are detected on pathological examination of surgical specimens resected for other diseases.

Recently, however, screening of the thyroid and carotid artery by ultraonography has facilitated the detection of small thyroid nodules measuring a minimum of 3 mm. These PMC can be diagnosed on cytologic examination of specimens obtained by ultrasonography-guided fine-needle aspiration biopsy (FNAB) [5]. Takebe et al., reported the detection of papillary carcinomas in 3.5% of otherwise healthy women aged 30 years or older by ultrasonography performed as a screening for breast and thyroid cancer and ultrasonography-guided FNAB, noting that 75% of these lesions measured 1.5 cm or smaller [6]. This incidence was not discrepant with that of latent PMC measuring 3.0–9.9 mm in autopsy specimens, which have been reported to range from 0.5 to 5.2% [7–9]. In contrast, however, the prevalence of clinical thyroid papillary carcinoma was 1.9–11.7 per 100,000 females of all ages [3, 10], which is about 1,000 times lower than that of PMC detected on ultrasonography. The marked difference between these prevalences suggests that PMC rarely grow and become clinically apparent, prompting the question of whether immediate surgery is mandatory for all PMC detected on mass screening, although PMC is also known to show multicentricity in 15–44% of lesions and regional lymph node metastasis in 14–64% of lesions [11–20].

Based on the above findings, we hypothesized that most PMC do not require immediate surgical treatment and that affected patients can be followed by observation in the outpatient clinic. In 1993, we initiated an observational trial of PMC. When we diagnosed nodules measuring 1 cm or less as papillary carcinoma by ultrasonography-guided FNAB, we propose two therapeutic alternatives, observation without surgery or surgical treatment, and we allowed the patient to choose. In 2003, we published our first report of the outcome of 162 patients with PMC, which indicated that over 70% of tumors did not change from their initial size and that novel lymph node metastasis appeared in only 1.2% of patients during follow-up (average follow-up was 47 months [range: 18–113 months]) [21]. In a review article published in 2007, we demonstrated that only 6.7% of tumors show enlargement by 3 mm or more during a 5-year follow-up [22]. In the present study, we present our most recent data from observation of PMC patients as a follow-up report.

Patients and methods

Diagnosis of PMC and recommendation of observation

Diagnosis of PMC and recommendation of observation were performed as described in our previous reports [20–23]. Briefly, when patients are diagnosed with nodules measuring 1 cm or less that showed as papillary carcinoma on ultrasonography-guided FNAB, we presented two therapy options: observation and surgical treatment. However, when the PMC demonstrated such unfavorable features (1) location adjacent to the trachea; (2) location on the dorsal surface of the thyroid lobe, possibly invading the recurrent laryngeal nerve; (3) FNAB findings suggesting high-grade malignancy; (4) presence of regional node metastasis; and/or (5) presence of signs of progression during follow-up, we recommend surgical treatment without observation. Regional lymph node metastasis was diagnosed on ultrasonography based on criteria described elsewhere [20, 21]. When patients choose observation, PMC is followed by ultrasonography once or twice per year to determine whether the tumor size has changed or lymph node metastasis newly appears. Between 1993 and 2004, 340 patients were diagnosed with PMC by ultrasonography-guided FNAB and underwent observation for 18 months or longer. These patients were enrolled in this study as the observation group. They consisted of 314 females and 26 males and their follow-up periods ranged from 18 to 187 months (average: 74 months). Twenty-seven patients underwent thyroid stimulating hormone (TSH) suppression treatment to the low normal or less than normal range by L-thyroxine based on the discretion of attending physicians. We routinely measured serum thyroglobulin at every follow-up. Antithyroid antibodies were positive for 93 patients. For the purposes of this study, tumor enlargement was defined by an increase in tumor size of 3 mm or more compared with the size at initiation of observation, but only when there was no change or a further increase on the next examination. We established this parameter because, in our experience, + 2 mm has been recognized as an observer variation. To date, 109 patients (102 females and 7 males) (32.1%) have undergone surgical treatment for various reasons. Intervals from initiation of observation to surgery ranged from 18 to 175 months (average: 51 months). Postoperative follow-up has included ultrasonography and chest roentgenography or CT scan more than once per year. Postoperative follow-up averaged 76 months (range: 1–198 months).

Immediate surgical treatment group

Between 1993 and 2004, 1,055 patients underwent surgery for PMC without follow-up. These patients were enrolled

in this study as the immediate surgical treatment group. Two patients with distant metastasis at diagnosis were excluded from the series. This patient group then consisted of 1,059 patients, 964 females and 95 males whose age ranged from 15 to 84 years (average: 52.0 years). After surgery, these patients were followed in our outpatient clinic in the same manner as the 109 patients of the observation group. Follow-up averaged 76 months (range: 1–183: months). Radioiodine whole body scan using 3–13 mCi radioiodine was performed for 52 patients, none of whom showed abnormal uptake except in the thyroid bed.

Statistical analyses

The Kaplan–Meier method and log-rank test were adopted to analyze time-dependent variables. The Cox regression model was also used for multivariate analysis. All analyses were performed using StatView-J 5.0. A p value less than 0.05 was regarded as significant.

Results

Outcome of PMC patients in observation group

Between 1993 and 2004, 340 patients underwent observation for periods ranging from 18 to 187 months (average 74 months). As indicated above in "Patients and methods" section, we made a judgment of tumor enlargement when the size increased by 3 mm or more compared to the size at initiation of observation and the increased size did not change or showed a further increase on the next examination. To date, PMC of 31 patients (9.1%) showed enlargement based on our criteria. Figure 1 shows the proportion of patients showing enlargement of PMC. On 5-year and 10-year follow-up, 6.4 and 15.9% of patients showed enlargement, respectively.

Seventeen patients (5.0%) were diagnosed as having familial papillary carcinoma, because they had one or more first-degree relatives who had undergone surgical treatment for papillary or follicular carcinoma in our hospital or other hospitals [24]. However, enlargement was not related to whether patients had familial or non-familial papillary carcinoma (Fig. 2a). We investigated the relationship between size enlargement and other various backgrounds of patients and clinical features such as gender, age, tumor size at diagnosis, multicentricity, and TSH suppression, but none of these parameters were related to enlargement (Fig. 2b–f). Although carcinomas of patients aged 45 years or younger tended to enlarge, the difference was not significant ($p = 0.0624$). Furthermore, tumor enlargement was not linked to change in serum throglobulin level or the presence of antithyroid antibodies (data not shown).

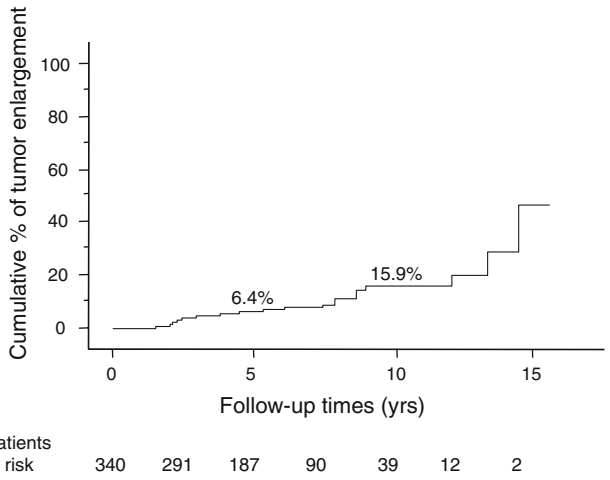


Fig. 1 Proportion of patients whose papillary microcarcinoma (PMC) showed enlargement by 3 mm or more

To date, 7 patients (2.1%) have shown the novel appearance of lymph node metastasis. All new metastases appeared in the lateral compartment ipsilateral to the primary lesion and were diagnosed on ultrasonography, ultrasonography-guided FNAB, and throglobulin measurement of wash-out from the needle used for FNAB [25]. As shown in Fig. 3, the proportion of patients showing novel node metastasis was 1.4 and 3.4% at 5-year and 10-year follow-up, respectively. None of the clinicopathological features described here were related to the novel appearance of lymph node metastasis (data not shown).

Of 340 patients, 109 (32%) underwent surgical treatment after observation. Table 1 summarizes the reasons for surgical treatment for 109 patients. The leading reason for surgery was tumor enlargement (32 patients). However, in 17 of those 32 patients, the tumor enlargement leading to operation was not based on the criteria of this study. Conversely, as indicated above, 31 patients were judged as showing carcinoma enlargement according to our criteria and 18 of these patients underwent surgery. Two of the 18 also showed novel appearance of lymph node metastasis. The remaining 13 were continuously followed without surgery, and tumor size was noted to decrease in 7 of these 13 patients. Novel appearance of lymph node metastasis was observed in 7 patients. Surgical treatment was recommended for these patients and 5 underwent surgery. However, the remaining 2 refused surgery and discontinued outpatient consultations. Seventeen patients underwent surgical treatment because of the location of tumor at the dorsal surface after observation. Four of these patients had been followed without diagnosis of malignancy and were recommended for immediate surgical treatment after the diagnosis of PMC. The remaining 13 were diagnosed as having PMC from the beginning but surgical treatment was recommended after a change in the policy of the attending

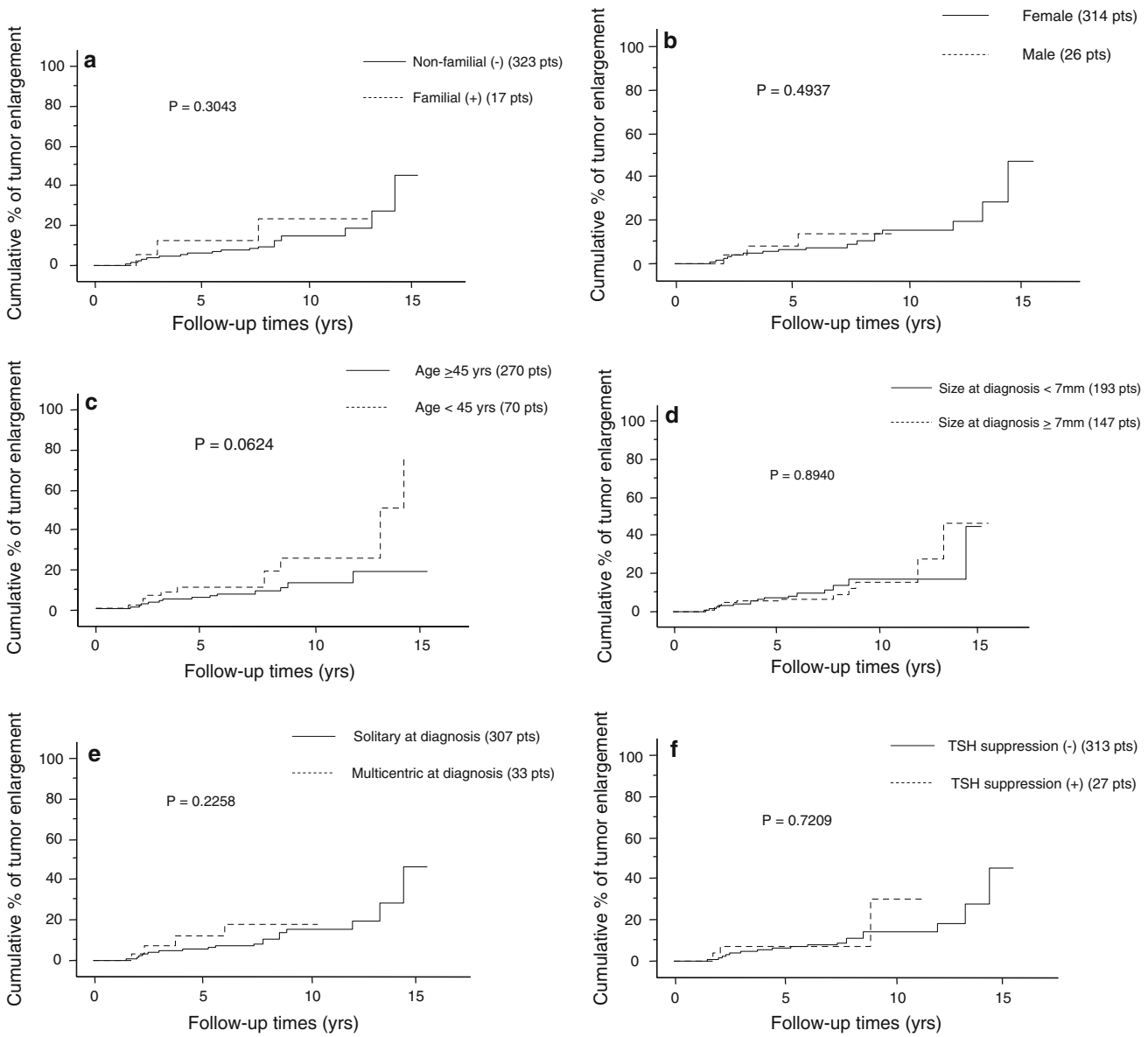


Fig. 2 **a** Proportion of patients with familial or non-familial PMC showing enlargement by 3 mm or more. **b** Proportion of male and female patients whose PMC showed enlargement by 3 mm or more. **c** Proportion of patients aged 45 years or older and those younger than 45 years whose PMC showed enlargement by 3 mm or more. **d** Proportion of patients whose PMC measured 7 mm or larger and those whose PMC was smaller than 7 mm at diagnosis and subsequently showed enlargement by 3 mm or more. **e** Proportion of patients whose solitary PMC and multiple PMC at diagnosis showed enlargement by 3 mm or more. **f** Proportion of patients whose PMC with thyroid stimulating hormone (TSH) suppression and without TSH suppression showed enlargement by 3 mm or more

physicians. Twelve patients underwent surgery after observation at their choice. Furthermore, one patient, a 15-year-old, was later diagnosed as having familial carcinoma and 7 others whose carcinomas were suspected of having multicentricity were recommended for and underwent surgery. The decision to proceed to operation in these 8 cases was not based on our present indications for surgery.

The extent of thyroidectomy and lymph node dissection in the 109 patients from the observation group who proceeded to operation is summarized in Table 2. The extent of resection in 2 patients is unknown because their surgery was performed at other hospitals. None of these patients showed carcinoma recurrence after surgery (average follow-up period: 76 months).

Outcome of PMC patients in the immediate surgical treatment group

We investigated the clinical outcomes of 1,055 patients with PMC in the immediate surgical treatment group. The

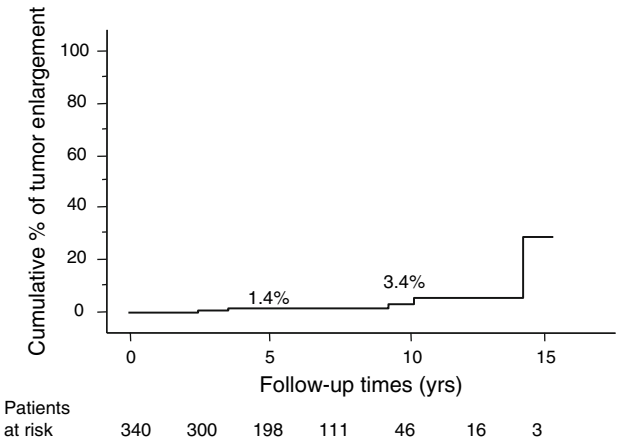


Fig. 3 Proportion of patients whose PMC showed the novel appearance of lymph node metastasis

Table 1 Reasons for surgical treatment in 109 patients with papillary microcarcinoma of the thyroid who initially underwent observation

Later diagnosed as having familial carcinoma[a]	1 patient
Tumor enlargement	32 patients[b]
Young age[a]	1 patient
Suspicion of multicentricity[a]	7 patients[c]
Tumor location near dorsal surface	17 patients[d]
Patients' choice	12 patients
Novel appearance of lymph node metastasis	5 patients
Coexistence of other thyroid diseases	10 patients
Unknown	25 patients

[a] They do not meet our criteria in the present study

[b] Seventeen patients were not recognized as showing tumor enlargement under the criteria for enlargement used in the present study

[c] One patient also showed tumor enlargement

[d] Including 4 patients who had been followed without diagnosis of PMC, who underwent immediate surgical treatment at diagnosis

extent of thyroidectomy and lymph node dissection is summarized in Table 3. To date, 32 patients showed carcinoma recurrence during postoperative follow-up. Table 4 summarizes organs in which PMC showed recurrence. The organ to which carcinoma most frequently recurred was the lymph node. We then investigated the prognostic implications of various clinicopathological parameters and patient background factors. As shown in Fig. 4a, patients with clinically apparent lateral node metastasis (N1b) showed significantly worse disease-free survival (DFS) than those with clinically apparent central node metastasis (N1a) or not having clinically apparent metastasis (N0) ($p < 0.0001$). The DFS of patients with N0 did not differ from that of patients with N1a. Male gender ($p < 0.0001$) (Fig. 4b) and pathologically confirmed lymph node metastasis (pN1) ($p = 0.0004$) also predicted a worse DFS. Our series included 25 patients (2.4%) having PMC with massive extrathyroid extension to the recurrent laryngeal nerve, trachea or esophagus, but none of these patients showed recurrence. Other clinicopathological features, such as age and multicentricity, did not affect DFS of patients (data not shown). We performed multivariate analysis for three features that did show prognostic significance on univariate analysis. N1b and male gender were recognized as independent prognostic factors for DFS (Table 5).

To date, two patients have died of carcinoma 79 and 94 months after the initial operation. Both patients were classified as having clinically apparent lateral node metastasis at presentation, and one also showed metastasis also in the mediastinal compartment.

Comparison between Rate of Novel Appearance of LN Metastasis in Patients Undergoing Observation and Recurrence Rate to the LN in Patients with N0 PMC Undergoing Immediate Surgical Treatment

Of 1,055 patients in the immediate surgical treatment group, 909 did not show clinically apparent lymph node metastasis in the central or lateral compartments (N0). Lymph node dissection was performed for 815 patients [central node dissection only for 525 and prophylactic modified neck dissection (MND) for 290]. To date, 5 of 525 patients who underwent central node dissection showed recurrence to lymph nodes in the lateral compartment. Of 290 patients who underwent MND, 4 showed recurrence to a lateral compartment (3 on the contralateral side and 1 in the ipsilateral compartment). As indicated

Table 2 Extent of thyroidectomy and lymph node dissection of 109 patients who underwent surgical treatment after observation

	Thyroidectomy	Lymph node dissection	
Total or near total	48 (44.0%)	CND only	79 (72.5%)
Subtotal	7 (6.4%)	Unilateral MND	26 (23.9%)
Lobectomy with isthmectomy	47 (43.1%)	Bilateral MND	2 (1.8%)
Isthmectomy	5 (4.6%)	Unknown[b]	2 (1.8%)
Unknown[a]	2 (1.8%)		

CND complete radical neck dissection, *MND* modified radical neck dissection

[a] These two patients underwent surgery at other hospitals

Table 3 Extent of thyroidectomy and lymph node dissection of 1,055 patients in the immediate surgical treatment group

	Thyroidectomy	Lymph node dissection	
Total or near total	432 (40.9%)	CND only	536 (50.8%)
Subtotal	101 (9.6%)	Unilateral MND	402[a] (38.1%)
Lobectomy with isthmectomy	490 (46.4%)	Bilateral MND	23 (2.2%)
		Not done	94 (8.9%)
Isthmectomy	25 (2.4%)		
Partial lobectomy	7 (0.7%)		

[a] One patient also underwent dissection of the mediastinal compartment

Table 4 Recurrence in 32 patients[a] (3.0%)

Lymph node	26 (2.5%)
Previously dissected compartments	11
Compartments that had not been dissected	13
Both compartments	2
Locoregional organs	
Thyroid	6 (0.6%)
Others	2 (0.2%)
Distant organs	
Lung	1 (0.1%)
Bone	1 (0.1%)

[a] Three patients showed recurrence in two or more organs

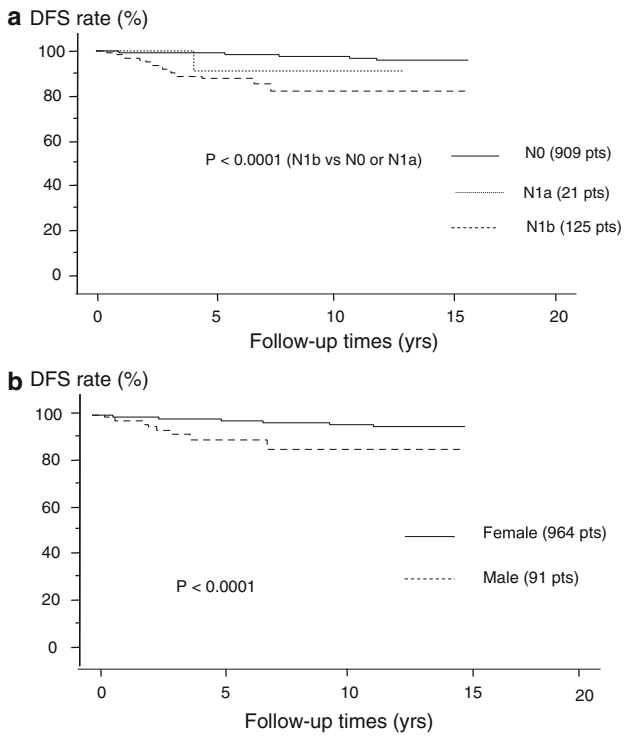


Fig. 4 **a** Kaplan–Meier curves for disease-free survival (DFS) of PMC patients with N1b, N1a, and N0. **b** Kaplan–Meier curves for DFS of male and female patients with PMC

Table 5 Multivariate analysis regarding disease-free survival (DFS) of PMC patients

Variables	*p* values	Hazard ratio (95% confidence interval)
N1b	0.0003	4.46 (2.00–10.00)
Male gender	0.0255	2.59 (1.12–5.95)
pN1	0.1283	2.08 (0.81–5.38)

above, 7 patients in the observation group showed novel lymph node metastasis during follow-up, and all of those lesions were in the lateral compartment ipsilateral to the primary lesion. We compared the rate of novel appearance of nodal metastasis or recurrence to the lymph node in these three subsets. As shown in Fig. 5, the rate of novel appearance of nodal metastasis in the observation group did not differ from that of recurrence to the lymph node in the immediate surgical treatment group. Furthermore, we could not find any significant difference in the rate of recurrence to the lymph node between patients who underwent central node dissection only and those who underwent prophylactic MND.

Discussion

This report is a continuation of our observation trial for PMC in patients without any unfavorable features and the prognosis for PMC patients who have undergone operation either immediately after diagnosis or after a period of observation. In this study, we enrolled 340 patients who were diagnosed as having PMC between 1992 and 2004 and subsequently underwent observation without immediate surgical treatment. The average follow-up period increased to 74 months, significantly longer than that in previous studies. However, the rate of carcinoma enlargement was 6.4% at 5-years follow-up, which was similar to that in our previous reports [22, 23]. At 10 years, 15.9% of tumors demonstrated enlargement, but the number of patients at risk at 10 years remained low at 39, indicating

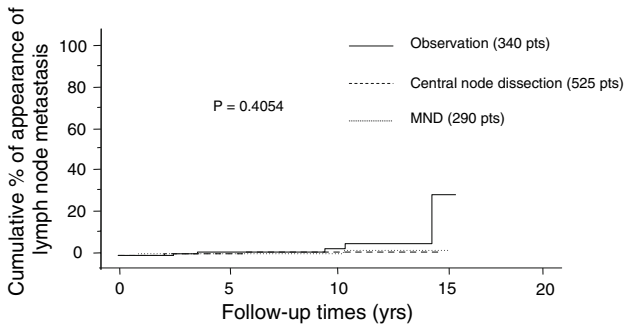


Fig. 5 Proportion of patients whose PMC showed novel appearance of lymph node metastasis during observation, those who underwent central node dissection only and those who underwent prophylactic modified neck dissection (MND) in the immediate surgical group showing recurrence to the node

that the incidence could change with further increases in the number of patients who have been observed for a long time. We also demonstrated the results of the novel appearance of lymph node metastasis during observation: 1.4% at 5 years and 3.4% at 10 years, indicating that the incidence is lower than enlargement of primary tumor.

In our observation series, 109 patients went on to surgical treatment after observation for various reasons. The most common reason was recorded as tumor enlargement. However, 17 of 32 patients whose tumors were judged by the attending physicians as showing enlargement did not meet the criteria for enlargement set forth in this study, indicating that the extent of enlargement in these cases was within-observer variation. Furthermore, 13 patients were recommended for surgery because of a dorsal tumor location, even though observation had been recommended at the initial diagnosis of PMC. More accurate evaluation of the tumor at the first examination and, if observation is decided, systematic evaluation of tumor size at each follow-up by the attending physician would be a more desirable approach. None of the 109 patients showed carcinoma recurrence or died of carcinoma during postoperative follow-up. It is important to note that, for patients whose tumor is under observation, it would not be too late to perform surgical treatment if there are signs of progression, such as tumor enlargement or novel appearance of lymph node metastasis.

We investigated whether patient backgrounds and clinical features are linked to PMC progression, tumor enlargement, and novel appearance of nodal metastasis. Male gender, multicentricity, and advanced age are known to be conventional prognostic factors of papillary carcinoma [3, 4], but these features did not affect PMC progression during observation. Furthermore, we failed to establish a relationship between carcinoma enlargement and tumor size at diagnosis. It is therefore suggested that all PMC without any unfavorable features can be candidates for observation regardless of patient background and clinical features. We could not find any evidence that TSH suppression effectively prevents carcinoma progression. However, there were only 27 patients who underwent TSH suppression in this series and further studies are necessary to draw a final conclusion on this issue. The incidence of familial carcinoma in our observation series was 5.0%, which is similar to that in previous reports from Japan with a large series of papillary carcinoma patients undergoing surgical treatment [26, 27]. We showed that the prognosis of familial papillary carcinoma after surgical treatment did not differ from that of non-familial carcinoma [27]. Also in this study, the rate of progression of familial PMC was the same as that of non-familial PMC in the observation group, indicating that immediate surgical treatment is not mandatory for familial PMC patients unless they have any unfavorable features or show progression during observation.

We previously demonstrated that PMC patients having clinically apparent lateral node metastasis (N1b) were more likely to show recurrence [20, 21]. This was confirmed on multivariate analysis in this study, indicating that N1b is an independent prognostic factor for DFS of PMC patients. The organ to which carcinoma most frequently shows recurrence is the lymph node, and recurrence to the compartment that had previously been dissected occurred with an incidence similar to that of recurrence to the compartment that had not previously been dissected. Even though the primary tumor is small, surgeons should carefully perform therapeutic lymph node dissection at first surgery for N1b PMC. Together with N1b, massive extrathyroid extension (pT4) also significantly affects the prognosis of papillary carcinoma [3, 4], but in our series, none of the patients with pT4 had carcinoma recurrence. The number of pT4 patients was small at 25, accounting only for 2.4% of this series, and the range of extension to adjacent organs is very limited for pT4 PMC, which may explain our findings.

In our previous study, we showed that in a subset of PMC patients without clinically apparent node metastasis, recurrence rate to the lymph node in patients who underwent central node dissection only did not differ from that in patients who underwent prophylactic MND [20, 21]. In addition, in this study, we demonstrated that these rates were similar to the rate of novel appearance of lymph node metastasis from PMC in the observation group. Our findings that the incidence of the novel appearance of lymph node metastasis in the observation group is as low as that of recurrence to the nodes in the immediate surgical treatment group, and that none of the patients showed recurrence even though they had undergone surgery after the appearance of nodal metastasis, further support the validity of observation for PMC from the perspective of lymph node metastasis.

In summary, we demonstrated that observation can be a therapeutic option for PMC without unfavorable features regardless of patient background and clinical features. It is not too late to perform surgical treatment after carcinoma shows signs of progression, such as tumor enlargement and/or the appearance of lymph node metastasis.

References

1. Sobin LH, Wittekind CH (eds) (2002) UICC: TNM classification of malignant tumors, 6th edn. Wiley-Liss, New York
2. Ito Y, Tomoda C, Uruno T et al (2005) Ultrasound-detectable and anatomopathologically detectable node metastasis in the lateral compartment as indicators of worse relapse-free survival in patients with papillary thyroid carcinoma. World J Surg 29:917–920
3. Ito Y, Tomoda C, Uruno T et al (2006) Prognostic significance of extrathyroid extension of papillary thyroid carcinoma: massive but not minimal extension affects the relapse-free survival. World J Surg 30:780–786
4. Ito Y, Miyauchi A, Jikuzono T et al (2007) Risk factors contributing to a poor prognosis of papillary thyroid carcinoma; validity of UICC/AJCC TNM classification and stage grouping. World J Surg 31:838–848
5. Yokozawa T, Miyauchi A, Kuma K et al (1995) Accurate and simple method of diagnosing thyroid nodules by the modified technique of ultrasound-guided fine needle aspiration biopsy. Thyroid 5:141–145
6. Takebe K, Date M, Yamamoto Y et al (1994) Mass screening for thyroid cancer with ultrasonography. KARKINOS 7:309–317 (in Japanese)
7. Harach HR, Franssila KO, Wasenius VM (1985) Occult papillary carcinoma of the thyroid: a "normal" finding in Finland. A systematic autopsy study. Cancer 56:531–538
8. Fukunaga FH, Yatani R (1975) Geographic pathology of occult thyroid carcinomas. Cancer 36:1095–1099
9. Samson RJ (1977) Prevalence and significant of occult thyroid cancer. In: DeGroot LJ (ed) Radiation-associated thyroid carcinoma. Grune & Stratton, New York, pp 137–153
10. Thorvaldsson SE, Tulinius H, Bjornsson J et al (1992) Latent thyroid carcinoma in Iceland at autopsy. Pathol Res Pract 188:747–750
11. Iida F, Sugenoya A, Muramatsu A (1991) Clinical and pathologic properties of small differentiated carcinomas of the thyroid gland. World J Surg 15:511–515
12. Hay ID, Grant CS, van Heerden JA et al (1992) Papillary thyroid microcarcinoma: a study of 535 cases observed in a 50-year period. Surgery 112:1139–1147
13. Rodriguez JM, Moreno A, Parrila P et al (1997) Papillary thyroid microcarcinoma: clinical study and prognosis. Eur J Surg 163:255–259
14. Lin KD, Lin JD, Huang MJ et al (1997) Clinical presentations and predictive variables of thyroid microcarcinoma with distant metastasis. Int Surg 82:378–381
15. Rassael H, Thompson LDR, Heffess CS (1998) A rationale for conservative management of microscopic papillary carcinoma of the thyroid gland: a clinicopathological correlation of 90 cases. Eur Arch Otorhinolaryngol 255:462–467
16. Sugitani I, Fumimoto Y (1999) Symptomatic versus asymptomatic papillary thyroid microcarcinoma: a retrospective analysis of surgical outcome and prognostic factors. Endocine J 46:209–216
17. Falvo L, D'Ercole C, Sorrenti S et al (2003) Papillary microcarcinoma of the thyroid gland: analysis of prognostic factors including histological subtype. Eur J Surg 168:28–32
18. Wada N, Duh QY, Sugino K et al (2003) Lymph node metastasis from 259 papillary thyroid microcarcinomas. Ann Surg 237:399–407
19. Chow SM, Law SCK, Chan JK (2003) Papillary microcarcinoma of the thyroid—prognostic significance of lymph node metastasis and multifocally. Cancer 98:31–40
20. Ito Y, Tomoda C, Uruno T et al (2004) Papillary microcarcinoma of the thyroid: how should it be treated? World J Surg 28:1115–1121
21. Ito Y, Uruno R, Nakano K et al (2003) An observation trial without surgical treatment in patients with papillary microcarcinoma of the thyroid. Thyroid 13:381–388
22. Ito Y, Miyauchi A (2007) A therapeutic strategy for incidentally detected papillary microcarcinoma of the thyroid. Nature Clin Pract Endocrinol Metab 3:240–248
23. Ito Y, Miyauchi A (2007) Appropriate treatment for asymptomatic papillary microcarcinoma of the thyroid. Exp Opin Pharmacother 8:3205–3215
24. Grossman RF, Tu SH, Duh QY (1995) Familial nonmedullary thyroid cancer. An emerging entity that warrants aggressive treatment. Arch Surg 130:892–897
25. Uruno T, Miyauchi A, Shimizu K et al (2005) Usefulness of thyroglobulin measurement in fine-needle aspiration biopsy specimens for diagnosing cervical lymph node metastasis in patients with papillary thyroid cancer. World J Surg 29:483–485
26. Uchino S, Noguchi S, Kawamoto H et al (2002) Familial nonmedullary thyroid carcinoma characterized by multifocality and a high recurrence rate in a large study population. World J Surg 26:897–902
27. Ito Y, Kakudo K, Hirokawa M et al (2009) Biological behavior and prognosis of familial papillary thyroid carcinoma. Surgery 145:100–105

THYROID
Volume 23, Number 9, 2013

DOI: 10.1089/thy.2012.0608

A Systematic Review and Meta-Analysis of Prophylactic Central Neck Dissection on Short-Term Locoregional Recurrence in Papillary Thyroid Carcinoma After Total Thyroidectomy

Brian Hung-Hin Lang,[1] Sze-How Ng,[2] Lincoln L.H. Lau,[3] Benjamin J. Cowling,[3] Kai Pun Wong,[1] and Koon Yat Wan[4]

Background: Prophylactic central neck dissection (pCND) at the time of total thyroidectomy (TT) remains controversial in clinically node-negative (cN0) papillary thyroid carcinoma (PTC). Despite occult central lymph node metastases being common, it is unclear if removing these metastases initially would reduce future locoregional recurrence (LRR). This systematic review and meta-analysis aimed at comparing the short-term LRR between patients who underwent TT with pCND and those who underwent TT alone.
Methods: A systematic review of the literature was performed to identify studies comparing LRR between patients with PTC who underwent TT + pCND (group A) and those who underwent TT alone (group B). Inclusion criteria were cN0 patients, with each comparative group containing >10 patients, and with the number of LRR and mean follow-up duration available. The pooled incidence rate ratio (IRR) was used for calculating the LRR rate between the two groups. Other parameters evaluated included postoperative radioiodine (RAI) ablation, surgically related complications, and overall morbidity. Meta-analysis was performed using a fixed-effects model.
Results: Fourteen studies matched the selection criteria. Of the 3331 patients, 1592 (47.8%) belonged to group A, while 1739 (52.2%) belonged to group B. Relative to group B, group A was significantly more likely to have postoperative RAI ablation (71.7% vs. 53.1%; odds ratio [OR] = 2.60 [95% confidence interval (CI) = 2.12–3.18]), temporary hypocalcemia (26.0% vs. 10.8%; OR = 2.56 [CI = 2.04–3.21]), and overall morbidity (33.2% vs. 17.7%; OR = 2.12 [CI = 1.75–2.57]). When temporary hypocalcemia was excluded, overall morbidity was similar between the two groups (7.3% vs. 6.8%; OR = 1.07 [CI = 0.78–1.47]). Group A had a significantly lower risk of LRR than group B (4.7% vs. 8.6%; IRR = 0.65 [CI = 0.48–0.86]).
Conclusions: Group A was more likely to have postoperative RAI ablation, temporary hypocalcemia, and overall morbidity than group B. Temporary hypocalcemia was the major surgical morbidity in pCND and, when excluded, the overall morbidity appeared similar between the two groups. Although our meta-analysis would suggest that those who undergo TT + pCND may have a 35% reduction in risk of LRR than those who undergo TT alone in the short term (<5 years), it remains unclear how much of this risk reduction is related to increased use of RAI ablation and potential selection bias in some of the studies examined.

Introduction

Papillary thyroid carcinoma (PTC) is the most common type of differentiated thyroid carcinoma, with its age-adjusted incidence doubling in the last 25 years (1). Despite its good prognosis, locoregional recurrence (LRR) is common (2). With recognition of the stepwise progression of lymph node metastasis (LNM) from the central (level VI) to lateral compartment (levels II–V), some surgeons have advocated routine prophylactic central neck dissection (pCND) at the time of total thyroidectomy for PTC (3). Although there is general agreement that formal lymph node dissection should be performed in the setting of imageable, biopsy-proven, or palpable nodal disease (cN1), it remains controversial in patients with no clinical evidence of nodal metastasis (cN0) (4). There is little evidence to suggest that patients

Departments of [1]Surgery and [4]Clinical Oncology, and [3]School of Public Health, The University of Hong Kong, Hong Kong, China.
[2]Breast and Endocrine Unit, Department of Surgery, Kuala Lumpur Hospital, Kuala Lumpur, Malaysia.

with cN0 undergoing a total thyroidectomy (TT) and pCND (TT + pCND) would reduce the risk of future LRR when compared to patients undergoing TT alone. Although the incidence of occult or microscopic LNM in patients with cN0 is relatively common, it is unclear whether removing these occult or microscopic LNM at the time of the primary operation could prevent LRR (5,6). Analysis of short-term surrogates for recurrence (such as postsurgical thyroglobulin level) would suggest that pCND may improve short-term outcomes, but this has not been fully resolved (4,7,8). Furthermore, patients undergoing pCND are at increased risk of temporary hypocalcemia (9–11).

One of the main reasons for the lack of evidence is that studies so far comparing TT + pCND with TT alone have not had the statistical power to detect a difference in LRR. A recent study estimated more than 5000 patients would be required to have sufficient statistical power to demonstrate a 25% reduction in LRR with pCND in patients with cN0 (12). To our knowledge, three meta-analyses have compared the outcomes between TT + pCND and TT alone. Two were not strictly relevant because one included patients with benign disease, while the other included patients who underwent therapeutic CND (9,10). Zetoune *et al.* pooled together five relevant studies and found a similar overall LRR rate between TT + pCND and TT alone (2.02% vs. 3.92%; odds ratio [OR] = 1.05 [95% confidence interval (CI) = 0.44–3.91]) (11). However, this study did not account for the difference in follow-up duration between the two groups. With an increasing number of new publications on this controversial subject in recent years, we conducted a systematic review and meta-analysis to compare the risk of LRR between TT + pCND and TT alone by reviewing the current literature.

Methods

This systematic review and meta-analysis was conducted in accordance with the PRISMA statement (13).

Search strategy

Studies comparing the rate of LRR between patients who underwent TT + pCND and TT alone were retrieved from the Scopus, Medline (PubMed), and Cochrane Library electronic databases on January 30, 2013. We used the following free-text search terms in "All fields": (i) "central neck dissection" *or* "level VI neck dissection" *or* "neck dissection"; (ii) "papillary thyroid carcinoma"; (iii) i *and* ii.

There was no language restriction and no methodological filters. The bibliographies of three previous meta-analyses were searched for other additional relevant references (9–11).

Study selection

All titles identified by the search strategy were independently screened by three authors (B.H.L., S.H.N., and K.P.W.). Search results were compared, and disagreements were resolved by consensus. Abstracts of potentially relevant titles were then reviewed for eligibility, and full-length articles were selected for closer examination if there was a specific description on CND in patients with PTC. The criteria for eligibility were as follows. First, any prospective or retrospective studies on patients with PTC only were included. Studies that analyzed differentiated thyroid carcinoma were considered if results of PTC were separately reported. Second, studies with two arms comparing LRR between TT + pCND and TT alone were included. Third, each study arm had to have >10 patients. Fourth, patients in either arm had to be cN0 by preoperative imaging and intraoperative examination; patients with cN1 or distant metastasis (M1) were not included. Finally, the number of LRR and the mean follow-up (in months) in each study arm had to be available. The reason for obtaining the mean follow-up period was because, in order to work out the pooled incidence rate ratio (IRR) for TT + pCND and TT alone groups, we had to first calculate the number of person-years in each respective arm. Studies that specifically reported the number of LRR and follow-up period in TT + pCND and TT alone as subgroups were included. Patients who underwent hemithyroidectomy with pCND or underwent simultaneous pCND and prophylactic lateral neck dissection were excluded. For studies that only provided the number of LRR without the mean follow-up duration or provided only the median and not the mean follow-up duration, the corresponding author of those studies was individually contacted for further information. Multiple reports of the same data set were assessed, and the most updated report of a study was included.

Data extraction

All data were extracted onto a standardized form. The primary data extracted from each article included type or design of study, first authorship, country of origin, year of publication, patient demographics, preoperative nodal assessment, method of selection for pCND, tumor characteristics, number of patients who underwent TT + pCND or TT alone, extent of pCND (unilateral vs. bilateral), number of normal and metastatic central LNs harvested, mean follow-up period, radioiodine (RAI) ablation given or not, number of LRR, operating time, volume of blood loss, and any surgically related morbidities. LRR was defined as a recurrence occurring in the thyroid bed, central and/or lateral compartments. A patient found to have distant recurrence only (i.e., without concomitant LRR) was not counted as a LRR, while a patient with concomitant LRR and distant recurrence was counted as a LRR. The percentage of recurrent laryngeal nerve (RLN) injury was calculated based on the number of patients. The overall morbidity rate was calculated by dividing the total number of patients who suffered one or more perioperative morbidity over the total number of patients. If a patient suffered from two or more morbidities, it was counted as one.

Statistical analysis

All the individual outcomes were integrated with the meta-analysis software Review Manager Software 5.0 (Cochrane Collaborative, Oxford, United Kingdom). LRR was assessed by IRR according to person-year of follow-up, and ORs were examined for the other surgical outcomes. All results were aggregated and analyzed using a fixed-effects model. A subgroup analysis of overall morbidity was performed excluding temporary postoperative hypocalcemia. Publication bias was estimated by Begg's rank correlation test and Egger's regression test (14,15). The meta-analyses in this study were

conducted using R version 2.15.1 (R Foundation for Statistical Computing, Vienna, Austria) and the metafor package (16).

Results

Figure 1 shows the flowchart of studies retrieved and excluded. Of the 1822 titles initially identified from the database search, 41 full-length articles were assessed for inclusion, of which 27 were excluded and 14 studies were determined to be eligible and were included in this systematic review (7,17–29). Appendix Table A1 lists these 27 articles (6,8,30–54) and the reason for their exclusion. No additional study was found from our search of the three bibliographies in previous meta-analyses (9–11). One study (8) was excluded, as it analyzed a subset of study subjects that were later recruited in a multicenter cohort study (25).

Baseline characteristics

Table 1 shows a comparison of the baseline characteristics between the 14 eligible studies. There was no randomized trial. Thirteen studies were retrospective, while one was prospective. Of the 3331 patients included, 1592 (47.8%) underwent TT + pCND (group A), while 1739 (52.2%) underwent TT only (group B). In terms of preoperative nodal assessment, ultrasonography (US) was used as the standard imaging modality in all studies, but only two studies specifically mentioned that both bilateral central and lateral neck compartments were examined (25,27).

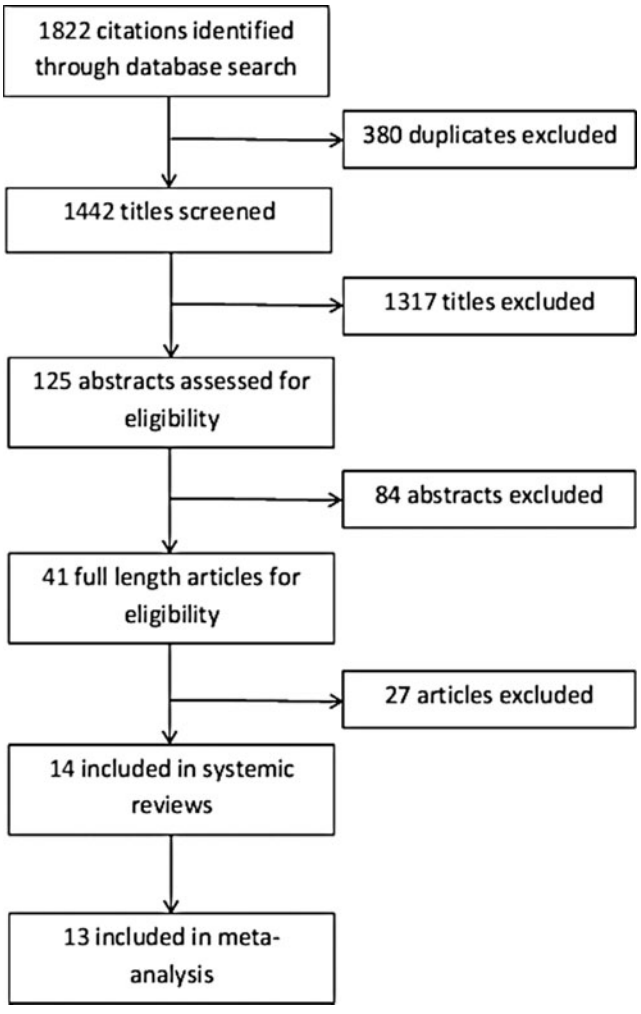


FIG. 1. Flow diagram for study selection.

In terms of selection for pCND, seven studies were based on individual surgeon's preference (7,18,22,24,25,27,28), while four studies did not specify their method of selection (17,19,23,26). Three studies used historical controls (TT alone) for comparison (20,21,29). Only 11 of 14 studies statistically compared age, sex ratio, tumor size, extrathyroidal extension, and tumor multifocality between the two groups (7,17,18,21–26,28,29). Of these, two studies found age to be significantly older in group B (21,25), and three studies found tumor size to be significantly different (7,21,23). Two studies found tumor size to be significantly larger in group A (7,21), while one study found tumor size to be significantly smaller in group A (23). Three of nine studies found the rate of extrathyroidal extension to be significantly higher in group A (7,22,24), and two of nine studies found the rate of tumor multifocality to be significantly higher in group A (22,26). Bilateral pCND was performed in eight studies (17,18,22–24,26,28,29), while the other six studies performed either unilateral or a combination of unilateral and bilateral pCND (7,19,20,21,25,27). Among the eight studies reporting bilateral pCND (17,18,22–24,26,28,29), the mean number of central lymph nodes harvested ranged between 5.6 and 9.6, while the one study reporting unilateral pCND harvested a median of five (7). The incidence of central LNM in group A ranged from 23.5% to 82.4%, while in group B it ranged from 0.9% to 9.7% with 9 of 14 studies not reporting the incidence of central LNM in group B.

Surgical outcomes

Table 2 shows a comparison of outcomes between the two groups. Only 9 of the 14 studies reported whether RAI ablation was given after surgery (7,18,21–24,26,27,29). Their dose ranged from 2.78 to 5.55 GBq. One study empirically gave the same dose of RAI, irrespective of the extent of LNM (7). The mean frequency of postoperative RAI ablation in groups A and B were 746/1041 (71.7%) and 498/937 (53.1%). Group A was significantly more likely to receive RAI ablation than group B (OR = 2.60 [CI = 2.12–3.18]). This was expected because of the higher incidence of central LNM (or N1a) in group A leading to tumor group upstaging in patients older than 45 years (28). Only 1 of 14 studies compared operating time between the two groups and found group B to have a significantly shorter operating time than group A (28).

Figure 2 shows the forest plot for temporary hypocalcemia. Of the 14 studies, 11 studies compared temporary postoperative hypocalcemia between the two groups, while 10 studies compared permanent postoperative hypocalcemia in the two groups. In eight studies, permanent hypocalcemia was defined as persistent hypocalcemia and/or need for calcium supplements for more than six months (7,20,23–28), while two studies defined it as more than 12 months (18,29). If one assumed all studies utilized a similar definition for temporary and permanent hypocalcemia, the overall temporary hypocalcaemia rate in group A was significantly higher than that in B (336/1294 (26.0%) and 144/1330 (10.8%), respectively; OR = 2.56 [CI = 2.04–3.21]) while the overall permanent hypocalcaemia was also similar between the group A and B (25/1254 (2.0%) and 15/1257 (1.2%), respectively; OR = 1.74 [CI = 0.87–3.50]).

Table 1. Summary and Comparison of Baseline Characteristics Between Total Thyroidectomy + Prophylactic Central Neck Dissection (Group A) and Total Thyroidectomy Alone (Group B)

First author (country)	No. of patients		Mean age at operation (years)	Sex ratio (male:female)	Mean tumor size (mm)	ETE (%)	Multifocality (%)	No significant difference[a] between A and B	No. of central LNs excised		Incidence of central LNM (%)	
	A	B							A	B	A	B
Roh (Korea) (17)	40	73	A = na B = 48.5	A = na B = 9:64	A = na B = 22	A = na B = 30.1	A = na B = 15.1	1,2,3,4,5	5.6	na	62.2	na
Choi (Korea) (18)	48	53	A = 52 B = 48	A = 6:42 B = 11:42	A = 6.8 B = 7.3	A = 45.8 B = 58.5	A = 31.3 B = 22.6	1,2,3,4,5	na	na	37.5	na
Bardet (France) (19)	36	161	na	na	na	na	na	na	na	na	23.5	na
Perrino (Italy) (20)	92	159	na	na	na	na	na	na	na	na	33–52	na
Costa (Italy) (21)	126	118	A = 46 B = 52	A = 26:100 B = 24:94	A = 17 B = 15	na	A = 45.2 B = 40.7	2,4,5	na	na	46.8	5.1
Zuniga (Columbia) (22)	136	130	A = 42.9 B = 41.5	A = 10:126 B = 13:117	na	A = 44.8 B = 27.3	A = 38.2 B = 13.1	1,2,3	na	na	82.4	na
Moo (United States) (23)	45	36	A = 45.7 B = 49.2	A = 10:35 B = 4:32	A = 14 B = 20	A = 24.4 B = 36.1	A = 55.6 B = 63.9	1,2,4,5	8.8	1.7	33.3	na
Hughes (United States) (24)	78	65	A = 46.8 B = 41.2	A = 17:61 B = 16:49	A = 19 B = 20	A = 41.0 B = 16.9	A = 33.3 B = 33.8	1,2,3,5	6.0	0.0	61.5	9.2
Popadich (Australia, United States, and United Kingdom) (25)	259	347	A = 44 B = 48	A = 52:207 B = 81:266	A = 23 B = 22	A = 27.8 B = 24.2	A = 48.3 B = 42.7	2,3,4,5	6.8	0.35	49.0	0.9
So (Korea) (26)	119	113	A = 49.2 B = 49.8	A = 21:98 B = 16:97	A = 6.6 B = 6.2	A = 51.3 B = 51.3	A = 37.8 B = 23.9	1,2,3,4	na	na	37.0	na
Lang (China) (7)	82	103	A = 52.0 B = 50.0	A = 18:64 B = 22:81	A = 15 B = 10	A = 26.8 B = 14.6	A = 36.6 B = 27.2	1,2,5	5	0	54.9	4.9
Wang (United States) (27)	49	37	na	na	s	na	na	na	9	na	40.8	na
Raffaelli (Italy) (28)	124	62	A = 42.7 B = 43.2	A = 24:100 B = 13:49	A = 12.9 B = 12.1	na	A = 51.6 B = 46.8	1,2,3,5	9.6	1.5	35.5	9.7
Barczynski (Poland) (29)	358	282	na	A = 75:283 B = 60:222	na	A = 13.1 B = 12.8	A = 37.7 B = 35.1	1,2,4,5	6.7	na	na	na

[a]Matching: 1 = age; 2 = sex ratio; 3 = tumor size; 4 = extrathyroidal extension; 5 = tumor multifocality.

s, significantly different ($p < 0.05$) between groups A and B; ns, not significantly different between groups A and B; na, not available or specified; LN, lymph node; LNM, lymph node metastasis; ETE, extrathyroidal extension.

Table 2. Comparison of Surgical Outcomes Between Total Thyroidectomy + Prophylactic Central Neck Dissection (Group A) and Total Thyroidectomy Alone (Group B)

First author	No. of postoperative RAI ablation (%)	No. of postoperative hypocalcemia (%)		No. of recurrent laryngeal nerve injury (%)		No. of hematoma formation (%)	No. of wound infection/seroma (%)	Total morbidity[a] (%)
		Temporary	Permanent	Temporary	Permanent			
Roh (17)	na	A = 13 (32.5) B = 7 (9.6)	na	A = 0 (0.0) B = 3 (4.1)	A = 0 (0.0) B = 2 (2.7)	A = 1 (2.5) B = 1 (1.4)	A = 0 (0.0) B = 0 (0.0)	A = 14 (35.0) B = 13 (17.8)
Choi (18)	A = 48 (100) B = 53 (100)	A = 8 (16.7) B = 6 (11.3)	A = 0 (0.0) B = 1 (1.9)	A = 1 (2.1) B = 0 (0.0)	na	A = 0 (0.0) B = 0 (0.0)	A = 0 (0.0) B = 0 (0.0)	A = 9 (18.8) B = 7 (13.2)
Bardet (19)	na	na	na	na	na	na	na	—
Perrino (20)	na	A = 8 (8.7) B = 11 (6.9)	A = 3 (3.2) B = 2 (1.3)	A = 3 (3.3) B = 5 (3.1)	A = 1 (1.1) B = 4 (2.5)	A = 0 (0.0) B = 0 (0.0)	A = 0 (0.0) B = 1 (0.6)	A = 15 (16.3) B = 23 (14.5)
Costa (21)	A = 87 (69.0) B = 62 (52.5)	na	na	na	na	na	na	—
Zuniga (22)	A = 79 (58.1) B = 55 (42.3)	na	na	na	na	na	na	—
Moo (23)	A = 31 (68.9) B = 26 (72.2)	A = 14 (31.1) B = 2 (5.6)	A = 0 (0.0) B = 2 (5.6)	A = 2 (4.4) B = 0 (0.0)	A = 0 (0.0) B = 0 (0.0)	na	na	A = 16 (35.6) B = 4 (11.1)
Hughes (24)	A = 72 (92.3) B = 56 (86.2)	A = 21 (26.9) B = 5 (7.7)	A = 2 (2.6) B = 0 (0.0)	A = 0 (0.0) B = 2 (3.1)	A = 0 (0.0) B = na	A = 1 (1.3) B = 2 (3.1)	A = 0 (0.0) B = 1 (1.5)	A = 24 (30.8) B = 10 (15.4)
Popadich (25)	na	A = 25 (9.7) B = 14 (4.0)	A = 2 (0.8) B = 2 (0.6)	A = 1 (0.4) B = 8 (2.3)	A = 1 (0.4) B = 6 (1.7)	A = 5 (1.9) B = 3 (0.9)	A = 3 (1.2) B = 5 (1.4)	A = 37 (14.3) B = 38 (11.0)
So (26)	A = 101 (84.9) B = 92 (81.4)	A = 49 (41.2) B = 38 (33.6)	A = 7 (5.9) B = 2 (1.8)	A = 4 (3.4) B = 4 (3.5)	A = 1 (0.8) B = 2 (1.8)	A = 1 (0.8) B = 0 (0.0)	A = 0 (0.0) B = 0 (0.0)	A = 64 (53.8)[b] B = 46 (40.7)
Lang (7)	A = 62 (75.6) B = 63 (61.2)	A = 15 (18.3) B = 9 (8.7)	A = 2 (2.4) B = 1 (1.2)	A = 3 (3.7) B = 0 (0.0)	A = 1 (1.2) B = 1 (1.0)	A = 0 (0.0) B = 1 (1.0)	A = 0 (0.0) B = 0 (0.0)	A = 21 (25.6) B = 12 (11.7)
Wang (27)	A = 35 (71.4) B = 12 (32.4)	A = 21 (42.9) B = 4 (10.8)	A = 0 (0.0) B = 3 (8.3)	A = 1 (2.0) B = 1 (2.7)	na	na	na	A = 22 (44.9) B = 8 (21.6)
Raffaelli (28)	na	A = 53 (42.7) B = 11 (17.7)	A = 1 (0.8) B = 0 (0.0)	A = 1 (0.8) B = 0 (0.0)	A = 1 (0.8) B = 0 (0.0)	A = 0 (0.0) B = 0 (0.0)	A = 0 (0.0) B = 0 (0.0)	A = 56 (45.2) B = 11 (17.7)
Barczynski (29)	A = 231 (64.5) B = 79 (28.0)	A = 109 (30.4) B = 37 (13.1)	A = 8 (2.2) B = 2 (0.7)	A = 26 (7.3) B = 18 (6.4)	A = 9 (2.5) B = 6 (2.1)	na	na	A = 152 (42.5) B = 63 (22.3)

[a]The sum of all complications; same patient with more than two complications was counted as one.
[b]Including two chyle leaks.

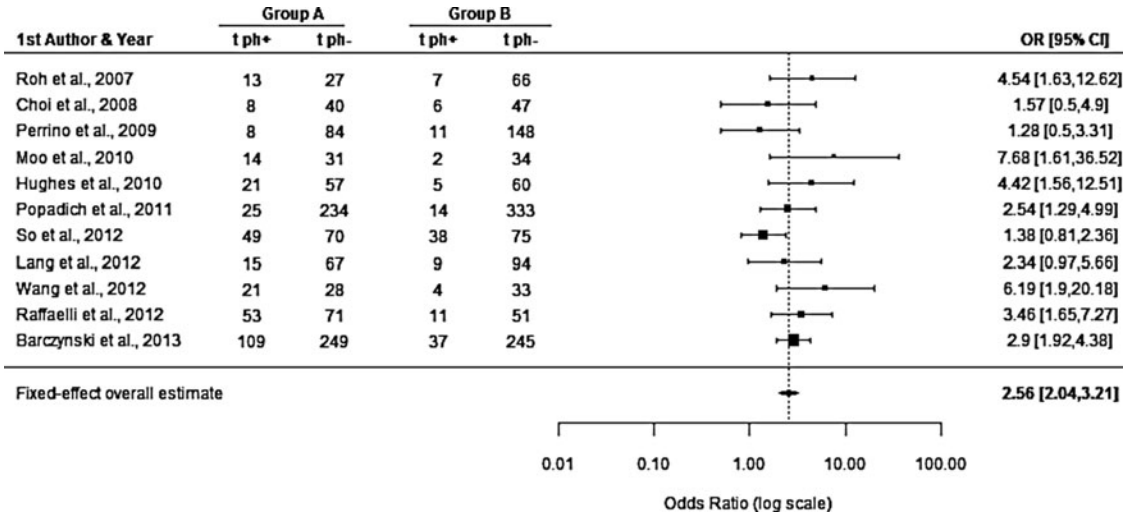


FIG. 2. Forest plot for temporary hypocalcemia (tph). OR, odds ratio; CI, confidence interval.

Similar to hypocalcemia, the definition for temporary and permanent RLN injury varied between studies. Routine perioperative DL was performed in five studies (7,25,26,28,29), and persistent impairment in vocal cord function for more than six months was defined as permanent RLN injury in seven studies (19,23–27). The cumulative temporary RLN palsy was comparable between group A and B (42/1294 (3.2%) and 41/1330 (3.1%), respectively; OR = 1.02 [CI = 0.64–1.64]). The cumulative permanent RLN palsy was also comparable between groups A and B (14/1197 (1.2%) and 21/1240 (1.7%), respectively; OR = 0.75 [CI = 0.37–1.55]).

The rate of hematoma was reported in eight studies. The cumulative hematoma rate was comparable between groups A and B (8/842 (1.0%) and 7/975 (0.9%), respectively; OR = 1.33 [CI = 0.53–3.35]). The wound infection/seroma rate was also similar between groups A and B (3/842 (0.4%) and 7/975 (0.9%), respectively; OR = 0.78 [CI = 0.28–2.07]). Figure 3 shows the forest plot for overall morbidity. The overall morbidity rate ranged between 14.3% and 53.8% in group A, while in group B it ranged between 11.0% and 40.7%. The overall morbidity after thyroid surgery in group A was significantly higher than in group B (430/1294 (33.2%) vs. 235/1330 (17.7%); OR = 2.12 [CI = 1.75–2.57]). However, after excluding temporary hypocalcemia, the overall morbidity in group A was not significantly different from group B (94/1294 (7.3%) vs. 90/1330 (6.8%); OR = 1.07 [CI = 0.78–1.47]). Figure 4 shows the forest plot for overall morbidity after excluding temporary hypocalcemia. The potential publication bias did not appear significant, as confirmed by the Begg analysis (Kendall's tau = −0.1636, $p = 0.5423$) and the Egger regression test ($z = -0.8921$, $p = 0.4167$).

LRR

Table 3 compares the LRR rate between the two groups. One study was excluded in the IRR calculation because the mean duration of follow-up was not available (19). Figure 5 shows the forest plot for LRR. The pooled mean follow-up in groups A and B were 45.2 and 50.8 months, respectively, while the pooled mean number of person-years in groups A

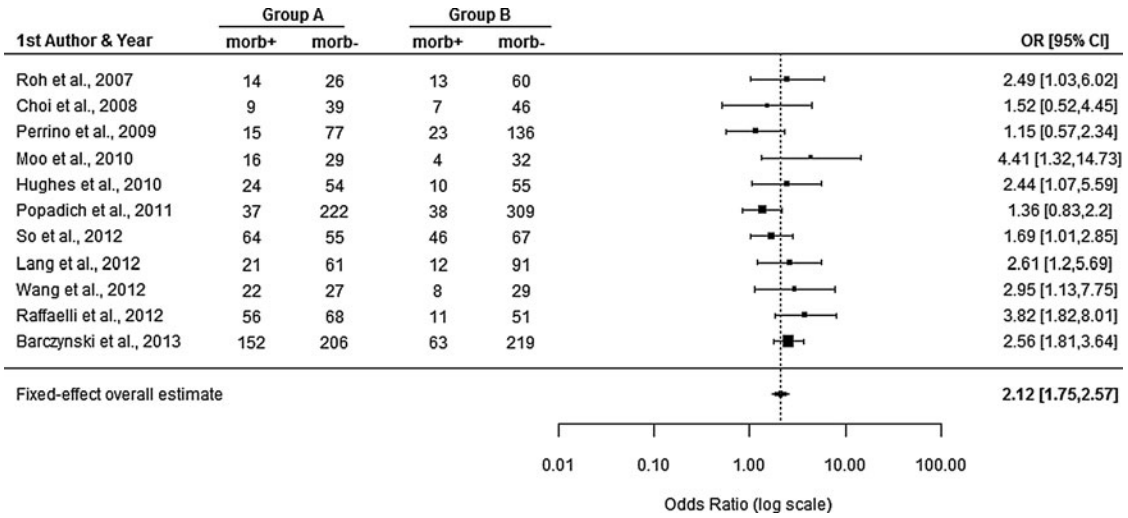


FIG. 3. Forest plot for overall morbidity (morb).

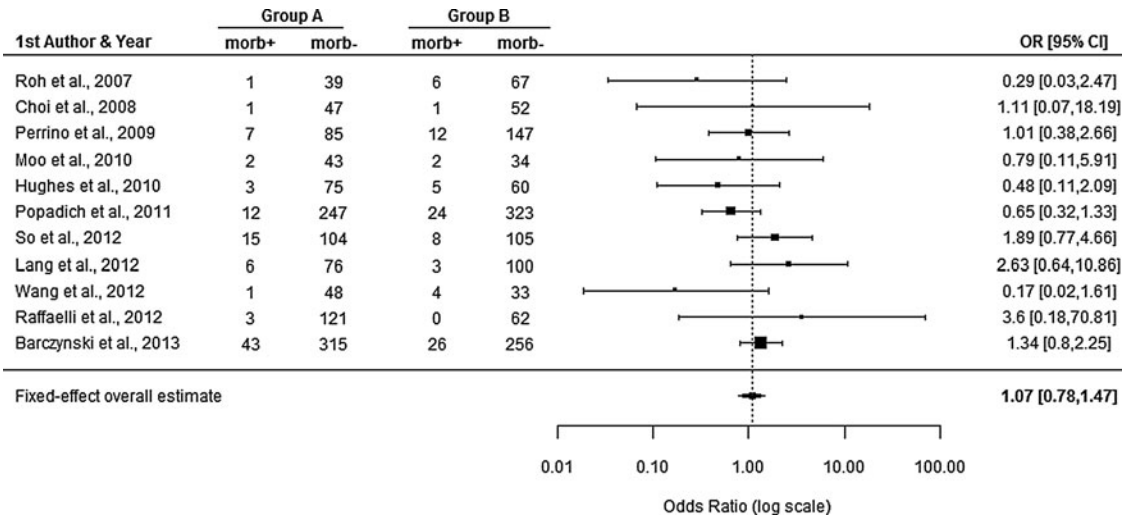


FIG. 4. Forest plot for overall morbidity after excluding temporary hypocalcemia.

and B were 598.9 and 662.3, respectively. Group A had significantly lower LRR than group B (75/1592 (4.7%) vs. 149/1739 (8.6%); IRR = 0.65 [CI = 0.48–0.86]). The potential publication bias was not significant, as confirmed by Begg analysis (Kendall's tau = −0.1677, $p = 0.4268$) and the Egger regression test ($z = 0.0984$, $p = 0.9216$).

Discussion

To our knowledge, this is to date one of the largest meta-analyses evaluating the impact of pCND on LRR in patients with clinically nodal negative PTC or cN0. With significantly more patients being included than in previous meta-analyses, our data suggest that those who undergo TT + pCND have a 35% reduction in risk of LRR than those who undergo TT alone. Although no significant publication bias was found in our meta-analysis, as shown by the Begg's rank correlation test and Egger's regression test, it is worth nothing that there was one particular large recent study that could have had a profound impact on the overall IRR (29). In fact, its number of person-years in groups A and B were almost two to three times of that of the next largest study (25). Nevertheless, on the funnel plot (data not shown), this particular study was just on the margin of the funnel, and therefore it was not excluded from the final meta-analysis.

Despite this important positive finding, we remain cautious in our conclusions, as there are a number of potential limitations. First, the mean follow-up period was relatively short in one study, having a mean follow-up period of only 10 months. In fact, the overall mean follow-up duration for groups A and B was only 45.2 and 50.8 months, respectively, and hence both groups had a mean follow-up of less than five years. Given the fact that PTC is a relatively slow-growing, indolent tumor, patients may not develop detectable LRR until many years

Table 3. Comparison of Locoregional Recurrence Rate Between Total Thyroidectomy + Prophylactic Central Neck Dissection (Group A) and Total Thyroidectomy Alone (Group B)

	Number of LRR (%)		*Mean follow-up (months)*		*Number of person-years*		
First author	*A*	*B*	*A*	*B*	*A*	*B*	*Incidence rate ratio [CI]*
Roh (17)	0 (0.0)	3 (4.1)	51	53	170	322	0.27 [0.01–5.25]
Choi (18)	1 (2.1)	2 (3.8)	24.4	24.4	98	108	0.55 [0.05–6.09]
Bardet (19)	4 (11.1)	6 (3.7)	na[a]	na[a]	—	—	—
Perrino (20)	5 (5.4)	22 (13.8)	69.2	69.2	531	917	0.39 [0.15–1.04]
Costa (21)	8 (6.3)	9 (7.6)	47	62	494	629	1.13 [0.44–2.94]
Zuniga (22)	19 (14.0)	26 (20.0)	73.44	95.52	832	1035	0.91 [0.50–1.64]
Moo (23)	2 (4.4)	6 (16.7)	37.2	37.2	140	112	0.27 [0.05–1.32]
Hughes (24)	2 (2.6)	2 (3.1)	19.1	27.5	124	149	1.20 [0.17–8.52]
Popadich (25)	13 (5.0)	29 (8.4)	32	50	691	1446	0.94 [0.49–1.81]
So (26)	2 (1.7)	4 (3.5)	44.7	45.4	443	428	0.48 [0.09–2.63]
Lang (7)	3 (3.7)	3 (2.9)	28.2	31.9	193	274	1.42 [0.29–7.04]
Wang (27)	0 (0.0)	0 (0.0)	10	10	41	31	0.76 [0.01–38.06]
Raffaelli (28)	1 (0.8)	0 (0.0)	25.0	25.5	258	132	1.53 [0.06–37.56]
Barczynski (29)	15 (5.8)	37 (13.1)	126.4	128.8	3771	3027	0.33 [0.18–0.59]
Overall	*75*	*149*	*45.2*	*50.8*	*598.9*	*662.3*	*0.65 [0.18–0.86]*

[a]Only medians were provided and therefore incidence rate ratio could not be calculated.
LRR, locoregional recurrence.

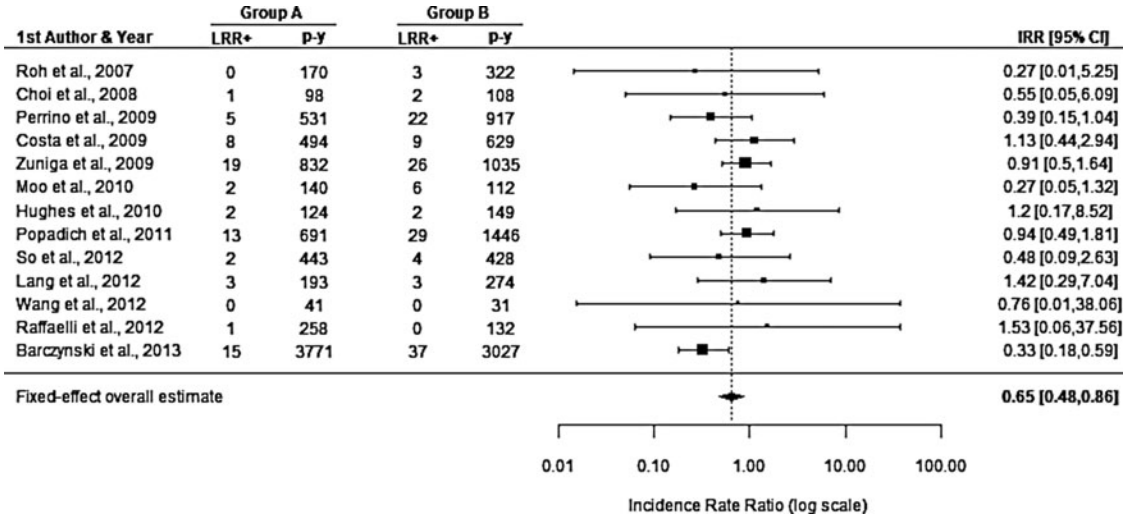


FIG. 5. Forest plot for locoregional recurrence (LRR).

after the initial operation. Therefore, a significant longer follow-up duration would be necessary to assess fully whether pCND could significantly reduce LRR at least in the medium to long term (12). Apart from this, 13 of 14 studies were retrospective analyses, and so they were subject to selection bias. Surgeon's preference or discretion was mentioned in 7 of 14 studies as their method of selecting pCND, while four studies did not clearly describe their method of selection. Three studies actually used historical controls for outcome comparison (20,21,29). These selection biases were evident by the fact that only one of the five baseline characteristics (i.e., sex ratio) was consistently comparable in all studies. The other baseline characteristics such as age, tumor size, presence of extrathyroidal extension, and tumor multifocality were not consistently comparable, and since some of these could also potentially influence the risk of LRR, it was difficult to assess the real impact of pCND on LRR. Accounting for these factors in the multivariate analysis may help, but not all these characteristics were readily available for analysis. Perhaps the best way to resolve this would be to conduct a prospective randomized trial in the future. Although all studies did mention using US as a method for preoperative nodal assessment, it was difficult to assess the quality and the comprehensiveness of the assessment. This issue was particularly relevant in the three studies where historical controls were analyzed because quality of imaging tended to change with time. Furthermore, it was unclear from these studies what US criteria were used for deciding on fine needle aspiration or surgery.

In terms of other outcomes, similar to previous studies (7,24,27), we found the rate of postoperative RAI ablation was significantly higher in group A than B (71.7% vs. 53.1%, respectively; OR = 2.60 [CI = 2.12–3.18]). This can likely be attributed to the higher incidence of central LNM in group A relative to B. However, it is interesting to note that the incidence of central LNM varied widely from 23.5% to 82.4% between studies. Perhaps this is also a reflection of the quality of preoperative US assessment, and might also be a result of differences in the extent of the pCND and quality of the histological examination between studies (5,55,56). Moreover, similar to previous meta-analyses (9–11), we found temporary hypocalcemia to be significantly higher in group A than B (26.0% vs. 10.8%, respectively; OR = 2.56 [CI = 2.04–3.21]). This would suggest that patients undergoing pCND during TT are 2.6 times more likely to develop temporary hypocalcemia than those undergoing TT alone. This is undoubtedly related to increased extent of surgical dissection leading to devascularization of parathyroid glands and/or inadvertent removal of parathyroid glands (7,17–27). However, it is worth noting that the rate of permanent hypocalcemia, temporary and permanent RLN injury, hematoma, and wound infection/seroma were not similar between the two groups. In addition, even though the overall morbidity was significantly higher in group A than B (OR = 2.12 [CI = 1.75–2.57]), when this analysis was repeated with temporary hypocalcemia excluded, the overall morbidity was similar between group A and B (OR = 1.07 [CI = 0.78–1.47]). This finding implied that the majority of morbidity arising from pCND was actually related to temporary hypocalcemia rather than other surgically related complications.

Conclusion

The addition of pCND to TT resulted in a greater likelihood of administering postoperative RAI ablation, temporary hypocalcemia, and overall morbidity. However, since temporary hypocalcemia accounted for the majority of overall morbidity in patients undergoing pCND, when temporary hypocalcemia was excluded from overall morbidity, it was similar between the two groups. Although our meta-analysis would suggest that those who undergo TT + pCND may have a 35% reduction in risk of LRR than those who undergo TT alone in the short term (< 5 years), it remains unclear how much of this risk reduction is related to increased use of RAI ablation and potential selection bias in some of the studies examined.

Author Disclosure Statement

All authors had nothing to disclose. No competing financial interests exist.

References

1. Hong Kong Cancer Registry, Hong Kong. Cancer incidence and mortality in Hong Kong 1983–2006. Available at www3.ha.org.hk/cancereg/statistics.html. Accessed September 15, 2012.
2. Lang BH, Lo CY, Chan WF, Lam KY, Wan KY 2007 Prognostic factors in papillary and follicular thyroid carcinoma: implications for cancer staging. Ann Surg Oncol **14:**730–738.
3. Machens A, Hauptmann S, Dralle H 2009 Lymph node dissection in the lateral neck for completion in central node-positive papillary thyroid cancer. Surgery **145:**176–181.
4. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM 2009 Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid **19:**1167–1214.
5. Lang BH, Tang AH, Wong KP, Shek TW, Wan KY, Lo CY 2012 Significance of size of lymph node metastasis on postsurgical stimulated thyroglobulin levels after prophylactic unilateral central neck dissection in papillary thyroid carcinoma. Ann Surg Oncol **19:**3472–3478.
6. Teixeira G, Teixeira T, Gubert F, Chikota H, Tufano R. 2011 The incidence of central neck micrometastatic disease in patients with papillary thyroid cancer staged preoperatively and intraoperatively as N0. Surgery **150:**1161–1167.
7. Lang BH, Wong KP, Wan KY, Lo CY 2012 Impact of routine unilateral central neck dissection on preablative and postablative stimulated thyroglobulin levels after total thyroidectomy in papillary thyroid carcinoma. Ann Surg Oncol **19:**60–67.
8. Sywak M, Cornford L, Roach P, Stalberg P, Sidhu S, Delbridge L 2006 Routine ipsilateral level VI lymphadenectomy reduces postoperative thyroglobulin levels in papillary thyroid carcinoma. Surgery **140:**1000–1007.
9. Chisholm EJ, Kulinskaya E, Tolley NS 2009 Systematic review and meta-analysis of the adverse effects of thyroidectomy combined with central neck dissection as compared with thyroidectomy alone. Laryngoscope **119:**1135–1139.
10. Shan CX, Zhang W, Jiang DZ, Zheng XM, Liu S, Qiu M 2012 Routine central neck dissection in differentiated thyroid carcinoma: a systematic review and meta-analysis. Laryngoscope **122:**797–804.
11. Zetoune T, Keutgen X, Buitrago D, Aldailami H, Shao H, Mazumdar M, Fahey TJ 3rd, Zarnegar R 2010 Prophylactic central neck dissection and local recurrence in papillary thyroid cancer: a meta-analysis. Ann Surg Oncol **17:**3287–3293.
12. Carling T, Carty SE, Ciarleglio MM, Cooper DS, Doherty GM, Kim LT, Kloos RT, Mazzaferri EL Sr, Peduzzi PN, Roman SA, Sippel RS, Sosa JA, Stack BC Jr, Steward DL, Tufano RP, Tuttle RM, Udelsman R; American Thyroid Association Surgical Affairs Committee 2012 American Thyroid Association design and feasibility of a prospective randomized controlled trial of prophylactic central lymph node dissection for papillary thyroid carcinoma. Thyroid **22:**237–244.
13. Moher D, Liberati A, Tetzlaff J, Altman DG 2009 PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med **6:** e1000097.
14. Begg CB, Mazumdar M 1994 Operating characteristics of a rank correlation test for publication bias. Biometrics **50:**1088–1101.
15. Egger M, Davey Smith G, Schneider M, Minder C 1997 Bias in meta-analysis detected by a simple, graphical test. BMJ **315:**629–634.
16. Viechtbauer W. 2010 Conducting meta-analyses in R with the metaphor package. J Stat Softw **36:**1–48.
17. Roh JL, Park JY, Park CI 2007 Total thyroidectomy plus neck dissection in differentiated papillary thyroid carcinoma patients: pattern of nodal metastasis, morbidity, recurrence, and postoperative levels of serum parathyroid hormone. Ann Surg **245:**604–610.
18. Choi SJ, Kim TY, Lee JC, Shong YK, Cho KJ, Ryu JS, Lee JH, Roh JL, Kim SY 2008 Is routine central neck dissection necessary for the treatment of papillary thyroid microcarcinoma? Clin Exp Otorhinolaryngol **1:**41–45.
19. Bardet S, Malville E, Rame JP, Babin E, Samama G, De Raucourt D, Michels JJ, Reznik Y, Henry-Amar M 2008 Macroscopic lymph-node involvement and neck dissection predict lymph-node recurrence in papillary thyroid carcinoma. Eur J Endocrinol **158:**551–560.
20. Perrino M, Vannucchi G, Vicentini L, Cantoni G, Dazzi D, Colombo C, Rodari M, Chiti A, Beck-Peccoz P, Fugazzola L 2009 Outcome predictors and impact of central node dissection and radiometabolic treatments in papillary thyroid cancers ≤ 2 cm. Endocrine-Related Cancer **16:**201–210.
21. Costa S, Giugliano G, Santoro L, Ywata De Carvalho A, Massaro MA, Gibelli B, De Fiori E, Grosso E, Ansarin M, Calabrese L. 2009 Role of prophylactic central neck dissection in cNO papillary thyroid cancer. Acta Otorhinolaryngologica Ital **29:**61–69.
22. Zuniga S, Sanabria A. 2009 Prophylactic central neck dissection in stage N0 papillary thyroid carcinoma. Arch Otolaryngol Head Neck Surg **135:**1087–1091.
23. Moo TA, McGill J, Allendorf J, Lee J, Fahey T 3rd, Zarnegar R. 2010 Impact of prophylactic central neck lymph node dissection on early recurrence in papillary thyroid carcinoma. World J Surg **34:**1187–1191.
24. Hughes DT, White ML, Miller BS, Gauger PG, Burney RE, Doherty GM 2010 Influence of prophylactic central lymph node dissection on postoperative thyroglobulin levels and radioiodine treatment in papillary thyroid cancer. Surgery **148:**1100–1106.
25. Popadich A, Levin O, Lee JC, Smooke-Praw S, Ro K, Fazel M, Arora A, Tolley NS, Palazzo F, Learoyd DL, Sidhu S, Delbridge L, Sywak M, Yeh MW 2011 A multicenter cohort study of total thyroidectomy and routine central lymph node dissection for cN0 papillary thyroid cancer. Surgery **150:**1048–1057.
26. So YK, Seo MY, Son YI 2012 Prophylactic central lymph node dissection for clinically node-negative papillary thyroid microcarcinoma: influence on serum thyroglobulin level, recurrence rate, and postoperative complications. Surgery **151:**192–198.
27. Wang TS, Evans DB, Fareau GG, Carroll T, Yen TW 2012 Effect of prophylactic central compartment neck dissection on serum thyroglobulin and recommendations for adjuvant radioactive iodine in patients with differentiated thyroid cancer. Ann Surg Oncol **19:**4217–4222.
28. Raffaelli M, De Crea C, Sessa L, Giustacchini P, Revelli L, Bellantone C, Lombardi CP 2012 Prospective evaluation of total thyroidectomy versus ipsilateral versus bilateral central neck dissection in patients with clinically node-negative papillary thyroid carcinoma. Surgery **152:**957–964.
29. Barczyński M, Konturek A, Stopa M, Nowak W 2013 Prophylactic central neck dissection for papillary thyroid cancer. Br J Surg **100:**410–418.

30. Yoo D, Ajmal S, Gowda S, Machan J, Monchik J, Mazzaglia P 2012 Level VI lymph node dissection does not decrease radioiodine uptake in patients undergoing radioiodine ablation for differentiated thyroid cancer. World J Surg **36:**1255–1261.
31. Giordano D, Valcavi R, Thompson GB, Pedroni C, Renna L, Gradoni P, Barbieri V 2012 Complications of central neck dissection in patients with papillary thyroid carcinoma: results of a study on 1087 patients and review of the literature. Thyroid **22:**911–917.
32. Henry JF, Gramatica L, Denizot A, Kvachenyuk A, Puccini M, Defechereux T 1998 Morbidity of prophylactic lymph node dissection in the central neck area in patients with papillary thyroid carcinoma. Langenbecks Arch Surg **383:**167–169.
33. Steinmüller T, Klupp J, Wenking S, Neuhaus P. 1999 Complications associated with different surgical approaches to differentiated thyroid carcinoma. Langenbecks Arch Surg **384:**50–53.
34. Wada N, Duh QY, Sugino K, Iwasaki H, Kameyama K, Mimura T, Ito K, Takami H, Takanashi Y. 2003 Lymph node metastasis from 259 papillary thyroid microcarcinomas: frequency, pattern of occurrence and recurrence, and optimal strategy for neck dissection. Ann Surg **237:**399–407.
35. Gemsenjäger E, Perren A, Seifert B, Schüler G, Schweizer I, Heitz PU 2003 Lymph node surgery in papillary thyroid carcinoma. J Am Coll Surg **197:**182–190.
36. Palestini N, Borasi A, Cestino L, Freddi M, Odasso C, Robecchi A. 2008 Is central neck dissection a safe procedure in the treatment of papillary thyroid cancer? Our experience. Langenbecks Arch Surg **393:**693–698.
37. Davidson HC, Park BJ, Johnson JT 2008 Papillary thyroid cancer: controversies in the management of neck metastasis. Laryngoscope **118:**2161–2165.
38. Hu W, Shi JY, Sheng Y, Li L 2008 Application of central lymph node dissection to surgical operation for clinical stage n0 papillary thyroid carcinoma. Chin J Cancer **27:**304–306.
39. Sadowski BM, Snyder SK, Lairmore TC 2009 Routine bilateral central lymph node clearance for papillary thyroid cancer. Surgery **146:**696–703; discussion 703–705.
40. Besic N, Zgajnar J, Hocevar M, Petric R 2009 Extent of thyroidectomy and lymphadenectomy in 254 patients with papillary thyroid microcarcinoma: a single-institution experience. Ann Surg Oncol **16:**920–928.
41. Rosenbaum MA, McHenry CR 2009 Central neck dissection for papillary thyroid cancer. Arch Otolaryngol Head Neck Surg **135:**1092–1097.
42. Giles Y, Tunca F, Boztepe H, Alagöl F, Terzioglu T, Tezelman S 2009 The long term outcome of papillary thyroid carcinoma patients without primary central lymph node dissection: expected improvement of routine dissection. Surgery **146:**1188–1195.
43. Bonnet S, Hartl D, Leboulleux S, et al. 2009 Prophylactic lymph node dissection for papillary thyroid cancer less than 2 cm: implications for radioiodine treatment. J Clin Endocrinol Metab **94:**1162–1167.
44. Lim YC, Choi EC, Yoon YH, Kim EH, Koo BS 2009 Central lymph node metastases in unilateral papillary thyroid microcarcinoma. Br J Surg **96:**253–257.
45. Shen WT, Ogawa L, Ruan D, Suh I, Duh QY, Clark OH 2010 Central neck lymph node dissection for papillary thyroid cancer: the reliability of surgeon judgment in predicting which patients will benefit. Surgery **148:**398–403.
46. Chung YS, Suh YJ 2010 Is central lymph node dissection mandatory in 2 cm or less sized papillary thyroid cancer? J Korean Surg Soc **79:**332–339.
47. Shindo M, Stern A 2010 Total thyroidectomy with and without selective central compartment dissection: a comparison of complication rates. Arch Otolaryngol Head Neck Surg **136:**584–587.
48. Bozec A, Dassonville O, Chamorey E, Poissonnet G, Sudaka A, Peyrottes I, Ettore F, Haudebourg J, Bussière F, Benisvy D, Marcy PY, Sadoul JL, Hofman P, Lassale S, Vallicioni J, Demard F, Santini J 2011 Clinical impact of cervical lymph node involvement and central neck dissection in patients with papillary thyroid carcinoma: a retrospective analysis of 368 cases. Eur Arch Otorhinolaryngol **268:**1205–1212.
49. Forest VI, Clark JR, Ebrahimi A, Cho EA, Sneddon L, Gao K, O'Brien CJ 2011 Central compartment dissection in thyroid papillary carcinoma. Ann Surg **253:**123–130.
50. Mitra I, Nichani JR, Yap B, Homer JJ 2011 Effect of central compartment neck dissection on hypocalcaemia incidence after total thyroidectomy for carcinoma. J Laryngol Otol **125:**497–501.
51. Kutler DI, Crummey AD, Kuhel WI 2012 Routine central compartment lymph node dissection for patients with papillary thyroid carcinoma. Head Neck **34:**260–263.
52. Hyun SM, Song HY, Kim SY, Nam SY, Roh JL, Han MW, Choi SH 2012 Impact of combined prophylactic unilateral central neck dissection and hemithyroidectomy in patients with papillary thyroid microcarcinoma. Ann Surg Oncol **19:**591–596.
53. Zhang L, Wei WJ, Ji QH, Zhu YX, Wang ZY, Wang Y, Huang CP, Shen Q, Li DS, Wu Y 2012 Risk factors for neck nodal metastasis in papillary thyroid microcarcinoma: a study of 1066 patients. J Clin Endocrinol Metab **97:**1250–1257.
54. Hartl DM, Leboulleux S, Al Ghuzlan A, Baudin E, Chami L, Schlumberger M, Travagli JP 2012 Optimization of staging of the neck with prophylactic central and lateral neck dissection for papillary thyroid carcinoma. Ann Surg **255:**777–783.
55. Lang B, Lo CY, Chan WF, Lam KY, Wan KY 2007 Restaging of differentiated thyroid carcinoma by the sixth edition AJCC/UICC TNM staging system: stage migration and predictability. Ann Surg Oncol **14:**1551–1559.
56. Lang BH, Yih PC, Shek TW, Wan KY, Wong KP, Lo CY 2012 Factors affecting the adequacy of lymph node yield in prophylactic unilateral central neck dissection for papillary thyroid carcinoma. J Surg Oncol **106:**966–971.

Address correspondence to:
Brian H.H. Lang, MS, FRACS
Division of Endocrine Surgery, Department of Surgery
Queen Mary Hospital
102 Pokfulam Road
Hong Kong SAR
China

E-mail: blang@hkucc.hku.hk

(***Appendix follows*** →)

Appendix Table A1. Articles That Were Excluded After Reviewing the Full-Length Text

First author	*Journal*	*Year, country*	*Title*	*Main reason(s) for excluding from analysis*
Henry (32)	Langenbeck's Archives of Surgery	1998, France	Morbidity of prophylactic lymph node dissection in the central neck area in patients with papillary thyroid carcinoma	The TT alone group had patients with benign thyroid disease
Steinmuller (33)	Langenbeck's Archives of Surgery	1999, Germany	Complications associated with different surgical approaches to differentiated thyroid carcinoma	Unable to separate some patients with follicular thyroid carcinoma and some who had therapeutic LND
Wada (34)	Annals of Surgery	2003, Japan	Lymph node metastasis from 259 papillary thyroid microcarcinomas: frequency, pattern of occurrence and recurrence, and optimal strategy for neck dissection	Unable to separate some patients who underwent therapeutic CND
Gemsenjäger (35)	Journal of the American College of Surgeons	2003, Switzerland	Lymph node surgery in papillary thyroid carcinoma	Unable to separate patients with therapeutic LND and lobectomy
Sywak (8)	Surgery	2006, Australia	Routine ipsilateral level VI lymphadenectomy reduces postoperative thyroglobulin levels in papillary thyroid cancer	Data from this study were included in a later study (25)
Palestini (36)	Langenbeck's Archives of Surgery	2008, Italy	Is central neck dissection a safe procedure in the treatment of papillary thyroid cancer? Our experience	Unable to exclude some patients with cN1; also no follow-up and recurrence data were available
Davidson (37)	Laryngoscope	2008, United States	Papillary thyroid cancer: controversies in the management of neck metastasis	Unable to separate patients who had nodal plucking, therapeutic CND, LND, or combination
Hu (38)	Chinese Journal of Cancer	2008, China	Application of central lymph node dissection to surgical operation for clinical stage N0 papillary thyroid carcinoma	No follow-up or recurrence data were reported
Sadowski (39)	Surgery	2009, United States	Routine bilateral central lymph node clearance for papillary thyroid cancer	All patients who underwent CND had cN1
Besic (40)	Annals of Surgical Oncology	2009, Slovenia	Extent of thyroidectomy and lymphadenectomy in 254 patients with papillary thyroid microcarcinoma: a single institution experience	Too few (i.e., <10) patients in the prophylactic arm
Rosenbaum (41)	Archives of Otorhinolaryngology Head Neck Surgery	2009, United States	Central neck dissection for papillary thyroid cancer	Unable to exclude patients with cN1
Giles (42)	Surgery	2009, Turkey	The long term outcome of papillary thyroid carcinoma patients without primary central lymph node dissection: expected improvement of routine dissection	Only included patients who underwent TT without CND
Bonnet (43)	Journal of Clinical and Endocrinology Metabolism	2009, France	Prophylactic lymph node dissection for papillary thyroid cancer less than 2cm: implications for radioiodine treatment	Prophylactic LND were included; also no TT alone group was available for comparison
Lim (44)	British Journal of Surgery	2009, Korea	Central lymph node metastases in unilateral papillary thyroid microcarcinoma	No TT alone group was available for comparison

(*continued*)

Appendix Table A1. (Continued)

First author	*Journal*	*Year, country*	*Title*	*Main reason(s) for excluding from analysis*
Shen (45)	Surgery	2010, United States	Central neck lymph node dissection for papillary thyroid cancer: the reliability of surgeon judgment in predicting which patients will benefit	All patients had therapeutic CND
Chung (46)	Journal of the Korean Surgical Society	2010, Korea	Is central lymph node dissection mandatory in 2 cm or less sized papillary thyroid cancer?	All patients had lobectomy
Shindo (47)	Archives of Otorhinolaryngology Head Neck Surgery	2010, United States	Total thyroidectomy with and without selective central compartment dissection	The TT alone group had benign thyroid disease
Bozec (48)	European Archives of Otorhinolaryngology	2011, France	Clinical impact of cervical lymph node involvement and central neck dissection in patients with papillary thyroid carcinoma: a retrospective analysis of 368 cases	TT alone group was not available; also some patients had therapeutic CND or LND
Forest (49)	Annals of Surgery	2011, Australia	Central compartment dissection in thyroid papillary carcinoma	No TT alone group was available for comparison
Mitra (50)	Journal of Laryngology and Otology	2011, United Kingdom	Effect of central compartment neck dissection on hypocalcaemia incidence after total thyroidectomy for carcinoma	No follow-up or recurrence data reported
Teixeira (6)	Surgery	2011, Brazil	The incidence of central neck micrometastatic disease in patients with papillary thyroid cancer staged preoperatively and intraoperatively as N0	No TT alone group for comparison
Kutler (51)	Head and Neck	2012, United States	Routine central compartment lymph node dissection for patients with papillary thyroid carcinoma	Unable to exclude patients who had hemithyroidectomy and CND; also some therapeutic LND were included
Hyun (52)	Annals of Surgical Oncology	2012, Korea	Impact of combined prophylactic unilateral central neck dissection and hemithyroidectomy in patients with papillary thyroid microcarcinoma	All patients underwent hemithyroidectomy
Zhang (53)	Journal of Clinical Endocrinology and Metabolism	2012, China	Risk factors for neck nodal metastasis in papillary thyroid microcarcinoma: a study of 1066 patients	All patients underwent hemithyroidectomy
Hartl (54)	Annals of Surgery	2012, France	Optimization of staging of the neck with prophylactic central and lateral neck dissection for papillary thyroid carcinoma	All patients routinely underwent prophylactic CND and ipsilateral LND
Yoo (55)	World Journal of Surgery	2012, United States	Level VI lymph node dissection does not decrease radioiodine uptake in patients undergoing radioiodine ablation for differentiated thyroid cancer	Unable to separate some patients who underwent therapeutic CND
Giordano (56)	Thyroid	2012, Italy	Complications of central neck dissection in patents with papillary thyroid carcinoma: results of a study on 1087 patients and review of the literature	No follow-up or recurrence data reported

TT, total thyroidectomy; cN1, clinically involved lymph node metastases; LND, lateral neck dissection; CND, central neck dissection.

Reprinted by permission of Surgery. 2012; 151(6):844-850.

Risk of second primary malignancy in differentiated thyroid carcinoma treated with radioactive iodine therapy

Brian Hung-Hin Lang, MS, FRACS,[a] Irene Oi Ling Wong, PhD,[b] Kai Pun Wong, MBBS, MRCS,[a] Benjamin J. Cowling, PhD,[b] *and* **Koon-Yat Wan, MBBS, FRCR,[c]** *Hong Kong, China*

Background. *Differentiated thyroid cancer survivors are at increased risk of nonsynchronous second primary malignancy, but the cause remains unclear. This study aimed to evaluate the association between radioiodine therapy and risk of nonsynchronous second primary malignancy and to examine whether the risk of nonsynchronous second primary malignancy in differentiated thyroid cancer survivors treated with radioiodine therapy is increased relative to the general population.*

Methods. *Among 895 radiation-naïve patients with differentiated thyroid cancer, 643 (71.8%) received ≥1 course of radioiodine therapy (radioiodine therapy–positive group) and 252 (28.2%) received no radioiodine therapy (radioiodine therapy–negative group). After a median follow-up of 93.5 months (range, 23.4–570.8), 64 (7.2%) patients developed ≥1 nonsynchronous second primary malignancy. Potential risk factors for nonsynchronous second primary malignancy were entered into a multivariable regression model and cancer incidence in the radioiodine therapy–positive and –negative groups were compared to that of the general population by estimating the standardized incidence ratios.*

Results. *The 20-year cumulative nonsynchronous second primary malignancy risk in radioiodine therapy–positive group was significantly higher than radioiodine therapy–negative group (13.5% vs 3.1%;* P = *.015). Cumulative radioiodine therapy activity of 3.0 to 8.9 GBq (relative risk, 2.77; 95% CI, 1.079–7.154;* P = *.034) was the only independent risk factor for nonsynchronous second primary malignancy after adjusting for age, sex, period of differentiated thyroid cancer diagnosis, and stage of differentiated thyroid cancer. For females, the standardized incidence ratio in the radioiodine therapy–positive group was 1.54 (95% CI, 1.11–2.08) and in the radioiodine therapy–negative group it was 0.92 (95% CI, 0.37–1.90).*

Conclusion. *Differentiated thyroid cancer female survivors treated by radioiodine therapy appeared to be at elevated risk of nonsynchronous second primary malignancy when compared to the general population and this risk was not apparent in those not previously treated by radioiodine therapy. (Surgery 2012;151:844-50.)*

From the Department of Surgery,[a] the School of Public Health,[b] and the Department of Clinical Oncology,[c] The University of Hong Kong, Hong Kong, China

Differentiated thyroid carcinoma (DTC) accounts for more than 90% of all follicular cell–derived thyroid malignancies and is the most common primary endocrine-related malignancy. In our locality, its age-adjusted incidence has doubled over the last 25 years, and a similar trend has been observed elsewhere.[1] Despite this, the disease-specific mortality remains low, with an overall 10-year disease-specific survival above 90%.[2] However, because this cancer affects mostly relatively young patients, the lifetime risk of developing nonsynchronous second primary malignancy (NSPM) poses a real concern.[3] In agreement with other studies, our previous analysis found that DTC survivors were at greater NSPM risk than that of the general population (relative risk [RR], 1.39; 95% confidence interval [CI], 1.09–1.73).[4-7] In addition, those survivors who eventually developed NSPM had a significantly poorer overall survival than those who did not.[4] While the occurrence of NSPM appeared to adversely affect the survival of DTC survivors,[4] the likely causes for the increased risk of NSPM in DTC survivors remain uncertain. Possibilities include the exposure to ionizing radiation from radioiodine therapy (RAI) or external local radiotherapy (ERT), common environmental or dietary factors, genetic predisposition,

Accepted for publication December 22, 2011.

Reprint requests: Brian Hung-Hin Lang, MS, FRACS, Division of Endocrine Surgery, Department of Surgery, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China. E-mail: blang@hkucc.hku.hk.

0039-6060/$ - see front matter

doi:10.1016/j.surg.2011.12.019

and surveillance bias.[5,8] A recent meta-analysis comprising more than 16,000 patients found that the risk of developing NSPM in DTC survivors treated with RAI therapy was increased compared to those not treated with RAI.[9] However, this meta-analysis, because of its relative strict inclusion criteria, only included 2 population-based studies, both of which revealed an increased risk of NSPM in those treated by RAI compared to those not treated by RAI.[5,6] There have been other studies, not included in the meta-analysis, that failed to reveal an association between cumulative RAI activity and risk of NSPM and no risk difference associated with RAI treatment.[8,10,11]

Given the existing conflicting evidence—and in light of our previous analysis showing an increased risk of NSPM in DTC survivors[4]—the aims of the present study were to evaluate the association between RAI therapy and development of NSPM and to examine whether the risk of NSPM in radiation-naïve DTC survivors treated with RAI is increased relative to that of the general population.

METHODS

Patients. Between 1971 and 2009, 1,122 patients with DTC were managed at our institution. Of these, 98 (8.7%) had clinically occult microcarcinoma, 41 (3.7%) had a documented history of radiotherapy or radiation exposure before the diagnosis of DTC, and 88 (7.8%) received ERT as adjuvant treatment for DTC. For the purpose of the present study, they were excluded, and therefore a total of 895 radiation-naïve patients were eligible for analysis. All eligible patients had at least 1 year of follow-up. There were 643 patients who received at least 1 course of RAI (RAI^+ group) and 252 who received no RAI at all during the study period (RAI^- group).

Methods. The present study protocol was approved by the local institutional review board. The protocol for I^{131} or RAI ablation remained unchanged throughout the study period, and details were described previously.[12] Patients with ≥1 risk factors, such as tumor size >1 cm, lymph node metastasis, age >40 years, presence of extrathyroidal extension, macroscopic postoperative residual disease in the neck, or distant metastasis, were considered for RAI ablation 8–10 weeks after thyroidectomy by either T4 withdrawal or the use of recombinant thyroid-stimulating hormone. Diagnostic whole body I^{131} scans were performed approximately 6 months after RAI therapy. Three giga-Becquerels (GBq) or 80 millicuries (mCi) I^{131} were administered as a standard ablative dose for all postsurgical patients, but higher doses were considered in the presence of extensive lymph node involvement or distant metastasis. Subsequent RAI therapy of 5.5 GBq (or 150 mCi) was administered periodically at 4- to 6-month intervals until uptake was no longer visible or disease progressed despite treatment. The cumulative RAI dose or activity for each individual patient was calculated. Although the above protocol was closely followed throughout the study period, individual patient preference was considered and respected.

To ensure an accurate and updated follow-up status of all patients, a careful manual search of all patients' status in the territory-wide Clinical Management System (CMS) was performed. The CMS is a computerized database linking all 41 public hospitals in Hong Kong that provides inpatient medical records corresponding to more than 90% of inpatient bed days in the region.[13] Specific variables including the latest date of follow-up or the date of death, date of birth, cause of death, diagnosis date, and type of second nonthyroidal primary malignancy were retrieved. Clinicopathologic data and management details relating to the DTC were prospectively collected since 1995. As of January 2011, 805 (81.9%) were still alive and being monitored. The other 178 patients died; in 66 cases, DTC was the cause of death.

Statistical analysis. For patients who developed ≥2 nonthyroidal primary malignancies after DTC, only the earliest occurred malignancy was recorded. The time to developing a second primary malignancy was calculated from the date of DTC diagnosis to the diagnosis date of the second malignancy. A second malignancy that occurred within 12 months of the date of DTC diagnosis was considered synchronous and was excluded from analysis. The time at risk for NSPM was calculated from the date of DTC to the date of NSPM, the date of death, or the date of last follow-up, whichever came first. To evaluate the relationship between patient characteristics, treatment, tumor stage, and risk of NSPM, 2 approaches were used. First, the cumulative proportion of NSPM as a function of time after DTC diagnosis was estimated using the Kaplan–Meier method. The relations between the time of NSPM occurrence and concomitant variables, such as sex, age, period of DTC diagnosis, tumor stage, ERT, and RAI therapy was assessed using the Cox proportional hazards regression model, which accounts for the length of follow-up. Second, the cancer incidence of both RAI^+ and RAI^- groups were compared to that of the general population by calculating the standardized incidence ratios (SIRs) for all sites/types of NSPMs. The SIRs of NSPM after DTC were

calculated by dividing observed incidence rates in our cohort by expected incidence rates in the general population. The observed number of each NSPM was calculated by compiling the person-years of observation according to 5-year age groups by sex from the diagnosis of DTC to the date of death, date of last follow-up, or date of diagnosis of NSPM, whichever occurred first. The expected number of each NSPM in 5-year age groups by sex were derived from cancer incidence statistics in 2008 reported by the Hong Kong Cancer Registry (http://www3.ha.org.hk/cancereg/) and were multiplied by the accumulated person-years of observation at risk to obtain the expected number of NSPM.

For the comparison for dichotomous variables between the RAI^+ and RAI^- groups, χ^2 and Fisher exact tests were used. The Mann–Whitney *U* test was used for the comparison of continuous variables between RAI^+ and RAI^- groups. All statistical analyses were conducted using SPSS software (version 18.0; SPSS, Inc., Chicago, IL).

RESULTS

Table I shows the baseline patient characteristics. Six hundred ninety-five (77.7%) patients had papillary thyroid carcinoma and 200 (22.3%) had follicular thyroid carcinoma. The majority were female (80.6%) and ethnic Chinese (94.1%). The median age of DTC diagnosis was 44.0 years (range, 7.1–90.6), and the median follow-up period was 93.5 months (range, 23.1–570.8). Seven hundred sixty-three (85.3%) patients underwent bilateral thyroid resection, and of these, 643 (84.3%) patients received at least 1 course of RAI therapy, whereas of the 132 (14.7%) patients who underwent unilateral thyroid resection, no patient received RAI therapy. There were 55 (6.1%) patients who received ≥2 doses of RAI therapy. Table II shows a comparison of demographics, period of DTC diagnosis, number and type of NSPM, major histologic types of DTC, and stage of DTC between the RAI^+ and RAI^- groups. Patients in the RAI^+ group were significantly older at the time of DTC diagnosis (47.5 vs 44.0; $P < .001$) and there was a significantly greater proportion of patients belonging to the ≥50-year-old age group (39.9% vs 28.2%; $P < .001$). There were a similar proportion of males in the 2 groups. When the periods of DTC diagnosis were compared, there were a significantly greater proportion of patients not receiving RAI treatment towards the later period. Tumor size was significantly larger in the RAI^+ group because size was an important criterion for RAI ablation. Similarly, the stages of DTC were more advanced in the RAI^+ group as reflected by the higher proportion of stage III and IV patients (18.8% vs 6.0% and 15.6% vs 5.6%, respectively; $P < .001$). In terms of the type of NSPM, primary breast, colon, and lung cancers were the 3 most common NSPMs in the RAI^+ and RAI^- groups.

Table I. Baseline patient demographics and characteristics ($n = 895$)

	Median	*Range or %*
Age at diagnosis of DTC, y	44.0	7.1–90.6
Sex		
Male	174	19.4
Female	721	80.6
Major histologic types of DTC		
Papillary	695	77.7
Follicular	200	22.3
Tumor stage of DTC by TNM		
I	586	65.5
II	59	6.6
III	136	15.2
IV	114	12.7
Follow-up time, months	93.5	23.4–570.8
No. of patients with NSPM* detected during follow-up	64	7.2
Latency period to NSPM* from time of DTC, months	189.5	22.8–531.1
Radiation exposure		
No RAI therapy given	252	28.2
RAI therapy given	643	71.8
1 course	588	65.7
2 courses	26	2.9
≥3 courses	29	3.2

*Only second primary malignancy which occurred >12 months after the diagnosis of DTC was included.
DTC, Differentiated thyroid carcinoma; *NSPM*, nonsynchronous second primary malignancy; *RAI*, radioactive iodine; *TNM*, American Joint Cancer Committee/Union Internationale Contre le Cancer tumor-nodes-metastasis staging system, 6th edition.

The Figure shows the cumulative risk of developing NSPM after the diagnosis of DTC in the RAI^+ and RAI^- groups. The 20-year cumulative risk of NSPM in the RAI^+ group was significantly higher than that in the RAI^- group (13.5% vs 3.1%; $P = .015$). The mean times to development of NSPM in the RAI^+ and RAI^- groups were 34.37 years (95% CI, 32.6–36.1) and 43.05 years (95% CI, 40.3–45.8).

Table III shows the Cox proportional hazards analysis of factors influencing the development of NSPM in patients with DTC. The analysis included variables that were significant in the comparison between RAI^+ and RAI^- groups and factors that might be linked to development of NSPM. Tumor size was not entered because the stages of DTC by

Table II. A bivariable comparison of demographics, type of second primary malignancies, histology of thyroid carcinoma, and TNM stages between those who did and did not receive radioiodine ablation

	RAI^+ group (n = *643*)	*RAI^- group* (n = *252*)	P *value**
Median age of DTC diagnosis	47.5 (19.3–88.8)	44.0 (7.1–90.6)	<.001
Age of DTC by groups, y			.002
<30	118 (18.4)	66 (26.2)	
30–49	280 (43.5)	118 (46.8)	
≥50	245 (38.1)	68 (27.0)	
Sex (male/female)	133/510	41/211	.133
Period of DTC diagnosis			<.001
Before 1980	72 (11.2)	42 (16.7)	
1980–1999	271 (42.1)	77 (30.6)	
After 2000	300 (46.7)	133 (52.8)	
Tumor size of DTC, cm	3.0 (0.1–11.0)	2.0 (0.1–7.0)	<.001
Histological type of DTC			.111
Papillary	505 (78.5)	190 (75.6)	
Follicular	138 (21.5)	62 (24.6)	
Stage of DTC by TNM			<.001
I	377 (58.6)	209 (82.9)	
II	45 (7.0)	14 (5.6)	
III	121 (18.8)	15 (6.0)	
IV	100 (15.6)	14 (5.6)	
Type/site of NSPM†			
All sites	56 (8.7)	8 (3.2)	.004
Breast	13 (2.0)	2 (0.8)	.120
Colon	9 (1.4)	1 (0.4)	.468
Lung	4 (0.6)	1 (0.4)	1.000
Liver	3 (0.5)	1 (0.4)	1.000
Corpus uteri	3 (0.5)	1 (0.4)	1.000
Stomach	3 (0.5)	0 (0.0)	.567
Non-Hodgkin lymphoma	3 (0.5)	0 (0.0)	.567
Rectum	2 (0.3)	1 (0.4)	1.000
Cervix	2 (0.3)	1 (0.4)	1.000

**P* values were generated by using bivariable tests including χ^2, Fisher exact, and Mann–Whitney *U* tests wherever appropriate.
†Only nonsynchronous second primary malignancy with a total number ≥3 was listed.
DTC, Differentiated thyroid carcinoma; *NSPM*, nonsynchronous second primary malignancy; *TNM*, American Joint Cancer Committee/Union Internationale Contre le Cancer tumor-nodes-metastasis staging system, 6th edition.

the American Joint Cancer Committee/Union Internationale Contre le Cancer tumor-node-metastasis staging system, 6th edition classification already incorporated tumor size. The following variables were entered in the final model: age groups, sex, period of DTC diagnosis, cumulative RAI activity, and stage of DTC. Variables that were significantly associated with an increased risk of NSPM were cumulative RAI activity equaled from 3.0 to 8.9 GBq (RR, 2.777; 95% CI, 1.079–7.145; P = .034). Cumulative RAI activity of >9.0 GBq was not significantly associated with risk of NSPM (RR, 3.149; 95% CI, 0.645–12.816; P = .131).

In this cohort, the total person-years of observation at risk were 10,414. After a median follow-up of 93.5 months (range, 23.4–570.8), 62 (6.9%) patients developed 1 NSPM and 2 (0.2%) patients developed 2 NSPMs (ie, 2 separate primary malignancies >12 months after DTC). Overall, 64 patients with NSPM were observed (15 males and 49 females). The median latency period from DTC to NSPM was 189.5 months (range, 22.8–531.1). A total of 15 of 64 (23.4%) patients developed NSPM in the 5 years of follow-up. The median (range) age of NSPM was 65.6 (23.0–95.5) years old. None had known hereditary or familial cancer syndromes. In males, the 3 most common types/sites for NSPM (in descending order of frequency) were colon (n = 3), prostate (n = 3), and liver (n = 2). In females, the 3 commonest types or sites of NSPM (in descending order of frequency) were breast (n = 13), colon (n = 7), and uterus (n = 4). Table IV shows the observed and expected number of cases and SIRs of NSPM in the RAI^+ and RAI^- groups for males, females, and both sexes. When compared to the incidence rate in the general population, after adjusting for age

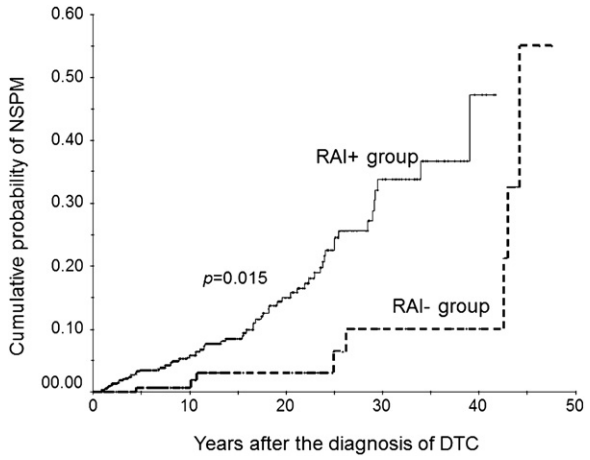


Figure. The cumulative risk curves of developing nonsynchronous second primary malignancy (NSPM) after the diagnosis of differentiated thyroid carcinoma (DTC) in those who received radioiodine therapy (RAI^+ group) and those who received no radioiodine therapy (RAI^- group).

and sex, the incidence or risk of developing NSPM in RAI^+ group was significantly higher, but in the RAI^- group, the incidence or risk appeared not significantly different to the general population. In the RAI^+ group, for both sexes, the SIR was 1.51 (95% CI, 1.14–1.96). For males in the RAI^+ group, the SIR was 1.41 (95% CI, 0.77–2.37), and for females the SIR was 1.54 (95% CI, 1.11–2.02). In contrast, for males in the RAI^- group, the SIR was 0.53 (95% CI, 0.01–2.96), and for females the SIR was 0.92 (95% CI, 0.37–1.90).

DISCUSSION

Previous studies found that the overall lifetime risk of developing NSPM was higher in DTC survivors by up to 30% to 40% more than that of the general population.[4,14] The present study, unlike our previous studies, was aimed at specifically evaluating whether RAI therapy was a potential risk factor for NSPM in a cohort of radiation-naïve DTC survivors.[4,15] Other proposed risk factors included exposure of ionizing radiation during DTC treatment, common environmental and dietary factors, genetic predisposition, and surveillance bias.[5,8] However, ionizing radiation exposure from treatment of DTC remains one of the most likely culprits for NSPM because RAI therapy is commonly administered in DTC either in the setting of thyroid remnant ablation after total or near-total thyroidectomy or in the setting of recurrence or metastasis.[5,6,16,17] As a result, more selective use of RAI therapy in

Table III. Cox proportional hazards analysis of factors for the development of nonsynchronous second primary malignancy in differentiated thyroid carcinoma

Covariates	*Relative risk*	*95% CI*	P *value*
Age of DTC by groups, y			
<30	Reference		
30–49	1.468	0.690–3.121	.319
≥50	1.704	0.910–5.085	.263
Sex			
Female	Reference		
Male	1.299	0.695–2.430	.413
Period of DTC diagnosis			
Before 1980	Reference		
1980–1999	1.676	0.840–3.346	.143
After 2000	1.717	0.601–4.910	.313
Cumulative RAI activity, GBq			
None	Reference		
3–8.9	2.777	1.079–7.145	.034
>9.0	3.149	0.645–12.816	.131
Stage of DTC by TNM			
I	Reference		
II	1.678	0.764–3.627	.162
III	1.513	0.681–3.364	.309
IV	1.760	0.781–3.969	.173

CI, Confidence interval; *DTC*, differentiated thyroid carcinoma; *NSPM*, nonsynchronous second primary malignancy; *RAI*, radioactive iodine (I^{131}); *TNM*, American Joint Cancer Committee/Union Internationale Contre le Cancer tumor-nodes-metastasis staging system, 6th edition.

DTC, particularly for low-risk tumors, has been increasingly advocated.[18] Although I^{131} is preferentially taken up by the normal and malignant thyroid follicular cells, particularly under the hypothyroid state, it is also taken up and accumulated into the stomach, salivary glands, colon, and bladders; these sites are often exposed to prolonged radiation. Interestingly, these sites were reported to the common sites for NSPM in DTC survivors.[5,15] In the present study, because only 2 female patients belonging to the RAI^+ group developed stomach and bladder cancers during the study period (data not shown), respectively, it was difficult to know whether these tumors were actually related to RAI therapy or occurred by chance. Calculating the SIR value for these 2 primary tumors was also not possible. Nevertheless, similar to previous studies, we did find that breast cancer was one of the most common NSPMs in DTC survivors. In addition, the relative frequency of the 3 most common NSPMs (ie, breast, colon, and lung) observed in the present study appeared similar to the frequency observed in other population studies.[5,6] Their frequencies ranged between 0.1% and 0.2% per 1 person-year of observation.[6] Possible explanations

Table IV. Observed and expected number of cases and standardized incidence ratio with the corresponding 95% CIs of nonsynchronous second primary malignancy for those who did and did not receive radioiodine in males, females, and both sexes*

	Observed no. of NSPMs	*Expected number of NSPMs*	*Standardized incidence ratio*	*95% CI*
RAI^+ group				
Males	14	9.90	1.41	0.77–2.37
Females	42	27.28	1.54	1.11–2.08
Both sexes	56	37.18	1.51	1.14–1.96
RAI^- group				
Male	1	1.88	0.53	0.01–2.96
Female	7	7.59	0.92	0.37–1.90
Both sexes	8	9.47	0.84	0.36–1.66

*Expected numbers of NSPM were based on population incidence of all sites in 2008 after adjusting for age.
CI, Confidence interval; *NSPM*, nonsynchronous second primary malignancy; *RAI*, radioactive iodine (I^{131}).

for the association between DTC and breast cancer included the presence of sodium iodide symporters in mammary tissue leading to cumulative doses of I^{131}, the female dominance in DTC survivors, genetic predisposition, and surveillance bias.[10,19]

Nevertheless, when all types/sites of NSPM were considered, our data are consistent with the hypothesis that RAI therapy increased the overall risk of developing NSPM in DTC survivors. In the multivariable analysis, after adjusting for other possible factors linked to the risk of NSPM, a cumulative RAI activity of 3–8.9 GBq was found to be an independent factor for the development of NSPM in DTC survivors, and the relative increased risk was approximately 2–3 times higher than survivors who received no RAI. However, because the RR was not significant in the group who had cumulative RAI activity >9.0, the present study was not able to establish a dose–effect relationship between cumulative RAI activity and risk of NSPM. This might have been because only 29 patients had cumulative RAI activity >9.0 GBq, and our study was therefore underpowered to identify an effect. Another explanation might have been related to the disease threshold phenomenon, where the risk of RAI activity ≥9.0 GBq far exceeded the dose threshold and so imparted the same risk as RAI activity of 3.0–8.9 GBq. To further confirm that RAI therapy was the factor responsible for the increased risk of NSPM and not factors that influenced the decision to prescribe RAI therapy in the first place, the present analysis compared factors that might have influenced the decision to prescribe RAI (such as the stage of DTC, period of DTC diagnosis, and age of DTC), and these significant factors were entered into the multivariable analysis together with other well-reported risk factors, such as sex.[17,20] Although the male sex was not an independent factor for NSPM, 2 previous studies found that male survivors were at increased risk of NSPM.[17,20] Therefore, male DTC survivors treated with RAI might be at even greater risk of NSPM. Nevertheless, our SIR analysis in NSPM did not support this finding with the males in the RAI^+ group having a slightly lower SIR value than females in the RAI^+ group after adjusting for age (1.41 vs 1.54). However, there appeared to be a strong association between breast cancer and DTC; the higher risk of NSPM observed in RAI^+ female patients might have been a result of this association.[4-6] Unlike other studies, our data showed that the risk of NSPM in the RAI^- group was similar to that of the general population.[6,8,11,17] Perhaps this further strengthened the association between RAI therapy and risk of NSPM in DTC survivors. However, because <30% of DTC patients received no RAI, our study might have been underpowered in this aspect.

One of the strengths of the present study was the complete patient follow-up status and data collection, and this was not possible without the establishment of the territory-wide CMS in 1995.[21] Unlike previous larger-scale analyses, all patients in the present study were managed under a standardized treatment protocol, and the histology of DTC was confirmed by the same group of pathologists at our institution. Given the completeness of the clinical data, the chance of NSPM misclassification would have been minimal, and our SIR results were already adjusted for both patient age and sex. However, similar to previous single-center analysis, the total number of NSPMs remained relatively small, which limited the power of our study to identify smaller effects. Also, the present analysis was potentially subject to a degree of institution and referral biases, because our center is an academic tertiary care institute to which more complex cases are often referred from other hospitals. Because there were a number of significant differences in baseline clinical and demographics between RAI^+ and RAI^- groups, they might not have been fully adjusted by the multivariate analysis. A future multicenter study involving other institutions in our territory would be desirable to further confirm our findings. Although none of the 64 patients with NSPM had known

hereditary/familial cancer syndromes, genetic testing for specific mutations, such as the CHEK2, PTEN, MLH1, and MSH2 genes, along with the consideration of other lifestyle factors, such as smoking, physical activity, and diet, would have been useful.[22,23] Because 15 of 64 (23.4%) of the NSPMs were diagnosed in the first 5 years after DTC, the authors could not exclude the possibility that some might have been present at or shortly after DTC and might have not been a result of the radiation effect of RAI (ie, surveillance bias).

From these data, after adjusting for potential risk factors in the multivariable analysis, such as age, gender, period of DTC diagnosis, and DTC stage, cumulative RAI activity of 3 to 8.9 GBq was the only independent risk factor for NSPM in radiation-naïve DTC survivors. The risk of developing NSPM in the female RAI^+ group was significantly higher than that of the general population, but this increased risk was not observed in the RAI^- group. Therefore, the authors concluded that female DTC survivors treated by RAI appeared to be at elevated risk of developing NSPM when compared to the general population and an excess risk was not apparent in those survivors not previously treated by RAI.

REFERENCES

1. Cancer incidence and mortality in Hong Kong 1983–2006. Hong Kong Cancer Registry, Hong Kong. Available from: http://www3.ha.org.hk/cancereg/e_stat.asp.
2. Lang BH, Lo CY, Chan WF, Lam KY, Wan KY. Prognostic factors in papillary and follicular thyroid carcinoma: implications for cancer staging. Ann Surg Oncol 2007;14:730-8.
3. Hay ID, Gonzalez-Losada T, Reinalda MS, Honetschlager JA, Richards ML, Thompson GB. Long-term outcome in 215 children and adolescents with papillary thyroid cancer treated during 1940 through 2008. World J Surg 2010;34:1192-202.
4. Lang BH, Lo CY, Wong IO, Cowling BJ. Impact of second primary malignancy on outcomes of differentiated thyroid carcinoma. Surgery 2010;148:1191-6.
5. Rubino C, de Vathaire F, Dottorini ME, Hall P, Schvartz C, Couette JE, et al. Second primary malignancies in thyroid cancer patients. Br J Cancer 2003;89:1638-44.
6. Brown AP, Chen J, Hitchcock YJ, Szabo A, Shrieve DC, Tward JD. The risk of second primary malignancies up to three decades after the treatment of differentiated thyroid cancer. J Clin Endocrinol Metab 2008;93:504-15.
7. Sandeep TC, Strachan MW, Reynolds RM, Brewster DH, Scelo G, Pukkala E, et al. Second primary cancers in thyroid cancer patients: a multinational record linkage study. J Clin Endocrinol Metab 2006;91:1819-25.
8. Berthe E, Henry-Amar M, Michels JJ, et al. Risk of second primary cancer following differentiated thyroid cancer. Eur J Nucl Med Mol Imaging 2004;31:685-91.
9. Sawka AM, Thabane L, Parlea L, et al. Second primary malignancy risk after radioactive iodine treatment for thyroid cancer: a systematic review and meta-analysis. Thyroid 2009;19:451-7.
10. Verkooijen RB, Smit JW, Romijn JA, Stokkel MP. The incidence of second primary tumors in thyroid cancer patients is increased, but not related to treatment of thyroid cancer. Eur J Endocrinol 2006;155:801-6.
11. Canchola AJ, Horn-Ross PL, Purdie DM. Risk of second primary malignancies in women with papillary thyroid cancer. Am J Epidemiol 2006;15:521-7.
12. Lang BH, Lo CY, Chan WF, Lam KY, Wan KY. Staging systems for papillary thyroid carcinoma: a review and comparison. Ann Surg 2007;245:266-78.
13. Wong IO, Chan WS, Choi S, Lo SV, Leung GM. Moral hazard or realised access to care? Empirical observation in Hong Kong. Health Policy 2006;75:251-61.
14. Subramanian S, Goldstein DP, Parlea L, et al. Second primary malignancy risk in thyroid cancer survivors: a systematic review and meta-analysis. Thyroid 2007;17:1277-88.
15. Ronckers CM, McCarron P, Ron E. Thyroid cancer and multiple primary tumors in the SEER cancer registries. Int J Cancer 2005;117:281-8.
16. Lang BH, Wong KP. Risk factors for nonsynchronous second primary malignancy and related death in patients with differentiated thyroid carcinoma. Ann Surg Oncol 2011;18:3559-65.
17. Chuang SC, Hashibe M, Yu GP, et al. Radiotherapy for primary thyroid cancer as a risk factor for second primary cancers. Cancer Lett 2006;238:42-52.
18. Cooper DS, Doherty GM, Hauger BR, Kloos RT, Lee SL, Mandel SJ, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2009;19:1167-214.
19. Kogai T, Taki K, Brent GA. Enhancement of sodium/iodide symporter expression in thyroid and breast cancer. Endocr Relat Cancer 2006;13:797-826.
20. Bhattacharyya N, Chien W. Risk of second primary malignancy after radioactive iodine treatment for differentiated thyroid carcinoma. Ann Otol Rhinol Laryngol 2006;115:607-10.
21. Presentation by Shane Solomon, the Chief Executive of Hospital Authority. Available from http://www.ha.org.hk/upload/presentation/47.pdf.
22. Cybulski C, Górski B, Huzarski T, et al. CHEK2 is a multiorgan cancer susceptibility gene. Am J Hum Genet 2004;75:1131-5.
23. Wei EK, Wolin KY, Colditz GA. Time course of risk factors in cancer etiology and progression. J Clin Oncol 2010;28:4052-7.

THYROID
Volume 23, Number 7, 2013

DOI: 10.1089/thy.2013.0045

The Increasing Incidence of Thyroid Cancer: The Influence of Access to Care

Luc G.T. Morris,[1] Andrew G. Sikora,[2] Tor D. Tosteson,[3] and Louise Davies[4,5]

Background: The rapidly rising incidence of papillary thyroid cancer may be due to overdiagnosis of a reservoir of subclinical disease. To conclude that overdiagnosis is occurring, evidence for an association between access to health care and the incidence of cancer is necessary.
Methods: We used Surveillance, Epidemiology, and End Results (SEER) data to examine U.S. papillary thyroid cancer incidence trends in Medicare-age and non–Medicare-age cohorts over three decades. We performed an ecologic analysis across 497 U.S. counties, examining the association of nine county-level socioeconomic markers of health care access and the incidence of papillary thyroid cancer.
Results: Papillary thyroid cancer incidence is rising most rapidly in Americans over age 65 years (annual percentage change, 8.8%), who have broad health insurance coverage through Medicare. Among those under 65, in whom health insurance coverage is not universal, the rate of increase has been slower (annual percentage change, 6.4%). Over three decades, the mortality rate from thyroid cancer has not changed. Across U.S. counties, incidence ranged widely, from 0 to 29.7 per 100,000. County papillary thyroid cancer incidence was significantly correlated with all nine sociodemographic markers of health care access: it was positively correlated with rates of college education, white-collar employment, and family income; and negatively correlated with the percentage of residents who were uninsured, in poverty, unemployed, of nonwhite ethnicity, non-English speaking, and lacking high school education.
Conclusion: Markers for higher levels of health care access, both sociodemographic and age-based, are associated with higher papillary thyroid cancer incidence rates. More papillary thyroid cancers are diagnosed among populations with wider access to healthcare. Despite the threefold increase in incidence over three decades, the mortality rate remains unchanged. Together with the large subclinical reservoir of occult papillary thyroid cancers, these data provide supportive evidence for the widespread overdiagnosis of this entity.

Introduction

Thyroid cancer is currently the third fastest rising cancer diagnosis in the United States. Estimates in the last decade placed the annual rate of increase at 3%, resulting in a doubling of thyroid cancer incidence in 30 years (1–4). Similar patterns of increase have been reported in Canada, Australia, and Western Europe (5–8). The causes of this so-called "thyroid cancer epidemic" are not completely understood (9).

The rising papillary thyroid cancer incidence rate may represent either a true increase in the occurrence of disease or an increasing number of diagnoses due to escalating levels of diagnostic scrutiny (1–3,10). With more widespread use of ultrasonography and fine-needle aspiration biopsy and with many radiographic "incidentalomas" discovered on nonthyroid imaging, a larger number of clinically occult, small thyroid nodules are being detected and investigated (1,9,11). These incidentalomas may exemplify the epidemiologic term "overdiagnosis," which postulates that the rising number of diagnoses reflects more effective detection of a subclinical reservoir of cancers, which would not have caused symptoms or death, if left undetected (12).

There are two prerequisites for concluding that overdiagnosis of a disease is occurring: there must be (i) a large reservoir of occult disease and (ii) increasing health care activities leading to the detection of the disease reservoir (12). There is strong evidence for the first condition, with the prevalence of occult papillary thyroid cancer at autopsy

[1]Head and Neck Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York.
[2]Department of Otolaryngology, Icahn School of Medicine at Mount Sinai, New York, New York.
[3]Section of Biostatistics and Epidemiology, Geisel School of Medicine at Dartmouth; [5]The Dartmouth Institute for Health Policy and Clinical Practice; Dartmouth University, Hanover, New Hampshire.
[4]The VA Outcomes Group, White River Junction Veterans' Affairs Medical Center, White River Junction, Vermont.

estimated as high as 8%–35% (13–15), but evidence for the second condition is limited. Our objective is to examine the strength of the association between health care activities and the incidence of papillary thyroid cancer.

We hypothesize that markers of increased access to health care will have a positive association with the incidence of papillary thyroid cancer. We test this hypothesis in two ways. First, we compare the trend in papillary cancer incidence over three decades, in two cohorts of patients with differing health insurance access: those age 65 years and older, who have near-universal (>95%) health care coverage through Medicare (16), and those under 65 years old, who have less certain health insurance coverage and among whom 18% are currently uninsured (17). We hypothesize that in recent years, incidence would increase faster in the Medicare-age cohort than in the non–Medicare-age cohort.

Second, we perform an ecologic analysis to determine the influence of county-level markers of health care access on papillary thyroid cancer incidence. We use nine widely accepted socioeconomic variables as markers of county-level healthcare access (18–25). We hypothesize that counties with higher levels of access to care have a higher incidence of papillary thyroid cancer. Here, we report that the incidence of papillary thyroid cancer is increasing more rapidly in the Medicare-age population and that markers of wider health care access are associated with a higher incidence of papillary thyroid cancer in U.S. counties.

Methods

Data sources

Data on thyroid cancer incidence, patient age, and county of residence are from the National Cancer Institute (NCI)'s Surveillance, Epidemiology, and End Results (SEER) program. Started in 1973, SEER has grown to capture 28% of the United States population. To form its socioeconomically representative cross-section of the U.S. population, SEER currently captures all cancers diagnosed in 18 geographic regions (26,27).

SEER collects details on demographics, tumor characteristics, therapy, and survival of cancer patients. Strict quality control is an integral part of the SEER program (26,28–30). Because SEER is a de-identified dataset, the NCI does not require institutional review board oversight; a data use agreement was signed. The SEER 18 and SEER 9 datasets were accessed using SEERStat, release 7.1.0 (released July 2012; NCI Division of Cancer Control and Population Sciences, Bethesda, MD). County-level socioeconomic data were obtained from the U.S. Census 2000 and Small Area Health Insurance Estimates programs (2005) (31,32). The nine variables used as indicators of health care access have been widely used in analyses of cancer incidence and sociodemographic markers (17–24): percentages of county population that are uninsured, below poverty, unemployed, employed in white collar occupations, of nonwhite ethnicity, non-English speaking (defined by the Census as "linguistic isolation"), without a high school education, with at least a bachelor's degree, and mean county-level family income.

Definitions

Papillary thyroid carcinomas were defined as tumors arising in the thyroid gland with papillary histology codes 8050, 8052, 8130, 8260, 8340–8344, 8450, 8452 (33). Incidence rates were calculated per 100,000 population, age-adjusted to the 2000 United States Census population (34). The Medicare-age cohort was defined as patients age 65 years or older at the time of cancer diagnosis; the non–Medicare-age cohort comprised patients under 65 years old.

Analysis

Papillary thyroid cancer incidence rates were calculated for Medicare-age and non–Medicare-age patients in the SEER 9 dataset, from 1973 to 2009 (the most recent year for which data are available). During these years, the percentage of Americans lacking health insurance has not appreciably changed (17). Because thyroid nodules and papillary thyroid cancer are more prevalent in older persons, Joinpoint log-linear regression analysis was used to identify inflection points in the incidence trend lines, and to compare annual percentage change. Joinpoint version 3.5.2 (NCI Surveillance Research, Bethesda, MD) was used to identify inflection points and to compare incidence trends using a permuted comparability test, in which the null hypothesis was that the regression lines for incidence in two cohorts are coincident or parallel.

For the ecologic analysis, county papillary thyroid cancer incidence in 2000–2005 was the dependent variable and the nine markers of county-level socioeconomic status were explanatory variables. We restricted the analysis to incidence data from 2000 to 2005 to maintain fidelity with the 2000 U.S. Census Data and Small Area Health Insurance Estimates Program data (31,32) and to minimize the effects of migration over time. We included only the 443 counties with a population >40,000. County-level data were expressed as mean values weighted by county population, with 5th and 95th percentile values. The nine socioeconomic variables were analyzed in univariate analysis, using Pearson correlation weighted by county population, and in multivariable regression. Because the variability of papillary thyroid cancer incidence rates is heteroscedastic, varying inversely with county population, a generalized least-squares regression model weighted by county population was used. All variables were entered into the regression model, to determine overall strength of the association, and to calculate the overall r^2 of the model.

To examine small area variation within states, a generalized linear mixed model was fitted to the rates with a log link and random effects for county (35). The correlation between counties was specified according to the distance between county centroids. Annual rates were combined for this analysis, and variables were included for year only. An autoregressive structure over time among repeated county rates was also specified. The empirical Bayes estimates for county random effects were plotted to obtain smoothed maps for assessing small area variation without including variability due to population sizes. These analyses were performed using SPSS 19 (IBM Corp., Armonk, NY) and PROC GLIMMIX in SAS 9.2 (SAS Institute, Cary, NC).

Results

Over 36 years, the incidence of papillary thyroid cancer in the United States increased to 3.6 times the 1973 rate—from 3.5 per 100,000 to 12.5 per 100,000 in 2009 ($p<0.001$; Fig. 1). During this time period, the majority of the increased incidence was attributable to cancers below palpable size: 65.1% of the increase was comprised of tumors <2.0 cm in size. The

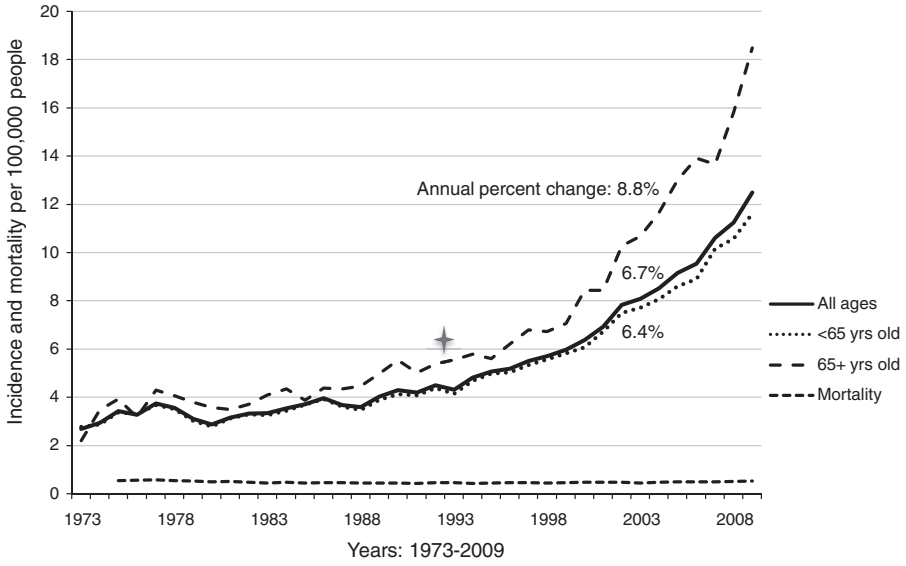


FIG. 1. Trends in incidence and mortality of papillary thyroid cancer, by patient age at diagnosis. Incidence data are from the Surveillance, Epidemiology and End Results (SEER) Program, SEER 9 Regs Research Data. Mortality data are from the National Center for Health Statistics. Incidence and mortality data are age-adjusted to year 2000 census, and reported per 100,000 people. Annual percent change calculation is for years 1993–2009, calculated in Joinpoint 3.5.2 (April 2011; Statistical Methodology and Applications Branch and Data Modeling Branch, Surveillance Research Program, National Cancer Institute).

majority of this increase occurred after 1993, when the incidence was 4.3 per 100,000. The annual percent change between 1993 and 2009 was 6.7%. Mortality has remained unchanged since data were first reported in 1975, near 0.5 per 100,000 (in 2009 [95% confidence interval (CI) 0.50–0.55]; annual percentage change since 1975, −0.11% [CI −0.24 to 0.018]) (36).

Papillary thyroid cancer incidence trends stratified by Medicare-eligible age

Before the early 1990s, the incidence rate of papillary thyroid cancer among persons of Medicare-eligible age (4–6 per 100,000) was marginally higher than among persons under 65 years old (3–5 per 100,000). However, in recent decades, incidence rates have diverged, with Joinpoint regression identifying an inflection point at 1993. In the Medicare-age cohort, papillary thyroid cancer incidence has increased more rapidly than in the population as a whole (from 1993 to 2009, annual percentage change 8.8%, $p < 0.001$). In 2009, the incidence in Medicare-age patients was 18.5 per 100,000, 67% higher than the nationwide incidence rate.

In the non–Medicare-age cohort, incidence more closely tracked the overall trend, increasing at an annual percent change of 6.4% between 1993 and 2009, a slower increase than in the population as a whole ($p < 0.001$). In 2009, the incidence in non–Medicare-age patients was 11.6 per 100,000 (Fig. 1).

Variation stratified by county and geographic area

Between 2000 and 2009, in the 18 geographic registries in SEER, incidence ranged widely from 5.9 per 100,000 among Alaska Natives to 12.0 per 100,000 in Connecticut—a twofold difference.

Among the 497 counties included in SEER, 10 counties had zero incident cases, including three counties with population greater than 40,000 (Howard County, IA; Martin County, KY; Trimble County, KY). The counties with population greater than 40,000 and the highest incidence rates were Los Alamos County, NM (29.7 per 100,000); Lucas County, IA (25.8 per 100,000); and Modoc County, CA (20.4 per 100,000). Figure 2 demonstrates the wide variability in incidence, even within geographically close areas within smaller states. Incidence data and mean county-level data (weighted by county population) for socioeconomic variables are summarized in Table 1.

All nine measures of county-level health care access were significantly correlated with the incidence of papillary thyroid cancer on univariate analysis (Table 2). Incidence was positively correlated with county-level mean family income ($p = 0.001$), county population with at least a bachelor's degree ($p = 0.001$), and county population employed in white collar occupations ($p = 0.003$). Papillary thyroid cancer incidence was inversely correlated with county unemployment rate ($p = 0.003$), poverty rate ($p < 0.001$), and population that

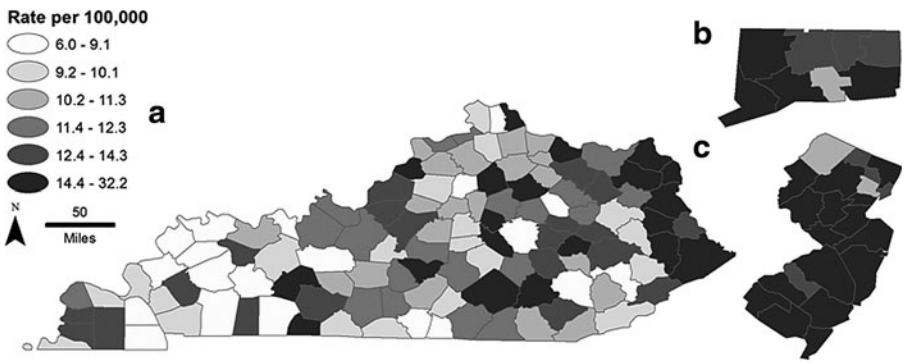


FIG. 2. Incidence of papillary thyroid cancer in 2009, by county, in Kentucky **(a)**, Connecticut **(b)**, and New Jersey **(c)**. Incidence data are from the SEER Program. Rates were smoothed by geographic distance using a generalized linear mixed model. Representative states were chosen to demonstrate the variability of thyroid cancer incidence within geographically close areas.

Table 1. County-Level Thyroid Cancer Incidence and Socioeconomic Data, 2000–2005

	Average	5th percentile	95th percentile
Median county population ($n=497$)	139,035	12,837	6,396,100
Measures of incidence (per 100,000)			
Incidence of PTC, all ages	7.39	1.50	13.16
Incidence of PTC, <65 years	4.96	0.00	10.20
Measures of socioeconomic status			
% uninsured	15.90	7.48	25.90
% below poverty	9.79	3.82	25.36
% with less than high school education	20.61	9.94	42.44
% with at least bachelor's degree	26.11	7.47	34.65
Median family income	53,679	26,136	66,808
% unemployed	6.91	3.82	25.36
% white collar employment	35.99	21.12	42.34
% with non-English primary language	6.46	0.00	9.45
% of nonwhite ethnicity	21.27	0.51	57.31

Data are presented as weighted means, except for those indicated as median values.
PTC, papillary thyroid cancer.

was non-English speaking ($p=0.016$), without high school education ($p<0.001$), of nonwhite ethnicity ($p<0.001$), and uninsured ($p<0.001$). Thus, areas with higher income and education were more likely to have higher incidence rates, while areas with more unemployment, poverty, and non-English speakers were more likely to have lower rates of papillary thyroid cancer incidence.

When analysis was limited to the non–Medicare-age population, several additional factors became independently significant on multivariable analysis: family income ($p=0.03$), unemployment rate ($p=0.03$), and population with white collar employment ($p=0.04$), non-English speaking ($p<0.001$), and without high school education ($p=0.012$).

When the regression model was limited to the non–Medicare-age population, these nine markers of health care access together explained 25% of the variability in county-level papillary thyroid cancer incidence ($r=0.50$, $r^2=0.25$, $F=15.32$, standard error of estimate=1630, $p<0.001$). When the regression model was expanded to include the Medicare-age population, only 14% of the variability in county-level incidence was explained by these nine markers ($r=0.38$, $r^2=0.14$, $F=7.94$, standard error=1912, $p<0.001$). This attenuated model is consistent with the leveling effect of near-universal health care access in the Medicare-age population, diminishing the ability of these nine markers to estimate the level of access to health care, once patients turn 65.

Discussion

Between 1973 and 2009, the incidence of papillary thyroid cancer more than tripled. Over the past two decades, the overall incidence rate has been increasing by >6% per year. Among patients with near-universal Medicare health care coverage at age 65, the annual rate of increase is higher, nearly 9% per year. Although thyroid cancer was marginally more prevalent among older persons before the 1990s, the incidence of thyroid cancer has accelerated at a faster rate in the Medicare-age cohort over the past two decades. Across the U.S. counties captured by the SEER cancer registry, markers of access to health care are strongly correlated with the incidence of papillary thyroid cancer. Incidence tends to be highest in counties with higher levels of income and with greater percentages of residents with white-collar employment and bachelor's degrees. Incidence rates tend to be lowest in counties with higher percentages of residents who are unemployed, uninsured, of nonwhite ethnicity, non-English speaking, in poverty, and without a high school education. Together, these findings illustrate an association between access to health care and the incidence of papillary thyroid cancer.

Seven years ago, we reported that the incidence of differentiated thyroid cancer had doubled between 1973 and 2002. We proposed that overdiagnosis may be the chief cause of this phenomenon (2). We and others had also previously observed that the incidence of thyroid cancer appeared to be rising

Table 2. Correlations Between County Health Care Access and County-Level Incidence of Papillary Thyroid Cancer

	Dependent variable					
	Incidence of papillary thyroid cancer (all ages)			Incidence of papillary thyroid cancer (age <65 years)		
Explanatory variable	*Correlation*	p *value (univariate)*	p *value (multivariable)*	*Correlation*	p *value (univariate)*	p *value (multivariable)*
Bachelor's degree	0.15	**0.001**	0.11	0.09	**0.03**	0.17
Family income	0.15	**0.001**	0.12	0.06	0.12	**0.03**
White collar employment	0.13	**0.003**	0.40	0.05	0.14	**0.04**
English not primary language	−0.10	**0.016**	0.18	−0.07	0.07	**<0.001**
Unemployment rate	−0.13	**0.003**	0.98	−0.04	0.22	**0.03**
No high school education	−0.23	**<0.001**	0.76	−0.23	**<0.001**	**0.012**
Uninsured	−0.25	**<0.001**	**0.02**	−0.26	**<0.001**	**<0.001**
Nonwhite ethnicity	−0.25	**<0.001**	**<0.001**	−0.29	**<0.001**	**<0.001**
Poverty rate	−0.27	**<0.001**	0.22	−0.25	**<0.001**	0.83

Values represent the Pearson correlation coefficient and p values, for both univariate and multivariable analyses. Significant values are presented in boldface.

fastest in more affluent regions of the country, and speculated that this may be attributable to wider access to healthcare (4,37). Consistent with this hypothesis, we and others had also reported differences in thyroid cancer incidence between ethnic groups, with incidence rates highest among non-Hispanic white individuals, again raising the possibility that thyroid cancer incidence may be correlated with access to health care. However, the variation in thyroid cancer incidence by ethnicity was attenuated in cases of nonpapillary histology, arguing against the presence of differences in diagnostic scrutiny (4,38). Therefore, the strength of the association between health care access and the incidence of thyroid cancer in the United States had been unclear.

The data in the present study now demonstrate that the rising incidence of differentiated thyroid cancer has continued unabated, and that the incidence of thyroid cancer is strongly associated with multidimensional measures of access to health care. These data therefore provide further support for the hypothesis of overdiagnosis.

Overdiagnosis is the identification of a disease which, if left undetected, would not cause symptoms or death for that patient during his or her lifetime. Before concluding that this phenomenon is occurring, two conditions must be satisfied. First, there must be evidence for a large reservoir of subclinical disease. Second, there must be a strong association between health care activity and the detection of the reservoir of subclinical cancers. There is robust evidence for a subclinical reservoir of papillary thyroid cancer. A meta-analysis of 24 autopsy series revealed a mean prevalence of occult papillary thyroid cancer of 7.6% (15). In two independent autopsy studies in which normal-appearing thyroid glands were thinly sectioned at 2–3 mm intervals, occult papillary thyroid cancers were identified in 33.3% and 35.6% of subjects (13,14). At these prevalence rates, the estimated subclinical reservoir in the United States is between 25 and 100 million Americans.

To date, there has been no direct evidence to satisfy the second condition for overdiagnosis: an association between health care activity and the incidence of papillary thyroid cancer. Here, we used a natural experiment design in a population-based U.S. registry to demonstrate a robust association between markers of health care access and the rate of papillary thyroid cancer diagnosis. A statistical model based on nine markers of access to care explained as much as 25% of the variability in the county-level incidence of papillary thyroid cancer. The model was most statistically robust when including only people under age 65, but was attenuated when Medicare-eligible persons (age 65 and older) were included. In the United States, at age 65, near-universal health care coverage provided by Medicare diminishes the ability to estimate the level of access to care with markers such as unemployment rate, poverty rate, income, and education. These findings are consistent with the hypothesis that papillary thyroid cancer diagnosis is highly dependent on access to health care.

Interestingly, the association between health care access and overdiagnosis has been shown in other cancers, such as prostate cancer, a disease known to be prone to overdiagnosis (39). Prostate cancer incidence has been robustly correlated with markers of access to care in multiple studies: regions with higher income and educational attainment have higher prostate cancer incidence, attributable to increased use of prostate-specific antigen testing (19,40–42). Because thyroid cancer is not a disease recommended for screening by the U.S. Preventive Services Task Force, a study specifically examining thyroid screening and thyroid cancer diagnoses is not possible.

Certainly, the association between access to care and papillary thyroid cancer incidence cannot rule out a coexistent true increase in the occurrence of thyroid cancer. It is possible that more thyroid cancers are developing, and that areas with increased access to care have been more successful at diagnosing these cases. However, in a scenario of increasing cancer incidence, thyroid cancer mortality rates would be expected to rise. Despite a 3.6-fold increase in papillary thyroid cancer incidence, nationwide papillary thyroid cancer mortality has not changed in 34 years, making this explanation less likely. Similar mortality data have been reported by others (10). Furthermore, a plausible biological explanation for an increase in papillary thyroid cancer cases is lacking. High levels of population exposure to the one known risk factor, ionizing radiation, have decreased over the past 50 years. In the United States, nuclear tests have not been performed since 1961 (43), and radiotherapy for benign conditions of the head and neck has not been routine since the late 1950s (44). Today, the main source of radiation exposure in the United States is background exposure to radon and thoron, followed by medical x-rays and computed tomography (CT) scans (45). CT scan radiation doses are much lower than these historical sources, with a very low estimated excess attributable cancer risk of $<0.01\%$–0.05% over a lifetime (46). Airplane travel results in radiation exposure, but at a dose several orders of magnitude below a CT scan (<0.1 mSv compared to 100 mSv for a full-body CT scan). Therefore, there is no biologically credible explanation that seems able to account for the tripling in papillary thyroid cancer incidence over the past 30 years.

Our study has important limitations, related to the fact that the available measures of health care access are necessarily crude and indirect. First, county-level measures of health care access are used as surrogates for more ideal measures, such as the number of practitioner-performed screening physical examinations or imaging studies of the neck and thyroid. Unfortunately, U.S. billing data, the ideal source for a large cohort, do not reliably capture incidences of physical examination of the neck or symptoms prompting neck imaging, making it impossible to test this association directly. Most importantly, billing databases, by their very nature, do not capture patients with other (or no) health insurance, and therefore do not allow the analysis of varying levels of access to care. For these reasons, a population-based registry is ideally suited for ecologic studies such as this one. A second caveat is that county levels of access to health care do not capture the individual experience of residents—many who live in affluent counties are unemployed, are of nonwhite ethnicity, or have less than a high school education. Given these limitations, the statistical tests we performed would tend to underestimate any association between health care access and the incidence of thyroid cancer.

In conclusion, these data demonstrate an association between levels of health care activity and the number of papillary thyroid cancers diagnosed in the United States. Together with the well-known large subclinical reservoir of disease, these results now provide evidence that overdiagnosis explains much of the thyroid cancer "epidemic." Current trends suggest that in coming years many more of these occult cancers will be detected and many more patients will undergo treatment for papillary thyroid cancer. The additional

treatment resulting from overdiagnosis is by definition of no benefit and only of potential harm, making thyroid cancer overdiagnosis a growing public health concern.

Acknowledgments

We thank the GeoSpatial Resource at The Dartmouth-Hitchcock Norris Cotton Center for assistance in map creation, with key work done by Heather Carlos. L.G.T.M. received funding from a National Institutes of Health grant (T32 CA009685). T.D.T. received funding from the National Institutes of Health (National Cancer Institute RC2 CA148259, and Cancer Center Support Grant 5P30 CA023108).

Disclosure Statement

The authors declare that no competing financial interests exist.

References

1. Davies L, Ouellette M, Hunter M, Welch HG 2010 The increasing incidence of small thyroid cancers: where are the cases coming from? Laryngoscope **120:**2446–2451.
2. Davies L, Welch HG 2006 Increasing incidence of thyroid cancer in the United States, 1973–2002. Jama **295:**2164–2167.
3. Morris LG, Myssiorek D 2010 Improved detection does not fully explain the rising incidence of well-differentiated thyroid cancer: a population-based analysis. Am J Surg **200:** 454–461.
4. Morris LG, Sikora AG, Myssiorek D, DeLacure MD 2008 The basis of racial differences in the incidence of thyroid cancer. Ann Surg Oncol **15:**1169–1176.
5. Burgess JR 2002 Temporal trends for thyroid carcinoma in Australia: an increasing incidence of papillary thyroid carcinoma (1982–1997). Thyroid **12:**141–149.
6. Colonna M, Grosclaude P, Remontet L, Schvartz C, Mace-Lesech J, Velten M, Guizard A, Tretarre B, Buemi AV, Arveux P, Esteve J 2002 Incidence of thyroid cancer in adults recorded by French cancer registries (1978–1997). Eur J Cancer **38:**1762–1768.
7. Kilfoy BA, Zheng T, Holford TR, Han X, Ward MH, Sjodin A, Zhang Y, Bai Y, Zhu C, Guo GL, Rothman N 2009 International patterns and trends in thyroid cancer incidence, 1973–2002. Cancer Causes Control **20:**525–531.
8. Liu S, Semenciw R, Ugnat AM, Mao Y 2001 Increasing thyroid cancer incidence in Canada, 1970–1996: time trends and age-period-cohort effects. Br J Cancer **85:**1335–1339.
9. Leenhardt L, Grosclaude P, Cherie-Challine L 2004 Increased incidence of thyroid carcinoma in france: a true epidemic or thyroid nodule management effects? Report from the French Thyroid Cancer Committee. Thyroid **14:**1056–1060.
10. Chen AY, Jemal A, Ward EM 2009 Increasing incidence of differentiated thyroid cancer in the United States, 1988–2005. Cancer **115:**3801–3807.
11. Grodski S, Brown T, Sidhu S, Gill A, Robinson B, Learoyd D, Sywak M, Reeve T, Delbridge L 2008 Increasing incidence of thyroid cancer is due to increased pathologic detection. Surgery **144:**1038-1043; discussion 1043.
12. Welch HG, Black WC 2010 Overdiagnosis in cancer. J Natl Cancer Inst **102:**605–613.
13. Harach HR, Franssila KO, Wasenius VM 1985 Occult papillary carcinoma of the thyroid. A "normal" finding in Finland. A systematic autopsy study. Cancer **56:**531–538.
14. Tanriover O, Comunoglu N, Eren B, Comunoglu C, Turkmen N, Dogan M, Gundogmus UN 2011 Occult papillary thyroid carcinoma: prevalence at autopsy in Turkish people. Eur J Cancer Prev **20:**308–312.
15. Valle LA, Kloos RT The prevalence of occult medullary thyroid carcinoma at autopsy. J Clin Endocrinol Metab **96:**E109–113.
16. Birnbaum M, Patchias EM 2010 Measuring coverage for seniors in Medicare Part A and estimating the cost of making it universal. J Health Polit Policy Law **35:**49–62.
17. DeNavas-Walt C, Proctor BD, Smith JC 2012 Income, Poverty, and Health Insurance Coverage in the United States: 2011. Vol P60-243. U.S. Government Printing Office, Washington, D.C.
18. Clarke CA, Moy LM, Swetter SM, Zadnick J, Cockburn MG 2010 Interaction of area-level socioeconomic status and UV radiation on melanoma occurrence in California. Cancer Epidemiol Biomarkers Prev **19:**2727–2733.
19. Mackillop WJ, Zhang-Salomons J, Boyd CJ, Groome PA 2000 Associations between community income and cancer incidence in Canada and the United States. Cancer **89:**901–912.
20. Clegg LX, Reichman ME, Miller BA, Hankey BF, Singh GK, Lin YD, Goodman MT, Lynch CF, Schwartz SM, Chen VW, Bernstein L, Gomez SL, Graff JJ, Lin CC, Johnson NJ, Edwards BK 2009 Impact of socioeconomic status on cancer incidence and stage at diagnosis: selected findings from the Surveillance, Epidemiology, and End Results: National Longitudinal Mortality Study. Cancer Causes Control **20:** 417–435.
21. Hao Y, Jemal A, Zhang X, Ward EM 2009 Trends in colorectal cancer incidence rates by age, race/ethnicity, and indices of access to medical care, 1995–2004 (United States). Cancer Causes Control **20:**1855–1863.
22. Harper S, Lynch J, Meersman SC, Breen N, Davis WW, Reichman MC 2009 Trends in area-socioeconomic and race-ethnic disparities in breast cancer incidence, stage at diagnosis, screening, mortality, and survival among women ages 50 years and over (1987–2005). Cancer Epidemiol Biomarkers Prev **18:**121–131.
23. Klassen AC, Curriero FC, Hong JH, Williams C, Kulldorff M, Meissner HI, Alberg A, Ensminger M 2004 The role of area-level influences on prostate cancer grade and stage at diagnosis. Prev Med **39:**441–448.
24. Schootman M, Lian M, Deshpande AD, Baker EA, Pruitt SL, Aft R, Jeffe DB 2010 Temporal trends in geographic disparities in small-area breast cancer incidence and mortality, 1988 to 2005. Cancer Epidemiol Biomarkers Prev **19:**1122–1131.
25. Singh GK, Miller BA, Hankey BF, Edwards BK 2004 Persistent area socioeconomic disparities in U.S. incidence of cervical cancer, mortality, stage, and survival, 1975–2000. Cancer **101:**1051–1057.
26. Clegg LX, Reichman ME, Hankey BF, Miller BA, Lin YD, Johnson NJ, Schwartz SM, Bernstein L, Chen VW, Goodman MT, Gomez SL, Graff JJ, Lynch CF, Lin CC, Edwards BK 2007 Quality of race, Hispanic ethnicity, and immigrant status in population-based cancer registry data: implications for health disparity studies. Cancer Causes Control **18:**177–187.
27. National Cancer Institute. SEER. Surveillance, Epidemiology, and End Results program data, 1973–2009. Available at http://seer.cancer.gov/data (accessed February 1, 2013).
28. Clegg LX, Feuer EJ, Midthune DN, Fay MP, Hankey BF 2002 Impact of reporting delay and reporting error on cancer incidence rates and trends. J Natl Cancer Inst **94:**1537–1545.

29. Clegg LX, Gail MH, Feuer EJ 2002 Estimating the variance of disease-prevalence estimates from population-based registries. Biometrics **58:**684–688.
30. Zippin C, Lum D, Hankey BF 1995 Completeness of hospital cancer case reporting from the SEER Program of the National Cancer Institute. Cancer **76:**2343–2350.
31. United States Census Bureau. Census 2000 Gateway. Available at www.census.gov/main/www/cen2000.html (accessed February 1, 2013).
32. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM, Whelan S (eds) 2000 International classification of diseases for oncology, 3rd ed. World Health Organization, Geneva.
33. O'Hara B 2008 Experimental health insurance estimates for low-income and demographic groups by state. Health Serv Res **43(5p1):**1693–1707.
34. Gigli A, Mariotto A, Clegg LX, Tavilla A, Corazziari I, Capocaccia R, Hachey M, Steve S 2006 Estimating the variance of cancer prevalence from population-based registries. Stat Methods Med Res **15:**235–253.
35. Breslow NE, Clayton DG 1993 Approximate inference in generalized linear mixed models. J Am Stat Assoc **88:**9–25.
36. Surveillance Research Program 2012 Age adjusted U.S. mortality rates and 95% confidence intervals by cancer site. National Cancer Institute, Bethesda, MD. Available online at http://seer.cancer.gov/statistics (accessed February 1, 2013).
37. Li N, Du XL, Reitzel LR, Xu L, Sturgis EM Impact of enhanced detection on the increase in thyroid cancer incidence in the United States: review of incidence trends by socioeconomic status within the surveillance, epidemiology, and end results registry, 1980–2008. Thyroid **23:**103–110.
38. Aschebrook-Kilfoy B, Ward MH, Sabra MM, Devesa SS Thyroid cancer incidence patterns in the United States by histologic type, 1992–2006. Thyroid **21:**125–134.
39. Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, Feuer EJ 2002 Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst **94:**981–990.
40. Cheng I, Witte JS, McClure LA, Shema SJ, Cockburn MG, John EM, Clarke CA 2009 Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California. Cancer Causes Control **20:**1431–1440.
41. Liu L, Cozen W, Bernstein L, Ross RK, Deapen D 2001 Changing relationship between socioeconomic status and prostate cancer incidence. J Natl Cancer Inst **93:**705–709.
42. Steenland K, Rodriguez C, Mondul A, Calle EE, Thun M 2004 Prostate cancer incidence and survival in relation to education (United States). Cancer Causes Control **15:**939–945.
43. 1976 Information for physicians on irradiation related thyroid cancer. CA Cancer J Clin **26:**150–159.
44. Ron E, Saftlas AF 1996 Head and neck radiation carcinogenesis: epidemiologic evidence. Otolaryngol Head Neck Surg **115:**403–408.
45. Sinnott B, Ron E, Schneider AB 2010 Exposing the thyroid to radiation: a review of its current extent, risks, and implications. Endocr Rev **31:**756–773.
46. Brenner DJ, Hall EJ 2007 Computed tomography—an increasing source of radiation exposure. N Engl J Med **357:**2277–2284.

Address correspondence to:
Luc G.T. Morris, MD, MSc
Memorial Sloan-Kettering Cancer Center
1275 York Ave., S-1210A
New York, NY 10065

E-mail: morrisl@mskcc.org

Louise Davies, MD, MS
VA Outcomes Group – 111B
215 North Main St.
White River Junction, VT 05009

E-mail: louise.davies@dartmouth.edu

Ann Surg Oncol (2015) 22:3708–3715
DOI 10.1245/s10434-015-4382-x

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – HEAD AND NECK ONCOLOGY

[^{99m}Tc]Tilmanocept Accurately Detects Sentinel Lymph Nodes and Predicts Node Pathology Status in Patients with Oral Squamous Cell Carcinoma of the Head and Neck: Results of a Phase III Multi-institutional Trial

Amit Agrawal, MD[1], Francisco J. Civantos, MD[2], Kevin T. Brumund, MD[3], Douglas B. Chepeha, MD[4], Nathan C. Hall, MD, PhD[5], William R. Carroll, MD[6], Russell B. Smith, MD[7], Robert P. Zitsch, MD[8], Walter T. Lee, MD[9], Yelizaveta Shnayder, MD[10], David M. Cognetti, MD[11], Karen T. Pitman, MD[12], Dennis W. King, PhD[13], Lori A. Christman, PhD[13], and Stephen Y. Lai, MD, PhD[14]

[1]Department of Otolaryngology—Head and Neck Surgery, Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, The Ohio State University Wexner Medical Center, Columbus, OH; [2]Department of Otolaryngology, University of Miami Hospital and Clinics/Sylvester Comprehensive Cancer Center, Miami, FL; [3]Division of Head and Neck Surgery, Department of Surgery, Moores UCSD Cancer Center and Veteran Affairs San Diego Medical Center, San Diego, CA; [4]Department of Otolaryngology, University of Toronto, Toronto, ON, Canada; [5]Division of Nuclear Medicine, Department of Radiology, The Ohio State University Wexner Medical Center, Columbus, OH; [6]Division of Otolaryngology—Head and Neck Surgery, Department of Surgery, University of Alabama at Birmingham, Birmingham, AL; [7]Department of Otolaryngology—Head and Neck Surgery, University of Nebraska Medical Center, Omaha, NE; [8]Department of Otolaryngology—Head and Neck Surgery, University of Missouri, Columbia, MO; [9]Division of Otolaryngology—Head and Neck Surgery, Department of Surgery, Duke University Medical Center, Durham, NC; [10]Department of Otolaryngology—Head and Neck Surgery, University of Kansas Medical Center, Kansas City, KS; [11]Department of Otolaryngology—Head and Neck Surgery, Thomas Jefferson University, Philadelphia, PA; [12]Banner MD Anderson Cancer Specialists, Gilbert, AZ; [13]STATKING Clinical Services, Fairfield, OH; [14]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX

ABSTRACT

Background. [^{99m}Tc]Tilmanocept, a novel CD206 receptor-targeted radiopharmaceutical, was evaluated in an open-label, phase III trial to determine the false negative rate (FNR) of sentinel lymph node biopsy (SLNB) relative to the pathologic nodal status in patients with intraoral or cutaneous head and neck squamous cell carcinoma (HNSCC) undergoing tumor resection, SLNB, and planned elective neck dissection (END). Negative predictive value (NPV), overall accuracy of SLNB, and the impact of radiopharmaceutical injection timing relative to surgery were assessed.

Methods and Findings. This multicenter, non-randomized, single-arm trial (ClinicalTrials.gov identifier NCT00911326) enrolled 101 patients with T1–T4, N0, and M0 HNSCC. Patients received 50 μg [^{99m}Tc]tilmanocept radiolabeled with either 0.5 mCi (same day) or 2.0 mCi (next day), followed by lymphoscintigraphy, SLNB, and END. All excised tissues were evaluated for tissue type and tumor presence. [^{99m}Tc]Tilmanocept identified one or more SLNs in 81 of 83 patients (97.6 %). Of 39 patients identified with any tumor-positive nodes (SLN or non-SLN), one patient had a single tumor-positive non-SLN in whom all SLNs were tumor-negative, yielding an FNR of 2.56 %; NPV was 97.8 % and

This study was presented in part at the 2013 Society of Nuclear Medicine and Molecular Imaging Annual Meeting, Vancouver, BC, Canada, 11 June 2013 (abstract no. 512), the 2013 American College of Surgeons Clinical Congress, Washington, DC, USA, 7 October 2013, and the 6th European Congress on Head and Neck Oncology, Liverpool, UK, 24 April 2014.

Amit Agrawal, Francisco J. Civantos, and Stephen Y. Lai are considered equal primary contributors.

First Received: 24 October 2014;
Published Online: 11 February 2015

A. Agrawal, MD
e-mail: amit.agrawal@osumc.edu

overall accuracy was 98.8 %. No significant differences were observed between same-day and next-day procedures.

Conclusions. Use of receptor-targeted [^{99m}Tc]tilmanocept for lymphatic mapping allows for a high rate of SLN identification in patients with intraoral and cutaneous HNSCC. SLNB employing [^{99m}Tc]tilmanocept accurately predicts the pathologic nodal status of intraoral HNSCC patients with low FNR, high NPV, and high overall accuracy. The use of [^{99m}Tc]tilmanocept for SLNB in select patients may be appropriate and may obviate the need to perform more extensive procedures such as END.

Head and neck squamous cell carcinoma (HNSCC) of both mucosal and cutaneous origin carries variable propensity to metastasize to regional cervical nodes. The presence of nodal metastases is the most important negative prognostic factor for long-term survival.[1–3] Thus, accurate identification and treatment of lymphatic metastases is important for this patient population.

As current methods, including physical examination and radiologic imaging, lack sufficient sensitivity and specificity,[4,5] elective neck dissection (END) has been the gold standard for assessing the presence or absence of lymphatic disease in patients without overt clinical or radiographic nodal metastases (cN0) undergoing surgical management of HNSCC.[6] However, END is associated with significant potential morbidity, including pain, contour changes, shoulder dysfunction, and lip paresis, as well as negative impact upon quality of life.[7–9] Furthermore, it may be argued that END is unnecessary in a large proportion of patients; for example, 70–80 % of patients initially presenting with early-stage oral cavity carcinoma (T1 or T2, cN0) ultimately prove to be free of lymphatic metastases.[8,10–12]

Sentinel lymph node biopsy (SLNB) has been advocated as a less invasive means of achieving accurate diagnostic assessment of regional metastatic tumor potential while reducing morbidity compared with more extensive procedures.[9]

Several studies have examined SLNB in HNSCC using radiolabeled colloid.[13–18] Despite excellent negative predictive values (NPV), the false negative rate (FNR) of SLNB for HNSCC (i.e. percentage of cases with overall positive END, SLN pathology-negative) appears variable and reached nearly 10 % in the two largest multicenter series.[14,18] Characteristics of radiolabeled colloid, including its particulate nature and lack of specific binding, may in part contribute to observed FNR when used for SLNB in HNSCC.

[^{99m}Tc]Tilmanocept, approved by the US FDA and recently granted marketing authorization by the European Medicine Agency's Committee for Medicinal Products for Human Use for breast cancer, melanoma, and oral HNSCC SLN detection, is a novel, receptor-targeted, non-particulate radiopharmaceutical that consists of multiple diethylenetriaminepentaacetic acid (DTPA) molecules for ^{99m}Tc chelation and mannose moieties for CD206 receptor binding tethered to a dextran scaffold. The small molecular size (7 nm diameter) of tilmanocept and its specific targeting to CD206 mannose-binding receptors located on reticuloendothelial cells within lymph nodes permit rapid injection site clearance and avid, stable binding within target nodes.[19]

This article describes the results of an open-label, FDA-designated, phase III trial to assess the accuracy of [^{99m}Tc]tilmanocept used in conjunction with lymphoscintigraphy and SLNB to detect SLNs, as well as predict pathologic nodal status (i.e. presence vs. absence of metastatic disease) in patients with oral or cutaneous HNSCC undergoing SLNB and END.

METHODS

Participants and Institutional Review/Consent

Eligibility criteria included T1–T4a, cN0, and M0 HNSCC located in the oral cavity or cutaneous head and neck region. Clinical nodal staging was confirmed by negative results from contrast-enhanced computed tomography (CT) scan, gadolinium-enhanced magnetic resonance imaging (MRI), or neck ultrasound. Patients with a history of neck dissection, gross injury to the neck, or radiotherapy to the neck or receiving systemic cytotoxic therapy were excluded from the trial.

Subject enrollment occurred across 13 centers. The protocol and informed consent were approved by the Institutional Review Boards of each center, and the study met all applicable regulatory and ethical requirements.

Procedures

Radiopharmaceutical Injection and Lymphoscintigraphy

Patients received 50 μg of [^{99m}Tc]tilmanocept radiolabeled with either 0.5 mCi (for surgeries on the same day as injection) or 2.0 mCi (for surgeries the day after injection). Timing of injection (i.e. day of surgery vs. day before surgery) was at the surgeon's discretion, except in patients with floor-of-mouth tumors. In such patients, day-before-surgery injection was required to allow for significantly reduced shine-through, whereby radioactivity at the primary site may obscure relevant SLNs. Following injection, all patients underwent preoperative lymphoscintigraphy imaging per institutional protocol, which

involved planar imaging (±dynamic) and/or fused single-photon emission computed tomography/CT (SPECT/CT).

Surgery/Sentinel Lymph Node Biopsy/Elective Neck Dissection Surgery was required either within 1–15 h (same day) or 15–30 h (next day) following injection. At surgery, excision of the primary tumor was performed prior to SLNB/END. Using a handheld gamma detector, the surgeon conducted an initial survey of the entire cervical lymph node basin at risk to identify the areas of increased radioactivity. An SLN was defined as a lymph node with a mean in vivo count >3 square roots of the mean normal tissue background count (i.e. three standard deviations) added to the mean normal tissue background count ('3σ rule') asserting 99.7 % certainty of the SLN signal. As each SLN was identified and dissected, radioactivity counts were recorded in vivo and ex vivo. SLNB was considered complete when no further hot nodes were detected. Following SLNB, END was then performed. Bilateral ENDs were performed when the primary lesion involved the midline, tumors <1 cm from midline with evidence of contralateral drainage on lymphoscintigraphy, or per surgical discretion.

Histopathology Assessment of Lymph Nodes All excised nodes (both SLNs and non-SLNs) underwent local routine histopathologic evaluation using hematoxylin and eosin (H&E) staining. After fixation, all SLNs were sectioned every 2 mm in transverse fashion along the longest axis and embedded into cassettes for sectioning, thus providing sections every 2–3 mm, producing at least three levels through the node for assessment. Additional staining was permitted locally based on institutional standards. All negative SLNs were sent to the study's central pathology laboratory for additional immunohistochemical staining for pancytokeratin markers (e.g. AE1/AE3, CK8/18, MNF 116, etc.). All locally positive SLNs had two unstained slides sent to the central laboratory for confirmation of pathology positivity.

Statistical Analyses

The primary endpoint was the FNR associated with assessment of [^{99m}Tc]tilmanocept-identified SLNs relative to the overall pathologic nodal status as determined by assessment of both SLNs and non-SLNs from the END. The FNR is the ratio of false negatives to the sum of true positives plus false negatives. The overall FNR point estimate was the observed rate and was made on a per-patient basis relative to all patients with pathology-positive nodes. The statistical hypotheses H_0: FNR $\geq$0.14 versus H_a: FNR $<$ 0.14, selected from an assessment of peer-reviewed publications of several prior studies examining SLNB in HNSCC, were tested using a one-sided significance level of 0.02486 such that if the upper limit of the 95.03 % confidence interval (CI) for the FNR was <0.14, the null hypothesis was rejected in favor of the alternative hypothesis. Exact binomial CIs were used.

Secondary patient-level measures of efficacy were NPV, overall accuracy of [^{99m}Tc]tilmanocept, and rate of SLN detection by [^{99m}Tc]tilmanocept. Point estimates for secondary endpoints were the observed rate; 95 % exact binomial CIs were calculated.

The intent-to-treat (ITT) population, consisting of all patients injected with [^{99m}Tc]tilmanocept who underwent surgery and had at least one lymph node (SLN or non-SLN) with known pathology status, was used for all efficacy analyses.

RESULTS

Demographics and Staging

Between June 2009 and November 2012, a total of 101 patients were enrolled. Of these, 16 patients withdrew from the study prior to drug administration or surgery—12 patients withdrew consent and four withdrew for other reasons. The remaining 85 patients were injected with [^{99m}Tc]tilmanocept. The majority of patients had oral tumors (92.9 %) and either T1 or T2 (84.7 %) clinical staging (Table 1).

Imaging

The preoperative SPECT/CT three-dimensional fused reconstruction cross-sectional images of a typical patient (image acquisition duration was 3–21 min) of [^{99m}Tc]tilmanocept are shown in Fig. 1. SPECT/CT imaging revealed four SLNs in this patient by 21 min post-injection of [^{99m}Tc]tilmanocept.

Efficacy Measures

Of 85 patients injected with [^{99m}Tc]tilmanocept, two patients did not undergo SLNB and END due to non-drug-related adverse events. Of note, there were no drug-related serious adverse events and no deaths on study. As such, 83 patients (78 intraoral and 5 cutaneous) injected with [^{99m}Tc]tilmanocept underwent SLNB/END and comprised the ITT population for efficacy analyses.

At least one SLN was identified in 81 of the 83 ITT patients yielding an SLN detection rate of 97.6 %. Table 2 shows lymph node statistics by pathology and node type, as well as statistics according to whether SLN pathology was positive or negative per subject. Among the 83 ITT

TABLE 1 Patient characteristics: ECOG status, tumor staging, and tumor location

Characteristic	No. of patients (%)		
	Cutaneous ($n = 6$)	Intraoral ($n = 79$)	Overall ($n = 85$)
Preoperative clinical T staging			
T1	0	26 (32.9)	26 (30.6)
T2	6 (100)	40 (50.6)	46 (54.1)
T3	0	7 (8.9)	7 (8.2)
T4	0	6 (7.6)	6 (7.1)
Preoperative clinical N staging			
N0	6 (100)	79 (100)	85 (100)
Preoperative clinical M staging			
M0	6 (100)	78 (98.7)	84 (98.8)
MX	0	1 (1.3)	1 (1.2)
ECOG performance status			
0	5 (83.3)	53 (67.1)	58 (68.2)
1	1 (16.7)	21 (26.6)	22 (25.9)
2	0	5 (6.3)	5 (5.9)

Data represent the [^{99m}Tc]tilmanocept-injected population ($N = 85$)

ECOG Eastern Cooperative Oncology Group

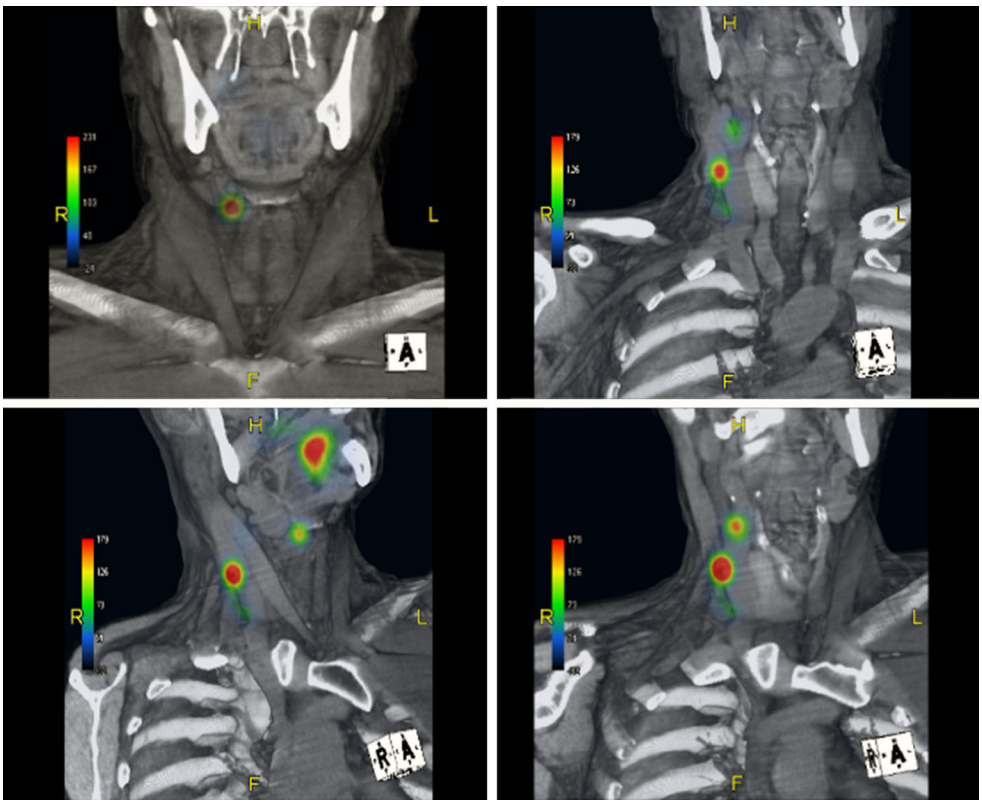


FIG. 1 SPECT/CT three-dimensional fused reconstruction cross-sectional images of a typical patient with floor-of-mouth tumor (duration of SPECT/CT acquisition was 3–21 min post-injection of [^{99m}Tc]tilmanocept. The *cube* in the lower right corner indicates the perspective of the image. *SPECT* single-photon emission computed tomography, *CT* computed tomography, *R* right, *L* left, *H* head, *F* feet, *A* anterior, *P* posterior

TABLE 2 Summary statistics for excised lymph nodes by pathology and per patient

Node type	Pathology status	Nodes per patient			
		Mean	95 % CI	Median	Range (min–max)
SLN ($n = 323$)	Overall	3.9	3.42–4.37	4	0–11
	Positive ($n = 67$)	0.8			
	Negative ($n = 255$)	3.1			
Non-SLN ($n = 2{,}823$)	Overall	34.0	30.02–38.01	30	0–82
	Positive ($n = 21$)	0.3			
	Negative ($n = 2{,}802$)	33.8			

Data represent the intent-to-treat population ($N = 83$)

min minimum, *max* maximum, *CI* confidence interval, *SLN* sentinel lymph node

patients, a mean of 3.9 SLNs (median 4) were removed per patient (range 0–11 nodes). Of the non-SLNs obtained via END (i.e. following SLNB), a mean of 34.0 non-SLNs were removed per patient (range 0–82 nodes).

In those subjects in whom one or more SLNs were pathology-positive for tumor, a mean of 4.5 SLNs (median 4.0) were removed per subject (range 2–11 nodes). In these same subjects, a mean of 32.5 non-SLNs (median 28.0) were removed via END (range 7–78 nodes).

Table 3 details SLN pathology status and overall nodal pathology status per subject, as well as efficacy metrics. Of the ITT patients, 39 (47.0 %), which were all intraoral patients, had at least one pathology-confirmed tumor-positive lymph node (SLN or non-SLN)—31 were staged T1–T2, and eight were staged T3–T4. The proportion of subjects identified with nodal tumor involvement was 44.3 % amongst patients with T1–T2 disease and 61.5 % amongst patients with T3–T4 disease. One patient (buccal mucosa tumor stage T2) in whom all SLNs identified by [^{99m}Tc]tilmanocept were negative for tumor, had one tumor-positive node (non-SLN) which was not detected via SLNB using [^{99m}Tc]tilmanocept ('false negative'). The overall FNR was 2.56 %, with a 95.03 % CI of 0.06–13.49; thus, the prospectively established null hypothesis was rejected in favor of the alternative hypothesis ($p = 0.0205$). To the extent that all cutaneous tumor patients would be excluded from the FNR analysis, the FNR remains unchanged. Thirty-eight patients had at least one SLN that was tumor positive ('true positives'). The FNR for the T1–T2 patients was 3.23 %, and 0 % for the T3–T4 patients. Forty-four of the patients in whom all SLNs were negative for tumor, as confirmed by the central laboratory, or in whom no SLNs were detected, also had all non-SLNs negative for tumor (both conditions included as 'true negatives'). These data yielded an NPV of 97.8 % (Table 3). For the ITT population, overall accuracy of SLN identified via [^{99m}Tc]tilmanocept in correctly determining the nodal pathology status of the neck was 98.8 %.

Pathology-positive and false-negative patients by tumor location and timing of surgery are shown in Table 4. No differences in FNR were observed between individual tumor subsites or between same-day and next-day procedures.

Data and Safety Monitoring

The current study was overseen by an independent Data and Safety Monitoring Committee (DSMC). The study was prospectively structured to include an interim analysis at 33.3 % ($N \geq 38$) of the targeted accrual cohort ($N \geq 114$) of node pathology-positive subjects. The trial was terminated early based on an interim review by the DSMC due to positive efficacy outcome. The DSMC noted that as the study achieved its primary efficacy endpoint, the added risk of END may not be justified in those situations where SLN assessment determined node-negative status.

DISCUSSION

Although routine in the management of breast cancer and melanoma, the use of SLNB procedures for HNSCC continues to evolve. Two large, multicenter, prospective trials to date have described SLNB for HNSCC using radiolabeled colloid with or without blue dye. A prospective trial at six centers in Europe followed 134 patients with T1–T2 N0 tumors of the oral cavity or oropharynx who either underwent SLNB alone or in SLNB in combination with END. In this trial, the FNR of SLNB after long-term follow-up was 9 %.[18,20] A prospective multi-institutional cooperative group trial (Z-0360) carried out in the US and sponsored by the American College of Surgeons Oncology Group (ACOSOG), involving 25 institutions over a 3-year period, assessed 140 patients with T1 and T2 oral cavity carcinoma. In this group, the NPV of SLNB was 96 %, with an observed FNR of 9.8 %.[14]

TABLE 3 Classification of patients according to pathology status of [^{99m}Tc]tilmanocept-identified SLNs, overall pathology nodal status, and calculated efficacy performance metrics

	Overall nodal pathology status (SLN and non-SLN), by patient	
	Positive (with one or more nodes)	Negative
Pathology status of SLN, by patient		
Positive (one or more nodes)	38 (true positive)	–
Negative (or no SLNs identified)	1 (false negative)	44 (true negative)
Performance metrics	Rate	95 % exact binomial CI[a]
False negative rate	0.0256	0.0006–0.1349
Negative predictive value	0.9778	0.8823– 0.9994
Overall accuracy	0.9880	0.9347– 0.9997

Data represent the intent-to-treat population ($N = 83$)

CI confidence interval, *SLN* sentinel lymph node

[a] The CI for the false negative rate is 95.03 %

TABLE 4 Summary of patients by tumor location and time of surgery

Variable	Total ITT patients	Patients with SLNs detected	All pathology-positive patients	False negative patients
Tumor location				
Buccal mucosa	8	8	4	1
Cutaneous	5	4	0	0
Floor of mouth	20	20	12	0
Lower alveolar ridge	3	3	2	0
Mucosal lip	1	1	0	0
Oral tongue	42	42	21	0
Retromolar gingiva	4	3	0	0
Time of surgery[a]				
Same day	40	40	22	1
Next day	42	40	16	0

Data represent the ITT population ($N = 83$)

ITT intent-to-treat, *SLNs* sentinel lymph nodes

[a] Time of surgery was missing for one patient and could therefore not be included in the time-of-surgery analyses

Despite the difference between studies in the number of subjects in the ITT population (ACOSOG Z-0360 study: 140 subjects; NEO3-06 study: 83 subjects), there was a similar number of node pathology-positive subjects (ACOSOG Z-0360: 41 subjects; NEO3-06: 39 subjects), which serves as the basis for the comparison of these studies.[14,21] In the current study, the FNR of [^{99m}Tc]tilmanocept (2.56 %) was statistically significantly lower than the upper limit of the FNR of [^{99m}Tc]sulfur colloid noted in the ACOSOG Z-0360 study (observed FNR of 9.8 %, 95 % CI 2.7–23.1; $p = 0.0005$). The accuracy of [^{99m}Tc]tilmanocept was also statistically significantly greater than the lower limit of the accuracy of [^{99m}Tc]-sulfur colloid as used in the Z-0360 study ($p = 0.0151$).[21]

Several contributing factors have been noted regarding the observed variable FNR for SLNB using radiolabeled colloid for HNSCC, including tumor location (floor-of-mouth tumors with higher FNR) and larger tumors (i.e. T2 vs. T1).[14,18] Due to its particulate nature and non-standardized preparation, radiolabeled colloids (100–1,000 nm particle diameter) are retained for prolonged periods within the injection site, which in turn contributes to the phenomenon of shine-through effect.[22] This is particularly problematic for floor-of-mouth tumors which, in previous studies, have been associated with significantly lower rates of SLN identification (88 %) and higher FNRs (20 %) compared with other oral sites.[18,20] In comparison, the current trial included 20 patients with floor-of-mouth tumors, of whom [^{99m}Tc]tilmanocept identified at least one SLN in all patients (100 %). Twelve of these patients were identified with metastatic nodal disease and, in all 12, at least one SLN was identified with metastatic disease. As such, the

observed NPV and overall accuracy of SLNB using [^{99m}Tc]tilmanocept in this group of patients was 100 %.

Criticism of the current study could focus on the inclusion of patients with larger tumors (higher expected nodal metastatic rate), as well as those with cutaneous HNSCC (lower expected nodal metastatic rate). Patients with larger tumors (T3, T4) comprised a relatively small group overall (13 patients, 15 %), but these patients were included as all patients were planned to undergo standard-of-care END. Given the high rate of occult nodal disease observed in these patients (8 of 13 patients, 61.5 %), one might reasonably forgo SLNB in favor of planned (i.e. therapeutic) END; however, in this study, the FNR for this subpopulation was 0 %. While the use of SLNB alone in patients with larger tumors is certainly controversial, lymphatic mapping procedures in such patients undergoing planned END (i.e. 'SLN-assisted END') might identify additional neck regions at risk, including the contralateral neck, not routinely encompassed during END alone. As such, the concept of SLNB procedures in this population may warrant further investigation. Patients with cutaneous HNSCC were a relatively small cohort (five patients, 6 %). None were found to have nodal disease following SLNB and END. The lack of observed nodal metastases in these patients limits the assessment of predictive utility of [^{99m}Tc]tilmanocept for SLNB (i.e. FNR, NPV) as related to cutaneous HNSCC, and also indicates the need for further study.

Of note, the specificity of tilmanocept for lymphatic tissues assessed via in vivo imaging and in vitro analysis of its receptor binding properties suggest that tilmanocept does not move downstream to distal lymph nodes, permitting high confidence that a hot node found during next-day procedures is in fact an SLN.[19] The present study supports that the SLN detection rate and FNR for nodal metastases were not significantly affected by the day of surgery relative to timing of [^{99m}Tc]tilmanocept injection. This attribute portends that the use of [^{99m}Tc]tilmanocept provides substantial leeway and scheduling flexibility with regard to time of injection and subsequent lymphoscintigraphy and SLNB procedures (i.e. next-day surgery) without compromising the reliability of results.

CONCLUSIONS

The current trial supports the use of [^{99m}Tc]tilmanocept in the setting of SLNB for HNSCC with a high rate of SLN identification. When used in conjunction with serial sectioning and immunohistochemistry, SLNB with [^{99m}Tc]tilmanocept accurately predicts the nodal pathology status of the neck in patients with oral HNSCC with low FNR, high NPV, and high overall accuracy. Given these results, the use of [^{99m}Tc]tilmanocept in this setting may help surgeons avoid the need to perform more extensive procedures, including END.

ACKNOWLEDGMENT The clinical trial described herein was supported by Navidea Biopharmaceuticals, Dublin, OH, USA. After the conclusion of this clinical trial, Dr. Lai became a Medical Affairs consultant for Navidea Biopharmaceutical, Inc.

CONFLICT OF INTEREST All other authors declare that they have no financial or other relevant conflicts of interests.

APPENDIX: INVESTIGATORS AND ENROLLING INSTITUTIONS

Amit Agrawal, MD
Department of Otolaryngology—Head and Neck Surgery, Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, The Ohio State University Wexner Medical Center, Columbus, OH, USA

Stephen Y. Lai, MD, PhD
Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Kevin T. Brumund, MD
Department of Surgery, Division of Head and Neck Surgery, Moores UCSD Cancer Center and Veteran Affairs San Diego Medical Center, San Diego, CA, USA

Francisco J. Civantos, MD
Department of Otolaryngology, University of Miami Hospital and Clinics/Sylvester Comprehensive Cancer Center, Miami, FL, USA

Douglas B. Chepeha, MD
Department of Otolaryngology, University of Michigan, Ann Arbor, MI, USA

William R. Carroll, MD
Department of Surgery, Division of Otolaryngology—Head and Neck Surgery, University of Alabama at Birmingham, Birmingham, AL, USA

Russell B. Smith, MD
Department of Otolaryngology—Head and Neck Surgery, University of Nebraska Medical Center, Omaha, NE, USA

Robert P. Zitsch, MD
Department of Otolaryngology—Head and Neck Surgery, University of Missouri, Columbia, MO, USA

Walter T. Lee, MD

Department of Surgery, Division of Otolaryngology—Head and Neck Surgery, Duke University Medical Center, Durham, NC, USA

Yelizaveta Shnayder, MD

Department of Otolaryngology—Head and Neck Surgery, University of Kansas Medical Center, Kansas City, KS, USA

David M. Cognetti, MD

Department of Otolaryngology—Head and Neck Surgery, Thomas Jefferson University, Philadelphia, PA, USA

Karen T. Pitman, MD

Department of Otolaryngology, University of Mississippi Medical Center, Jackson, MS, USA

REFERENCES

1. Alvi A, Johnson JT. Extracapsular spread in the clinically negative neck (N0): implications and outcome. *Otolaryngol Head Neck Surg*. 1996;114:65–70.
2. Mamelle G, Pampurik J, Luboinski B, et al. Lymph node prognostic factors in head and neck squamous cell carcinomas. *Am J Surg*. 1994;168:494–8.
3. Rassekh CH, Johnson JT, Myers EN. Accuracy of intraoperative staging of the N0 neck in squamous cell carcinoma. *Laryngoscope*. 1995;105:1334–6.
4. de Bondt RB, Nelemans PJ, Hofman PA, et al. Detection of lymph node metastases in head and neck cancer: a meta-analysis comparing US, USgFNAC, CT and MR imaging. *Eur J Radiol*. 2007; 64:266–72.
5. Kyzas PA, Evangelou E, Denaxa-Kyza D, Ioannidis JP. 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. *J Natl Cancer Inst*. 2008;100:712–20.
6. Yuen AP, Wei WI, Wong YM, et al. Elective neck dissection versus observation in the treatment of early oral tongue carcinoma. *Head Neck*. 1997; 19:583–8.
7. Chepeha DB, Taylor RJ, Chepeha JC, et al. Functional assessment using Constant's Shoulder Scale after modified radical and selective neck dissection. *Head Neck*. 2002;24:432–6.
8. Rogers SN, Ferlito A, Pellitteri PK, et al. Quality of life following neck dissections. *Acta Otolaryngol*. 2004;124:231–6.
9. Schiefke F, Akdemir M, Weber A, et al. Function, postoperative morbidity, and quality of life after cervical sentinel node biopsy and after selective neck dissection. *Head Neck*. 2009;31:503–12.
10. Schmitz S, Machiels JP, Weynand B, et al. Results of selective neck dissection in the primary management of head and neck squamous cell carcinoma. *Eur Arch Otorhinolaryngol*. 2009;266:437–43.
11. Chone CT, Magalhes RS, Etchehebere E, et al. Predictive value of sentinel node biopsy in head and neck cancer. *Acta Otolaryngol*. 2008;128:920–4.
12. Bilde A, von Buchwald C, Therkildsen MH, et al. Need for intensive histopathologic analysis to determine lymph node metastases when using sentinel node biopsy in oral cancer. *Laryngoscope*. 2008;118:408–14.
13. Civantos FJ, Moffat FL, Goodwin WJ. Lymphatic mapping and sentinel lymphadenectomy for 106 head and neck lesions: contrasts between oral cavity and cutaneous malignancy. *Laryngoscope*. 2006;112:1–15.
14. Civantos FJ, Zitsch RP, Schuller DE, et al. Sentinel lymph node biopsy accurately stages the regional lymph nodes for T1–T2 oral squamous cell carcinomas: results of a prospective multi-institutional trial. *J Clin Oncol*. 2010;28:1395–400.
15. Shoaib T, Soutar DS, MacDonald DG, et al. The accuracy of head and neck carcinoma sentinel lymph node biopsy in the clinically N0 neck. *Cancer*. 2001;91:2077–83.
16. Zitsch RP 3rd, Todd DW, Renner GJ, et al. Intraoperative radiolymphoscintigraphy for detection of occult nodal metastasis in patients with head and neck squamous cell carcinoma. *Otolaryngol Head Neck Surg*. 2000;122:662–6.
17. Civantos FJ, Gomez C, Duque C, et al. Sentinel node biopsy in oral cavity cancer: correlation with PET scan and immunohistochemistry. *Head Neck*. 2003;25:1–9.
18. Alkureishi LW, Ross GL, Shoaib T, et al. Sentinel node biopsy in head and neck squamous cell cancer: 5-year follow-up of a European multicenter trial. *Ann Surg Oncol*. 2010;17:2459–64.
19. Vera DR, Wallace AM, Hoh CK, et al. A synthetic macromolecule for sentinel node detection: (99m)Tc-DTPA-mannosyldextra. *J Nucl Med*. 2001;42:951–9.
20. Ross GL, Soutar DS, MacDonald GD, et al. Sentinel node biopsy in head and neck cancer: preliminary results of a multicenter trial. *Ann Surg Oncol*. 2004;11:690–6.
21. Civantos FJ, Agrawal A, Lai SY. Comparison of false negative rates and overall accuracy of sentinel lymph node biopsy in phase 3 ^{99m}Tc-tilmanocept vs ACOSOG Z-0360 ^{99m}Tc-sulfur colloid in head/neck squamous cell cancer. In: Society of nuclear medicine and molecular imaging annual meeting, 11 Jun 2013, Vancouver.
22. Wallace AM, Hoh CK, Limmera KK, et al. Sentinel lymph node accumulation of Lymphoseek and Tc-99m-sulfur colloid using a "2-day" protocol. *Nucl Med Biol*. 2009;36:687–692.

Reprinted by permission of JAMA Otolaryngol Head Neck Surg. 2016; 142(12):1171-1176.

Original Investigation

Sentinel Lymph Node Biopsy for Cutaneous Squamous Cell Carcinoma on the Head and Neck

Alison B. Durham, MD; Lori Lowe, MD; Kelly M. Malloy, MD; Jonathan B. McHugh, MD; Carol R. Bradford, MD; Heather Chubb, MS; Timothy M. Johnson, MD; Scott A. McLean, MD, PhD

IMPORTANCE Metastasis of cutaneous squamous cell carcinoma (SCC) to the nodal basin is associated with a poor prognosis. The role of sentinel lymph node biopsy (SLNB) for regional staging in patients diagnosed with SCC is unclear.

OBJECTIVE To evaluate a single institution's experience with use of SLNB for regional staging of SCC on the head and neck.

DESIGN, SETTING, AND PARTICIPANTS A retrospective review of 53 patients who were diagnosed with SCC on the head and neck, at high risk for nodal metastasis based on National Comprehensive Cancer Network (NCCN) risk factors, and treated with wide local excision (WLE) and SLNB from December 1, 2010, through January 30, 2015, in a single academic referral center was performed. The follow-up period ended November 5, 2015. Sentinel lymph node biopsy paraffin blocks were retrieved and processed retrospectively with serial sectioning and immunohistochemical analysis (IHC) in cases with nodal recurrence following a negative SLNB.

MAIN OUTCOMES AND MEASURES Sentinel node (SN) identification rate, SLNB positivity rate, local recurrence, regional nodal recurrence, and distant recurrence.

RESULTS In 53 patients with 54 tumors, the SN identification rate was 94%. The SLNB positivity rate was 11.3%. On more thorough tissue processing and IHC, metastatic SCC was identified in 2 of 5 (40%) cases previously deemed negative. After reclassification of these cases, the adjusted SLNB positivity rate was 15.1%. The adjusted rate of false omission was 7.1% (95% CI, 2%-19%). Nodal disease developed in 20.8% overall. Angiolymphatic invasion (Cohen d, 3.52; 95% CI, 1.83-5.21), perineural invasion (Cohen d, 0.81; 95% CI, 0.09-1.52), and clinical size (Cohen d, 0.83; 95% CI, 0.05-1.63) were associated with the presence of nodal disease.

CONCLUSIONS AND RELEVANCE Rigorous study of SLNB for cutaneous SCC incorporating prospectively-collected comprehensive data sets based on standardized treatment algorithms is justified with potential to modify clinical practice. Our study demonstrates the critical importance of serial sectioning and IHC of the SLNB specimen for accurate diagnosis. Use of the NCCN guidelines may facilitate identification of patients with SCC at high risk for nodal metastasis.

JAMA Otolaryngol Head Neck Surg. doi:10.1001/jamaoto.2016.1927
Published online July 20, 2016.

Author Affiliations: Department of Dermatology, University of Michigan Medical School and Comprehensive Cancer Center, Ann Arbor (Durham, Lowe, Chubb, Johnson); Department of Pathology, University of Michigan Medical School and Comprehensive Cancer Center, Ann Arbor (Lowe, McHugh); Department of Otolaryngology, University of Michigan Medical School and Comprehensive Cancer Center, Ann Arbor (Malloy, Bradford, Johnson, McLean); Division of Plastic Surgery, Department of Surgery, University of Michigan Medical School and Comprehensive Cancer Center, Ann Arbor (Johnson).

Corresponding Author: Alison B. Durham, MD, University of Michigan Health System, 1910 A. Alfred Taubman Center, SPC 5314, 1500 E Medical Center Dr, Ann Arbor, MI 48109 (ambates@med.umich.edu).

Squamous cell carcinoma (SCC) is the second most common skin cancer type with a continually increasing incidence and a predilection for chronically sun exposed sites including the head and neck.[1] Although the majority of cutaneous SCC is diagnosed early and treatment is curative, metastasis and death occurs. The regional lymph node basin is the site of first metastasis in roughly 85% of cases. The 5-year survival rate decreases from more than 90% for local disease to roughly 30% when regional node metastasis occurs.[2] The estimated number of annual nodal metastases ranges from 5604 to 12 572; annual deaths from 3932 to 8791.[3] Sentinel lymph node biopsy (SLNB) is standard care for staging the regional nodal basin for melanoma and Merkel cell carcinoma in appropriate patients.[4,5] Accurate staging drives treatment and treatment options. For melanoma, microscopic detection with SLNB and early completion lymph node dissection (CLND) results in improved regional control, fewer adverse effects, fewer overall number of positive nodes, and potential for small but improved survival in node-positive patients.[6] For Merkel cell carcinoma, microscopic detection with SLNB drives primary and adjuvant surgery and radiation decision making.[5] In contrast, it is unclear if SLNB has any benefit for high-risk cutaneous SCC. Our purpose was to report our series utilizing SLNB in the management of cutaneous SCC on the head and neck, and add unique data to contemporary reports for optimal design of future studies.

Key Points

Question Should patients with cutaneous squamous cell carcinoma (SCC) on the head and neck be considered for staging with sentinel lymph node biopsy (SLNB)?

Findings In this retrospective review of 53 patients, nodal metastasis was identified in 15.1% by SLNB and the rate of false omission was 7.1%. The importance of histologic processing of SLNB specimens was demonstrated.

Meaning Our findings indicate that there may be a role for SLNB in the treatment of SCC on the head and neck for patients at high risk of nodal metastasis as defined by the National Comprehensive Cancer Network guidelines.

Methods

Following University of Michigan institutional review board approval, a database was created to identify patients with head and neck cutaneous SCC treated at our institution with wide local excision (WLE) and SLNB for potential retrospective analysis. Written consent for inclusion in the database was obtained from patients at their consultation visit, and participants were not compensated. Patients treated from December 2010 to January 2015 were identified. Demographic, clinical, and histopathological data were obtained via the electronic medical record and by telephone contact with the patient if data was missing. The follow up period ended November 5, 2015. Patients with multiple or prominent National Comprehensive Cancer Network (NCCN) risk factors for regional lymph node metastasis were considered for SLNB. Risk factors included: Breslow depth of 2 mm or more or Clark level of IV or V; rapid growth; locally recurrent; occurrence in a prior radiation or chronic inflammation and/or ulcer site; perineural invasion (PNI), angiolymphatic invasion (ALI); immunosuppression; size of 1 cm or more on the cheek, forehead, scalp, neck, or 0.6 cm or more on the face mask area; and poorly differentiated histologic pattern.[7]

Patients underwent preoperative lymphoscintigraphy using a mean dose of 2.3 μCi technetium Tc 99m sulfur colloid (CIS-US Inc) injected intradermally at the primary lesion site. Single photon emission computed tomography (SPECT-CT) imaging was performed 15 to 30 minutes following injection. Approximately 1 mL of vital blue dye (methylene blue or indigo carmine) was subsequently injected intradermally at the lesion site. Wide local excision was performed first to minimize shine-through from radiocolloid. Following WLE, a handheld gamma probe (Navigator GPS; RMD Instruments) was used to interrogate the nodal basins transcutaneously, using SPECT-CT as a guide. Each SN was dissected through small incisions from surrounding tissue using blunt dissection, taking care to identify and preserve nearby neurovascular structures. Tissue (WLE and SLNB) was processed using formalin-fixed permanent sections. Depending on size, SNs were bivalved or serially sectioned and stained with hematoxylin-eosin (H&E). Cytokeratin immunohistochemical (IHC) staining was variably performed per pathologist preference. Patients with a positive SLNB were counseled to undergo CLND. Adjuvant radiation or chemoradiation was individually considered under the auspices of the Multidisciplinary Head and Neck Tumor Board.

Demographic and clinical variables abstracted included: age, gender, primary vs recurrent, SCC arising within an area of prior radiation or chronic ulcer, immunosuppression, rapid growth, location, and clinical size. Treatment data included: excision margin size (cm) and adjuvant therapy if performed. Histopathologic factors from the initial biopsy and WLE included: histologic pattern, PNI, and ALI. Sentinel lymph node biopsy factors included: number of SNs, positive or negative, extracapsular extension (ECE), and IHC staining. Completion lymph node dissection factors included: number of nodes, positive or negative, and ECE. Outcome measures included: SN identification rate, SLNB positivity rate, local recurrence, regional nodal recurrence, and distant recurrence.

Sentinel lymph node biopsy paraffin blocks were retrieved for retrospective processing in cases with nodal recurrence in the basin following a negative SLNB. Slides were processed with 3 levels deeper in the tissue block separated by 50 to 80 μm. Four consecutive slides were stained at each level as: (1) H&E, (2) pancytokeratin (Cam 5.2 BD Biosciences, clone 5.2, dilution 1:40 and AE1/AE3 EMD Millipore, clone AE1/AE3, dilution 1:200;), (3) cytokeratin MNF-116 (DAKO, clone MNF 116, dilution 1:100), and (4) unstained. Initial and newly processed slides were reviewed independently by 2 pathologists (L.L. and J.B.M.).

All clinical and laboratory assessments were summarized with standard descriptive statistics. Continuous variables were

summarized using mean, standard error, and range. Categorical variables were summarized by frequency and percentage for each response category (N, %). Standard strategies for assessing diagnostic test accuracy were employed. A t test was used to determine if continuous assessments were significantly different between the groups based on nodal disease status. A Wilcoxon-Mann-Whitney test with exact *P* values was used for ordinal assessments or when normality was violated. Fisher exact or χ^2 tests assessed group differences for categorical data. The standardized mean difference effect size, Cohen d, and corresponding 95% CIs were computed using means, standard deviations, and χ^2 φ coefficients. All data was analyzed using SAS statistical software (SAS Institute, Inc; version 9.3) and the Practical Meta-Analysis Effect Size Calculator.[8]

Results

Fifty-three patients with 54 tumors treated with WLE and SLNB were identified. Mean age was 73 years (range, 47-90 years). Nine (17%) were women; 44 (83%) were men. Twenty-four (44.4%) tumors were located on the cheek, temple, or forehead; 14 (25.9%) on the scalp; 9 (16.7%) on the ear; 4 (7.4%) on the lip; 2 (3.7%) on the neck; and 1 (1.9%) on the nose. Six (11.1%) were recurrent. One (1.9%) developed within an area of radiation and 1 (1.9%) within a chronic ulcer. Fourteen tumors (25.9%) exhibited rapid growth. Mean lesion clinical diameter was 2.56 cm. Ten (18.5%) initial biopsies showed a well differentiated histologic pattern, 23 (42.6%) were moderately differentiated, 15 (27.8%) were poorly differentiated, 2 (3.7%) were sarcomatoid, and 4 (7.4%) did not have a histologic pattern reported. Fourteen (26.4%) patients were immunosuppressed; 9 had an organ transplant, 2 had chronic lymphocytic leukemia, 1 had non-Hodgkin lymphoma, and 2 patients were on immunosuppressive medication for ulcerative colitis and rheumatoid arthritis, respectively. A WLE was performed and SLNB attempted for all 54 lesions. The mean WLE margin was 1.3 cm. The tumor in the WLE specimen exhibited higher grade tumor differentiation compared with the diagnostic biopsy in 9 (17%) lesions: 6 graded initially as well differentiated were changed to moderate and 3 went from moderate to poor.

Although PNI and ALI were inconsistently reported, PNI was noted in 19 (35.2%) tumors and ALI in 5 (9.3%). Eleven (57.9%) tumors with PNI were poorly differentiated, 7 (36.8%) were moderately differentiated, and 1 (5.3%) was well differentiated. Three (60%) of the tumors with ALI were poorly differentiated, 2 (40%) were moderately differentiated. Four tumors with ALI also had PNI.

The SN was identified in 50 (94%) of 53 patients. Tracers failed to migrate in 1 failed SLNB, low radioactivity counts minimally elevated over background with no identifiable blue node were noted in 1, and no nodal tissue was identified by histological examination in the third failed SLNB. The average number of SNs identified per case was 3 (range 1-8). Six (11.3%) of the 53 patients had a positive SLNB, prior to retrospective reanalysis with more thorough tissue processing as below. Five had 1 positive node and 1 had 2 positive nodes, with ECE noted in 2 (33%) of the 6 positive SLNB cases. Immunohistochemical analysis was performed in 29 (58%) of 50 patients where SNs were identified. Of the 6 patients who had a positive SN, 3 had IHC performed. In 1 case, the SN was noted to be positive only on IHC. Five of the 6 patients with a positive SLNB underwent CLND. One patient was diagnosed with multiple comorbidities following SLNB, obviating CLND. Two (40%) of the 5 who underwent CLND had additional positive nodes (1/21 and 13/26 nodes, respectively).

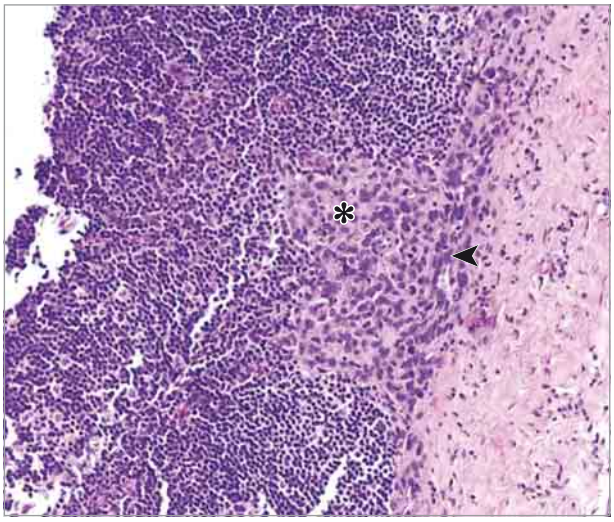
Mean follow up time for the entire group was 25.5 months (range, 2-57 months). Local recurrence occurred in 5, with an average time of 11 months (range, 3-24 months). In 3, SCC invaded the central nervous system, causing death. Regional nodal recurrence occurred in 6 patients; 5 following a negative SLNB and 1 following a positive SLNB treated with CLND. Two of these patients first developed a local recurrence (2 and 4 months prior to nodal recurrence, respectively). On retrospective review of the SLNB specimens (as detailed below), 1 of these patients was found to have a positive SLNB. Because of this finding and because we did not want to underestimate the development of nodal disease in this high-risk population, we did not exclude patients from the study analysis if they had a clinical local recurrence prior to clinical nodal recurrence. Average time to nodal recurrence was 7.5 months (range, 2-22 months). Two patients developed distant metastasis. One had a failed SLNB with bone metastasis 17 months later. The other developed lung metastases 4 years after WLE and negative SLNB, however, in the interim had developed many other primary cutaneous SCCs.

Thus, in this patient cohort, there were 5 false-negative SLNB results. The false-negative rate was 45.5% (5 false negatives/[5 false negatives +6 true positives]), 95% CI, 21% to 72%. The false-omission rate (patients with a negative SLNB that failed in the nodal basin) was 11.4% (5 false negatives/[5 false negatives +39 true negatives]), 95% CI, 5% to 24%.

Overall, 11 (20.8%) patients had nodal disease identified by SLNB or palpable recurrence. Angiolymphatic invasion (Cohen d, 3.52; 95% CI, 1.83-5.21), perineural invasion (Cohen d, 0.81; 95% CI, 0.09-1.52), and clinical size (Cohen d, 0.83; 95% CI, 0.05-1.63) were associated with the presence of nodal disease. All patients with nodal disease were referred for adjuvant therapy; 1 declined. Two completed radiation to the nodal basin. Eight had radiation to the primary site and nodal basin, 2 of these 8 had concurrent chemotherapy, with carboplatin in 1 and cisplatin in the other.

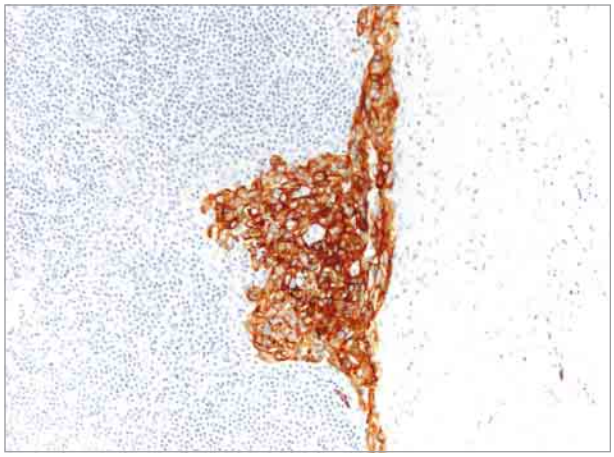
The 5 original SLNB tissue blocks from patients with a negative SLNB and nodal recurrence in the negative basin were retrieved and processed with more thorough serial sectioning and IHC. On independent review by 2 pathologists, metastatic SCC was identified in deeper sections by both pathologists in 2 of 5 cases (40%). In 1, deeper sections revealed SCC evident on both H&E and IHC (**Figure 1** and **Figure 2**). In the other, SCC was only identified by IHC. The original H&E and IHC (performed in 4 cases) slides were confirmed negative by both pathologists. After reclassification of these 2 cases as positive, our adjusted false-negative rate was 27.3% (3 false negatives/[3 false negatives +8 true positives]), 95% CI, 10% to 57%. The adjusted false omission rate was 7.1% (3 false negatives/[3 false negatives +39 true negatives]), 95% CI, 2% to 19%.

Figure 1. Histopathologic Image



Deeper section into the block demonstrates a focus of metastatic squamous cell carcinoma involving the subcapsular sinus (black arrowhead) and parenchyma (asterisk) of sentinel lymph node. Hematoxylin-eosin stain (original magification ×200).

Figure 2. Immunostain



Focus of metastatic squamous cell carcinoma in sentinel lymph node staining with pancytokeratin immunostain (original magnification ×100).

Sentinel lymph node biopsy after prior wide local excision, at least theoretically, may be less accurate owing to prior surgery at the primary site. In this cohort, 1 patient with recurrent SCC as an indication for mapping was found to have a positive SLNB. After reclassification of the SLNB status in 2 cases, as above, no patients with recurrent SCC as an indication for staging with SLNB had a nodal recurrence following negative SLNB.

Discussion

We present data on 53 patients with cutaneous SCC on the head and neck treated with WLE and SLNB, the largest single-institution cohort reported to date. Our results and previous data form a foundation and validate the need for rigorous prospective study of SLNB for cutaneous SCC, with potential to modify clinical practice. Our results confirm feasibility of SLNB for head and neck cutaneous SCC identifying a SN in 94% of cases with the combined use of radiocolloid, vital blue dye, and SPECT-CT. We uniquely demonstrate the critical importance of serial sectioning and IHC of the SLNB specimen for accurate diagnosis.

The data, including our own, pertaining to SLNB for cutaneous head and neck SCC is globally limited by heterogeneous risk factor reporting; inconsistent data, surgical details, and study design; relatively small numbers, limited follow-up, and most of the data are retrospective in nature.[9-22]

Several factors may lead to higher rates of nodal recurrence after a negative SLNB including: surgeon, pathologist, and nuclear medicine experience and/or technique; prior surgery in the area with scar tissue affecting migration of the tracers; accuracy of tracer injection sites; and specimen processing. The increased accuracy of SLNB on the head and neck for melanoma with the use of SPECT-CT is documented.[23] Our work underscores the importance of standardizing SLNB technique and histopathological tissue processing protocols for cutaneous SCC. Numerous studies document enhanced detection of small tumor deposits by use of comprehensive serial sectioning and IHC for melanoma.[24-30] Sentinel lymph node biopsy processing for SCC is limited by a paucity of data. One study[9] of SLNB for mucosal SCC utilizing IHC staining reported an approximately 10% higher detection rate of metastatic deposits in the SN with IHC compared with use of H&E alone. While the use of frozen sections for analysis of the SLNB for SCC guides proceeding to an immediate CLND, reliability data are absent with clinically significant consequences for false-positive and false-negative results, which both occur. Based on our experience, optimal histopathological evaluation of the SLNB for cutaneous SCC includes formalin-fixed, permanent section processing with serial sectioning with H&E and IHC staining.

A systematic literature review analyzing SLNB for cutaneous SCC on the head and neck was published in 2014. Eleven publications with 73 total patients met the authors' inclusion criteria (range 1-15 patients/report, median 5). The overall rate of SLNB positivity was 13.5%. The rate of regional nodal recurrence in the same basin following a negative SLNB was 4.76% (range 0%-33%).[31] A more rigorous multi-center prospective study of SLNB for high-risk cutaneous SCC on the head and neck involving 57 patients was published in 2015. Patients had at least 1 high-risk factor defined as tumor size larger than 2 cm, poorly differentiated histology, perineural invasion, lymphovascular invasion, invasion into the subcutaneous fat or thickness of more than 5mm, local recurrence, location on the ear or lip, immunosuppression, and SCC arising in a scar. Seven (12.3%) of 57 had a positive SLNB. The SLNB specimens were processed with formalin-fixed permanent sections stained with H&E and IHC in 55, with 2 processed with frozen sections because the SN was deemed suspicious for metastatic SCC intraoperatively. No nodal recurrences were reported following a negative SLNB; mean follow up was 19.4

months. One nodal recurrence occurred after a positive SLNB, another after a failed SLNB. The overall rate of nodal disease was 14% (7 positive SLNB, 1 nodal recurrence). Predictors of nodal disease were multiple high-risk factors (P = .008), PNI (P = .05), and ALI (P = .05).[32]

The lack of a cutaneous SCC National Tumor Registry impedes large retrospective multi institutional analysis of prognostic factors. Risk factors associated with a higher rate of local recurrence and metastases are currently defined based on low-moderate evidence and expert consensus.[7,33,34] We evaluated our data using effect size to aid in comparison of the relative size of effect of each NCCN high-risk feature with regard to the presence of nodal disease and found that presence of ALI, presence of PNI, and a large clinical size had a large effect on the development of nodal disease. The large width of the CIs around the estimates of the false-negative and false-omission rates, however, exposes the small sample size and demonstrates the variability of these estimates. Until higher level evidence is produced, our results, which are relatively consistent with the literature, suggest that utilization of the NCCN guidelines may facilitate appropriate patient selection for future study design and current consideration for SLNB.[7]

Limitations

Limitations of our study include a retrospective design associated with missing data of some variables of interest, relatively short follow up including some patients lost to follow up after the immediate postoperative period, and overall small numbers despite being the largest single institution report. The purpose of our study was to review our institutional experience utilizing SLNB for cutaneous SCC on the head and neck to provide a basis to optimize future prospective analyses over a long period of time with long-term follow-up. We included outcomes data, although not complete, for all patients to add to the current body of literature on the subject, acknowledging that, owing to the limited follow up for some of our patients, the rates of recurrence and false-omission may be underestimates. Despite these limitations, our study provides unique data, particularly with regard to histologic processing of the SLNB specimens, and additional evidence to justify future investigation incorporating prospectively-collected, homogeneous, comprehensive data sets based on standardized treatment algorithms.

Conclusions

Rigorous study with optimal methodology is necessary to improve surgical and histopathologic protocols for SLNB for cutaneous SCC and to advance our understanding of what role SLNB may play with respect to improved staging for patients at high risk of nodal metastasis. Further work will be necessary to determine if early identification and intervention leads to improved outcomes for these patients.

ARTICLE INFORMATION

Accepted for Publication: June 3, 2016.

Published Online: July 20, 2016. doi:10.1001/jamaoto.2016.1927.

Author Contributions: Alison B. Durham 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: Durham, Lowe, Malloy, Bradford, Johnson, McLean.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Durham, Lowe, Malloy, Chubb.
Critical revision of the manuscript for important intellectual content: Durham, Lowe, Malloy, McHugh, Bradford, Johnson, McLean.
Statistical analysis: Chubb.
Administrative, technical, or material support: Durham, McLean.
Study supervision: Lowe, Malloy, Bradford, Johnson, McLean.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Previous Presentation: This study was presented at the American Head and Neck Society Ninth International Conference on Head and Neck Cancer; July 20, 2016; Seattle, Washington.

REFERENCES

1. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. *JAMA Dermatol*. 2015;151(10):1081-1086.

2. Stratigos A, Garbe C, Lebbe C, et al; European Dermatology Forum (EDF); European Association of Dermato-Oncology (EADO); European Organization for Research and Treatment of Cancer (EORTC). Diagnosis and treatment of invasive squamous cell carcinoma of the skin: European consensus-based interdisciplinary guideline. *Eur J Cancer*. 2015;51(14):1989-2007.

3. Karia PS, Han J, Schmults CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *J Am Acad Dermatol*. 2013;68(6):957-966.

4. Coit DG, et al. NCCN Clinical Practice Guidelines in Oncology: Melanoma. Version I.2016. https://www.nccn.org/. Accessed June 1, 2016.

5. Bichakjian CK, et al. NCCN Clinical Practice Guidelines in Oncology: Merkel Cell Carcinoma. Version I.2016. https://www.nccn.org/. Accessed June 1, 2016.

6. Morton DL, Thompson JF, Cochran AJ, et al; MSLT Group. Final trial report of sentinel-node biopsy versus nodal observation in melanoma. *N Engl J Med*. 2014;370(7):599-609.

7. Bichakjian CK, et al NCCN Clinical Practice Guidelines in Oncology: Squamous Cell Skin Cancer. Version I.2016. https://www.nccn.org/. Accessed June 1, 2016.

8. Wilson DB. Practical Meta-Analysis Effect Size Calculator. http://www.campbellcollaboration.org/resources/effect_size_input.php. Accessed June 1, 2016.

9. Ross AS, Schmults CD. Sentinel lymph node biopsy in cutaneous squamous cell carcinoma: a systematic review of the English literature. *Dermatol Surg*. 2006;32(11):1309-1321.

10. Eastman AL, Erdman WA, Lindberg GM, Hunt JL, Purdue GF, Fleming JB. Sentinel lymph node biopsy identifies occult nodal metastases in patients with Marjolin's ulcer. *J Burn Care Rehabil*. 2004;25(3):241-245.

11. Hatta N, Morita R, Yamada M, Takehara K, Ichiyanagi K, Yokoyama K. Implications of popliteal lymph node detected by sentinel lymph node biopsy. *Dermatol Surg*. 2005;31(3):327-330.

12. Stadelmann WK, Javaheri S, Cruse CW, Reintgen DS. The use of selective lymphadenectomy in squamous cell carcinoma of the wrist: a case report. *J Hand Surg Am*. 1997;22(4):726-731.

13. Weisberg NK, Bertagnolli MM, Becker DS. Combined sentinel lymphadenectomy and mohs micrographic surgery for high-risk cutaneous squamous cell carcinoma. *J Am Acad Dermatol*. 2000;43(3):483-488.

14. Weber F, Bauer JW, Sepp N, et al. Squamous cell carcinoma in junctional and dystrophic epidermolysis bullosa. *Acta Derm Venereol*. 2001;81(3):189-192.

15. Ardabili M, Gambichler T, Rotterdam S, Altmeyer P, Hoffmann K, Stücker M. Metastatic cutaneous squamous cell carcinoma arising from a previous area of chronic hypertrophic lichen planus. *Dermatol Online J*. 2003;9(1):10.

16. Ozçelik D, Tatlidede S, Hacikerim S, Uğurlu K, Atay M. The use of sentinel lymph node biopsy in squamous cell carcinoma of the foot: a case report. *J Foot Ankle Surg*. 2004;43(1):60-63.

17. Yamada M, Hatta N, Sogo K, Komura K, Hamaguchi Y, Takehara K. Management of squamous cell carcinoma in a patient with recessive-type epidermolysis bullosa dystrophica. *Dermatol Surg*. 2004;30(11):1424-1429.

18. Perez-Naranjo L, Herrera-Saval A, Garcia-Bravo B, Perez-Bernal AM, Camacho F. Sentinel lymph node biopsy in recessive dystrophic epidermolysis bullosa and squamous cell carcinoma. *Arch Dermatol*. 2005;141(1):110-111.

19. Navarrete-Dechent C, Veness MJ, Droppelmann N, Uribe P. High-risk cutaneous squamous cell carcinoma and the emerging role of sentinel lymph node biopsy: A literature review. *J Am Acad Dermatol*. 2015;73(1):127-137.

20. Renzi C, Caggiati A, Mannooranparampil TJ, et al. Sentinel lymph node biopsy for high risk cutaneous squamous cell carcinoma: case series and review of the literature. *Eur J Surg Oncol*. 2007;33(3):364-369.

21. Takahashi A, Imafuku S, Nakayama J, Nakaura J, Ito K, Shibayama Y. Sentinel node biopsy for high-risk cutaneous squamous cell carcinoma. *Eur J Surg Oncol*. 2014;40(10):1256-1262.

22. Fukushima S, Masuguchi S, Igata T, et al. Evaluation of sentinel node biopsy for cutaneous squamous cell carcinoma. *J Dermatol*. 2014;41(6):539-541.

23. Stoffels I, Boy C, Pöppel T, et al. Association between sentinel lymph node excision with or without preoperative SPECT/CT and metastatic node detection and disease-free survival in melanoma. *JAMA*. 2012;308(10):1007-1014.

24. Spanknebel K, Coit DG, Bieligk SC, Gonen M, Rosai J, Klimstra DS. Characterization of micrometastatic disease in melanoma sentinel lymph nodes by enhanced pathology: recommendations for standardizing pathologic analysis. *Am J Surg Pathol*. 2005;29(3):305-317.

25. Yu LL, Flotte TJ, Tanabe KK, et al. Detection of microscopic melanoma metastases in sentinel lymph nodes. *Cancer*. 1999;86(4):617-627.

26. Gershenwald JE, Colome MI, Lee JE, et al. Patterns of recurrence following a negative sentinel lymph node biopsy in 243 patients with stage I or II melanoma. *J Clin Oncol*. 1998;16(6):2253-2260.

27. Li LX, Scolyer RA, Ka VS, et al. Pathologic review of negative sentinel lymph nodes in melanoma patients with regional recurrence: a clinicopathologic study of 1152 patients undergoing sentinel lymph node biopsy. *Am J Surg Pathol*. 2003;27(9):1197-1202.

28. Lobo AZ, Tanabe KK, Luo S, et al. The distribution of microscopic melanoma metastases in sentinel lymph nodes: implications for pathology protocols. *Am J Surg Pathol*. 2012;36(12):1841-1848.

29. Cochran AJ, Wen DR, Herschman HR. Occult melanoma in lymph nodes detected by antiserum to S-100 protein. *Int J Cancer*. 1984;34(2):159-163.

30. Karimipour DJ, Lowe L, Su L, et al. Standard immunostains for melanoma in sentinel lymph node specimens: which ones are most useful? *J Am Acad Dermatol*. 2004;50(5):759-764.

31. Ahmed MM, Moore BA, Schmalbach CE. Utility of head and neck cutaneous squamous cell carcinoma sentinel node biopsy: a systematic review. *Otolaryngol Head Neck Surg*. 2014;150(2):180-187.

32. Gore SM, Shaw D, Martin RC, et al. Prospective study of sentinel node biopsy for high-risk cutaneous squamous cell carcinoma of the head and neck. *Head Neck*. 2016;38(suppl 1):E884-E889.

33. Thompson AK, Kelley BF, Prokop LJ, Murad MH, Baum CL. Risk Factors for Cutaneous Squamous Cell Carcinoma Recurrence, Metastasis, and Disease-Specific Death: A Systematic Review and Meta-analysis. *JAMA Dermatol*. 2016;152(4):419-428.

34. Schmults CD, Karia PS, Carter JB, Han J, Qureshi AA. Factors predictive of recurrence and death from cutaneous squamous cell carcinoma: a 10-year, single-institution cohort study. *JAMA Dermatol*. 2013;149(5):541-547.

Reprinted by permission of Laryngoscope. 2016; 126(1):9-10.

The Laryngoscope

The Triological Society
Best Practice

What is the Role of Sentinel Lymph Node Biopsy in Early-Stage Oral Cavity Carcinoma?

Vikas Mehta, MD, MPH, FACS; Cherie-Ann Nathan, MD, FACS

BACKGROUND

Lymphatic spread in oral cavity squamous cell carcinoma (OSCC) remains a critical factor for staging, treatment, and prognosis. Involvement of the regional lymphatics portends approximately 50% decrease in survival. OSCC–cervical metastases remain common. Due to the inaccuracy of the physical exam and imaging to reliably detect occult disease, elective neck dissections (ENDs) have become the standard of care for the majority of clinically node-negative (cN0) patients. However, many patients (55%–76%)[1–3] with T1/2 cN0 OSCC disease will not have pathologically positive cervical metastases, and are being subjected to overtreatment with unnecessary morbidity by an END. Sentinel lymph node biopsy (SLNB) has emerged as a powerful tool for advancing minimally invasive surgical management of many cancers. SLNB has been proven to be highly sensitive, cost-effective, and beneficial to patient quality of life. The data supporting the use of SLNB in early-stage OSCC, a brief description of the SLNB method, and recent technical advances are the focus of this article.

LITERATURE REVIEW

A prospective multi-institutional trial was conducted with 25 institutions and 34 surgeons.[1] The study enrolled patients with newly diagnosed T1/T2 cN0 OSCC over 3 years. All patients underwent SLNB during the primary resection followed by END. Of the 140 eligible patients, there were 52 T1 (37.1%) and 88 T2 lesions (62.9%). Forty-one patients (29%) had positive nodes after sentinel lymph node (SLN) sectioning and immunohistochemistry (IHC), with 21 having the SLN as the sole positive node. Of the 106 negative SLNBs, 100 were classified as truly negative on final pathology of the neck dissection (ND) specimen, which corresponded to a 0.94 negative predictive value (NPV) (95% confidence interval [CI]: 0.88-0.98). Step-sectioning and IHC increased the NPV to 0.96, with T1 and T2 lesions having an NPV of 1.0 and 0.94, respectively. The false-negative rate (FNR) was 9.8% overall (four false negatives out of 41 positives, Table I). With an overall NPV of 0.96 in a population of 30% with metastatic disease, a negative SLNB would thus demonstrate a regional recurrence in only 4% of patients. Of the 140 patients, 100 could have been spared END.

A recent meta-analysis of 26 studies looking at SLNB for head and neck cancer combined 593 early-stage (T1/2) OSCC patients who had undergone SLNB and a concurrent END.[3] The SLNB was positive in 177 patients (29.8%) and true negative in 408 patients (68.8%). The overall sensitivity and NPV of SLNB in the OSCC cohort were 94% (95% CI: 89–98%) and 96% (95% CI: 93–99%), respectively (Table I). This included an additional 38 OSCC patients with T3/4 patients (n = 631). There were only 12 patients (<2%) misclassified as N0 on SLNB who had a positive concurrent END, with eight patients having early T1/T2 oral cavity tumors and four patients with T3/T4 OSCC. A separate analysis was done that included five studies, which examined regional recurrence in oral cavity and oropharyngeal SCC patients who did not receive END following negative SLNB. There were 11 documented regional recurrences from 200 total patients (5.5%) with a follow-up of ≥2 years.

Once a positive SLN is identified, the finding has both therapeutic and prognostic implications. A retrospective study of 109 cT1/T2 N0 OSCC patients with positive SLNB from 15 centers was conducted.[4] All patients had subsequently undergone ND at the time of the SLNB or within 3 weeks (I–III, 13%; I–IV, 23%; I–V, 64%) for a total of 122 ND specimens with additional (+non-SLN) metastases in 42/122 (34.4%). In those patients with +non-SLN, 18/42 patients (42.9%) had

From the Department of Otolaryngology/Head and Neck Surgery, Louisiana State University Health–Shreveport, Shreveport, Louisiana, U.S.A.

Editor's Note: This Manuscript was accepted for publication July 6, 2015.

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

Send correspondence to Vikas Mehta, MD, Co-Director of Head and Neck Surgical Oncology, Feist-Weiller Cancer Center, 1501 Kings Highway, Rm 9-203, Shreveport, LA 71130.
E-mail: dr.vikasmehta@gmail.com

DOI: 10.1002/lary.25541

TABLE I.
False Negative Rates and NPV of SLNB for OSCC.

Study	No. of Patients	False-Negative Rate	Negative Predictive Value (95% CI)
Thompson et al.	631	1.9%	96% (93%–99%)
Civantos et al.	140	9.8%	96% (90%–98%)
Agrawal et al.	83	2.6%	98% (88%–99%)

CI = confidence interval.

disease only in the same level as the positive SLNs, 21/42 patients (50%) had additional disease in an adjacent nodal level (7/21 higher and 14/21 lower) from the positive SLN, and 3/42 patients (7.1%) had disease in a non-adjacent level. Only one ND yielded neck nodes in levels other than I to III. The three factors that predicted +non-SLN in multivariate analysis were lymphovascular invasion, positive margins, and non-SLN extracapsular spread. Only 15 patients (13.7%) developed recurrence, with six of those being regional. Kaplan-Meier and log-rank analyses showed only two variables to be significant for nodal recurrence: positive lymph nodes in addition to the SLN and +non-SLN in levels outside of the SLN ($P = .04$ and .01, respectively). In breast disease, those patients with +non-SLNs in different fields receive more aggressive adjuvant therapy, and nomograms have been developed to predict the presence of +non-SLNs in these fields. Similar approaches could be taken with OSCC.

The SLNB is traditionally performed by injecting the primary site with unfiltered ^{99}Tc-sulfur colloid within 18 hours of the procedure. Dosages are adjusted based on the timing. Serial nuclear imaging is then performed. Some authors advocate injecting methylene blue at the time of resection. After removal of the primary, the SLNB is performed utilizing a small incision within the planned END incision, with the SLNs identified using the gamma probe. Any lymph node exhibiting >10% of the radioactivity of the most active node are removed. The SLNs are then sectioned from hilum to periphery, longitudinally, at 2- to 3-mm thickness and hematoxylin and eosin stained for immediate analysis. If the lymph nodes are not grossly positive, the central laboratory evaluates the nodes in permanent section and stains the slides for cytokeratin using IHC. Any IHC cytokeratin-positive clusters are further reviewed for morphology consistent with metastatic SCC. Two novel methods have recently emerged that could improve the SLNB process. The use of [^{99}Tc]tilmanocept, a novel radiopharmaceutical that specifically target CD206 mannose-binding receptors on reticuloendothelial cells within lymph nodes, was recently investigated in a phase III multi-institutional trial.[2] Incorporating tilmanocept resulted in an NPV of 97.8%, an FNR of 2.56%, and an overall accuracy in correctly determining the nodal status of 98.8% (Table I). Quantitative real-time polymerase chain reaction (qRT-PCR) has also shown potential to increase the sensitivity of SLNB in detecting carcinoma microdeposits. Ferris et al. demonstrated in a validation set of 102 nodes that a multiplexed assay using two markers for squamous cell carcinoma demonstrated excellent reproducibility, linearity, and accuracy (96% NPV) for identifying positive and negative nodal status.[5]

BEST PRACTICE

SLNB has emerged as a powerful adjunct to END in early-stage OSCC to identify cervical metastases, which can have significant therapeutic and prognostic implications. The method has shown excellent NPV that can be even more effective with novel radiopharmaceuticals and qRT-PCR. This technique, when properly conducted, can reliably be done in lieu of an END for cT1/2 N0 OSCC, thereby avoiding unnecessary morbidity and cost.

LEVEL OF EVIDENCE

Recommendations for SLNB for early stage OSCC is based on level II evidence, with a meta-analysis conducted of level II studies.

BIBLIOGRAPHY

1. Civantos FJ, Zitsch RP, Schuller DE, et al. Sentinel lymph node biopsy accurately stages the regional lymph nodes for T1-T2 oral squamous cell carcinomas: results of a prospective multi-institutional trial. *J Clin Oncol* 2010;28:1395–1400.
2. Agrawal A, Civantos FJ, Brumund KT, et al. [Tc]Tilmanocept accurately detects sentinel lymph nodes and predicts node pathology status in patients with oral squamous cell carcinoma of the head and neck: results of a phase III multi-institutional trial [published on line February 11, 2015. *Ann Surg Oncol*. doi: 10.1245/s10434-015-4382-x.
3. Thompson CF, St John MA, Lawson G, Grogan T, Elashoff D, Mendelsohn AH. Diagnostic value of sentinel lymph node biopsy in head and neck cancer: a meta-analysis. *Eur Arch Otorhinolaryngol* 2013;270:2115–2122.
4. Gurney BA, Schilling C, Putcha V, et al. Implications of a positive sentinel node in oral squamous cell carcinoma. *Head Neck* 2012;34:1580–1585.
5. Ferris RL, Stefanika P, Xi L, Gooding W, Seethala RR, Godfrey TE. Rapid molecular detection of metastatic head and neck squamous cell carcinoma as an intraoperative adjunct to sentinel lymph node biopsy. *Laryngoscope* 2012;122:1020–1030.

European Journal of Cancer 51 (2015) 2777–2784

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Sentinel European Node Trial (SENT): 3-year results of sentinel node biopsy in oral cancer

Clare Schilling [a], Sandro J. Stoeckli [b], Stephan K. Haerle [c], Martina A. Broglie [b], Gerhard F. Huber [d], Jens Ahm Sorensen [e], Vivi Bakholdt [e], Annelise Krogdahl [f], Christian von Buchwald [g], Anders Bilde [g], Lars R. Sebbesen [g], Edward Odell [h], Benjamin Gurney [a], Michael O'Doherty [i], Remco de Bree [j], Elisabeth Bloemena [k], Geke B. Flach [j], Pedro M. Villarreal [l], Manuel Florentino Fresno Forcelledo [m], Luis Manuel Junquera Gutiérrez [l], Julio Alvarez Amézaga [n], Luis Barbier [n], Joseba Santamaría-Zuazua [n], Augusto Moreira [o], Manuel Jacome [o], Maurizio Giovanni Vigili [p], Siavash Rahimi [q], Girolamo Tartaglione [r], Georges Lawson [s], Marie-Cecile Nollevaux [s], Cesare Grandi [t], Davide Donner [u], Emma Bragantini [v], Didier Dequanter [w], Philippe Lothaire [w], Tito Poli [x], Enrico M. Silini [y], Erinco Sesenna [x], Giles Dolivet [z], Romina Mastronicola [z], Agnes Leroux [aa], Isabel Sassoon [ab], Philip Sloan [ac], Mark McGurk [a,*]

[a] *Department of Head and Neck Surgery, Guys and St Thomas NHS Trust, London, UK*
[b] *Department of Otorhinolaryngology, Head and Neck Surgery Kantonsspital St Gallen, Switzerland*
[c] *Department of Head and Neck Surgery, University of Basel, Switzerland*
[d] *Department of Otolaryngology University Hospital Zurich, Switzerland*
[e] *Department of Plastic and Reconstructive Surgery, Odense University Hospital, Denmark*
[f] *Department of Pathology, Odense University Hospital, Denmark*
[g] *Department of Otolaryngology Head and Neck Surgery and Audiology, Rigshospitalet, Copenhagen, Denmark*
[h] *Head and Neck/Oral Pathology, King's College London, Guys and St Thomas NHS Trust, London, UK*
[i] *Department of Nuclear Medicine, Guys and St Thomas NHS Trust, London, UK*
[j] *Department of Otolaryngology Head and Neck Surgery, VU University Medical Center, Amsterdam, The Netherlands*
[k] *Department of Pathology, VU University Medical Centre and Academic Centre of Dentistry Amsterdam, The Netherlands*
[l] *Department of Maxillofacial Surgery, Hospital Universitario Central de Asturias, Oviedo, Spain*
[m] *Department of Pathology, Hospital Universitario Central de Asturias, Oviedo, Spain*
[n] *Department of Maxillofacial Surgery, BioCruces, Hospital Universitario De Cruces, Universidad del Pais Vasco (UPV/EHU), Bilbao, Spain*
[o] *Department of Head and Neck Surgery, Instituto Portugues de Oncologia do Porto, Portugal*
[p] *Department of Otorhinolaryngology, San Carlo Hospital Rome, Italy*

* *Corresponding author*: Department of Oral and Maxillofacial Surgery, Head and Neck Unit, 3rd Floor Bermondsey Wing, Guy's Hospital, Great Maze Pond, London, SE1 9RT, UK. Tel.: +44 2071884349, fax: +44 2071884360.
E-mail address: mark.mcgurk@kcl.ac.uk (M. McGurk).

http://dx.doi.org/10.1016/j.ejca.2015.08.023

[q] Department of Histopathology, San Carlo Hospital Rome, Italy
[r] Department of Nuclear Medicine, Cristo Re Hospital, Rome, Italy
[s] Department of Head and Neck Surgery, CHU Dinant Godinne, Université Catholique de Louvain, Belgium
[t] Department of Otolaryngology, Ospedale S. Chiara, Trento, Italy
[u] Department of Nuclear Medicine, Ospedale S. Chiara, Trento, Italy
[v] Department of Surgical Pathology, Ospedale S. Chiara, Trento, Italy
[w] Department of Maxillofacial Surgery CHU de Charleroi Belgium, Belgium
[x] Department of Maxillofacial Surgery, Azienda Ospedaliera, Universitaria of Parma, Italy
[y] Department of Pathology Azienda Ospedaliera Universitaria of Parma, Italy
[z] Department of Head and Neck Surgery Centre Alexis Vautrin, Vandoeuvre Les Nancy, France
[aa] Department of Pathology Centre Alexis Vautrin, Vandoeuvre Les Nancy, France
[ab] Department of Informatics, Kings' College London, UK
[ac] Department of Cellular Pathology, Newcastle University Hospital, UK

Received 13 July 2015; received in revised form 22 August 2015; accepted 23 August 2015
Available online 18 November 2015

KEYWORDS
Oral Cancer;
Sentinel lymph node biopsy;
Metastasis;
Recurrence;
Micrometastasis;
Lymphoscintigraphy

Abstract ***Purpose:*** Optimum management of the N0 neck is unresolved in oral cancer. Sentinel node biopsy (SNB) can reliably detect microscopic lymph node metastasis. The object of this study was to establish whether the technique was both reliable in staging the N0 neck and a safe oncological procedure in patients with early-stage oral squamous cell carcinoma.
Methods: An European Organisation for Research and Treatment of Cancer-approved prospective, observational study commenced in 2005. Fourteen European centres recruited 415 patients with radiologically staged T1–T2N0 squamous cell carcinoma. SNB was undertaken with an average of 3.2 nodes removed per patient. Patients were excluded if the sentinel node (SN) could not be identified. A positive SN led to a neck dissection within 3 weeks. Analysis was performed at 3-year follow-up.
Results: An SN was found in 99.5% of cases. Positive SNs were found in 23% (94 in 415). A false-negative result occurred in 14% (15 in 109) of patients, of whom eight were subsequently rescued by salvage therapy. Recurrence after a positive SNB and subsequent neck dissection occurred in 22 patients, of which 16 (73%) were in the neck and just six patients were rescued. Only minor complications (3%) were reported following SNB. Disease-specific survival was 94%. The sensitivity of SNB was 86% and the negative predictive value 95%.
Conclusion: These data show that SNB is a reliable and safe oncological technique for staging the clinically N0 neck in patients with T1 and T2 oral cancer.
EORTC Protocol 24021: Sentinel Node Biopsy in the Management of Oral and Oropharyngeal Squamous Cell Carcinoma.

1. Introduction

Head and neck squamous cell carcinoma is the eighth most common cancer worldwide in males and is increasing significantly amongst females [1]

Approximately half the patients with oral cancer present with stage I/II disease and up to 33% [2,3] have occult cervical disease undetectable by current imaging techniques (computed tomography [CT]/magnetic resonance imaging [MRI]/ultrasound/positron-emission tomography) [4,5]. Cervical metastasis is associated with a 50% reduction in cure. Consequently, if the estimated chance of metastasis exceeds 20% [6], current practice is to offer an elective neck dissection (END) rather than 'wait and see' policy [7]. The corollary of this strategy is that up to 80% of stage I/II patients undergo an unnecessary neck dissection.

Sentinel node biopsy (SNB) is capable of detecting occult metastases in head and neck cancer [8–11] and is becoming established in a range of other cancers [12,13]. SNB offers a potential solution for management of the N0 neck but at the present time it is not widely offered. There is a paucity of data on the expected success of the technique, particularly with respect to the accuracy of sentinel node (SN) detection, disease recurrence and survival. The Sentinel European Node Trial (SENT) study population is the largest cohort of oral cancer patients in which SNB was performed as a sole staging procedure without

concurrent END. The end-points of this study were SN identification rate, false-negative rate (FNR) and disease-free survival (DFS) at 3 years post-recruitment. Because of the lack of contemporaneous SNB data, we have used comparable data from patients treated by conventional END as well as SNB data in similar tumour groups who routinely use SNB in management of the neck to contextualise the results. The aim of this investigation was to assess whether SNB is a safe and reliable therapeutic technique in T1–T2 oral squamous cell carcinoma.

2. Patients and methods

A European multicentre prospective study (October 2005–October 2010) was approved by the European Organisation for Research and Treatment of Cancer and local ethics committee, with patients providing informed consent. Eligible patients had 0.5- to 4-cm squamous cell carcinoma with an N0 neck on CT and/or MRI (<1.1 cm or up to 1.5 cm in level II and no atypical features) ± ultrasound-guided fine needle aspiration cytology. SENT was principally designed for oral squamous cell carcinoma; however, tumour-bordering structures of the oropharynx which were transorally accessible and resectable (without mandibular split or robotic techniques) were also included. Tumour location was recorded according to Systematized Nomenclature of Medicine (SNOMED) topography code [14] apart from 'oral tongue posterior 1/3' which related to tumours of the posterior part of the oral tongue and not tongue base tumours.

Patients with a previous malignant neoplasm of the head and neck or any disease that might have altered lymphatic drainage were excluded. Patients had to be fit enough to tolerate a completion neck dissection if the SNB proved positive.

A total of 480 cases were recruited prospectively from 14 European centres. The criterion for unit participation was completion of at least 10 successful training SNB procedures (validated against neck dissection) prior to recruiting to SENT. Sixty-five patients (14%) were excluded from the final analysis (Table 1).

Table 1
Cases excluded from SENT trial.

Reason for exclusion	Cases	% Total recruited (n = 481)
Failed lymphoscintigraphy	N = 1	0.5
Failed identification of SN at surgery	N = 5	1
Obvious nodal disease at surgery	N = 1	0.5
Breach of protocol (neck dissection despite negative SNB (n = 5), or no neck dissection after positive SNB (n = 20)	N = 25	5
Lost to follow-up	N = 16	3.5
Radiotherapy for close margin or second primary tumour	N = 17	4.5
Total	N = 66	14

SENT, Sentinel European Node Trial; SN, sentinel node; SNB, sentinel node biopsy.

When adjuvant treatment (radiotherapy [RT] or chemoradiotherapy) was given during the follow-up period for close margins or a metachronous primary tumour (N = 17), patients were excluded on the premise that radiation fields extended into the upper neck and could theoretically extinguish missed metastases, thereby erroneously reducing the SNB FNR.

Pre-operative lymphoscintigraphy was performed within 24 hours of surgery after Tc-99m nanocolloid (Nanocoll/Nanocis®) was injected using a standardised technique [15] at four points around the tumour (median dose 57 MBq–interquartile range 60 MBq).

The position of the SNs was marked on the neck. At surgery, the SNs were detected by a hand-held gamma probe and in 164 of 415 patients (39%), peritumoural injection of blue dye was given (SN recorded by colour, radiation count and site in neck). Lymph nodes with radiation count more than three times the background activity were considered SNs. If a radiation hot spot was in more than one neck level (SN versus second or third echelon nodes), then the primary SN was decided by maximum radiation count.

The SNs were fixed in 10% neutral-buffered formalin and a validated protocol for analysis was followed [16]. Five serial sections were cut every 150 μm through the block and one from the centre of each series was stained with haematoxylin and eosin (H&E). If metastasis was still not detected, an adjacent section at each level was stained with anti-pan cytokeratin antibody AE1/3. If cytokeratin was detected but the viability of the cells was in question, the adjacent serial sections were examined stained with H&E. One center (67 in 415: 16% of patients) cut a single frozen section (FS) from the midline of the node, with remaining specimen examined as above. This allowed on-table diagnosis [17] and immediate neck dissection if the FS was positive.

SENT recorded metastasis as viable or non-viable deposits sized in terms of percentage of the total node. For the purposes of this report, SNB+ nodes were retrieved where possible (75 in 94 cases, 80%) and re-graded according to the Union for International Cancer Control (UICC) Seventh Edition guidelines [18]. Deposits were re-classified as isolated tumour cells (ITC, <200 cells or <0.2 mm deposit with no stromal reaction), micrometastasis (0.2–2 mm), and macrometastasis (>2 mm). In the SENT cohort, ITC was treated as a positive neck (completion neck dissection performed within 3 weeks).

Tumours were excised aiming for pathological clear margin of >4 mm, and all defects were closed without free flap reconstruction. Neck specimens were pinned out maintaining alignment and fixed in neutral-buffered formalin. They were examined macroscopically and by routine H&E with cervical metastasis mapped by neck

Table 2
SENT patient demographics and tumour characteristics.

Characteristic	Overall n=	Negative SNB	Positive SNB	Effect of characteristic on sentinel node status
Total patients	415	321 (77%)	94 (23%)	
Male	247 (60%)	194 (60%)	53 (56%)	Gender p = 0.48
Female	168 (40%)	127(40%)	41 (44%)	
Median age (years, range)	61 (28-92)	61(28–92)	62 (29–91)	Age p = 0.74
Primary tumour site				
Oral tongue anterior 2/3	213 (51%)	157/213 (74%)	56/213(26%)	Topographic site p = 0.1
Oral tongue posterior 1/3	43 (10.5%)	32/43(74%)	11/43 (26%)	
Buccal mucosa	17 (4%)	13/17 (76%)	4 /17(24%)	
Floor of mouth	101 (25%)	87/101 (86%)	14/101 (14%)	
Hard palate	3 (1%)	3/3 (100%)	0	
Lower alveolus/gingival	8 (2%)	4/8 (50%)	4/8(50%)	
Lower lip	6 (1.5%)	6 /6 (100%)	0	
Upper alveolus/gingival	5 (1%)	5/5 (100%)	0	
Retromolar	9 (2%)	6/9 (66%)	3/9 (33%)	
Soft palate	10 (2%)	8/10 (80%)	2/10 (20%)	
T stage				
T1	296 (71%)	239 (74%)	57 (61%)	T stage p = 0.032
T2	119 (29%)	82 (26%)	37(39%)	
Neck dissection (for SNB+ or recurrent disease)				
Yes	121(29%)	27(8%)	94 (100%)	N/A
No	294(71%)	294 (92%)	0	
Radiotherapy (n = 36) or chemoradiotherapy (n = 12) (>1 positive node, ECS or recurrence)				
Yes	48 (12%)	23 (7%)	25 (27%)	N/A
No	367 (88%)	298(93%)	69 (73%)	

Statistical testing by chi-square or two-sample t-test depending upon characteristic.
SENT, Sentinel European Node Trial; SN, sentinel node; SNB, sentinel node biopsy; N/A, not applicable; ECS, extracapsular spread.

level. Demographic data, pathological features, location of SN, and survival data were collected for each patient.

Statistical analysis was performed using R survival package [19]. Univariate survival analysis models were built using Kaplan–Meier product-limit estimator for overall survival (OS), disease-specific survival (DSS) and DFS, and multivariate and models with univariate continuous covariates were built using Cox proportional hazards model. Table analysis on 3-year outcomes (such as recurrence within 3 years of SNB) was performed using either chi-square or Fisher' exact to test significance, depending upon the distribution of the variable in question.

3. Results

The patient and carcinoma characteristics are shown in Table 2.

3.1. Lymphatic drainage characteristics

A total of 483 neck sides were examined from 415 patients with 1342 SNs harvested. There were a mean of 2.75 SN per neck or 3.2 SN per patient (range 1–10), with an average size of 11.8 mm (range 3–30 mm). The primary tumour was positioned in the midline in 11.4% (N = 46) and laterally in 88.6% (N = 369) of cases. Lateral tumours drained ipsilaterally in 87% of cases (320 in 369) but in 10% (40 cases) they drained bilaterally and in 2.4% (9 cases) exclusively to the contralateral neck. Sixty percent (28 in 46) of midline lesions drained bilaterally.

3.2. Occult cervical disease

SNB detected metastasis in 94 patients (23%), 16 of whom had extra-capsular spread (17%). Of the 75 cases classified by the UICC guidelines, 12 (16%) contained ITC, 36 (48%) contained micrometastasis and 27 (36%) macrometastasis.

Fifteen patients with a negative SNB subsequently developed isolated cervical metastasis with a negative primary tumour site (one with concomitant distant metastasis) and these were recorded as a false-negative biopsy. Therefore, of 415 patients, 109 had occult metastasis. SNB had a sensitivity, negative predictive value and FNR of 86%, 95%, and 14%, respectively. The FNR, sensitivity and negative predictive value for the three most common tumour sites are shown in Table 3.

In the 49 patients with unexpected bilateral or contralateral drainage from a lateral carcinoma, a positive SN was identified in seven (two bilateral and five solely contralateral).

All 94 patients with a positive SNB underwent neck dissection. In seven cases, dissection was bilateral, giving a total of 101 neck dissections, of which 47% (47 in 101) were selective, and the remainder modified radical.

In 85% of cases, no further positive nodes were found in the completion specimen. Of the patients with additional

Table 3
NPV, sensitivity and FNR by tumour location where a false-negative result is recorded as isolated neck recurrence following a negative sentinel node biopsy.

Tumour	False-negative rate	Sensitivity	NPV
Anterior tongue	14% (9/65)	85%	94%
Posterior tongue	21% (3/14)	79%	91%
Floor of mouth	13% (2/16)	87.5%	98%
Total SENT group	14% (15/109)	86%	95%
Fisher' exact test	p = 1		

NPV, negative predictive value; SENT, Sentinel European Node Trial.

positive non-sentinel nodes, 13 in 15 (87%) were located in the same neck level as the SN or an adjacent neck level.

3.3. Outcome

In this cohort of patients, 3-year figures for OS, DFS and DSS were 88%, 92% and 94%, respectively. Disease recurred in 56 patients (Table 4).

Univariate analysis of the factors that affected outcome (overall survival) was investigated with Kaplan–Meier survival analysis for categorical variables and Cox proportional hazards for continuous variables (such as age of patient) (Table 5, Fig. 1–3). A multivariate Cox proportional hazards model was then run with all variables that showed a univariate p value <0.25. The resulting multivariate Cox proportional hazards model found that the grouped number of positive nodes (p = 0.0008) and SN status (p = 0.003) were the only significant factors.

3.4. Complications

Morbidity of SNB was minimal. Minor complications were seroma [1], haematoma [8], local infection [3], and lymphoedema [1]. There were two notable complications: one phrenic nerve palsy and one patient had a cerebellar stroke secondary to surgery.

Mean hospital stay following SNB and primary tumour resection was 5.7 d (range 0–30) with 161 patients discharged within 3 d of their surgery. Lengths of stay varied considerably by country (average of 9 d in Belgium compared to 3 d in Denmark).

Table 4
Recurrences at 3 years.

	Total (n = 415)	SNB negative (n = 321)	SNB positive (n = 94)
Local (±distant)	18 (4.3%)	13 (4.0%)	5 (5.3%)
Local and neck	9 (2.2%)	8 (2.5%)	1 (1.0%)
Neck (±distant)	29 (7.5%)	15 (4.7%)	14 (15%)
Distant	0	0	0
Outcome following recurrence			
Dead with disease		14 (4.3%)	16 (17%)
Alive no disease		19 (5.9%)	2 (2.1%)
Dead with no disease		2 (0.6%)	2 (2.1%)
Alive with disease		2 (0.6%)	0

SNB, sentinel node biopsy.

Table 5
Univariate Kaplan–Meier (for categorical) or Cox PH (for continuous) analysis of factors influencing overall survival following SNB (significance levels *0.05, **0.01, ***0.001).

Factor	Overall survival (OS) p value
Age of patient	0.003* (Cox PH)
Site of tumour (grouped per location anterior versus posterior oral cavity)	0.755
T size (T1 versus T2)	0.465
Depth of invasion (>4 versus ≤4 mm)	0.142
Degree of differentiation (well versus moderate, poor)	0.004*
Margin (≤1 versus >1 mm)	0.741
Sentinel node status (SNB+ versus SNB−)	0.000083***
Metastasis type (ITC versus Mi, Ma)	0.032*
Total positive nodes (0, 1, 2, >2)	0.000016***
Extra-capsular spread (no versus yes)	0.029*
RT (no, yes)	0.93

PH, proportional hazards; SNB, sentinel node biopsy; ITC, isolated tumour cells; Mi, micrometastasis; Ma, macrometastasis; RT, radiotherapy.

3.5. Adjuvant therapy

Adjuvant therapy (RT or chemoradiotherapy) was given to 12% (48 in 415) of patients. In the SNB-positive group, 27% (25 in 94) received adjuvant therapy (more than one positive node or extracapsular spread – ECS) but was used more freely (80%) in the false-negative SNB group (12 in 15) to help salvage patients. There was no significant survival difference between those with and without adjuvant radiotherapy (p = 0.67).

4. Discussion

The results of the study demonstrate clearly the value and safety of SNB for staging the N0 neck in routine clinical practice. The principal aim of the study was to establish whether SNB is a safe oncological procedure. This has been confirmed with DFS of 92% at 3 years following treatment.

The second objective was to determine, in the context of oral and oropharyngeal cancer, whether SNB was an effective diagnostic test for microscopic deposits of metastatic cancer. The study showed conclusively that the SNB technique works effectively in the oral cavity. The injection of radiotracer (lymphoscintigraphy) will define an SN in the vast majority of patients (>99%).

In this cohort of patients with a 3:1 distribution of T1:T2 oral squamous carcinoma and radiologically N0 neck, it transpired that 26% (109 in 415) had occult cervical disease. The SNB technique failed to detect occult metastasis in 14% (15 in 109) of patients, only half of whom (53.3%: 8 in 15) were amenable to salvage.

This is somewhat counterbalanced through identification of unexpected contralateral lymphatic drainage by SNB. This occurred in 12% (49 in 369) of cases and in seven instances, the contralateral SN was positive. Thus, 6% (7 in 109) of occult cervical metastasis would have

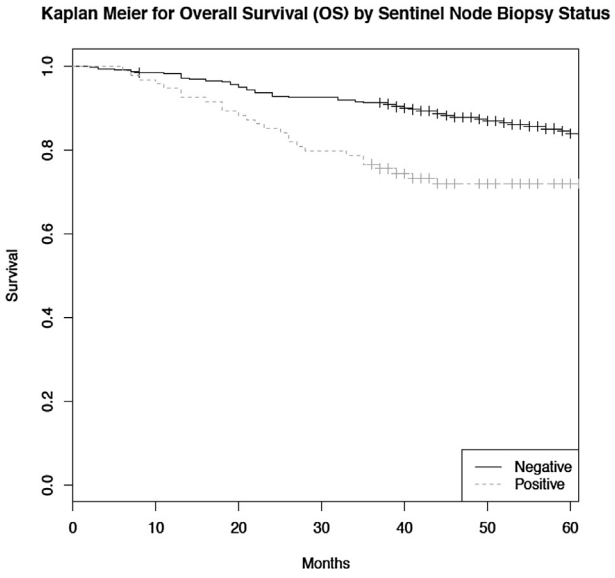


Fig. 1. Overall survival for SNB+ versus SNB− biopsy ($p = 0.00083$). SNB, sentinel node biopsy.

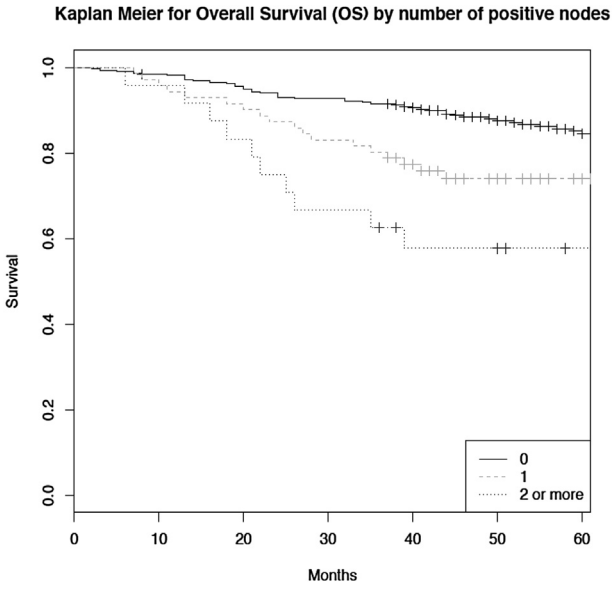


Fig. 2. Overall survival for 0, 1, or ≥2 positive sentinel nodes ($p = 0.000016$).

been missed by conventional treatment of ipsilateral neck dissection.

In head and neck cancer, historically, there has been concern that biopsy of suspected neck metastasis would facilitate dissemination of tumour in the neck. A systematic review [20] of 109 papers calculated regional recurrence rates of 13% in surgically treated early-stage oral cancer. A further review of 164 [21] patients with pT1–T2 tongue SCC staged pN0 after END reported a regional recurrence rate of 18%. The results of SENT when reported in an identical way show the neck recurrence rate for SNB− and SNB+ and the total group were 5%, 15% and 7.5%. The low rate of regional recurrence argues against

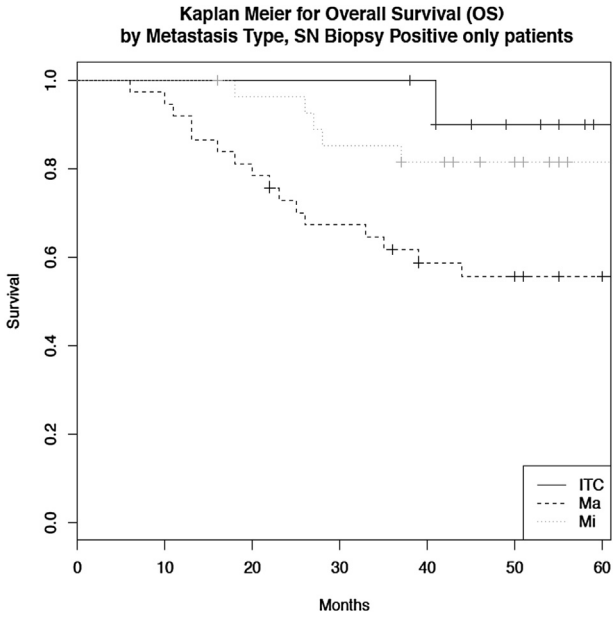


Fig. 3. Overall survival by metastasis type: isolated tumour cells (I) versus macrometastasis (Ma) versus micrometastasis (Mi) ($p = 0.0318$).

SNB causing tumour spillage and in turn neck recurrence.

Two- and 5-year overall survival in early oral and oropharyngeal carcinoma is in the region of 82% and 76%, respectively [20,22]. In this study, overall crude (88%) and DSS (94%) are unlikely to change significantly and suggest strongly that SNB does not adversely affect outcome. An FNR of 14% is similar to that reported in a meta-analysis of 25,000 melanoma patients (12.5%) [23] and 20% FNR in 10-year follow-up of the Multicenter Selective Lymphadenectomy Trial (MSLT) trial in melanoma [24]. However, this is on the borderline of acceptability and we should aim to reduce this to the 7% FNR accepted in breast cancer [25]. Further analysis of the factors associated with a false-negative biopsy is warranted but initial review of our data suggests that operator factors are principally responsible for the FNR. It is well established that there is a learning curve to the SN technique [11].

It is of particular note that previous studies [9,11] indicated that SNB was less reliable for tumours in the floor of mouth presumably due to the close proximity of the injection site to the primary draining nodes. The same association was not found in this study.

The major positive patient benefit of SNB is that in this series 71% of patients were spared neck dissection with consequent improved function and reduced morbidity [26,27]. There were also 47 patients with midline tumours who by convention would have received bilateral neck dissection. In this group, only eight underwent bilateral and eight unilateral dissection based on positive SN. A low complication rate as well as

a reduced in-patient stay supports the economic argument for SNB over END [28].

The identification of aberrant drainage patterns is a huge advantage of SNB and will also have application in patients with second primary tumours where a neck dissection has already been performed and drainage has been disturbed. The disadvantage of blanket ipsilateral END is illustrated in the study of pN0 necks treated by END [21]. In this series, the regional recurrence rate of 18% seems high but it is worth noting that in over one-third of patients (39%) recurrence occurred in the contralateral neck. One further advantage of SNB is that because the tissues have not been significantly disturbed, comprehensive salvage surgery is possible if a recurrence is detected promptly. Our results have also shown that the metastasis type (ITC, micrometasis, macrometastasis) was a prognostic value for overall survival. This confirms recent findings [29,30] and may be important for stratifying personalised treatment in the future.

At the present time, SNB is not widely recognised as standard care in early oral and oropharyngeal cancer. However, increasingly it is gaining utility in Europe and in some countries, such as Denmark, it is integrated into the standard care pathway. Data emerging from this study are relevant to the evolving therapeutic use of SNB technique and provide data to support further investigation by prospective randomised trials. The drive towards patient-specific and minimally invasive surgery is further refining the SN technique and we expect that the use of intraoperative 3D navigation [31], new tracers [32] and fluorescent markers [33] will improve the ease and accuracy of sampling sentinel lymph nodes. SNB potentially offers the solution to the dilemma 'How do you manage the N0 neck?'

Presentations of this work

Part of this work has been presented by members of the Sentinel European Node Trial (SENT) group at meetings of the European Association of Craniomaxillofacial Surgery, The American Head and Neck Society, the European Congress of Pathology, the International Association of Oral and Maxillofacial Surgery, and the British Association of Oral and Maxillofacial Surgeons as well as our annual SENT group meetings, part of the International Symposium on metastasis in head and neck cancer.

Disclaimers

None.

Conflict of interest statement

The authors confirm that there is no conflict of interest in relation to this publication.

References

[1] Chaturvedi AK, Anderson WF, Lortet-Tieulent J, Curado MP, Ferlay J, Franceschi S, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol 2013; 31(36):4550–9.

[2] Shimizu K, Inoue H, Saitoh M, Ohtsuki N, Ishida H, Makino K, et al. Distribution and impact of lymph node metastases in oropharyngeal cancer. Acta Otolaryngol 2006;126(8):872–7.

[3] De Zinis LO, Bolzoni A, Piazza C, Nicolai P. Prevalence and localization of nodal metastases in squamous cell carcinoma of the oral cavity: role and extension of neck dissection. Eur Arch Otorhinolaryngol 2006;263(12):1131–5.

[4] de Bondt RB, Hoeberigs MC, Nelemans PJ, Deserno WM, Peutz-Kootstra C, Kremer B, et al. Diagnostic accuracy and additional value of diffusion-weighted imaging for discrimination of malignant cervical lymph nodes in head and neck squamous cell carcinoma. Neuroradiology 2009;51(3):183–92.

[5] Stoeckli SJ, Haerle SK, Strobel K, Haile SR, Hany TF, Schuknecht B. Initial staging of the neck in head and neck squamous cell carcinoma: a comparison of CT, PET/CT, and ultrasound-guided fine-needle aspiration cytology. Head Neck 2012;34(4):469–76.

[6] Weiss MH, Harrison LB, Isaacs RS. Use of decision analysis in planning a management strategy for the stage N0 neck. Arch Otolaryngol Head Neck Surg 1994;120(7):699–702.

[7] D'Cruz AK, Vaish R, Kapre N, Dandekar M, Gupta S, Hawaldar R, et al. Elective versus therapeutic neck dissection in node-negative oral cancer. N Engl J Med 2015.

[8] Thompson CF, St John MA, Lawson G, Grogan T, Elashoff D, Mendelsohn AH. Diagnostic value of sentinel lymph node biopsy in head and neck cancer: a meta-analysis. Eur Arch Otorhinolaryngol 2013;270(7):2115–22.

[9] Civantos FJ, Zitsch RP, Schuller DE, Agrawal A, Smith RB, Nason R, et al. Sentinel lymph node biopsy accurately stages the regional lymph nodes for T1–T2 oral squamous cell carcinomas: results of a prospective multi-institutional trial. J Clin Oncol 2010; 28(8):1395–400.

[10] Broglie MA, Haile SR, Stoeckli SJ. Long-term experience in sentinel node biopsy for early oral and oropharyngeal squamous cell carcinoma. Ann Surg Oncol 2011;18(10):2732–8.

[11] Alkureishi LW, Ross GL, Shoaib T, Soutar DS, Robertson AG, Thompson R, et al. Sentinel node biopsy in head and neck squamous cell cancer: 5-year follow-up of a European multicenter trial. Ann Surg Oncol 2010;17(9):2459–64.

[12] Jeschke S, Beri A, Grull M, Ziegerhofer J, Prammer P, Leeb K, et al. Laparoscopic radioisotope-guided sentinel lymph node dissection in staging of prostate cancer. Eur Urol 2008;53(1): 126–32.

[13] Altgassen C, Hertel H, Brandstadt A, Kohler C, Durst M, Schneider A. Multicenter validation study of the sentinel lymph node concept in cervical cancer: AGO Study Group. J Clin Oncol 2008;26(18):2943–51.

[14] Lussier YA, Rothwell DJ, Cote RA. The SNOMED model: a knowledge source for the controlled terminology of the computerized patient record. Methods Inf Med 1998;37(2):161–4.

[15] Alkureishi LW, Burak Z, Alvarez JA, Ballinger J, Bilde A, Britten AJ, et al. Joint practice guidelines for radionuclide lymphoscintigraphy for sentinel node localization in oral/oropharyngeal squamous cell carcinoma. Ann Surg Oncol 2009;16(11): 3190–210.

[16] Ross GL, Soutar DS, Gordon MacDonald D, Shoaib T, Camilleri I, Roberton AG, et al. Sentinel node biopsy in head and neck cancer: preliminary results of a multicenter trial. Ann Surg Oncol 2004;11(7):690–6.

[17] Tschopp L, Nuyens M, Stauffer E, Krause T, Zbaren P. The value of frozen section analysis of the sentinel lymph node in clinically

N0 squamous cell carcinoma of the oral cavity and oropharynx. Otolaryngol Head Neck Surg 2005;132(1):99–102.
[18] Sobin Leslie H, G MK, Wittekind Christian, editors. TNM classification of malignant tumours. 7th ed. Wiley-Blackwell; 2009.
[19] T T. A package for survival analysis in R. 2015. http://CRAN.R-project.org/package=survival.
[20] Brown JS, Shaw RJ, Bekiroglu F, Rogers SN. Systematic review of the current evidence in the use of postoperative radiotherapy for oral squamous cell carcinoma. Br J Oral Maxillofac Surg 2012;50(6):481–9.
[21] Ganly I, Goldstein D, Carlson DL, Patel SG, O'Sullivan B, Lee N, et al. Long-term regional control and survival in patients with "low-risk," early stage oral tongue cancer managed by partial glossectomy and neck dissection without postoperative radiation: the importance of tumor thickness. Cancer 2013;119(6):1168–76.
[22] Woolgar JA, Rogers S, West CR, Errington RD, Brown JS, Vaughan ED. Survival and patterns of recurrence in 200 oral cancer patients treated by radical surgery and neck dissection. Oral Oncol 1999;35(3):257–65.
[23] Valsecchi ME, Silbermins D, de Rosa N, Wong SL, Lyman GH. Lymphatic mapping and sentinel lymph node biopsy in patients with melanoma: a meta-analysis. J Clin Oncol 2011;29(11):1479–87.
[24] Morton DL, Thompson JF, Cochran AJ, Mozzillo N, Nieweg OE, Roses DF, et al. Final trial report of sentinel-node biopsy versus nodal observation in melanoma. N Engl J Med 2014;370(7):599–609.
[25] Pesek S, Ashikaga T, Krag LE, Krag D. The false-negative rate of sentinel node biopsy in patients with breast cancer: a meta-analysis. World J Surg 2012;36(9):2239–51.
[26] Schiefke F, Akdemir M, Weber A, Akdemir D, Singer S, Frerich B. Function, postoperative morbidity, and quality of life after cervical sentinel node biopsy and after selective neck dissection. Head Neck 2009;31(4):503–12.
[27] Murer K, Huber GF, Haile SR, Stoeckli SJ. Comparison of morbidity between sentinel node biopsy and elective neck dissection for treatment of the n0 neck in patients with oral squamous cell carcinoma. Head Neck 2011;33(9):1260–4.
[28] O'Connor R, Pezier T, Schilling C, McGurk M. The relative cost of sentinel lymph node biopsy in early oral cancer. J Craniomaxillofac Surg 2013;41(8):721–7.
[29] Broglie MA, Haerle SK, Huber GF, Haile SR, Stoeckli SJ. Occult metastases detected by sentinel node biopsy in patients with early oral and oropharyngeal squamous cell carcinomas: impact on survival. Head Neck 2013;35(5):660–6.
[30] Den Toom IJ, Heuveling DA, Flach GB, van Weert S, Karagozoglu KH, van Schie A, et al. Sentinel node biopsy for early-stage oral cavity cancer: the VU University Medical Center experience. Head Neck 2014.
[31] Heuveling DA, van Weert S, Karagozoglu KH, de Bree R. Evaluation of the use of freehand SPECT for sentinel node biopsy in early stage oral carcinoma. Oral Oncol 2015;51(3):287–90.
[32] Agrawal A, Civantos FJ, Brumund KT, Chepeha DB, Hall NC, Carroll WR, et al. [Tc]Tilmanocept accurately detects sentinel lymph nodes and predicts node pathology status in patients with oral squamous cell carcinoma of the head and neck: results of a phase iii multi-institutional trial. Ann Surg Oncol 2015.
[33] Brouwer OR, Klop WM, Buckle T, Vermeeren L, van den Brekel MW, Balm AJ, et al. Feasibility of sentinel node biopsy in head and neck melanoma using a hybrid radioactive and fluorescent tracer. Ann Surg Oncol 2012;19(6):1988–94.

The Laryngoscope

Factors That Impact Health-Related Quality of Life Over Time for Individuals With Head and Neck Cancer

Bryce B. Reeve, PhD; Jianwen Cai, PhD; Hongtao Zhang, PhD; Mark C. Weissler, MD; Kathy Wisniewski, BS; Heather Gross, MEd; Andrew F. Olshan, PhD

Objectives/Hypothesis: To identify sociodemographic, behavioral, and clinical factors associated with health-related quality of life (HRQOL) for head and neck cancer (HNC) patients over time.

Study Design: A population-based longitudinal cohort study.

Methods: Newly diagnosed HNC patients (N = 587) were administered the Functional Assessment of Cancer Therapy–Head and Neck questionnaire at baseline (median 3 months postdiagnosis) and two follow-up assessments (median 22 and 42 months). Linear mixed-effect models were used with backward variable selection to identify factors associated with HRQOL over time ($P < .05$). Adjusted means reported at 2 years postdiagnosis.

Results: African Americans reported better Functional Well-Being than whites (mean of 20.01 vs. 18.53) and fewer HNC symptoms over time. Older patients (75+ years) reported better HRQOL than younger patients (< 50 years). Current tobacco use compared to no tobacco use had worse Physical (20.20 vs. 21.50), Emotional (17.55 vs. 19.06), Social (21.28 vs. 22.88), and Functional (17.32 vs. 19.29) Well-Being and more HNC symptoms (21.50 vs. 23.71). Radiation therapy was associated with worse Physical and Functional Well-Being and more head and neck symptoms over time, but HRQOL was similar to those who were not irradiated by 2 to 4 years postdiagnosis.

Conclusion: This study identified key factors for individuals at risk for poorer HRQOL that may help clinicians and caregivers find solutions to address these decrements. Smoking cessation programs can be encouraged for survivors who use tobacco. Psychological and social support and medications may help for dealing with emotional distress and dealing with the physical symptoms from treatment.

Key Words: Health-related quality of life, symptoms, head and neck cancer, radiation therapy, longitudinal study.

Level of Evidence: 4.

***Laryngoscope*, 00:000–000, 2016**

INTRODUCTION

Overall, the incidence of head and neck cancer (HNC) (oral cavity, pharynx, larynx) began decreasing in about 1991 but has stabilized since about 2003.[1] In 2015, an estimated 59,340 individuals in the United States were diagnosed, and approximately 12,290 died from a HNC.[2] Incidence rates are twice as high in men as women. Incidence rates in African Americans have also declined over the past two decades and are currently lower than in non-Hispanic whites. The standard for treating HNC has been surgery; however, since the early 1990s there has been an increase in the use of primary radiation and chemotherapy.[3,4] This has been partially driven by the increase in the number of human papilloma virus (HPV)-associated oropharyngeal cancers.

Both the cancer and its treatment can have profound effects on an individual's health-related quality of life (HRQOL). Physically, they may experience facial disfigurement and problems with eating, breathing, and speaking.[5] Mentally, they may experience negative body image, depression, anxiety, and fatigue.[5–8] Socially, they may experience impaired communication, disrupted social relationships, social isolation, stigmatism, and work impairment.[5,9] Although some symptoms improve over time after treatment, some HRQOL decrements persist, including functional limitations and psychosocial impact.[5,9]

Several factors have been found to be associated with HRQOL in HNCs. Behavioral factors include use of tobacco[10–15] and alcohol.[16–20] Sociodemographic variables include age,[9,13,21,22] race,[23] education,[13,18] and employment status.[9] Clinical factors include disease stage and the use of radiation therapy.[9,24–31] However, much of the literature is either more than 10 years old, based on small sample sizes, or shows no or conflicting relationships with HRQOL.

This study builds on a previous published study of the HRQOL of individuals receiving care for HNC in

From the Department of Health Policy and Management (B.B.R.); the Department of Biostatistics (J.C., H.Z.); the Department of Epidemiology (K.W., A.F.O.), Gillings School of Global Public Health; the Lineberger Comprehensive Cancer Center (B.B.R., A.F.O.); the Sheps Health Services Research Center (B.B.R., H.G.); and the Department of Otolaryngology–Head & Neck Surgery (M.C.W., A.F.O.), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, U.S.A.

Editor's Note: This Manuscript was accepted for publication April 6, 2016.

This study was funded by a grant from the Lance Armstrong Foundation and in part by the National Cancer Institute (R01-CA90731). The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Bryce B. Reeve, PhD, University of North Carolina at Chapel Hill, 1101-D McGavran-Greenberg Building, 135 Dauer Drive, CB 7411, Chapel Hill, NC 27599-7411.
E-mail: bbreeve@email.UNC.edu

DOI: 10.1002/lary.26073

North Carolina.[32] This previous cross-sectional study found that at approximately 3 months postdiagnosis, African Americans reported higher physical well-being than non-Hispanic whites. The current study includes the median 3-month survey (baseline) and two additional assessment points at approximately 2 and 3 years postdiagnosis. The strengths of this study relative to other published studies are the long-term HRQOL follow-up assessments, the large population-based study group, and inclusion of a large African American cohort. The overall goal of this study is to identify the sociodemographic, behavioral, and clinical characteristics associated with HRQOL over time.

MATERIALS AND METHODS

Participants and Study Design

A detailed description of the parent study, the Carolina Head and Neck Cancer Study (CHANCE), can be found in Divaris et al.[33] Briefly, the CHANCE study—a population-based case-control study of risk factors for HNC—included patients aged 20 to 80 years who were residents of a 46-county region in North Carolina, and who had a newly diagnosed, first primary invasive squamous cell carcinoma between January 1, 2002, and February 28, 2006. Patients were excluded if they had cancer of the lip, salivary glands, nasopharynx, nasal cavity, and nasal sinuses. In addition, individuals with carcinomas of other histologies, carcinomas at other head and neck sites, or a history of recurrent or second primary tumors were also not eligible. Cases were identified by means of a rapid case ascertainment process in conjunction with the North Carolina Central Cancer Registry in which newly diagnosed cancer cases were identified and reported to the study office every month. This case information was sent to the study office, and initial eligibility was confirmed. With physician's permission, patients were approached and invited to participate. After consent at a baseline in-person interview, information was obtained on demographics, risk factors, medical history, diet, and other factors. At the baseline interview, permission to obtain medical records was requested. The medical records were collected and abstracted in order to obtain information on each patient's first course of treatment and pertinent comorbid conditions. All procedures were performed in accordance with the ethical standards of the University of North Carolina at Chapel Hill (Chapel Hill, NC) Institutional Review Board and in accordance with the Helsinki Declaration of 1975, as revised in 1983.

A subcohort of CHANCE cases was targeted for a multiphase prospective assessment of HRQOL. These phases included data collection at up to three time points postdiagnosis. The first phase of HRQOL data collection (baseline) was conducted in 2005 as part of the overall study baseline in-person interview that collected HRQOL data on cases with a diagnosis of HNC between 2002 and 2005. The median time between diagnosis and the baseline HRQOL data collection was 3 months. Phase 2 (follow-up 1 [FU1]) was conducted using a mailed questionnaire between 2005 and 2008. The median time between diagnosis and the FU1 data collection was 22 months. The third HRQOL data collection phase (follow-up 2 [FU2]) using a mailed questionnaire was conducted between 2007 and 2009. The median time between diagnosis and the FU2 data collection was 42 months. A total of 587 patients were included in this longitudinal analysis, of which 133 patients had HRQOL data from all three phases (baseline, FU1, FU2); 37 had baseline and FU1 data; 236 had FU1 and FU2 data; 52 had only baseline data; 127 had only FU1 data; and two had only FU2 data. The response rate was 86% for baseline, 91% for FU1, and 95% for FU2.

Measures

The baseline in-person interview collected data on individuals' sociodemographic characteristics, including age, sex, race, education, marital status, and insurance status. Participants reported their health behaviors, including tobacco and alcohol use. Individuals also reported their height and weight (for calculating body mass index) and comorbid conditions, including anemia, liver disease, hepatitis, cardiovascular disease, pulmonary disease, renal disease, diabetes, and cancers other than head and neck. The comorbid conditions variable was categorized as none versus one or more chronic conditions. Patients also self-reported the type of treatment that they received (e.g., surgery, radiation, chemotherapy) and if they were currently using a feeding tube.

Health-related quality of life was measured by the Functional Assessment of Cancer Therapy–Head and Neck (FACT-H&N) questionnaire.[34,35] The FACT-H&N includes the FACT-General scales of Physical Well-Being (7 questions; score range 0–28; minimally important difference, [MID] = 2–3 points), Social Well-Being (7 questions; range 0–28; MID not available), Emotional Well-Being (6 questions; range 0–24; MID = 2 points), and Functional Well-Being (7 questions; range 0–28; MID = 2–3 points).[36,37] The FACT-H&N includes 12 additional questions on symptoms and issues specific to HNCs, including ability to eat; dry mouth; difficulty breathing; difficulty swallowing; voice quality; ability to communicate; body image; pain in the mouth, throat, or neck; and alcohol and tobacco use. Following the FACT-H&N scoring guidelines, we summed across nine of the 12 questions in the FACT-H&N (excluding alcohol and tobacco use and pain) to create the Head and Neck Cancer Subscale score (score range 0–36; MID = 3–4 points). All FACT-H&N questions have five response options: "not at all," "a little bit," "somewhat," "quite a bit," and "very much." Higher summed scale scores represent better HRQOL.[37] Minimally important differences are defined as the smallest difference in scores between groups that patients perceive as important or personally meaningful, either beneficial or harmful.[37,38] The MIDs were determined in a separate study conducted by the FACT developers using anchor- and distribution-based methods.[37]

Statistical Analyses

Demographic and clinical characteristics were summarized using proportions. Linear mixed effect models were used on each of the FACT scale scores to account for the longitudinal nature of the data. Time since diagnosis (in months) was used to track time. Each participant may have up to three measurements of HRQOL scores, at baseline and follow-up visits 1 and 2. The correlation between any two HRQOL measurements of a participant was specified as $\exp\{-d_{ij}/\theta\}$, where d_{ij} was the time between the two measurements. This specification reflected stronger correlation between two HRQOL scores when they were measured closer in time and weaker correlation when they were measured farther apart in time.

To identify factors associated with the FACT scales, backward variable selection procedures were used with stay-in P value set to 0.05. The initial model included time of measurement (linear and quadratic) and all the main effects of the demographic and clinical characteristics listed in Table I, plus their interactions with time (linear and quadratic), to allow for the effect of these covariates to vary over time. Among the covariates, feeding tube usage and tobacco/alcohol usage were time-dependent. Once

TABLE I.
Distribution of Demographics and Clinical Characteristics of Individuals With Head and Neck Cancer.*

Demographic and Clinical Characteristics	Total (N = 587)
Race	
White	79.90%
African-American	20.10%
Age at Diagnosis (years)	
< 50	19.08%
50–64	49.23%
65–74	24.53%
75+	7.16%
Income	
$0 to < $20,000	29.30%
$20,000+	66.44%
Number Supported in Household	
1 to 3 individuals	85.35%
4 or more individuals	14.14%
Sex	
Male	75.81%
Female	24.19%
Education	
High school or less	56.39%
Some college	25.72%
College degree or higher	17.89%
Marital Status	
Living with spouse or partner	64.57%
Living alone	35.43%
Health Insurance Coverage	
None	10.39%
Private	43.78%
Government	32.71%
Multiple	13.12%
Tobacco Use (baseline, n = 222)	
Not at all	74.32%
A little bit	10.36%
Somewhat to very much	15.32%
Tobacco Use (follow-up 1, n = 528)	
Not at all	74.43%
A little bit	8.90%
Somewhat to very much	16.67%
Tobacco Use (follow-up 2, n = 370)	
Not at all	76.76%
A little bit	7.30%
Somewhat to very much	15.95%
Alcohol Use (baseline, n = 222)	
Not at all	75.23%
A little bit	11.26%
Somewhat to very much	13.51%
Alcohol Use (follow-up 1, n = 528)	
Not at all	57.20%
A little bit	25.76%
Somewhat to very much	17.05%
Alcohol Use (follow-up 2, n = 370)	
Not at all	52.43%
A little bit	26.49%
Somewhat to very much	21.08%
Body Mass Index	
Underweight/normal	36.97%
Overweight	34.92%
Obese	28.11%
Comorbid Conditions (sum)	
0	44.63%
1+	55.37%
Cancer Type	
Oral	52.64%
Laryngeal	37.65%
Pharyngeal	9.71%
Stage of Cancer	
I	26.24%
II	17.04%
III	17.38%
IV	39.35%
Lymph Nodes	
No	70.53%
Yes	28.79%
Feeding Tube	
Yes (baseline, n = 222)	24.32%
Yes (follow-up 1, n = 521)	9.02%
Yes (follow-up 2, n = 369)	4.88%
Received Radiation Treatment	
No	22.83%
Yes	76.66%
Received Surgery	
No	42.59%
Yes	57.24%
Received Chemotherapy	
No	57.92%
Yes	41.23%

*At diagnosis and follow-up visits where indicated.

the final model was determined, the adjusted means for each categorical variable in the final model were obtained. The procedure was repeated for each of the five subscale scores. The analysis was performed in SAS Version 9.4 (SAS Institute Inc., Cary, NC).

RESULTS

The demographic and clinical characteristics of the 587 participants in this study are summarized in Table I. The baseline survey included 222 participants; the first follow-up survey included 528 participants; and the second follow-up survey included 370 participants. The mean age at diagnosis was 59.35 years (standard deviation = 10.16). Approximately 20% of participants were African American, 76% were male, and 56% had a

high school degree or less. Approximately 63% were categorized as overweight or obese, and 55% had one or more comorbid conditions. Head and neck cancers included 53% oral, 38% laryngeal, and 10% pharyngeal cancer, with 17% in stage III and 39% in stage IV. Approximately 57% received surgery, 77% radiation, and 41% chemotherapy. Given that only 10% of the cancers were pharyngeal, HPV status is not likely to play a major role in this group of patients.

Table II includes a summary of the adjusted mean FACT HRQOL scores for those factors that were statistically significantly related to the FACT measure for each HRQOL domain at the significance level of 0.05.

Physical Well-Being

Statistically significant factors associated with better physical well-being included older age, higher education, private insurance, no current tobacco use, somewhat or very much current alcohol use, no comorbidities, early stage cancer, currently without a feeding tube, and received surgery. The interaction between receiving radiation therapy and time was statistically significant. Although radiation therapy had a strong effect on HRQOL early after completion of therapy, QOL returned to that of patients not receiving radiation therapy by 4 years after treatment.

Emotional Well-Being

Statistically significant factors associated with better emotional well-being included being of African-American race, older age, higher education, not currently using tobacco, no comorbid conditions, oral cancer (compared to laryngeal cancer), and early stage of cancer. Both the interaction between feeding tube and time and between lymph node status and time were statistically significant (quadratic). Figure 1 shows that the emotional well-being for those who continued to need a feeding tube decreased over time, whereas those who never needed a feeding tube gradually increased. For lymph node status, emotional well-being gradually decreased over time, whereas those without lymph node involvement had slightly better well-being. This may be because of the late fibrosis seen in the necks of patients, with advanced nodal disease treated with multimodality therapy.

Social Well-Being

Statistically significant factors associated with better social well-being included being female, living with a spouse or significant other, not currently using tobacco, and received surgery.

Functional Well-Being

Statistically significant factors associated with better functional well-being include African American race, older age, higher education level, private insurance coverage, not currently using tobacco, no comorbid condition, oral (compared to laryngeal) cancer, early stage of cancer, and not currently having a feeding tube. The interaction between receiving radiation and time was statistically significant. Those receiving radiation therapy had poorer functional status early in treatment, and then over time returned to similar functional levels as those who did not receive radiation treatment approximately 2 years after diagnosis.

Head and Neck Cancer Symptoms

Statistically significant factors associated with reduced HNC symptoms include older age, higher education, having private insurance options, not currently using tobacco, somewhat or very much current alcohol use, no comorbid conditions, early stage cancer, and no feeding tube currently. There were significant time interactions with race (quadratic), income (quadratic), number supported in household (linear), and received radiation therapy (quadratic). Symptoms for African Americans improved more quickly over time than non-Hispanic whites, who only reported gradual improvement in symptoms. Figure 2 shows that those who received radiation therapy experienced more symptoms during and posttreatment with improvement over time compared to those who did not undergo radiation therapy. The interactions of time with income and time with number supported in the household were not clinically meaningful in terms of their differences (results not shown).

DISCUSSION

Having a HNC negatively impacts an individual's health-related quality of life. This study followed HNC survivors over 5 years and found a number of demographic, behavioral, and clinical factors that were associated with different levels of HRQOL. Many of the HRQOL initial negative impacts, however, improve over time.

There were several demographic factors that were consistently associated with HRQOL. Older patients reported better physical, emotional, and functional well-being, and fewer symptoms than younger HNC patients, even after adjusting for comorbidity status and treatment (when a significant factor). This conflicts with Hammerlid et al.[21] and Penedo et al.'s [9] finding that younger individuals had better posttreatment HRQOL, and the Ronis et al.[15] study that found no relationship between age and HRQOL. Our population-based study findings, with a larger number of participants, could reflect that younger patients often receive more intense multimodality therapy[39] or are less accepting of the cosmetic and functional repercussions of therapy. Higher educational status was also associated with a better long-term HRQOL, which could reflect that education (as an indicator of socioeconomic status) is associated with better access to care and support networks. Studies by Fang et al.[13] and Kugaya et al.[18] also supported the positive association between education level and HRQOL; however, Ronis et al.[15] found no relationship in a sample from Michigan. Private insurance was associated with better physical and functional well-being and less symptoms compared with no or government insurance. Lack of insurance is associated with greater

TABLE II.
Adjusted Mean FACT HRQOL Scores by Significant Predictors.*

Demographic or Clinical Characteristic	Physical Well-being Mean (SE)	Emotional Well-Being Mean (SE)	Social Well-Being Mean (SE)	Functional Well-Being Mean (SE)	H&N Cancer Mean (SE)
Race					
African American		19.95 (0.39)		20.01 (0.57)	Quad time
White		18.43 (0.18)		18.53 (0.26)	Quad time
Income					
$0 to < $20,000					Quad time
$20,000 +					Quad time
Race x Income			Not a significant predictor		
No. Supported in House					
1 to 3 individuals					Linear time
4 or more individuals					Linear time
Age (years)					
< 50	20.25 (0.48)	17.98 (0.38)		18.05 (0.58)	23.12 (0.63)
50–64	20.37 (0.30)	18.50 (0.24)		17.87 (0.37)	22.22 (0.39)
65–74	22.93 (0.51)	19.47 (0.34)		20.65 (0.62)	24.36 (0.67)
75+	23.79 (0.80)	19.79 (0.63)		21.26 (0.97)	27.03(1.05)
Gender					
Female			23.23 (0.38)		
Male			22.33 (0.22)		
Education Level					
High school or less	20.65 (0.27)	18.27 (0.23)		18.20 (0.33)	22.36 (0.35)
Some college	21.30 (0.38)	18.92 (0.32)		19.03 (0.46)	23.19 (0.49)
College degree+	22.77 (0.47)	19.82 (0.39)		20.37 (0.56)	25.91 (0.61)
Marital Status					
Living alone			21.70 (0.32)		
Living with spouse			23.02 (0.23)		
Health Insurance					
None	20.29 (0.67)			16.82 (0.80)	21.59 (0.85)
Private	22.52 (0.35)			20.83 (0.44)	24.78 (0.49)
Government	19.87 (0.40)			17.10 (0.49)	21.66 (0.56)
Multiple	20.70 (0.63)			17.65 (0.77)	22.88 (0.82)
Tobacco Use					
Not at all	21.50 (0.22)	19.06 (0.19)	22.88 (0.21)	19.29 (0.26)	23.71 (0.28)
A little bit	20.69 (0.54)	18.12 (0.45)	22.29 (0.51)	17.71 (0.61)	22.51 (0.65)
Somewhat/very much	20.20 (0.44)	17.55 (0.37)	21.28 (0.42)	17.32 (0.51)	21.50 (0.54)
Alcohol Use					
Not at all	20.76 (0.25)				22.67 (0.32)
A little bit	21.74 (0.36)				23.87 (0.44)
Somewhat/very much	21.95 (0.43)				24.08 (0.52)
Body Mass Index			Not a significant predictor		
Comorbid Condition					
0	21.71 (0.30)	19.25 (0.25)		19.41 (0.36)	23.86 (0.38)
1+	20.81 (0.27)	18.31 (0.23)		18.35 (0.32)	22.72 (0.35)
Cancer Type					
Laryngeal		18.02 (0.28)		17.64 (0.40)	
Oral		19.19 (0.23)		19.69 (0.34)	
Pharyngeal		19.04 (0.56)		18.74 (0.80)	
Stage of Cancer					
I, II	21.80 (0.31)	19.22 (0.26)		19.65 (0.39)	25.03 (0.40)
III, IV	20.76 (0.28)	18.35 (0.23)		18.18 (0.34)	21.84 (0.35)

TABLE II.
(*Continued*)

Demographic or Clinical Characteristic	Physical Well-being Mean (SE)	Emotional Well-Being Mean (SE)	Social Well-Being Mean (SE)	Functional Well-Being Mean (SE)	H&N Cancer Mean (SE)
Time (Diagnosis to Survey)	Not a significant predictor				
Lymph Nodes					
No		Quad time			
Yes		Quad time			
Feeding Tube					
No	21.47 (0.20)	Quad time		19.16 (0.24)	23.77 (0.26)
Yes	18.52 (0.51)	Quad time		15.37 (0.57)	17.79 (0.60)
Received Radiation					
No	Quad time			Linear time	Quad time
Yes	Quad time			Linear time	Quad time
Received Surgery					
No	20.54 (0.33)		21.97 (0.29)		
Yes	21.69 (0.27)		22.99 (0.25)		
Received Chemotherapy	Not a significant predictor				

*Adjusted mean scores are only provided for variables significantly related to the outcome measure ($P < .05$). Final model was obtained based on backward variable selection procedures. Adjusted means were calculated at 2 years postdiagnosis, with the covariates taking the values at the proportions presented in Table I. For the time-dependent feeding tube usage and tobacco/alcohol usage, the proportions were held at the follow-up 1 visit values. Linear or quad time indicates that the interaction with time is significant in linear or quadratic form, respectively. Higher scores on all FACT scales represent better health-related quality of life.

FACT = Functional Assessment of Cancer Therapy; H&N = head and neck; HRQOL = Health-Related Quality of Life; quad = quadratic form; SE = Standard Error.

financial burden from cancer care and has been found to be associated with poorer HRQOL.[40] African Americans reported better emotional and functional well-being than non-Hispanic whites. For HNC symptoms, African Americans had similar symptom levels during the treatment period, but symptoms improved more dramatically over time compared to whites. These findings are consistent with our initial study of the 3-month postdiagnosis data.[32] Although not all outcomes can be explained, better mental well-being may be the result of better coping strategies, including close-knit friends and family and more spirituality.[23,41,42]

Both current tobacco and alcohol use are risk factors for HNC, and both appear to be factors associated with HRQOL for survivors but have different relationships. Current tobacco use was associated with decrements in all measures of HRQOL and increased symptom experiences. This finding is consistent with the literature,[11–15] including a study by Duffy of 81 HNC patients in which smoking was negatively associated with Health Survey Short Form-36 measures of Physical Functioning, General Health, Vitality, Social Functioning, and Role-Emotional health.[10] Together, our studies reinforce the need for smoking cessation services for HNC patients who continue to smoke after diagnosis. Current alcohol use was statistically associated with better physical well-being and HNC symptoms; however, the differences in means between drinkers and nondrinkers did not exceed minimally important differences threshold. Several studies [10,12,13,15,17] found no association between alcohol use and HRQOL, whereas some studies found alcohol abusers had poorer HRQOL.[17,18] Allison et al.[16] did find an association between alcohol use and better physical and role functioning; better global HRQOL; and fewer symptoms of fatigue, pain, problems swallowing, dry mouth, and feelings of illness. Without an in-depth study assessing frequency and quantity of alcohol drinking, it is difficult to speculate if moderate alcohol drinking is promoting better HRQOL or resulting from it.[16]

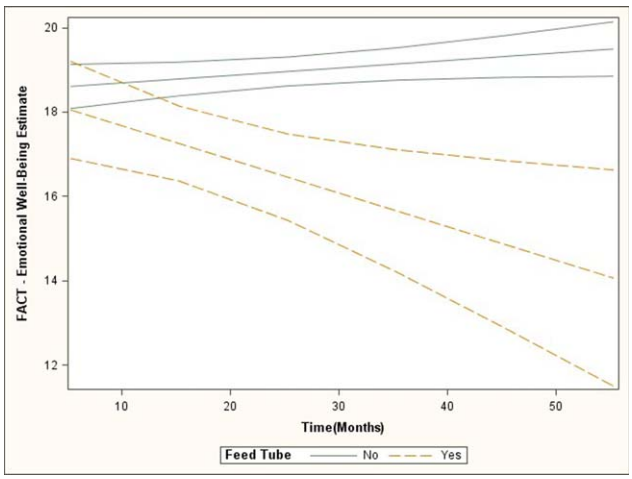


Fig. 1. FACT-G Emotional Well-Being scores over time for those who needed a feeding tube and those who never needed a feeding tube. The adjusted means were calculated with the covariates taking the values at the proportions presented in Table I. For the time-dependent feeding tube usage and tobacco/alcohol usage, the proportions were held at the follow-up 1 visit values. The outside lines are 95% confidence interval lines.
FACT-G = Functional Assessment of Cancer Therapy-General.

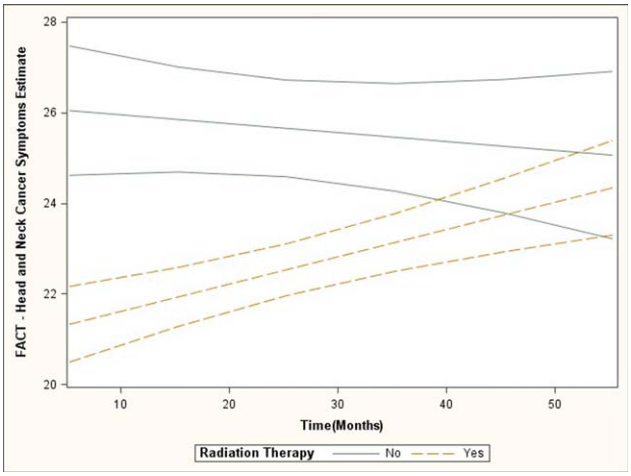


Fig. 2. FACT-G Head and Neck Cancer Symptoms scores over time for those who received radiation therapy and those who did not receive radiation therapy. The adjusted means were calculated with the covariates taking the values at the proportions presented in Table I. The outside lines are 95% confidence interval lines. FACT-G = Functional Assessment of Cancer Therapy-General.

There were several notable clinical factors associated with HRQOL. Consistently, individuals with comorbid conditions, higher stages of cancer, and continued need for a feeding tube reported poorer HRQOL on all domains except social well-being. This result is consistent with a large study by Terrell et al.,[26] who found that feeding tube status and comorbid conditions were strong predictors of HRQOL, and that stage of cancer was a moderate predictor of HRQOL. The acute effects of radiation therapy had a negative impact on physical and functional well-being and increased HNC-associated symptoms. The literature supports this finding: irradiation has been associated with several side effects, including dry mouth, difficulty swallowing, mouth and gum sores, fatigue, nausea, and lymphedema.[9,28–31,43] Our study found that the detrimental effects of radiation therapy lessened over the years such that Functional HRQOL was similar to those who were not irradiated by 2 years, and physical well-being was similar between those who did and didn't receive radiation therapy by four or five years postdiagnosis.

Limitations

There were several limitations of this study. All participants were receiving care in North Carolina. Although the sample was heterogeneous with respect to race, socioeconomic status, and clinical factors, it may not be generalizable to the U.S. population. The measures of HRQOL were added to the study after the parent study had begun, thus resulting in lower sample sizes at the baseline 3-month (median) assessment point. We did not have a pretreatment measure of HRQOL to know to what extent HRQOL differences existed prior to treatment. Future studies are recommended to better understand the factors associated with HRQOL in HNC patients. Strengths include the population-based relatively large sample size and diverse population with multiple follow-up surveys. Human papillomavirus infection has been shown to be a risk factor for HNC, especially cancer of the base of tongue, tonsils, and oropharynx.[44] We did not have HPV data on all cases included in this analysis, and because its potential association with HRQOL is mediated only through treatment, we did not consider HPV as a covariate.

CONCLUSION

Despite these limitations, this study contributes to the literature by identifying sociodemographic, behavioral, and clinical factors associated with poorer HRQOL for HNC survivors. Understanding these factors will help to identify those at risk for decrements in HRQOL. Some factors, such as tobacco use, can be modifiable by encouraging patients to participate in smoking cessation program. For other factors, awareness of these factors can help to identify those at risk and provide more long-term care and surveillance to reduce the deleterious effects of the cancer and its treatment. The acute ill effects of radiation therapy on HRQOL seem to decrease significantly over time and return to levels seen in non-irradiated patients by 2 to 4 years posttherapy. Enhanced psychological and social support may help with issues of depression and dealing with the physical effects of surgery and radiation.[45]

BIBLIOGRAPHY

1. Brown LM, Gridley G, Devesa SS. Descriptive epidemiology: U.S. patterns. In: Olshan AF, ed. Epidemiology, Pathogenesis, and Prevention of Head and Neck Cancer. New York, NY: Springer; 2010:23–39.
2. American Cancer Society. Cancer Facts & Figures 2015. Atlanta, GA: American Cancer Society; 2015.
3. Beadle BM, Liao KP, Elting LS, et al. Improved survival using intensity-modulated radiation therapy in head and neck cancers: a SEER-Medicare analysis. *Cancer* 2014;120:702–710.
4. Chera BS, Eisbruch A, Murphy BA, et al. Recommended patient-reported core set of symptoms to measure in head and neck cancer treatment trials. *J Natl Cancer Inst* 2014;106. pii: dju127. doi: 10.1093/jnci/dju127.
5. List MA, Ritter-Sterr CA, Baker TM, et al. Longitudinal assessment of quality of life in laryngeal cancer patients. *Head Neck* 1996;18:1–10.
6. Mochizuki Y, Matsushima E, Omura K. Perioperative assessment of psychological state and quality of life of head and neck cancer patients undergoing surgery. *Int J Oral Maxillofac Surg* 2009;38:151–159.
7. Neilson KA, Pollard AC, Boonzaier AM, et al. Psychological distress (depression and anxiety) in people with head and neck cancers. *Med J Aust* 2010;193(5 suppl):S48–S51.
8. Sawada NO, de Paula JM, Sonobe HM, Zago MM, Guerrero GP, Nicolussi AC. Depression, fatigue, and health-related quality of life in head and neck cancer patients: a prospective pilot study. *Support Care Cancer* 2012;20:2705–2711.
9. Penedo FJ, Traeger L, Benedict C, et al. Perceived social support as a predictor of disease-specific quality of life in head-and-neck cancer patients. *J Support Oncol* 2012;10:119–123.
10. Duffy SA, Terrell JE, Valenstein M, Ronis DL, Copeland LA, Connors M. Effect of smoking, alcohol, and depression on the quality of life of head and neck cancer patients. *Gen Hosp Psychiatry* 2002;24:140–147.
11. Bloom EL, Oliver JA, Sutton SK, Brandon TH, Jacobsen PB, Simmons VN. Post-operative smoking status in lung and head and neck cancer patients: association with depressive symptomatology, pain, and fatigue. *Psychooncology* 2015;24:1012–1019.
12. Duffy SA, Ronis DL, Valenstein M, et al. Depressive symptoms, smoking, drinking, and quality of life among head and neck cancer patients. *Psychosomatics* 2007;48:142–148.
13. Fang FM, Chiu HC, Kuo WR, et al. Health–related quality of life for nasopharyngeal carcinoma patients with cancer–free survival after treatment. *Int J Radiat Oncol Biol Phys* 2002;53:959–968.
14. Morse DE, Psoter WJ, Baek LS, et al. Smoking and drinking in relation to depressive symptoms among persons with oral cancer or oral epithelial dysplasia. *Head Neck* 2010;32:578–587.
15. Ronis DL, Duffy SA, Fowler KE, Khan MJ, Terrell JE. Changes in quality of life over 1 year in patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2008;134:241–248.
16. Allison PJ. Alcohol consumption is associated with improved health-related quality of life in head and neck cancer patients. *Oral Oncol* 2002;38:81–86.

17. Potash AE, Karnell LH, Christensen AJ, Vander Weg MW, Funk GF. Continued alcohol use in patients with head and neck cancer. *Head Neck* 2010;32:905–912.
18. Kugaya A, Akechi T, Okuyama T, et al. Prevalence, predictive factors, and screening for psychologic distress in patients with newly diagnosed head and neck cancer. *Cancer* 2000;88:2817–2823.
19. Aarstad AK, Lode K, Larsen JP, Bru E, Aarstad HJ. Choice of psychological coping in laryngectomized, head and neck squamous cell carcinoma patients versus multiple sclerosis patients. *Eur Arch Otorhinolaryngol* 2011;268:907–915.
20. Elani HW, Allison PJ. Coping and psychological distress among head and neck cancer patients. *Support Care Cancer* 2011;19:1735–1741.
21. Hammerlid E, Silander E, Hornestam L, Sullivan M. Health-related quality of life three years after diagnosis of head and neck cancer—a longitudinal study. *Head Neck* 2001;23:113–125.
22. de Graeff A, de Leeuw JR, Ros WJ, Hordijk GJ, Blijham GH, Winnubst JA. Long-term quality of life of patients with head and neck cancer. *Laryngoscope* 2000;110:98–106.
23. Agarwal M, Hamilton JB, Crandell JL, Moore CE. Coping strategies of African American head and neck cancer survivors. *J Psychosoc Oncol* 2010;28:526–538.
24. Hammerlid E, Bjordal K, Ahlner-Elmqvist M, et al. A prospective study of quality of life in head and neck cancer patients. Part I: at diagnosis. *Laryngoscope* 2001;111:669–680.
25. Bjordal K, Ahlner-Elmqvist M, Hammerlid E, et al. A prospective study of quality of life in head and neck cancer patients. Part II: Longitudinal data. *Laryngoscope* 2001;111:1440–1452.
26. Terrell JE, Ronis DL, Fowler KE, et al. Clinical predictors of quality of life in patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2004;130:401–408.
27. Abendstein H, Nordgren M, Boysen M, et al. Quality of life and head and neck cancer: a 5 year prospective study. *Laryngoscope* 2005;115:2183–2192.
28. Hughes PJ, Scott PM, Kew J, et al. Dysphagia in treated nasopharyngeal cancer. *Head Neck* 2000;22:393–397.
29. Chambers MS, Garden AS, Kies MS, Martin JW. Radiation-induced xerostomia in patients with head and neck cancer: pathogenesis, impact on quality of life, and management. *Head Neck* 2004;26:796–807.
30. Jensen K, Jensen AB, Grau C. A cross sectional quality of life study of 116 recurrence free head and neck cancer patients. The first use of EORTC H&N35 in Danish. *Acta Oncol* 2006;45:28–37.
31. Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. *J Clin Oncol* 2008;26:3770–3776.
32. Reeve BB, Cai J, Zhang H, et al. Health-related quality of life differences between African Americans and non-Hispanic whites with head and neck cancer. *Head Neck* 2013;35:1255–1264.
33. Divaris K, Olshan AF, Smith J, et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. *Cancer Causes Control* 2010;21:567–575.
34. D'Antonio LL, Zimmerman GJ, Cella DF, Long SA. Quality of life and functional status measures in patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 1996;122:482–487.
35. List MA, D'Antonio LL, Cella DF, et al. The Performance Status Scale for Head and Neck Cancer Patients and the Functional Assessment of Cancer Therapy-Head and Neck Scale. A study of utility and validity. *Cancer* 1996;77:2294–2301.
36. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol* 1993;11:570–579.
37. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes* 2003;1:79.
38. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002;77:371–383.
39. VanderWalde NA, Fleming M, Weiss J, Chera BS. Treatment of older patients with head and neck cancer: a review. *Oncologist* 2013;18:568–578.
40. Fenn KM, Evans SB, McCorkle R, et al. Impact of financial burden of cancer on survivors' quality of life. *J Oncol Pract* 2014;10:332–338.
41. Rao D, Debb S, Blitz D, Choi SW, Cella D. Racial/Ethnic differences in the health-related quality of life of cancer patients. *J Pain Symptom Manage* 2008;36:488–496.
42. Utsey SO, Chae MH, Brown CF, Kelly D. Effect of ethnic group membership on ethnic identity, race-related stress, and quality of life. *Cultur Divers Ethnic Minor Psychol* 2002;8:366–377.
43. Epstein JB, Robertson M, Emerton S, Phillips N, Stevenson-Moore P. Quality of life and oral function in patients treated with radiation therapy for head and neck cancer. *Head Neck* 2001;23:389–398.
44. Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiol Biomarkers Prev* 2005;14:467–475.
45. Penner JL. Psychosocial care of patients with head and neck cancer. *Semin Oncol Nurs* 2009;25:231–241.

Reprinted by permission of Cancer. 2016; 122(12):1861-1870.

Original Article

Health-Related Quality of Life Before and After Head and Neck Squamous Cell Carcinoma: Analysis of the Surveillance, Epidemiology, and End Results–Medicare Health Outcomes Survey Linkage

Eleni M. Rettig, MD[1]; Gypsyamber D'Souza, PhD[1,2]; Carol B. Thompson, MS, MBA[3]; Wayne M. Koch, MD[1]; David W. Eisele, MD[1]; and Carole Fakhry, MD, MPH[1,2]

BACKGROUND: Understanding health-related quality of life (HRQOL) is crucial to providing high-quality survivorship care for patients with head and neck squamous cell carcinoma (HNSCC). Trends in and prognostic significance of HRQOL before and after HNSCC have not been well described. **METHODS:** HRQOL for older individuals with HNSCC was examined using the linked Surveillance, Epidemiology, and End Results–Medicare Health Outcomes Survey database. Surveys assessing HRQOL from 5 years prediagnosis to 10 years postdiagnosis were included. HRQOL over time was modeled using multilevel linear regression with restricted cubic splines and was reported as either total HRQOL or change in HRQOL (denoted Δ). The association of prediagnosis HRQOL with survival was examined. **RESULTS:** In total, 1653 individuals were included; of these, 61% completed 1 survey, and 39% completed multiple surveys. Overall HRQOL decreased progressively until 13 months postdiagnosis, then recovered toward baseline between 2 and 5 years. However, after stratification by survival group, the postdiagnosis recovery was not observed. Individuals with shorter survival had lower HRQOL prediagnosis (<2-year survivors, 87.3; >5-year survivors, 96.4; $P = .004$) with a steeper decline in HRQOL during diagnosis and treatment (<2-year survivors: Δ, −16.6; 95% confidence interval [CI], −23.8, −9.4; >5-year survivors: Δ, −0.9; 95% CI, −1.8, 0.08). Radiotherapy and advanced stage were associated with greater declines in HRQOL during diagnosis and treatment ($P < .001$). Higher prediagnosis HRQOL was independently associated with improved overall survival (adjusted hazard ratio for 10-point increase, 0.91; 95% CI, 0.85-0.97). **CONCLUSIONS:** HRQOL declines before and after HNSCC, whereas any observed posttreatment recovery is likely an artifact of shorter survival among individuals with the lowest HRQOL. The prognostic implication of prediagnosis HRQOL may inform patient counseling. ***Cancer* 2016;122:1861-70.** © *2016 American Cancer Society.*

KEYWORDS: head and neck cancer, oropharynx, quality of life, radiotherapy, survival.

INTRODUCTION

Cancer survivorship has emerged as a health care priority in the United States (US) since the Institute of Medicine consensus report in 2005.[1] Historically, health-related quality of life (HRQOL) was not a component of either medical education curricula or cancer treatment and surveillance guidelines. However, the aging population and improving cancer treatment outcomes have highlighted the need to describe the spectrum and determinants of cancer survivors' HRQOL and to educate providers accordingly. In addition, the need to accurately measure HRQOL has been underscored.[1] The growing appreciation that HRQOL is paramount to improved cancer survivorship care is manifested by a growing literature of rigorous HRQOL evaluation in prostate and breast cancer survivors.[2,3]

HRQOL is particularly important in the current era of head and neck cancer, because the incidence of new diagnoses among young individuals with relatively few comorbidities and favorable expected long-term survival has increased significantly in recent years and is expected to continue increasing.[4-6] In the context of these incidence and survival trends and the short-term and long-term morbidities of head and neck cancer therapy, which range from speech and swallowing

Corresponding author: Carole Fakhry, MD, MPH, Johns Hopkins University School of Medicine, 601 N. Caroline Street, Sixth Floor, Baltimore, MD 21287; Fax: (410) 955-6526; cfakhry1@jhmi.edu

[1]Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins University School of Medicine, Baltimore, Maryland; [2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; [3]Biostatistics Center, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

The content of this article is solely the authors' responsibility and does not necessarily represent the official view of the National Institutes of Health.

Additional supporting information may be found in the online version of this article

DOI: 10.1002/cncr.30005, **Received:** September 30, 2015; **Revised:** January 31, 2016; **Accepted:** March 4, 2016, **Published online** May 16, 2016 in Wiley Online Library (wileyonlinelibrary.com)

disruption to cosmetic deformity, understanding HRQOL is increasingly essential. Indeed, the prospective evaluation of HRQOL was recently identified as a priority in head and neck cancer clinical trials.[7]

Despite an expanding HRQOL literature for head and neck cancer, HRQOL trends before diagnosis have not been explored, and our understanding of the HRQOL trajectory after treatment is limited. In the current study, trends in HRQOL over time relative to head and neck cancer diagnosis, its determinants, and its prognostic significance were examined using population-based HRQOL data from older individuals with head and neck squamous cell carcinoma (HNSCC).

MATERIALS AND METHODS

SEER-MHOS Database

The Medicare Health Outcomes Survey (MHOS) has been administered yearly since 1998 to a nationwide sample of individuals aged ≥65 years enrolled in US Medicare Advantage Organizations (MAOs). Baseline and 2-year follow-up MHOS surveys are administered by mail or telephone to 1000 randomly selected enrollees from each participating MAO.[8] MHOS response rates are 66% for the baseline survey and 81% for the follow-up surveys.[9]

MHOS has been linked to the Surveillance, Epidemiology, and End Results (SEER) national cancer registry to create the SEER-MHOS database for use in studying the HRQOL of older cancer survivors. The database has been described in detail elsewhere.[8,10] External investigators may access the data through an application process (available at: http://healthcaredelivery.cancer.gov/seer-mhos/; accessed March 20, 2016).

Data Collected

MHOS contains demographic, socioeconomic, health, and HRQOL data. HRQOL was measured using the Medical Outcomes Study Short Form 36 (SF-36)[11] from 1998 to 2005. The SF-36 has been used extensively in HRQOL research[12] and yields a Physical Component Summary (PCS) and a Mental Component Summary (MCS). The Veterans RAND 12-item health survey (VR-12), which includes PCS and MCS scores and is highly correlated with the SF-36, was used from 2006 through 2009.[13] In the SEER-MHOS database, PCS and MCS scores are each normalized to the general US population with a mean score of 50 and a standard deviation of 10.[8] Higher scores represent better HRQOL. PCS and MCS scores were combined into a summary score for this analysis, hereinafter referred to as "HRQOL," to reflect global health status encountered in the clinical setting. Two questions about depressive symptoms from the MHOS that were included in all cohorts but did not contribute to the HRQOL score were combined in this analysis as a screen for recent depression (questions 38 and 39, MHOS-1998).[14] Smoking status was determined from the question, "Do you now smoke every day, some days, or not at all?" (question 43, MHOS-1998).[14] The 12 comorbid conditions assessed in all MHOS cohorts were included in this analysis (hypertension, angina pectoris/coronary artery disease, congestive heart failure, myocardial infarction, other heart conditions, stroke, chronic lung conditions, inflammatory bowel disease, hip/knee arthritis, hand/wrist arthritis, sciatica, and diabetes).[14]

The SEER database contains information about incident cancers in certain areas of the United States. Cancer site, stage, treatment (radiotherapy and surgery), and vital status, with date of death, are included. For this analysis, we used the SEER staging system[15] rather than that of the American Joint Committee on Cancer,[16] because the SEER system has been used since the database began. SEER stages include: in situ (no basement membrane penetration or stromal invasion), localized (limited to the organ of origin), regional (extension beyond the organ of origin, by direct extension, and/or regional lymph node involvement), and distant (tumor cells have broken away from the primary tumor and grown at a new location).[15] This analysis did not include surgical treatment data because of the limited information available and the heterogeneity of HNSCC surgical treatment.

Study Population

Individuals eligible for this analysis were diagnosed with HNSCC, participated in an MHOS cohort from 1998 through 2009, and completed ≥ 1 MHOS survey within 5 years before and/or 10 years after HNSCC diagnosis. If multiple surveys were available for any individual, then up to 4 were included and were restricted to those most proximal to the time of cancer diagnosis. The HNSCC sites included were (with *International Classification of Disease for Oncology* codes): larynx (C32.0-C32.03, C32.8-C32.9, D02.0), oral cavity (C02.0-C02.3, C02.8-C02.9, C03.0-C03.1, C03.9, C04.0-C04.1, C04.8-C04.9, C05.0, C05.8-C05.9, C06.0-C06.2, C06.8-C06.9), oropharynx (C01, C02.4, C05.1-C05.2, C09.0-C09.1, C09.8-C09.9, C10.0, C10.2-C10.3, C10.8-C10.9, C14.2), hypopharynx (C12, C13.0-C13.2, C13.8-C13.9), lip (C00.0-C00.6, C00.8-C00.9), nasopharynx (C11.0-C11.3, C11.8-C11.9), and nasal cavity/paranasal sinuses (C30.0, C31.0-C31.3, C31.8-C31.9). The nasopharynx, nasal cavity, and paranasal sinuses were

combined for this analysis. Individuals who had histology other than squamous cell carcinoma (SEER histology codes 8050-8089) or multiple primary head and neck cancers were excluded. Oropharyngeal squamous cell carcinoma (OPSCC) was compared with nonoropharyngeal HNSCC (non-OPSCC) for some analyses.

Statistical Analysis

Summary statistics were reported as the number and percentage, mean and standard deviation, or median and interquartile range. For individuals with multiple surveys, time-varying demographic information (eg, age) was taken from the survey most proximal to HNSCC diagnosis for summarizing population characteristics. In the absence of a noncancer control group, HRQOL was described in relation to the time from HNSCC diagnosis, and HRQOL scores at various times relative to diagnosis were compared. Multilevel linear regression[17] with restricted cubic splines[18] was used to model trends in HRQOL as a dependent variable over time relative to HNSCC diagnosis, accounting for clustering by individual and MAO with random-effects intercepts. Spline terms for the time from diagnosis were selected as recommended by Harrell.[19] Models were selected by a comparison of fit to lowess smoothing functions, residuals, and Akaike Information Criteria. Similar models were applied after stratification for survival group, with 3 knots each for $<$ 2-year and 2-year to 5-year survivors and 6 knots for $>$ 5-year survivors. HRQOL at different times and absolute changes in HRQOL over given time intervals (denoted by "Δ"), were reported with 95% confidence intervals (CIs).

The overall model was then used to examine the association of variables of interest with HRQOL, accounting for the time from diagnosis. Differences in HRQOL relative to the reference group were reported with 95% CIs. For variables that interacted significantly with time from diagnosis, changes in HRQOL relative to the reference group at 3 different time points ($-$24, 13, and 60 months) were reported. A multivariable model also was constructed using variables that were significantly associated with HRQOL after adjustment for other factors and/or were deemed clinically relevant.

Survival analysis was limited to individuals who completed surveys within 5 years before HNSCC diagnosis. Survival rates were estimated using the Kaplan-Meier method.[20] Overall survival (OS) was calculated as the time from diagnosis to death, with censoring at the last known vital status. Survival curves were compared using log-rank tests. Risk factors for mortality were explored using univariable and multivariable Cox proportional

TABLE 1. Study Population Characteristics[a]

Characteristic	No. of Individuals (%)
Total no.	1653
No. of surveys per individual	
1	1006 (61)
2	546 (33)
3	58 (4)
4	43 (3)
MHOS cohort	
1998-1999	456 (28)
2000-2001	379 (23)
2002-2003	183 (11)
2004-2005	156 (9)
2006-2007	179 (11)
2008-2009	300 (18)
Age at diagnosis: Mean $\pm$ SD, y	71.7 $\pm$ 8.5
Sex	
Men	1175 (71)
Women	478 (29)
Smoking status	
Former/never	1103 (74)
Current	387 (26)
Marital status	
Married	917 (57)
Divorced/separated/never married	318 (20)
Widowed	372 (23)
Education	
$<$High school	498 (31)
High school graduate/GED	506 (32)
$>$High school	597 (37)
Household income, US$	
$<$$19,999	595 (36)
$20,000-49,999	573 (35)
$\geq$$50,000	178 (11)
Do not know/missing	307 (19)
Race	
White	1299 (79)
Other	354 (21)
Recent depressive symptoms	
No	1134 (70)
Yes	483 (30)
No. of comorbidities	
0-1	557 (35)
2-3	573 (36)
$\geq$4	478 (30)
Calendar period of diagnosis	
1988-1997	380 (23)
1998-2000	367 (22)
2001-2003	265 (22)
2004-2006	311 (19)
2007-2009	230 (14)
Primary site	
Larynx	625 (38)
Oral cavity	385 (23)
Lip	198 (12)
Oropharynx	295 (18)
Hypopharynx	71 (4)
Nasopharynx, nasal cavity, paranasal sinuses	79 (5)
SEER stage	
In situ	94 (6)
Local	822 (56)
Regional	426 (29)
Distant	119 (8)
Radiotherapy	
No	661 (41)
Yes	963 (59)

Abbreviations: GED, General Educational Development; MHOS, Medicare Health Outcomes Survey; SD, standard deviation; SEER, Surveillance, Epidemiology, and End Results.

[a] For individuals who completed multiple surveys, time-varying demographic information is from the survey most proximal to the time of head and neck cancer diagnosis.

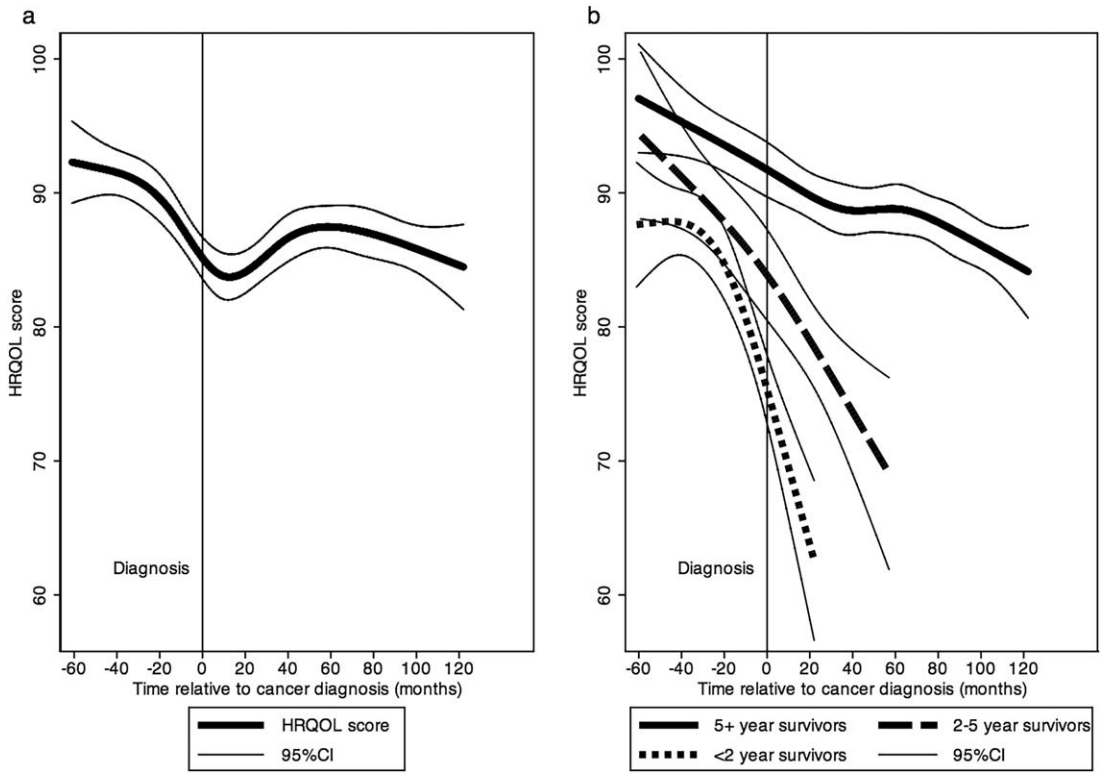


Figure 1. Health-related quality of life (HRQOL) is illustrated over time from the diagnosis of head and neck cancer for (a) overall study population and (b) each survival group. CI indicates confidence interval. Vertical line at 0 months indicates time of diagnosis.

hazards models.[21] Log(-log) plots were inspected, and statistical tests of the proportional hazards assumption were used to ensure validity of the proportional hazards model. To examine the association of HRQOL with survival, HRQOL from the survey most proximal to HNSCC diagnosis was considered both as a categorical variable by quartile and as a continuous variable per 10-point increase, to obtain hazard ratios (HRs) associated with clinically relevant differences in HRQOL ($^1/_2$ of 1 standard deviation, consistent with previous research[22]). The multivariate model had 80% power to detect an 8% reduction in hazard of death per 10-point increase in the HRQOL score.

Statistical analyses were performed using STATA version 11.2 (Stata Corporation, College Station, Tex.). Two-sided P values $\leq .05$ were considered statistically significant. This study was exempted from review by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board.

RESULTS

Study Population

The study population consisted of 1653 individuals. A single survey was available for 61% of individuals (N = 1006), and 2 or more surveys were available for 39% (N = 647). Characteristics of the study population are summarized in Table 1.

Trends in HRQOL Over Time From HNSCC Diagnosis

Overall, HRQOL decreased slowly in the 2 to 5 years before HNSCC diagnosis (Δ, −2.1; 95% CI, −5.4, 1.3) (Fig. 1a, Table 2). A steep decline was then observed beginning approximately 24 months prediagnosis and culminating in a nadir at 13 months postdiagnosis (Δ,−6.5; 95% CI, −8.9, −4.1). This was followed by an increase in HRQOL from 13 months until approximately 5 years postdiagnosis (Δ, + 3.9; 95% CI, 2.0-5.9), and finally a steady decline for the remaining 5 years of the study period (Δ,−3.4; 95% CI, −6.9, 0.1).

When considering MCS and PCS scores separately (Supporting Fig. 1; see online supporting information), trends over time relative to HNSCC diagnosis were similar to overall HRQOL scores, with the exception that the increase in the MCS score was not significant in the 5 to 10 years postdiagnosis (Δ, + 0.6; 95% CI, −1.6, 2.9).

TABLE 2. Trends in Health-Related Quality of Life at Various Times Relative to Head and Neck Cancer Diagnosis, Overall and Stratified by Survival Group

Time	Overall, N = 1653		<2 Year Survivors, N = 296		2-5 Year Survivors, N = 209		>5-Year Survivors, N = 1081	
	HRQOL Score (95% CI)	Change From Baseline: Δ (95% CI)	HRQOL Score (95% CI)	Change From Baseline: Δ (95% CI)	HRQOL Score (95% CI)	Change From Baseline: Δ (95% CI)	HRQOL Score (95% CI)	Change From Baseline: Δ (95% CI)
Prediagnosis*								
5 y (Baseline)	92.3 (89.3, 95.2)	–	87.3 (92.7, 91.9)	–	92.8 (85.1, 100.5)	–	96.4 (91.8, 100.9)	–
2 y	90.2 (88.4, 92.0)	−2.1 (−5.4, 1.3)	86.3 (83.4, 89.2)	−1.0 (−6.7, 4.8)	89.8 (85.3, 94.2)	−3.1 (−12. 4,6.3)	94.5 (91.9, 97.1)	−1.9 (−6.9, 3.1)
Diagnosis: 0 y	85.0 (83.4, 86.6)	−7.2 (−10.4, −4.0)	73.9 (70.3, 77.6)	−13.4 (−18.6, −8.2)	82.9 (79.0, 86.9)	−9.9 (−18.3, −1.5)	91.5 (89.4, 93.5)	−4.9 (−9.8, 0.04)
Postdiagnosis^								
13 mo	83.7 (82.0, 85.4)	−8.5 (−11.6, −5.4)	69.7 (62.8, 76.7)	−17.6 (−26.1, −9.0)	79.9 (76.1,83.7)	−12.9 (−20.9, −4.9)	90.1 (87.9, 92.2)	−6.3 (−11.1, −1.6)
2 y	84.1 (82.4, 85.8)	−8.2 (−11.2, −5.2)	63.8 (35.9, 91.7)	−23.5 (−51.6, 4.5)	78.0 (73.6, 82.5)	−14.8 (−23.2, −6.3)	89.2 (87.2, 91.2)	−7.2 (−11.7, −2.7)
5 y	88.0 (86.2, 89.7)	−4.3 (−7.5, −1.0)	–	–	52.1 (14.9, 89.3)	−40.7 (−79.1, −2.3)	88.6 (86.8, 90.3)	−7.8 (−12.5, −3.2)
10 y	84.6 (81.6, 87.6)	−7.7 (−11.9, −3.5)	–	–	–	–	84.2 (81.4, 87.1)	−12.1 (−17.5, −6.8)

Abbreviations: Δ, absolute changes in health-related quality of life over given time intervals; CI, confidence interval; HRQOL, health-related quality of life; *Time interval prediagnosis; ^Time interval postdiagnosis.

The MCS score was significantly higher than the PCS score at all time points ($P < .001$).

To explore whether HRQOL is a proxy for survival, individuals were stratified by OS (Fig. 1b, Table 2). HRQOL trends were evaluated among <2-year survivors (N = 296), 2-year to 5-year survivors (N = 209), and >5-year survivors (N = 1081). HRQOL declined over time for each group. Individuals with shorter survival had lower HRQOL 5 years before HNSCC diagnosis (87.3 for < 2-year survivors vs 92.8 for 2-year to 5-year survivors [$P = .16$], and 96.4 for > 5-year survivors [$P = .004$], respectively) and experienced greater declines in HRQOL around the time of diagnosis than those with longer survival (from 24 months prediagnosis to 13 months postdiagnosis: Δ, −16.6 [95% CI, −23.8, −9.4] for < 2-year survivors; Δ, −9.9 [95% CI, −15.8, −3.9] for 2-year to 5-year survivors; Δ, −0.9 [95% CI, −1.8, 0.08] for > 5-year survivors). It is noteworthy that the significant increase in HRQOL observed in the overall study population between 13 months and 5 years postdiagnosis was not observed in any group after stratification by survival.

Determinants of HRQOL

Associations between characteristics of interest and HRQOL were examined (Tables 3 and 4). In univariate analysis, factors that were significantly associated with HRQOL included age at diagnosis, sex, smoking, marital status, education, household income, race, comorbidities, depression, tumor site, disease stage, and radiotherapy. In multivariable analysis, variables that were independently associated with HRQOL included age, household income, comorbidities, depression, disease stage, and radiotherapy ($P < .01$ for all).

Interactions between characteristics of significance and time from HNSCC diagnosis in association with HRQOL also were explored. In univariate analysis, significant interactions with time from HNSCC diagnosis were observed for SEER stage ($P < .001$) and radiotherapy ($P = .001$). Individuals with late-stage disease experienced greater decreases in HRQOL around the time of diagnosis (from 24 months prediagnosis to 13 months postdiagnosis) compared with individuals who had early stage disease (Δ, −13.0 [95% CI, −16.8, −9.2] vs Δ, −2.6 [95% CI, −5.9, 0.7]; $P < .001$). Similarly, individuals who received radiotherapy had greater decreases in HRQOL than those who did not (Δ,−10.3 [95% CI, −13.4, −7.2] vs Δ, −0.8 [95% CI, −4.6, 3.0]; $P < .001$). After adjustment for other factors, significantly greater declines in HRQOL around the time of diagnosis were still observed for individuals with late-stage disease ($P = .002$) or who received radiotherapy ($P < .001$).

TABLE 3. Characteristics Associated With Health-Related Quality of Life (HRQOL) Among Individuals Diagnosed With Head and Neck Cancer in Which the Association With HRQOL Does Not Vary Significantly by Time From Diagnosis

	Difference in HRQOL: Δ (95% CI)			
Characteristic	Univariate Analysis[a]	*P*	Multivariate Analysis[a,b]	*P*
Age by quartile at diagnosis, y		<.001		.01
<67	REF		REF	
68-72	3.9 (1.5, 6.3)		1.9 (−0.1, 3.9)	
73-77	2.6 (−0.1, 5.1)		1.0 (−1.0, 3.0)	
>78	−2.0 (−4.6, 0.5)		−1.3 (−3.4, 0.9)	
Sex		.006		
Men	REF			
Women	−2.6 (−4.4, −0.7)			
Smoking status		<.001		
Former/never	REF			
Current	−3.6 (−5.5, −1.8)			
Marital status		.001		
Married	REF			
Not married	−2.7 (−4.4, −1.1)			
Education		<.001		
<High school	REF			
High school graduate/GED	3.8 (1.9, 5.8)			
>High school	7.1 (5.1, 9.1)			
Household income, US$		<.001		<.001
<$19,999	REF		REF	
$20,000-49,999	5.6 (3.9, 7.3)		3.7 (2.3, 5.2)	
≥$50,000	12.3 (9.8, 14.9)		8.5 (6.3, 10.7)	
Do not know/missing	3.5 (1.6, 5.4)		3.9 (2.1, 5.8)	
Race		.008		
White	REF			
Other	−2.8 (−4.9, −0.7)			
No. of comorbidities		<.001		<.001
0-1	REF		REF	
2-3	−9.3 (−10.9, −7.7)		−7.3 (−8.8, −5.8)	
≥4	−17.9 (−19.6, −16.1)		−14.9 (−16.5, −13.2)	
Depression in past year		<.001		<.001
No	REF		REF	
Yes	−20.2 (−21.6, −18.8)		−16.8 (−18.2, −15.4)	
Primary site		.03		
Larynx/hypopharynx	REF			
Oral cavity	−0.9 (−3.0, 1.3)			
Oropharynx	−0.6 (−2.9, 1.8)			
Other[c]	2.9 (0.4, 5.3)			

Abbreviations: Δ, absolute difference in HRQOL compared with reference category; CI, confidence interval; REF, reference category.
[a] Models include cubic spline terms for time relative to head and neck squamous cell carcinoma diagnosis.
[b] Model includes all characteristics for which multivariate analysis results are reported in Tables 3 and 4, survey by proxy, and calendar year of diagnosis.
[c] Other sites include the nasopharynx, nasal cavity, paranasal sinuses, and lip.

MCS and PCS scores were analyzed separately (Supporting Tables 1 and 2; see online supporting information). Although similar patterns were observed for HRQOL overall, radiotherapy was associated with MCS scores ($P < .001$) but not PCS scores ($P = .36$), whereas sex was associated with PCS scores ($P < .001$) but not MCS scores ($P = .91$).

Survival Analysis

Given the distinct trends in HRQOL by survival group, the prognostic implication of HRQOL for HNSCC was explored. The analysis was limited to individuals with surveys before HNSCC diagnosis (N = 664). The median OS was 48.9 months (95% CI, 38.5-54.6 months). The 2-year survival rate was 62.5% (95% CI, 58.7%-66.1%), and the 5-year survival rate was 43.9% (95% CI, 40.1%-47.7%).

OS was compared by quartiles of prediagnosis HRQOL. Higher HRQOL was significantly associated with improved OS ($P < .001$, $P_{trend} < .001$) (Fig. 2). The median OS for the highest HRQOL quartile was 79.1 months (95% CI, 65.4-100.7 months) compared with only 22.5 months (95% CI, 14.8-36.3 months) for the lowest quartile. Higher prediagnosis PCS and MCS scores were also associated with improved OS (PCS: $P < .001$, $P_{trend} < .001$; MCS: $P = .003$, $P_{trend} < .001$).

TABLE 4. Characteristics Associated With Health-Related Quality of Life (HRQOL) Among Individuals Diagnosed With Head and Neck Cancer in Which the Association With HRQOL Varies Significantly by Time From Diagnosis

	Difference in HRQOL: Δ (95% CI)							
	Univariate Analysis[a]				Multivariate Analysis[a,b]			
Characteristic	2 Years Prediagnosis	13 Months Postdiagnosis	5 Years Postdiagnosis	*P*	2 Years Prediagnosis	13 Months Postdiagnosis	5 Years Postdiagnosis	*P*
Stage				< .001				< .001
Early	REF	REF	REF		REF	REF	REF	
Late	0.3 (−3.2, 3.7)	−10.1 (−13.5, −6.7)	−0.5 (−4.1, 3.0)		0.2 (−2.8, 3.1)	−6.8 (−9.8, −3.8)	1.5 (−1.5, 4.4)	
Radiotherapy				.001				.01
No	REF	REF	REF		REF	REF	REF	
Yes	2.2 (−1.2, 5.6)	−7.3 (−10.6, −4.0)	−0.8 (−3.9, 2.2)		1.9 (−1.1, 4.9)	−5.1 (−8.1, −2.2)	−1.2 (−3.9, 1.5)	

Abbreviations: Δ, absolute difference in HRQOL compared with reference category; CI, confidence interval; REF, reference category.
[a] Models include cubic spline terms for time relative to head and neck squamous cell carcinoma diagnosis.
[b] Model includes all characteristics for which multivariate analysis results are reported in Tables 3 and 4, survey by proxy, and calendar year of diagnosis.

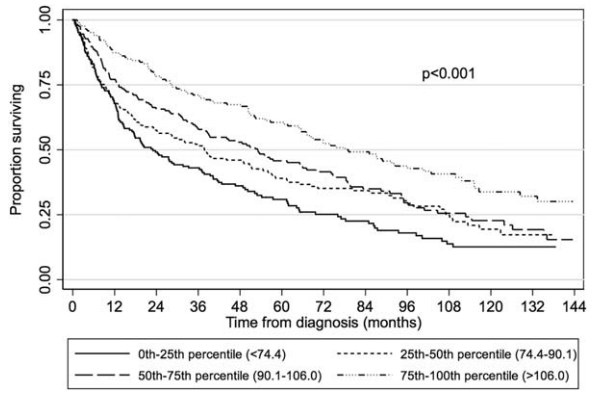


Figure 2. Overall survival is illustrated according to health-related quality-of-life quartiles assessed within 5 years before head and neck cancer diagnosis.

To account for the observed decline in HRQOL before diagnosis, survival was evaluated according to the timing of surveys before HNSCC diagnosis. Prediagnosis HRQOL quartile was significantly associated with OS after HNSCC diagnosis after limiting analysis to surveys from 1 to 5 years (N = 428; $P<.001$), 2 to 5 years (N = 239; $P<.001$), and 3 to 5 years (N = 155; $P = .006$) before diagnosis.

The prognostic significance of clinically relevant 10-point changes in HRQOL was also evaluated. A 10-point increase in prediagnosis HRQOL was associated with a 14% reduced risk of death (HR, 0.86; 95% CI, 0.82-0.91) (Table 5). Even after adjustment for other factors associated with survival (age, smoking status, marital status, comorbidities, household income, tumor site, stage, and radiotherapy), each 10-point increase in the prediagnosis HRQOL score was associated with a 9% reduction in the risk of death (adjusted HR, 0.91; 95% CI, 0.85-0.97).

OPSCC

Most OPSCCs in the United States are human papillomavirus (HPV)-related, and HPV-related OPSCC (HPV-OPSCC) is considered a distinct disease from HPV-negative HNSCC.[6] Therefore, we examined trends in HRQOL and the prognostic significance of HRQOL among individuals who had OPSCC compared with those who had non-OPSCC. Two years before HNSCC diagnosis, the HRQOL of individuals with OPSCC (N = 295) and non-OPSCC (N = 1358) was similar (Supporting Fig. 2, Supporting Table 3; see online supporting information). However, 13 months after diagnosis, HRQOL for individuals with OPSCC was significantly lower than for those with non-OPSCC (difference: −5.4; 95% CI, −9.0, −1.2). At 5 and 10 years after diagnosis, HRQOL was again similar between the 2 groups. Overall, HRQOL was not significantly different for individuals with OPSCC and those with non-OPSCC ($P = .13$). Higher prediagnosis HRQOL quartile was nonsignificantly associated with improved survival after OPSCC among 131 individuals who had prediagnosis surveys available (HR, 0.95; $P = .32$).

DISCUSSION

To our knowledge, this study is the first to examine both prediagnosis and postdiagnosis HRQOL among individuals with HNSCC. A significant and progressive decline in HRQOL is observed before and after HNSCC diagnosis, and the magnitude of decline in HRQOL differs by survival group. HRQOL in the years leading up to

TABLE 5. Risk Factors for Mortality Among 664 Individuals for Whom Health-Related Quality of Life Was Assessed Within 5 Years Before Head and Neck Cancer Diagnosis

Characteristic	No. of Patients (%)	HR [95% CI]	
		Univariate Analysis	Multivariate Analysis[a]
Prediagnosis HRQOL score per 10-point increase: Median/IQR	90.1/74.4-106.0	0.86 [0.82-0.91]	0.91 [0.85-0.97]
Age at diagnosis: Mean ± SD, y	75.2 ± 7.5	1.04 [1.03-1.06]	1.07 [1.05-1.09]
Sex			
Men	466 (70)	REF	
Women	198 (30)	1.12 [0.92-1.37]	
Smoking status			
Former/never	405 (68)	REF	REF
Current	190 (32)	1.24 [1.02-1.52]	1.48 [1.15-1.91]
Marital status			
Married	369 (57)	REF	REF
Not married	279 (43)	1.41 [1.17-1.69]	1.18 [0.92-1.50]
Education			
<High school	205 (32)	REF	
High school graduate/GED	204 (32)	0.86 [0.68-1.08]	
>High school	235 (36)	0.77 [0.61-0.96]	
Household income, US$			
<$19,999	241 (36)	REF	REF
$20,000-49,999	220 (33)	0.74 [0.60-0.92]	1.08 [0.83-1.42]
≥$50,000	73 (11)	0.50 [0.35-0.71]	0.64 [0.42-0.97]
Do not know/missing	130 (20)	0.86 [0.66-1.11]	0.83 [0.59-1.17]
Race			
White	530 (80)	REF	
Other	134 (20)	1.14 (0.89-1.44]	
No. of comorbidities			
0-1	215 (33)	REF	REF
2-3	237 (37)	1.11 [0.88-1.39]	0.99 [0.76-1.31]
≥4	196 (30)	1.46 [1.16-1.82]	1.19 [0.89-1.59]
Primary site			
Larynx, hypopharynx	267 (40)	REF	REF
Oral cavity	170 (26)	1.24 [0.99-1.56]	0.98 [0.73-1.30]
Oropharynx	131 (20)	1.59 [1.25-2.03]	1.08 [0.79-1.46]
Other[b]	96 (14)	1.48 [0.90-2.41]	0.88 [0.61-1.28]
Stage			
Early	338 (54)	REF	REF
Late	284 (46)	2.30 [1.90-2.79]	2.50 [1.95-3.19]
Radiotherapy			
No	257 (39)	REF	REF
Yes	397 (61)	1.20 [1.00-1.45]	1.20 [0.94-1.54]

Abbreviations: CI, confidence interval; GED, General Educational Development; HR, hazard ratio; HRQOL, health-related quality of life; IQR, interquartile range; REF, reference category; SD, standard deviation.

[a] The model includes all characteristics for which multivariate analysis results are reported and survey by proxy.

[b] Other sites include the nasopharynx, nasal cavity, paranasal sinuses, and lip.

HNSCC diagnosis has prognostic significance that was not previously appreciated; these large population-based data provide compelling evidence that prediagnosis HRQOL independently predicts survival.

These findings have relevance to patient care. Although most studies to date have reported a decline in HRQOL after diagnosis followed by a recovery toward baseline after treatment,[23-25] these data demonstrate that such recovery is not observed after stratification by survival group and actually may represent an artifact of earlier deaths among individuals with lower HRQOL. Therefore, to counsel patients that overall HRQOL will certainly improve after therapy, as suggested by previous studies, is likely inaccurate; instead, perhaps providers should emphasize the importance of acclimatization to a new standard of emotional and physical health. Furthermore, the finding that individuals with the lowest prediagnosis HRQOL suffer a worse prognosis independent of other prognostic indicators should lend added gravity to treatment decisions for this subset of patients, such as when considering potentially morbid, life-prolonging interventions versus high-quality palliative care.

The decline in HRQOL before an impending HNSCC diagnosis has not been described previously but is consistent with 2 longitudinal cohort studies that described significant declines in self-reported overall health before the diagnosis of other cancer types.[26,27] We therefore conclude that, for a given individual, HRQOL at diagnosis probably is significantly lower than their true baseline HRQOL. Studies that report baseline HRQOL as assessed at diagnosis, including most studies of patients with head and neck cancer to date,[23-25] should be interpreted with this discrepancy in mind. It is worth noting that this also suggests that, in the clinical setting, a rapid decline in HRQOL should be regarded with caution, because it may herald a future diagnosis of malignancy or another significant health event.[26]

Depressive symptoms, income, number of comorbidities, and age were associated with HRQOL in a constant manner over time, whereas radiotherapy and advanced disease stage were both associated with a significantly increased magnitude of decline in HRQOL during HNSCC diagnosis and treatment. In ours and other studies, it has been demonstrated that global HRQOL decreases precipitously during the radiotherapy treatment period, during which acute mucositis and other toxicities are common.[25,28] In long-term survivors, however, although many irradiated patients experience chronic toxicities, such as xerostomia and dysphagia,[29] global HRQOL scores were similar whether or not radiotherapy was received. It appears that, although radiotherapy affects specific functions, such as eating, among long-term survivors of HNSCC, it does not have a durable impact on individuals' perceptions of their global health status, which is consistent with conclusions drawn from other groups.[25,30] Advanced disease stage appears to impact HRQOL in a manner similar to that of radiotherapy in ours and other studies.[30,31]

Emerging evidence suggests that patients with HPV-OPSCC, who are generally younger and of higher socioeconomic status than their HPV-negative counterparts,[6,32] have a unique HRQOL trajectory. Despite higher prediagnosis scores, patients with HPV-OPSCC experience a greater decline in HRQOL during treatment.[33] Although the HRQOL trends observed for OPSCC compared with non-OPSCC followed this general pattern in our study, they did not reach statistical significance. However, HPV tumor status was unavailable, so the proportion of HPV-negative OPSCCs in this study was unknown. It is noteworthy that prediagnosis HRQOL was not associated with survival among individuals with OPSCC. Further investigation is required to determine whether the prognostic value of HRQOL is modified by HPV tumor status.

Limitations and Strengths

This study is a secondary analysis of population-based data and, as such, has several limitations and barriers to clinical application. The HRQOL assessment tools were not specific to head and neck cancer. No chemotherapy data and limited surgery data were available, so treatment-related changes in HRQOL and survival differences could not be fully examined. The study population was heterogenous, encompassing multiple sites and stages of head and neck cancer, and did not include a control group of individuals without cancer. Surveys were taken by adults aged $\geq$65 years, which limits generalizability to younger populations, although the majority of head and neck cancers arise in individuals in their 60s and 70s.[34] Selection bias is inevitable when analyzing data from voluntary questionnaires. Finally, approximately 66% of individuals completed only 1 MHOS survey, restricting our analysis of their HRQOL trajectory over time. However, the unique access to prediagnosis HRQOL scores and the large study population lend significance to the findings reported herein despite these drawbacks.

Conclusion

Understanding HRQOL is crucial to providing high-quality survivorship care for the growing population of HNSCC survivors. The prognostic implication of prediagnosis HRQOL should inform patient counseling. Additional research is needed to further clarify trends in and determinants of HRQOL and to examine the potential for targeted interventions to optimize HRQOL for HNSCC survivors.

FUNDING SUPPORT

This research was supported by a National Institute of Dental and Craniofacial Research/National Institutes of Health Research Training in Otolaryngology grant 2T32DC000027-26 (Eleni M. Rettig). Statistical support was provided in part by the Johns Hopkins Institute for Clinical and Translational Research.

CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

AUTHOR CONTRIBUTIONS

Eleni M Rettig: Conceptualization, methodology, formal analysis, investigation, writing–original draft, writing–review and editing, and visualization. **Gypsyamber D'Souza**: Conceptualization and writing–review and editing. **Carol B Thompson**: Formal analysis and writing–review and editing.

Wayne Koch: Validation, resources, writing–review and editing, and supervision. **David Eisele**: Writing–original draft and supervision. **Carole Fakhry**: Conceptualization, investigation, resources, writing–original draft, writing–review and editing, visualization, supervision, project administration, and funding acquisition.

REFERENCES

1. Hewitt ME, Greenfield S, Stovall E, eds; Committee on Cancer Survivorship: Improving Care and Quality of Life, National Cancer Policy Board. From Cancer Patient to Cancer Survivor: Lost in Transition. Washington, DC: The National Academies Press; 2006.
2. Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. *N Engl J Med*. 2008;358:1250-1261.
3. Howard-Anderson J, Ganz PA, Bower JE, Stanton AL. Quality of life, fertility concerns, and behavioral health outcomes in younger breast cancer survivors: a systematic review. *J Natl Cancer Inst*. 2012; 104:386-405.
4. Pulte D, Brenner H. Changes in survival in head and neck cancers in the late 20th and early 21st century: a period analysis. *Oncologist*. 2010;15:994-1001.
5. Fakhry C, Andersen KK, Eisele DW, Gillison ML. Oropharyngeal cancer survivorship in Denmark, 1977-2012. *Oral Oncol*. 2015;51: 982-984.
6. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol*. 2011;29:4294-4301.
7. Chera BS, Eisbruch A, Murphy BA, et al. Recommended patient-reported core set of symptoms to measure in head and neck cancer treatment trials [serial online]. *J Natl Cancer Inst*. 2014;106;dju127.
8. Ambs A, Warren JL, Bellizzi KM, Topor M, Haffer SC, Clauser SB. Overview of the SEER-Medicare Health Outcomes Survey linked dataset. *Health Care Financ Rev*. 2008;29:5-21.
9. Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health. The SEER-MHOS data file: response rates. Available at: http://healthcaredelivery.cancer.gov/seer-mhos/aboutdata/table.response.rates.html. Accessed April 30, 2015.
10. Reeve BB, Potosky AL, Smith AW, et al. Impact of cancer on health-related quality of life of older Americans. *J Natl Cancer Inst*. 2009;101:860-868.
11. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30:473-483.
12. Ware JE Jr. SF-36 health survey update. *Spine (Phila Pa 1976).* 2000;25:3130-3139.
13. Selim AJ, Rogers W, Qian SX, Brazier J, Kazis LE. A preference-based measure of health: the VR-6D derived from the veterans RAND 12-Item Health Survey. *Qual Life Res*. 2011;20:1337-1347.
14. Medicare Health Outcomes Survey (HOS) Program, Centers for Medicare and Medicaid Services, National Institutes of Health. Health of Seniors Survey. Available at: http://hosonline.org/globalassets/hos-online/survey-instruments/hos_1998_survey.pdf. Accessed March 20, 2016.
15. Young JL Jr, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA, eds. SEER Summary Staging Manual-2000: Codes and Coding Instructions. Bethesda, MD: National Cancer Institute; 2001.
16. Edge S, Byrd DR, Compton CC, Fritz AG, Greene FL, Trittotti A, eds. AJCC Cancer Staging Manual. 7th ed. New York: Springer; 2010.
17. Rabe-Hesketh S, Skrondal A. Multilevel and Longitudinal Modeling Using Stata. College Station, TX: Stata Press; 2008.
18. Marrie RA, Dawson NV, Garland A. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. *J Clin Epidemiol*. 2009;62:511-517, e511.
19. Harrell FE Jr. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2001.
20. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457-481.
21. Cleves M. An Introduction to Survival Analysis Using Stata. College Station, TX: Stata Press; 2008.
22. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-592.
23. So WK, Chan RJ, Chan DN, et al. Quality-of-life among head and neck cancer survivors at 1 year after treatment—a systematic review. *Eur J Cancer*. 2012;48:2391-2408.
24. Bottomley A, Tridello G, Coens C, et al. An international phase 3 trial in head and neck cancer: quality of life and symptom results: EORTC 24954 on behalf of the EORTC Head and Neck and the EORTC Radiation Oncology Group. *Cancer*. 2014;120:390-398.
25. Klein J, Livergant J, Ringash J. Health related quality of life in head and neck cancer treated with radiation therapy with or without chemotherapy: a systematic review. *Oral Oncol*. 2014;50:254-262.
26. Diehr P, Williamson J, Patrick DL, Bild DE, Burke GL. Patterns of self-rated health in older adults before and after sentinel health events. *J Am Geriatr Soc*. 2001;49:36-44.
27. Petrick JL, Foraker RE, Kucharska-Newton AM, et al. Trajectory of overall health from self-report and factors contributing to health declines among cancer survivors. *Cancer Causes Control*. 2014;25: 1179-1186.
28. Egestad H, Emaus N. Changes in health related quality of life in women and men undergoing radiation treatment for head and neck cancer and the impact of smoking status in the radiation treatment period. *Eur J Oncol Nurs*. 2014;18:339-346.
29. Machtay M, Moughan J, Trotti A, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. *J Clin Oncol*. 2008;26:3582-3589.
30. Funk GF, Karnell LH, Christensen AJ. Long-term health-related quality of life in survivors of head and neck cancer. *Arch Otolaryngol Head Neck Surg*. 2012;138:123-133.
31. Hammerlid E, Silander E, Hornestam L, Sullivan M. Health-related quality of life 3 years after diagnosis of head and neck cancer—a longitudinal study. *Head Neck*. 2001;23:113-125.
32. Gillison ML, D'Souza G, Westra W, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst*. 2008;100:407-420.
33. Sharma A, Mendez E, Yueh B, et al. Human papillomavirus-positive oral cavity and oropharyngeal cancer patients do not have better quality-of-life trajectories. *Otolaryngol Head Neck Surg*. 2012;146: 739-745.
34. Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2011. Available at: http://seer.cancer.gov/csr/1975_2011/. Accessed October 15, 2014.

Oral Oncology 59 (2016) 80–85

Contents lists available at ScienceDirect

Oral Oncology

journal homepage: www.elsevier.com/locate/oraloncology

Impact of prophylactic gastrostomy or reactive NG tube upon patient-reported long term swallow function following chemoradiotherapy for oropharyngeal carcinoma: A matched pair analysis

Brinda Sethugavalar [a,1], Mark T. Teo [a,1], Catriona Buchan [b], Ekin Ermiş [a], Gillian F. Williams [c], Mehmet Sen [a], Robin J.D. Prestwich [a,*]

[a] Department of Clinical Oncology, Leeds Cancer Centre, St. James's Institute of Oncology, Leeds, UK
[b] Department of Radiotherapy, Leeds Cancer Centre, St. James's Institute of Oncology, Leeds, UK
[c] Dietetic Department, Leeds Cancer Centre, St. James's Institute of Oncology, Leeds, UK

ARTICLE INFO

Article history:
Received 29 March 2016
Received in revised form 6 June 2016
Accepted 9 June 2016

Keywords:
Head and neck cancer
Oropharynx cancer
Radiotherapy
Chemotherapy
Swallow
Late toxicity
Quality of life
Gastrostomy
Nasogastric tube

SUMMARY

Objectives: The purpose of this matched pair analysis is to assess patient-reported long term swallow function following chemoradiotherapy for locally advanced oropharyngeal cancer in relation to the use of a prophylactic gastrostomy or reactive nasogastric (NG) tube.

Materials and methods: The MD Anderson Dysphagia Inventory (MDADI) was posted to 68 consecutive patients with stage III/IV oropharyngeal squamous cell carcinoma who had completed parotid sparing intensity modulated radiotherapy with concurrent chemotherapy between 2010 and 2012, had not required therapeutic enteral feeding prior to treatment, minimum 2 years follow up post treatment, and who were disease free. 59/68 replies were received, and a matched pair analysis (matching for T and N stage) was performed for 52 patients, 26 managed with a prophylactic gastrostomy and 26 with an approach of an NG tube as needed.

Results: There were no significant differences in patient demographics, pre-treatment diet and treatment factors between the two groups. Patient-reported swallowing function measured using the MDADI was superior for patients managed with an NG tube as required compared with a prophylactic gastrostomy: overall composite score 68.1 versus 59.4 ($p = 0.04$), global score 67.7 versus 60 ($p = 0.04$), emotional subscale 73.5 versus 60.4 ($p < 0.01$), functional subscale 75.4 versus 61.7 ($p < 0.01$), and physical subscale 59.6 versus 57.1 ($p = 0.38$).

Conclusions: Compared with an approach of an NG tube as required, the use of a prophylactic gastrostomy was associated with inferior long term patient-reported long term swallow outcomes.

Introduction

Long term dysphagia remains a major treatment-related morbidity of organ preserving approaches to the treatment of head and neck cancers [1–5], with the use of concurrent chemotherapy identified as a significant clinical factor associated with risk of long term dysphagia [6,7]. Dysphagia has a major detrimental effect upon health-related quality-of-life, with multiple studies reporting an association between health-related quality of life and dysphagia [4,8,9]. In a patient questionnaire study, swallowing was rated by a majority of patients as a priority concern 12 months following completion of (chemo)radiotherapy [8].

The timing, route and duration of enteral feeding during and after treatment may have an important influence upon the severity of late dysphagia. During concurrent chemoradiotherapy, the majority of patients require enteral tube feeding support either during or soon after treatment. Rates of enteral tube feeding vary widely between institutions between around 50–100% [10–13]. The chosen route of enteral tube feeding is generally either with a nasogastric (NG) tube or a gastrostomy (percutaneous endoscopic gastrostomy (PEG) or radiologically guided gastrostomy (RIG)). The choice of placement of a prophylactic feeding tube (usually a gastrostomy) prior to definitive chemoradiotherapy or a reactive approach (often with an NG tube) remains an area of highly variable practice. Reported outcomes are variable and in

* Corresponding author at: Level 4, Bexley Wing, Leeds Cancer Centre, St. James's Institute of Oncology, Beckett St., Leeds LS9 7TF, UK.
E-mail address: Robin.Prestwich@nhs.net (R.J.D. Prestwich).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.oraloncology.2016.06.007
1368-8375/

general prophylactic PEG tubes have been advocated for reduced weight loss [14–17] (albeit a small difference in several series [11,18]), lower rates of hospitalisation [11,16,17] and improved quality of life [18,19]. However, the duration of enteral feeding with a prophylactic gastrostomy has been shown to be consistently longer than with a reactive approach [11,20]. There is concern raised in some [21,22] but not all series [1,18] that prophylactic gastrostomy feeding may have a detrimental impact upon long term swallow function. It is hypothesised that prophylactic tube placement may promote a reliability upon enteral feeding, whilst NG tubes are hypothesised to promote swallowing, discourage protracted tube dependence and consequently reduce late fibrosis [23]. The potential of the choice of timing and route of enteral feeding tube to influence long term swallow outcomes remains highly controversial [10].

Dysphagia can be evaluated by a multitude of different tools, including physician reported and patient reported outcomes [4]. However, clinician and patient reported outcomes do not necessarily correlate, with the observation that patients may rate dysphagia more severely than clinicians [24]. Patient reported outcome measures are hence a key tool in assessing long term outcomes in relation to the route and timing of enteral feeding. We examined long term swallow outcomes in our previously reported cohort [11] of patients treated with chemoradiotherapy for oropharynx carcinoma [1]. We compared MDADI scores in 43 patients managed with a prophylactic PEG and 13 with a reactive NG tube; there was no difference between the two groups in any domain of the MDADI. However, the interpretation of this study is limited by the small number of patients managed with a reactive NG tube and by the use of non-parotid sparing 3D-conformal radiotherapy.

The aim of this study is to used a matched pair analysis to assess patient-reported long term swallow outcomes with the MDADI tool in patients with oropharyngeal carcinoma treated with chemoradiotherapy and parotid-sparing IMRT, in relation to the approach of using a prophylactic PEG tube or reactive NG tube if required.

Methods

Study design

The study was registered with the Institutional Quality Improvement Board.

Consecutive patients with locally advanced squamous cell carcinoma treated with concurrent chemoradiotherapy between October 2010 and December 2012 were identified from electronic records. The inclusion criteria were: oropharynx primary, squamous cell carcinoma pathology, stage III or stage IV, non-surgical treatment with curative intent, delivery of concurrent chemotherapy, use of IMRT, radiotherapy target included the bilateral neck, no prior therapeutic surgery, disease free on follow up for at least 2 years from last day of radiotherapy treatment. Patients were excluded if treatment was for recurrence, prior neck dissection, or if therapeutic enteral feeding was commenced prior to treatment.

During this period of time the policy at St. James's Institute of Oncology regarding a prophylactic or reactive approach to enteral nutritional support was to consider either a prophylactic gastrostomy or reactive NG tube approach based upon clinician ± patient preference. Gastrostomy tubes were either RIG or PEG tubes depending upon disease factors and local practice.

Patients included in the study who had completed treatment over two years previously were sent an explanatory letter inviting them to complete and return an enclosed copy of the MDADI questionnaire [25]. In the event of a non-response a follow up letter and a further copy of the questionnaire was sent after an interval of one month. The MDADI is a validated self-administered questionnaire designed for patients with head and neck cancer [25]. The questionnaire comprises 20 questions which are scored using a 5-point scale ranging from 'strongly agree' to 'strongly disagree', and is subdivided into global, emotional, functional and physical subscales. For each subscale the scores are summed and the mean score multiplied by 20 to provide a score in the range of 0–100. A higher score indicates superior swallowing quality of life and level of functioning.

Pre-treatment dietary data categorising oral intake into five categories (nil by mouth, sips, pureed, soft, normal) was prospectively collected during pre-treatment nursing and dietetic assessments as part of routine clinical care. These data were extracted by review of electronic and paper records.

Treatment details

Induction chemotherapy

Induction chemotherapy was used based upon clinician preference, patient and tumour factors; in general induction chemotherapy was considered for patients with bulky disease. Standard induction chemotherapy (ICT) consisted of either TPF (docetaxel 75 mg/m^2 day 1, cisplatin 75 mg/m^2 day 1 and 5-fluorouracil (5FU) 750 mg/m^2 days 2–5 three weekly) for selected fit patients [26], or PF (cisplatin 80 mg/m^2 day 1 and 5-fluorouracil (5FU) 800 mg/m^2 days 2–5, three weekly) [27].

Concurrent chemotherapy

Patients <70 years old were considered for concurrent chemotherapy. Standard concurrent chemotherapy was cisplatin 100 mg/m^2 days 1 and 29. Carboplatin AUC 4 was substituted for cisplatin if creatinine clearance was <55 ml/min.

Radiation treatment

Patients were treated supine with a 5 point thermoplastic mask. Planning CT scans were acquired with intravenous contrast with 2 mm slices. The planning CT dataset was transferred to the treatment planning system (Monaco®, Electa). A compartmental approach to target volume delineation was adopted as previously described [28]. Gross tumour volume (GTV) was outlined as primary tumour and clinically and/or radiologically involved lymph nodes. A primary tumour clinical target volume (CTV) was created to include at least GTV+10 mm and the anatomical compartment, modified to anatomical boundaries to exclude air and/or bone without evidence of invasion. The high dose nodal CTV was constructed to include the whole involved nodal level. Nodal levels which did not include a radiologically abnormal lymph node were treated at an intermediate or lower dose level according to clinician preference. The lymph node target routinely included levels 1b-V in the node positive neck; nodal levels in a node negative neck were selectively irradiated depending upon tumour site and disease extent according to published recommendations [29]. Retropharyngeal lymph nodes were routinely included in the target volume in cases with positive level II lymph nodes and involvement of the pharyngeal wall. The planning target volume (PTV) was created by auto-expansion of the CTV by 4 mm. Standard radical dose was 70 Gy in 35 fractions to high dose planning target volume (PTV), 63 Gy in 35 fractions to the intermediate risk PTV, and 57 Gy in 35 fractions to the elective PTV. Organ at risk (OAR) constraints were spinal canal maximum 48 Gy, brainstem maximum 54 Gy, larynx mean <45 Gy (excluding parts of larynx within PTV), contralateral parotid mean <26 Gy. Treatment was delivered with a 5–7 angle step and shoot IMRT technique.

Pre-treatment and post-treatment care

All patients underwent pre-treatment nursing, dietetic and speech and language therapy pre-treatment assessments as part of routine practice. During chemoradiotherapy treatment, all patients were routinely reviewed twice weekly by medical and nursing teams, with additional reviews with the dietetic and speech and language therapy teams as required. Post-treatment, all patients were offered ongoing dietetic and speech and language therapy support in post-treatment rehabilitation clinics.

Statistical analysis

Statistical analyses were performed using STATA software version 12 (Statacorp, Texas, USA) and GraphPad Prism version 5. Prophylactic gastrostomy versus NG tube as required matching was performed using T stage (grouped as T1/2 and T3/4) and N stage (grouped as N0/1, N2a/b, and N2c/3). Unmatched variables were compared using the Student *t*-test, Fisher's exact test and test for trend. A *t*-test and chi-squared test were used as appropriate to test for differences in the subgroups analysed. MDADI scores were compared using a non-parametric Mann-Whitney *U* test. Univariate non-parametric analyses were performed to determine if there was any correlation between MDADI scores with clinical parameters. Variables included were: age, T stage, N stage, use of induction chemotherapy, number of induction chemotherapy cycles, type of concurrent chemotherapy, number of concurrent chemotherapy cycles, mean contralateral parotid dose, pre-radiotherapy diet. Statistical significance was declared at $p < 0.05$. A multivariate MANOVA was performed on factors with a p-value < 0.20 on univariate analysis (N-stage, number of induction chemotherapy cycles, concurrent chemotherapy type, number of concurrent chemotherapy cycles, pre-treatment oral intake and mean contralateral parotid gland dose) The natural logarithm of MDADI scores were used for the multivariate analysis to normalise the data.

Results

From review of electronic records, 94 patients with a diagnosis of squamous cell carcinoma of the oropharynx who had received definitive concurrent chemoradiotherapy with curative intent were identified. Of these patients the following were excluded: 18 patients had experienced disease recurrence, 3 patients had died post-treatment with no evidence of disease recurrence, 5 had commenced therapeutic enteral feeding pre-treatment. Of the remaining 68 patients, completed MDADI questionnaires were received from 59 (87%) patients. 31/59 (53%) had been managed with a prophylactic gastrostomy tube, and 28/59 (47%) had been managed with a reactive approach with insertion of a nasogastric tube if required during treatment.

A retrospective matched pair analysis was performed, matching 26/31 patients managed with a prophylactic gastrostomy with 26/28 managed with a nasogastric tube if required, on the basis of T and N stage. On review of case notes, 24 of these matched 56 patients (43%) were documented as being offered a choice of either a prophylactic gastrostomy or a reactive approach; 12/24 (50%) opted for a prophylactic gastrostomy. Of the 26 patients managed with a prophylactic gastrostomy tube, the tube was documented as being used for at least supplemental nutrition in all patients. Within the reactive NG tube if required group, 17/26 (65%) patients had an NG tube inserted and commenced enteral feeding during treatment; no gastrostomies were used in this group. Mean follow up from the last day of radiotherapy was 36 months (range 24–59) and 34 months (range 24–59) for patients managed with a prophylactic gastrostomy or NG tube as required respectively.

The baseline matched and unmatched characteristics, along with treatment details and pre-treatment diet and shown in Table 1. With regard to the matched factors, the T stage distribution was similar between the two groups. For the purposes of matching analysis, N0 and N1 were grouped together; Table 1 shows that there was a small excess of patients with N0 disease in the group of patients managed with a prophylactic gastrostomy, although there is no statistically significant difference in N stage between the two groups. With regard to unmatched factors, Table 1 shows that there were no statistically significant differences in baseline characteristics between the matched groups, including

Table 1
Patient, tumour and treatment details.

	Prophylactic gastrostomy ($N = 26$)	NG as needed ($N = 26$)	p-value
Median follow up	34	30	0.31
Age (Mean, range)	56 (36–66)	55 (38–68)	0.55
Sex			
Male	20 (76.9%)	22 (84.6%)	0.48
Female	6 (23.1%)	4 (15.4%)	
WHO PS			
0	14 (53.8%)	20 (76.9%)	0.22
1	4 (15.4%)	2 (7.7%)	
Not recorded	8 (30.8%)	4 (15.4%)	
Smoking			
Never	10 (38.5%)	10 (38.5%)	0.87
Ex	13 (50%)	11 (42.3%)	
Current	2 (7.7%)	3 (11.5%)	
Not recorded	1 (3.8%)	2 (7.7%)	
Oropharynx subsite			
Tonsil	13 (50.0%)	15 (57.7%)	0.55
BOT	12 (46.2%)	11 (42.3%)	
Posterior pharynx	1 (3.8%)	0 (0%)	
T stage			
T1	5 (19.2%)	6 (23.1%)	0.94
T2	13 (50%)	12 (46.2%)	
T3	2 (7.7%)	3 (11.5%)	
T4	6 (23.1%)	5 (19.2%)	
Nodal stage			
N0	5 (19.2%)	0 (0%)	0.11
N1	1 (3.8%)	5 (19.2%)	
N2a	2 (7.7%)	2 (7.7%)	
N2b	15 (57.7%)	16 (61.5%)	
N2c	3 (11.5%)	2 (7.7%)	
N3	0 (0%)	1 (3.8%)	
Induction chemotherapy			
None	18 (69.2%)	24 (92.3%)	0.09
TPF	6 (23.1%)	2 (7.7%)	
PF	2 (7.7%)	0 (0%)	
RT dose			
65 Gy in 30 fractions	1 (3.8%)	1 (3.8%)	0.6
70 Gy in 35 fractions	24 (92.3%)	24 (92.3%)	
Mean contralateral parotid dose (range)/Gy	37 (24–54)	33 (21–57)	0.06
Concurrent chemotherapy			
Cisplatin	22 (96.2%)	24 (92.3%)	0.39
Carboplatin	4 (3.8%)	2 (7.7%)	
No. of concurrent chemotherapy cycles			
1	4 (15.4%)	6 (23.1%)	0.45
2	22 (84.6%)	19 (73.1%)	
3	0 (0%)	1 (3.8%)	
Pre-treatment oral intake			
NBM	0 (0%)	0 (0%)	0.38
Sips	0 (0%)	0 (0%)	
Pureed	2 (7.7%)	0 (0%)	
Soft	3 (11.5%)	3 (11.5%)	
Normal	21 (80.8%)	21 (80.8%)	
Not recorded	0 (0%)	2 (7.7%)	

Table 2
MDADI scores according to intended enteral feeding route.

	Prophylactic gastrostomy ($N = 26$)	NG as needed ($N = 26$)	p-value
Total: Mean (SD)	59.4 (16.8)	68.1 (12.9)	**0.04**
Global: Mean (SD)	60 (26.5)	67.7 (24.7)	**0.04**
Physical: Mean (SD)	57.1 (14.9)	59.6 (11.4)	0.38
Emotional: Mean (SD)	60.4 (19.1)	73.5 (15.1)	**<0.01**
Functional: Mean (SD)	61.7 (20.1)	75.4 (15.5)	**<0.01**

Values which are statistically significant are shown in bold.

Table 3
Univariate analysis of predictors of MDADI score.

Factor	p-value				
	Total	Global	Physical	Emotional	Functional
Age	0.34	0.78	0.5	0.53	0.22
T stage	0.71	0.83	0.74	0.52	0.68
N stage	0.11	0.42	0.06	0.09	0.16
Induction chemo	0.37	0.39	0.52	0.44	0.26
No. induction chemo cycles	0.11	0.45	0.33	0.16	0.05
Conc chemo type	0.1	0.15	0.08	0.06	0.11
No. conc chemo cycles	0.19	0.84	0.27	0.18	0.30
Pre-treatment diet	**0.02**	0.16	**<0.01**	**0.02**	**0.03**
Mean contralateral parotid dose	0.16	0.46	0.60	0.07	0.06

Values which are statistically significant are shown in bold.

age, sex, smoking status and baseline diet. There was a non-significant difference in the number of patients who received induction chemotherapy between the two groups, although the total number of cycles of concurrent chemotherapy delivered was similar. The difference in mean contralateral parotid dose between the two groups was not significant.

Patient-reported swallowing function measured using the MDADI was superior for patients managed with an NG tube as required compared with a prophylactic gastrostomy: overall composite score 68.1 versus 59.4 ($p = 0.04$), global score 67.7 versus 60 ($p = 0.04$), emotional subscale 73.5 versus 60.4 ($p < 0.01$), functional subscale 75.4 versus 61.7 ($p < 0.01$), and physical subscale 59.6 versus 57.1 ($p = 0.38$). The composite total, global and domain-specific (physical, emotional and functional) subscales for each group are detailed in Table 2. Each domain is scored 0–100 with higher scores indicating better swallow function. A univariate analysis was performed to explore the relationship between patient, tumour and treatment factors with MDADI scores (Table 3). Age, T stage, N stage, and treatment factors including use of induction chemotherapy, number of concurrent chemotherapy cycles and mean contralateral parotid dose, did not correlate with MDADI scores in any domain. By contrast, the quality of pre-treatment diet according to a simple scale of consistency was found to be significantly associated with all domains of the MDADI other than the global subscale. This was confirmed on a multivariate analysis, with only having a more normal pre-treatment oral intake was significantly associated with higher MDADI emotional ($p = 0.02$), physical ($p = 0.01$), total ($p = 0.02$) and possibly functional ($p = 0.06$) scores.

Discussion

Long term swallow function is a major survivorship issue [4]. There have been conflicting reports regarding whether the use of a prophylactic gastrostomy may have a detrimental impact upon long term swallow function compared with a reactive approach. Chen et al. [22] reported an increased risk of late oesophageal strictures requiring dilatation (30% versus 6% for the prophylactic versus reactive approach). A recent retrospective study reported a 5-year incidence of severe late dysphagia in 30.8% of the reactive NG tube cohort ($n = 36$), and 60.9% in the prophylactic PEG cohort ($n = 25$) with a PEG being associated with an increased rate of severe late dysphagia on a multivariate analysis [21]; however, the prophylactic PEG cohort were a historically earlier cohort prior to a shift in the institutional approach to enteral feeding. Mekhail et al. reported a 30% versus 8% dysphagia rate at the relatively early time point of 6 months post treatment for gastrostomy versus NG feeding [30]. In a prospective study Corry [31] similarly found an increase in grade 3 dysphagia with the use of a gastrostomy (25% versus 8%) at the same 6 month timepoint post-treatment. Oozeer et al. [32] used patient-reported swallow outcomes obtained using the validated MD Anderson Dysphagia Inventory (MDADI) questionnaire to perform a matched pair analysis in a group of 31 patients who completed the questionnaire at least 2 years post-treatment; the MDADI scores for all domains were significantly superior for the reactive NG group compared with the prophylactic gastrostomy group. By contrast, one prospective randomised study [18] reported that the prophylactic PEG group had a lower rate of long term grade 3 dysphagia (3% versus 9%) and a higher proportion of patients who resumed a normal diet (93% versus 80%) [18]. Recent systematic reviews have reported that the impact of prophylactic PEG use on swallowing and swallow-related outcomes remains unclear and an area of clinical equipoise [33,34]. This remains an area of controversy and wide variation in practice [10]. An ongoing randomised trial [35] may prove to be informative if recruitment can be successfully completed, although previous randomised studies have failed to complete recruitment [36].

We have previously reported long term patient-reported swallow outcomes in a cohort of patients treated in the era of 3D-conformal radiotherapy [1]. However, advances in radiotherapy techniques may impact upon swallow function, and conclusions from studies performed in the 3D-conformal radiotherapy era are not necessarily applicable to the IMRT era. Advances such as the introduction of parotid sparing IMRT have reduced xerostomia [37] and may benefit swallow function; by contrast, the move from 3D-conformal radiotherapy with a matched anterior neck field with midline shielding to whole field IMRT has led to an increase in the midline neck dose, including the larynx and pharyngo-oesophageal axis. The impact of this remains unclear, with an uncertain dose response for swallow dysfunction. An alternative technique is matching IMRT with a neck field, and it remains unclear whether this has a favourable impact upon swallow function compared with whole field IMRT [2]. There is currently interest in developing 'swallow-sparing' IMRT, although at present the efficacy of this approach remains uncertain [2].

Assessment of swallow outcomes is complex, with multiple potential tools, including physician assessed toxicity scores, patient reported function, and physical outcomes including stricture rates [2]. Patient and clinician reported outcomes may both be valuable, although it is recognised that clinicians may underestimate dysphagia compared with patients [38]. The MDADI is a validated tool for assessing patient reported swallowing outcomes [25]. In addition to the method of swallow assessment, the timing is likely to be a critical factor influencing outcome. For example MDADI scores have been found to significantly improve at 12 months post treatment when compared with earlier timepoints [4]. These data suggest that swallow function is continuing to improve 12 months post-treatment and may have yet to plateau. This is consistent with the observation that salivary recovery does not plateau until two years post treatment [37].

This study has addressed the important clinical question of whether the use of a prophylactic gastrostomy or a reactive NG

approach is preferable for long term swallow outcomes. We have examined a retrospective cohort of patients who all received concurrent chemo-radiotherapy using whole-field parotid-sparing IMRT for locally advanced oropharyngeal carcinoma, using the MDADI as a validated patient-reported tool, with long term follow up of at least 2 years post-treatment. These data suggest that the use of a prophylactic gastrostomy results in statistically inferior overall MDADI scores, as well as in the global, emotional and functional subscales, with a small statistically non-significant benefit in the physical domain.

The clinical significance of these results is dependent upon the extent to which these two groups of patients are comparable. All patients received concurrent chemotherapy and bilateral neck IMRT and did not require therapeutic enteral feeding prior to treatment. The matched pair analysis was performed to minimise differences in swallow outcome which may have been due to tumour stage, which are recognised to influence long term swallow function [2,7]. The selection of feeding route was dependent upon clinician and patient preference, and all patients were entered into routine programmes of dietetic and speech and language therapy support during and after treatment. There was no difference in baseline swallow function between these two groups of patients, measured using a simple dietetic consistency scale. In addition there were no significant differences in patient demographics, tumour stage, treatment details between the two groups. Despite this, it is not possible to completely exclude the possibility that baseline factors may have influenced the choice of approach to enteral feeding and consequently confound possible associations with swallow function. The T stage and N stage match involved grouping stages together e.g. N0 and N1, to allow an adequate number of patients to be matched for subsequent analysis. There was a higher number of N0 patients within the prophylactic gastrostomy group, although there was no significant difference between T and N stages, and N0 nodal stage might be expected to be associated with superior swallowing outcomes. A slightly higher proportion of patients in the prophylactic gastrostomy group received induction chemotherapy, possibly reflecting a perceived clinical preference for using a gastrostomy to support patients through treatment involving induction and concurrent chemotherapy.

There are some limitations to this study. We do not have human papilloma virus (HPV) status available for this a useful proportion of this historical cohort of patients as it was not being routinely tested at our institution in this era. However, it seems likely that HPV status is balanced across the two groups as the proportion with current or previous smoking status was similar, as was the proportion of patients with advanced nodal disease. There is no data to suggest that there are differences in the impact of chemoradiotherapy upon late dysphagia risk depending upon HPV status. It should also be noted that the mean parotid doses achieved with the compartmental outlining methods are considerably higher than we would currently expect with current volumetric outlining and more advanced IMRT delivery techniques; it is possible that this may have impacted upon the overall swallow function.

Conclusion

Many factors may influence long term swallow recovery post-chemoradiotherapy, including patient characteristics, baseline swallow function, tumour factors, smoking status, and swallowing support and rehabilitation provided during and after treatment [2,10]. The timing and route of enteral feeding tube may be an important factor. This is an area in which previous randomised trials [36] have failed to adequately recruit, and institutional outcomes are important to inform practice. This matched pair analysis reinforces concern over the potential for a prophylactic gastrostomy to negatively impact upon long term swallow recovery.

Conflict of interest statement

We have no conflicts of interest.

References

[1] Prestwich RJ, Teo MT, Gilbert A, Williams G, Dyker KE, Sen M. Long-term swallow function after chemoradiotherapy for oropharyngeal cancer: the influence of a prophylactic gastrostomy or reactive nasogastric tube. Clin Oncol (R Coll Radiol) 2014;26:103–9.

[2] Batth SS, Caudell JJ, Chen AM. Practical considerations in reducing swallowing dysfunction following concurrent chemoradiotherapy with intensity-modulated radiotherapy for head and neck cancer. Head Neck 2014;36:291–8.

[3] Machtay M, Moughan J, Trotti A, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. J Clin Oncol 2008;26:3582–9.

[4] Roe JW, Drinnan MJ, Carding PN, Harrington KJ, Nutting CM. Patient-reported outcomes following parotid-sparing intensity-modulated radiotherapy for head and neck cancer. How important is dysphagia? Oral Oncol 2014;50:1182–7.

[5] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol 2008;26:3770–6.

[6] Caudell JJ, Schaner PE, Meredith RF, et al. Factors associated with long-term dysphagia after definitive radiotherapy for locally advanced head-and-neck cancer. Int J Radiat Oncol Biol Phys 2009;73:410–5.

[7] Langendijk JA, Doornaert P, Rietveld DH, Verdonck-de Leeuw IM, Leemans CR, Slotman BJ. A predictive model for swallowing dysfunction after curative radiotherapy in head and neck cancer. Radiother Oncol 2009;90:189–95.

[8] Wilson JA, Carding PN, Patterson JM. Dysphagia after nonsurgical head and neck cancer treatment: patients' perspectives. Otolaryngol Head Neck Surg 2011;145:767–71.

[9] Lovell SJ, Wong HB, Loh KS, Ngo RY, Wilson JA. Impact of dysphagia on quality-of-life in nasopharyngeal carcinoma. Head Neck 2005;27:864–72.

[10] Koyfman SA, Adelstein DJ. Enteral feeding tubes in patients undergoing definitive chemoradiation therapy for head-and-neck cancer: a critical review. Int J Radiat Oncol Biol Phys 2012;84:581–9.

[11] Williams GF, Teo MT, Sen M, Dyker KE, Coyle C, Prestwich RJ. Enteral feeding outcomes after chemoradiotherapy for oropharynx cancer: a role for a prophylactic gastrostomy? Oral Oncol 2012;48:434–40.

[12] Clavel S, Fortin B, Despres P, et al. Enteral feeding during chemoradiotherapy for advanced head-and-neck cancer: a single-institution experience using a reactive approach. Int J Radiat Oncol Biol Phys 2011;79:763–9.

[13] Nguyen NP, North D, Smith HJ, et al. Safety and effectiveness of prophylactic gastrostomy tubes for head and neck cancer patients undergoing chemoradiation. Surg Oncol 2006;15:199–203.

[14] Wiggenraad RG, Flierman L, Goossens A, et al. Prophylactic gastrostomy placement and early tube feeding may limit loss of weight during chemoradiotherapy for advanced head and neck cancer, a preliminary study. Clin Otolaryngol 2007;32:384–90.

[15] Lee H, Havrila C, Bravo V, et al. Effect of oral nutritional supplementation on weight loss and percutaneous endoscopic gastrostomy tube rates in patients treated with radiotherapy for oropharyngeal carcinoma. Support Care Cancer 2008;16:285–9.

[16] Lee JH, Machtay M, Unger LD, et al. Prophylactic gastrostomy tubes in patients undergoing intensive irradiation for cancer of the head and neck. Arch Otolaryngol Head Neck Surg 1998;124:871–5.

[17] Piquet MA, Ozsahin M, Larpin I, et al. Early nutritional intervention in oropharyngeal cancer patients undergoing radiotherapy. Support Care Cancer 2002;10:502–4.

[18] Silander E, Nyman J, Bove M, Johansson L, Larsson S, Hammerlid E. Impact of prophylactic percutaneous endoscopic gastrostomy on malnutrition and quality of life in patients with head and neck cancer: a randomized study. Head Neck 2012;34:1–9.

[19] Salas S, Baumstarck-Barrau K, Alfonsi M, et al. Impact of the prophylactic gastrostomy for unresectable squamous cell head and neck carcinomas treated with radio-chemotherapy on quality of life: prospective randomized trial. Radiother Oncol 2009;93:503–9.

[20] Corry J, Poon W, McPhee N, et al. Prospective study of percutaneous endoscopic gastrostomy tubes versus nasogastric tubes for enteral feeding in patients with head and neck cancer undergoing (chemo)radiation. Head Neck 2009;31:867–76.

[21] Ward MC, Bhateja P, Nwizu T, et al. Impact of feeding tube choice on severe late dysphagia after definitive chemoradiotherapy for human papillomavirus-negative head and neck cancer. Head Neck 2015;38(S1):E1054–60.

[22] Chen AM, Li BQ, Lau DH, et al. Evaluating the role of prophylactic gastrostomy tube placement prior to definitive chemoradiotherapy for head and neck cancer. Int J Radiat Oncol Biol Phys 2010;78:1026–32.

[23] Paleri V, Roe JW, Strojan P, et al. Strategies to reduce long-term postchemoradiation dysphagia in patients with head and neck cancer: an evidence-based review. Head Neck 2014;36:431–43.
[24] Frowen JJ, Perry AR. Swallowing outcomes after radiotherapy for head and neck cancer: a systematic review. Head Neck 2006;28:932–44.
[25] Chen AY, Frankowski R, Bishop-Leone J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M.D. Anderson dysphagia inventory. Arch Otolaryngol Head Neck Surg 2001;127:870–6.
[26] Prestwich RJ, Oksuz DC, Dyker K, Coyle C, Sen M. Feasibility and efficacy of induction docetaxel, cisplatin, and 5-fluorouracil chemotherapy combined with cisplatin concurrent chemoradiotherapy for nonmetastatic stage IV head-and-neck squamous cell carcinomas. Int J Radiat Oncol Biol Phys 2011;81:e237–43.
[27] Prestwich RJ, Kancherla K, Oksuz DC, et al. A single centre experience with sequential and concomitant chemoradiotherapy in locally advanced stage IV tonsillar cancer. Radiat Oncol 2010;5:121.
[28] Bayman E, Prestwich RJ, Speight R, et al. Patterns of failure after intensity-modulated radiotherapy in head and neck squamous cell carcinoma using compartmental clinical target volume delineation. Clin Oncol (R Coll Radiol) 2014;26:636–42.
[29] Gregoire V, Levendag P, Ang KK, et al. CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG consensus guidelines. Radiother Oncol 2003;69:227–36.
[30] Mekhail TM, Adelstein DJ, Rybicki LA, Larto MA, Saxton JP, Lavertu P. Enteral nutrition during the treatment of head and neck carcinoma: is a percutaneous endoscopic gastrostomy tube preferable to a nasogastric tube? Cancer 2001;91:1785–90.
[31] Corry J. Feeding tubes and dysphagia: cause or effect in head and neck cancer patients. J Med Imaging Radiat Oncol 2009;53:431–2.
[32] Oozeer NB, Corsar K, Glore RJ, Penney S, Patterson J, Paleri V. The impact of enteral feeding route on patient-reported long term swallowing outcome after chemoradiation for head and neck cancer. Oral Oncol 2011;47:980–3.
[33] Shaw SM, Flowers H, O'Sullivan B, Hope A, Liu LW, Martino R. The effect of prophylactic percutaneous endoscopic gastrostomy (PEG) tube placement on swallowing and swallow-related outcomes in patients undergoing radiotherapy for head and neck cancer: a systematic review. Dysphagia 2015;30:152–75.
[34] Paleri V, Patterson J. Use of gastrostomy in head and neck cancer: a systematic review to identify areas for future research. Clin Otolaryngol 2010;35:177–89.
[35] A feasibility randomised controlled trial of pre-treatment gastrostomy tube versus oral feeding plus as-needed nasogastric tube feeding in patients undergoing chemoradiation for head and neck cancer (TUBE trial). <http://wwwnetsnihracuk/projects/hta/123532> [accessed 06/06/2016].
[36] Corry J, Poon W, McPhee N, et al. Randomized study of percutaneous endoscopic gastrostomy versus nasogastric tubes for enteral feeding in head and neck cancer patients treated with (chemo)radiation. J Med Imaging Radiat Oncol 2008;52:503–10.
[37] Nutting CM, Morden JP, Harrington KJ, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol 2011;12:127–36.
[38] Gluck I, Feng FY, Lyden T, et al. Evaluating and reporting dysphagia in trials of chemoirradiation for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2010;77:727–33.

Reprinted by permission of JAMA Otolaryngol Head Neck Surg. 2015; 141(6):499-504.

Original Investigation

Transoral Robotic Surgery Alone for Oropharyngeal Cancer
Quality-of-Life Outcomes

Garret W. Choby, MD; Jeehong Kim, BS; Diane C. Ling, BA; Shira Abberbock, MS; Rajarsi Mandal, MD; Seungwon Kim, MD; Robert L. Ferris, MD, PhD; Umamaheswar Duvvuri, MD, PhD

IMPORTANCE Few studies have examined quality-of-life (QOL) outcomes in patients who undergo transoral robotic surgery (TORS) alone (ie, without adjuvant radiotherapy or chemoradiotherapy).

OBJECTIVE To report QOL outcomes of patients with oropharyngeal squamous cell carcinoma who receive only TORS.

DESIGN, SETTING, AND PARTICIPANTS Medical records for all patients undergoing TORS for treatment of primary oropharyngeal squamous cell carcinoma from May 1, 2010, to March 31, 2014, at a tertiary care academic cancer center were examined from June through September 2014. Thirty-four patients who did not receive adjuvant therapy after TORS were included in the study.

INTERVENTION Primary surgical resection via TORS.

MAIN OUTCOMES AND MEASURES The University of Washington Quality of Life, version 4, questionnaire was completed by patients preoperatively and at 1-, 6-, 12-, and 24-month intervals after TORS. Demographic, clinicopathologic, and follow-up data were collected.

RESULTS Mean follow-up time was 14 months (May 1, 2010, to April 30, 2014). Most patients had T1 (20 [59%]) or T2 (13 [38%]) and N0 (13 [38%]) or N1 (16 [47%]) disease. Statistically significant improvement in QOL outcomes was noted in the following postoperative domains: chewing from 1 month (median, 50 [IQR, 50-100]) to 12 months (100 [IQR, 100-100]; $P = .048$), swallowing from 1 month (70 [IQR, 30-85]) to 6 months (100 [IQR, 70-100]; $P = .047$) and 1 to 24 months (100 [IQR, 70-100]; $P = .048$), pain from 1 month (38 [IQR, 25-75]) to 6 months (88 [IQR, 75-100]; $P = .006$) and 1 to 12 months after surgery (100 [IQR, 75-100]; $P = .01$), and activity from 1 month (63 [IQR, 50-88]) to 24 months (100 [IQR, 75-100]; $P = .03$). Two participants (6%) died during the follow-up period: 1 because of disease and 1 because of a myocardial infarction. Two patients (6%) required temporary gastrostomy tube placement, but none required tracheostomy.

CONCLUSIONS AND RELEVANCE Appropriately selected patients who undergo TORS alone for oropharyngeal squamous cell carcinoma experience acceptable short- and long-term QOL outcomes.

JAMA Otolaryngol Head Neck Surg. 2015;141(6):499-504. doi:10.1001/jamaoto.2015.0347
Published online April 2, 2015.

Author Affiliations: Department of Otolaryngology–Head and Neck Surgery, University of Pittsburgh Medical Center Pittsburgh, Pennsylvania (Choby, Mandal, S. Kim, Ferris, Duvvuri); medical student, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania (J. Kim, Ling); Biostatistics Facility, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania (Abberbock); Department of Otolaryngology, Veterans Affairs Pittsburgh Health System, Pittsburgh, Pennsylvania (Duvvuri).

Corresponding Author: Umamaheswar Duvvuri, MD, PhD, Department of Otolaryngology–Head and Neck Surgery, University of Pittsburgh Medical Center, 200 Lothrop St, Eye and Ear Institute 500, Pittsburgh, PA 15213 (duvvuriu@upmc.edu).

Large shifts in treatment recommendations for oropharyngeal squamous cell carcinoma (OPSCC) have occurred over the past 3 decades resulting from technological advances in all treatment modalities. The use of primary chemoradiotherapy (CRT) for OPSCC doubled between 1985 and 2001, and use of primary radiotherapy (RT) and primary surgical therapy decreased.[1] However, acute and late tissue toxic effects are a limiting factor for treatment success with RT and CRT. Common adverse effects include mucositis, xerostomia, dysgeusia, and increased risk of oral infections, all of which impair posttreatment quality of life (QOL).[2]

Over the past decade, the use of transoral robotic surgery (TORS) as a treatment option for OPSCC has been increasing. Multiple studies[3-5] have demonstrated that TORS, with or without adjuvant therapy, offers excellent long-term oncologic and survival outcomes. The use of TORS has been associated[6] with decreased length of hospitalization, tracheostomy tube requirement during treatment, and permanent gastrostomy tube requirement. Faster postoperative recovery after TORS may decrease treatment duration and toxic effects associated with adjuvant RT and CRT.[7] Even so, patients who undergo TORS followed by adjuvant therapy appear to score lower on QOL indexes compared with those who receive TORS alone up to 1 year after treatment, especially in the swallowing and diet domains.[8-10]Overall, few studies have examined QOL outcomes in patients who undergo TORS alone.

Herein, we report our single-institutional experience with the use of TORS alone for patients with early-stage OPSCC and describe patient-reported QOL outcomes up to 2 years after treatment. We hypothesized that, for select patients with low-risk features, TORS alone would be an effective treatment algorithm that allows for acceptable short- and long-term QOL outcomes in the absence of adjuvant therapy.

Methods

Patient Selection

This retrospective review of medical records was conducted at the University of Pittsburgh Medical Center, a tertiary referral center. Surgical scheduling records were reviewed from June through September 2014 to identify all patients who underwent TORS between May 1, 2010, and March 31, 2014. In total, 172 patients received TORS for oncologic resection during that time. Thirty-four patients met the criteria for inclusion. All patients underwent TORS as the primary treatment modality for OPSCC. At our institution, adjuvant therapy following TORS is generally not recommended if patients lack adverse prognostic pathologic features, such as extracapsular spread, multiple involved lymph nodes, perineural invasion, or positive or close margins. Few patients (11) in the present study were recommended to receive adjuvant therapy following TORS for high-risk pathologic features but refused. We excluded patients who received any postoperative adjuvant therapy including RT or CRT, those who received TORS for an unknown primary tumor or salvage purposes, and those with a primary tumor site other than the oropharynx. Demographic data (ie, age, sex, race, alcohol use, and smoking status), rates of tracheostomy and gastrostomy tube insertion, and oncologic data (ie, tumor markers, tumor staging, extracapsular spread, tumor grade, surgical margin status, histologic characteristics, and tumor recurrence) were collected.

Approval for the study was obtained from the University of Pittsburgh Medical Center Office of Quality and Research. The requirement for informed consent was waived and the data were deidentified.

QOL Assessment

The University of Washington Quality of Life (UW-QOL), version 4, questionnaire is a previously validated 12-item survey that scores pain, appearance, activity, recreation, swallowing, chewing, speech, shoulder function, taste, saliva, mood, and anxiety.[11,12] The survey also includes 3 global QOL scores. Scores for each domain range from 0 to 100, with 100 being the best functional outcome reported by the patient. El-Deiry et al[13] demonstrated that a 7-point difference in the score on this scale is sensitive to predict for a clinically meaningful difference in QOL.

The UW-QOL questionnaires were routinely completed by patients during clinic visits preoperatively and at 1-month (±1 month), 6-month (±2 months), 12-month (±3 months), and 24-month (±3 months) postoperative visits from the date of TORS (followed up through April 30, 2014). Surveys were pooled by time from TORS into 4 categories (1, 6, 12, and 24 months after surgery) for analysis.

Statistical Analysis

Demographic and clinical oncologic data were summarized with proportions for categorical data and with means (SDs) for continuous data. Medians and interquartile ranges were used to summarize the UW-QOL survey scores. The overall distribution of the UW-QOL scores at 1 month after surgery was compared with that of each subsequent QOL time point with the Wilcoxon Mann-Whitney test. Overall trends in QOL scores over time were assessed with simple linear regression. Individual statistical tests were not adjusted for multiple comparisons. All reported P values are 2-sided, and significance was set at $P < .05$. Statistical analyses were performed using SAS/STAT, version 9.4 (SAS Institute Inc) and R, version 3.0.1 (R Foundation for Statistical Computing).

Results

Patient and Disease Characteristics

A total of 34 patients were included in this analysis. Patient characteristics are reported in **Table 1**. The most common oropharyngeal primary subsite was the tonsil (16 patients [47%]), followed by the base of tongue (15 [44%]). Cancer in most patients was category T1 (20 [59%]) or T2 (13 [38%]) and category N0 (13 [38%]) or N1 (16 [47%]). One patient (3%) had a positive margin, 4 patients (12%) had confirmed nodal extracapsular spread, and 4 individuals (12%) had perineural invasion. A synopsis of disease data can be found in Table 1. Advanced oncologic data analysis from this patient cohort will be included in an upcoming multi-institutional report (not included here to prevent reporting duplication of data).

Table 1. Characteristics of the Study Population

Characteristic	No. (%)
Patient	
Male sex	26 (76)
Age, mean (SD), y	59 (8)
Race	
White	32 (94)
African American	2 (6)
History	
Smoking	24 (70)
Alcohol use[a]	20 (59)
Disease	
Primary site	
Tonsil	16 (47)
Tongue base	15 (44)
Soft palate	2 (6)
Pharyngeal wall	1 (3)
Extracapsular spread	
Yes	4 (12)
No	15 (44)
Not evaluated	15 (44)
T category	
T1	20 (59)
T2	13 (38)
T3	1 (3)
N category	
N0	13 (38)
N1	16 (47)
N2a	3 (9)
N2b	2 (6)
p16 Status	
Positive	25 (74)
Negative	8 (24)
Not evaluated	1 (3)
Perineural invasion	4 (12)
Positive margins	1 (3)

[a] *History of alcohol use* was defined as any "regular use of alcohol" on the self-reported University of Washington Quality of Life, version 4, survey.

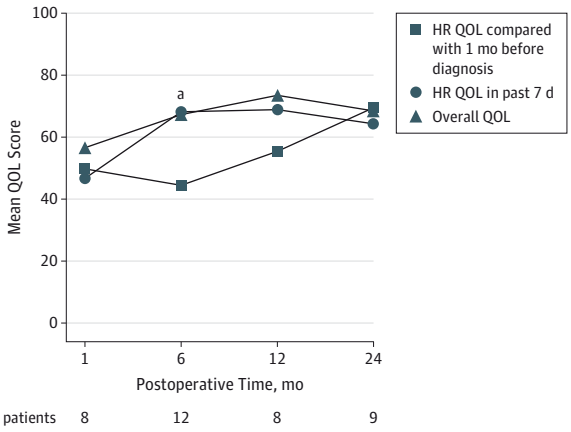
Follow-up

The follow-up period for overall survival was defined as the number of months from the date of TORS to the date of the last follow-up determined by clinic visit, telephone survey, or death. Mean follow-up for this cohort was 14 months (range, 13 days to 38 months; from May 1, 2010, to April 30, 2014). Two patients (6%) died during the follow-up period: 1 due to disease and 1 due to a myocardial infarction. There were no intraoperative complications. Two patients (6%) required temporary gastrostomy tube placement, but no patients required tracheostomy. Among all the completed UW-QOL forms, 4 forms were completed preoperatively, 8 at 1 month after surgery, 12 at 6 months, 8 at 12 months, and 9 at 24 months.

Quality of Life

The scores for the 3 global QOL survey questions ("health-related QOL compared to 1 month before cancer," "health-related QOL during the past 7 days," and "overall QOL including personal well-being over the past 7 days") showed a tendency to improve throughout follow-up (**Figure 1**). One interval reached statistically significant improvement ("health-related QOL during the past 7 days" 6 months after surgery) (Figure 1B and **Table 2**); improvements were observed in several other domains, although these were not statistically significant (Figure 1 and Table 2) compared with 1-month follow-up scores.

Figure 1. Trends in 3 Global Quality-of-Life (QOL) Scale Scores Across 24 Months



Health-related (HR) QOL compared with 1 month before cancer diagnosis, during the past 7 days ($P = .01$ at 6 months), and overall QOL, including personal well-being, during the past 7 days.

[a] $P < .05$ compared with 1 month after surgery.

Scores for the QOL domains of pain, swallowing, activity, and chewing also tended to improve throughout follow-up (**Figure 2**). Statistically significant improvement in chewing scores was noted from 1 to 12 months after surgery ($P = .048$) (Figure 2B). A positive trend was observed for chewing scores over time ($P = .05$). Pain scores improved from 1 to 6 months ($P = .006$) and 12 months ($P = .01$) after surgery (Figure 2C). However, there was no evidence that the median pain score continued to improve over time ($P = .10$). Swallowing scores improved from 1 to 6 months ($P = .047$) and 24 months ($P = .048$) after surgery (Figure 2D). There was an overall positive trend in swallowing scores ($P = .01$). In addition, the median activity score improved over time ($P = .03$) (Figure 2A). No other specific symptom domains showed statistical evidence of improvement or deterioration from 1 month after surgery over time (**Table 3**).

Discussion

Increasing recognition of the adverse effects of CRT and their negative effect on QOL has provided the rationale for TORS as a primary treatment modality option for OPSCC. The present study is especially timely in the current era of human papilloma virus–positive OPSCC, with younger and healthier patients seeking treatment modalities with less long-term treatment-related morbidity. There is, however, a paucity of literature describing the long-term QOL of patients who receive TORS

Table 2. Global QOL Domains

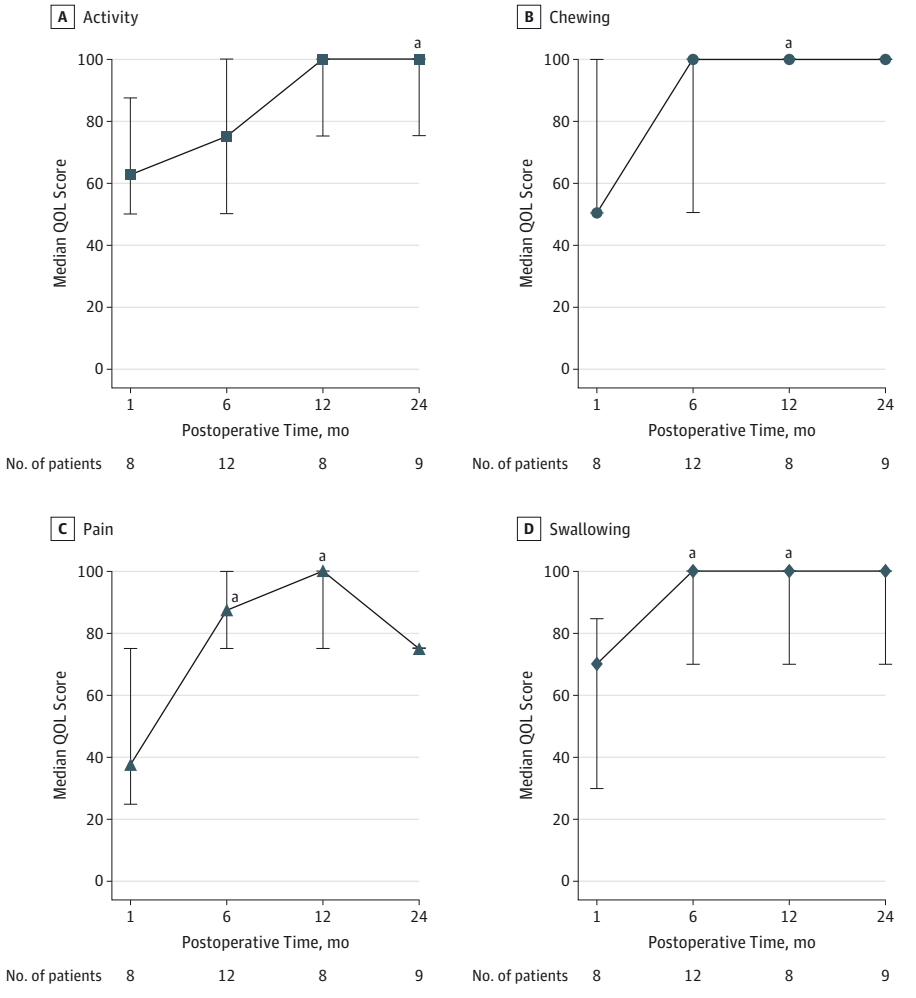
QOL Question	Postoperative Months, Median (IQR)[a]			
	1	6	12	24
Patients, No. (%)	8	12	8	9
Health-related QOL vs 1 mo before cancer	50 (50-50)	50 (25-75)	50 (50-75)	75 (50-100)
P value		>.99	.62	.27
Health-related QOL during the past 7 d	40 (40-60)	80 (60-100)	70 (50-100)	60 (40-80)
P value		.01[b]	.12	.22
Overall QOL during the past 7 d	50 (40-80)	80 (60-80)	80 (50-100)	60 (60-80)
P value		.12	.18	.33

Abbreviations: IQR, interquartile range; QOL, quality of life.

[a] Quality-of-life scores were compared with QOL scores at 1 month after baseline using the Wilcoxon Mann-Whitney test. No adjustments were made for multiple testing.

[b] Statistically significant at $P < .05$.

Figure 2. Trends in Symptom-Specific Quality of Life (QOL) Domains



Changes in median scores for activity (A), chewing (B), pain (C), and swallowing (D). Error bars indicate the interquartile range.

[a] $P < .05$ compared with 1 month after surgery.

alone. To our knowledge, this is the largest study with the longest follow-up period investigating QOL in patients who receive only TORS.

Our study suggests that selected patients with OPSCC treated with TORS alone experience continued improvement in QOL in multiple domains soon after surgery, as well as in the long term. Statistically significant improvements were noted when compared with QOL 1 month after surgery in the following domains: swallowing and pain at 6 months, chewing and pain at 12 months, and activity and swallowing at 24 months (Table 3). No domain demonstrated decreases of QOL that were statistically significant at any time. These findings are in contrast to those of previous studies[10] showing that patients who received adjuvant RT or CRT experienced deterioration in QOL scores to a nadir at approximately 3 months after TORS. Although it is possible that patients have not recovered completely from surgery at the start of adjuvant therapy in these previous studies, it has been suggested[9,10,14,15] that this lack of improvement could be secondary to substantial adjuvant treatment-related toxic effects.

Table 3. Symptom-Specific QOL Domains

QOL Domain	Postsurgery QOL Score, Median (IQR)[a]			
	1 Month	6 Months	12 Months	24 Months
Patients, No.	8	12	8	9
Activity	63 (50-88)	75 (50-100)	100 (75-100)	100 (75-100)
P value		.43	.10	.03[b]
Anxiety	70 (30-70)	70 (70-100)	85 (50-100)	70 (70-100)
P value		.19	.37	.33
Appearance	88 (75-100)	100 (75-100)	100 (75-100)	100 (75-100)
P value		.35	.67	>.99
Chewing	50 (50-100)	100 (50-100)	100 (100-100)	100 (100-100)
P value		.40	<.05[b]	.11
Mood	75 (75-100)	75 (50-100)	100 (75-100)	75 (75-100)
P value		.66	.45	.83
Pain	38 (25-75)	88 (75-100)	100 (75-100)	75 (75-75)
P value		<.01[b]	.01[b]	.06
Recreation	75 (63-100)	88 (75-100)	100 (88-100)	100 (75-100)
P value		.45	.20	.40
Saliva	85 (70-100)	100 (70-100)	100 (70-100)	70 (70-100)
P value		.85	.55	.75
Shoulder function	85 (70-100)	85 (30-100)	100 (85-100)	100 (30-100)
P value		.71	.43	.67
Speech	100 (85-100)	100 (70-100)	100 (85-100)	100 (100-100)
P value		.25	>.99	.51
Swallowing	70 (30-85)	100 (70-100)	100 (70-100)	100 (70-100)
P value		.05[b]	.07	.05[b]
Taste	100 (50-100)	70 (70-85)	100 (70-100)	100 (70-100)
P value		.43	.86	>.99

Abbreviations: IQR, interquartile range; QOL, quality of life.

[a] Quality-of-life scores were compared with QOL scores at 1 month after baseline using the Wilcoxon Mann-Whitney test. No adjustments were made for multiple testing.

[b] Statistically significant at $P < .05$.

It is especially notable in our study that speech function was minimally affected 1 month after surgery, and patients were able to maintain similar levels of function throughout follow-up. A similar result was reported for patients with OPSCC treated with TORS alone by Leonhardt et al,[9] although with smaller numbers (N = 9). This minimal effect on speech only in patients who underwent surgery is not surprising since studies[9,10] have shown that adjuvant RT is significantly correlated with lower speech function and speech attitude scores at 12 months following TORS.

To our knowledge, the present study is the first to report a statistically significant improvement in pain in the short term with lasting long-term relief among patients who undergo TORS without adjuvant therapy. Pain scores at 1 month were initially low (mean score, 47), but they improved at 6 months (mean score, 83) ($P = .006$) and remained stable at 12 months. This finding is in contrast to that in patients who received adjuvant therapy after TORS and experienced a significant deterioration at 6 months in the bodily pain domain of the Short Form 8 Health Survey.[9] It appears that the addition of RT or CRT following TORS hampers recovery from pain associated with surgery, but TORS alone is associated with short-term pain and good long-term recovery.

Similarly, patients reported relatively low scores in chewing and swallowing at 1 month following surgery (median score, 50 and 70, respectively). This difficulty was followed by statistically significant recovery to a higher level of function with long-term follow-up (chewing: $P = .048$ at 12 months; swallowing: $P = .047$ at 6 months and stable at 24 months; $P = .048$), confirming a previous finding in a small number of patients receiving TORS alone.[9] This recovery is not unexpected; previous studies[8-11,16,17] suggested that RT and CRT cause substantial deterioration in short-term and long-term patient-perceived swallowing function, with slow recovery.

Our study included 2 patients (6%) who had recurrences, both of whom did not adhere to recommendations for adjuvant therapy. These patients demonstrated high-risk features after TORS (extracapsular spread, positive margin, or perineural invasion); adjuvant therapy was recommended, but the patients declined. At a 2-year follow-up, 1 patient demonstrated regional failure, and 1 had both local and regional failure. None of the 34 patients experienced distant metastasis or failure in the retropharyngeal nodal basin. The patient with local and regional failure showed a sharp decrease of QOL score, but the other patient with regional failure maintained a high QOL score at the time of recurrence. Overall, excluding these 2 patients did not affect the statistical significance of QOL scores found in our original analysis.

Our patients had a good rate of survival throughout the 2-year follow-up period. Based on scores for the 2 global health-related QOL items, patients experienced a trend toward increasing health-related QOL during 2 postoperative years. At the 6-month follow-up evaluation, significant improvement in health-related QOL over the past 7 days was recognized com-

pared with 1 month after surgery (P = .01). No significant decline in QOL was noted at any time during the follow-up period. None of the patients required tracheostomy, and only 2 patients required transient gastrostomy tube at any time point.

Our study has limitations. Although this cohort included 34 patients, fewer individual patients (8-12 patients per time point) provided UW-QOL responses at each postoperative visit. Because of this small cohort and the large number of comparisons, there exists the possibility that some of the statistical significance that was achieved could have been by chance. Similarly, the QOL scores were compared in a pooled fashion and not on an individual basis. There was also no comparison arm for patients who received adjuvant CRT after TORS, which would have allowed direct evaluation of the effect of adjuvant therapy on QOL in patients who undergo TORS. The patients included in our study had early T category (category T1-T2; 97%), light nodal burden (category N0-N1; 85%) and few high-risk features (12%, extracapsular spread; 3%, positive margin; and 12%, perineural invasion). Although this cohort was comparable to that reported in a previous review of patients who underwent only TORS,[18] it should be noted that our patients had a much smaller percentage of T3/T4 tumors and N3 disease compared with previously reported CRT series.[18,19]

Conclusions

Optimizing posttreatment QOL for patients with head and neck cancer is important in early T-stage disease with good prognosis. Our study suggests that appropriately selected patients who undergo TORS alone for OPSCC experience acceptable short- and long-term QOL outcomes.

ARTICLE INFORMATION

Submitted for Publication: October 2, 2014; final revision received January 27, 2015; accepted February 15, 2015.

Published Online: April 2, 2015. doi:10.1001/jamaoto.2015.0347.

Author Contributions: Drs Choby and Duvvuri 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.
Study concept and design: Choby, Ferris, Duvvuri.
Acquisition, analysis, or interpretation of data: J. Kim, Ling, Abberbock, Mandal, S. Kim, Ferris.
Drafting of the manuscript: Choby, J. Kim, Ling, Abberbock, Mandal, Ferris.
Critical revision of the manuscript for important intellectual content: Choby, J. Kim, Ling, Abberbock, Mandal, S. Kim, Duvvuri.
Statistical analysis: Choby, J. Kim, Abberbock, Ferris.
Obtained funding: Ferris, Duvvuri.
Administrative, technical, or material support: Mandal, S. Kim, Ferris, Duvvuri.
Study supervision: Mandal, Ferris, Duvvuri.

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was funded in part by the Department of Veterans Affairs Career Development Award and the PNC Foundation (Dr Duvvuri).

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: This manuscript does not represent the views of the US government or the Department of Veterans Affairs.

REFERENCES

1. Chen AY, Schrag N, Hao Y, Stewart A, Ward E. Changes in treatment of advanced oropharyngeal cancer, 1985-2001. *Laryngoscope*. 2007;117(1):16-21.

2. Logan RM. Advances in understanding of toxicities of treatment for head and neck cancer. *Oral Oncol*. 2009;45(10):844-848.

3. Moore EJ, Olsen SM, Laborde RR, et al. Long-term functional and oncologic results of transoral robotic surgery for oropharyngeal squamous cell carcinoma. *Mayo Clin Proc*. 2012;87(3):219-225.

4. White HN, Moore EJ, Rosenthal EL, et al. Transoral robotic-assisted surgery for head and neck squamous cell carcinoma: one- and 2-year survival analysis. *Arch Otolaryngol Head Neck Surg*. 2010;136(12):1248-1252.

5. Weinstein GS, Quon H, Newman HJ, et al. Transoral robotic surgery alone for oropharyngeal cancer: an analysis of local control. *Arch Otolaryngol Head Neck Surg*. 2012;138(7):628-634.

6. Moore EJ, Hinni ML. Transoral laser microsurgery and robotic-assisted surgery for oropharynx cancer including human papillomavirus-related cancer. *Int J Radiat Oncol Biol Phys*. 2013;85(5):1163-1167.

7. Carpenter TJ, Kann B, Buckstein MH, et al. Tolerability, toxicity, and temporal implications of transoral robotic surgery (TORS) on adjuvant radiation therapy in carcinoma of the head and neck. *Ann Otol Rhinol Laryngol*. 2014;123(11):791-797.

8. Hurtuk AM, Marcinow A, Agrawal A, Old M, Teknos TN, Ozer E. Quality-of-life outcomes in transoral robotic surgery. *Otolaryngol Head Neck Surg*. 2012;146(1):68-73.

9. Leonhardt FD, Quon H, Abrahão M, O'Malley BW Jr, Weinstein GS. Transoral robotic surgery for oropharyngeal carcinoma and its impact on patient-reported quality of life and function. *Head Neck*. 2012;34(2):146-154.

10. Dziegielewski PT, Teknos TN, Durmus K, et al. Transoral robotic surgery for oropharyngeal cancer: long-term quality of life and functional outcomes. *JAMA Otolaryngol Head Neck Surg*. 2013;139(11):1099-1108.

11. Durmus K, Patwa HS, Gokozan HN, et al. Functional and quality-of-life outcomes of transoral robotic surgery for carcinoma of unknown primary. *Laryngoscope*. 2014;124(9):2089-2095.

12. Rogers SN, Gwanne S, Lowe D, Humphris G, Yueh B, Weymuller EA Jr. The addition of mood and anxiety domains to the University of Washington Quality of Life scale. *Head Neck*. 2002;24(6):521-529.

13. El-Deiry MW, Futran ND, McDowell JA, Weymuller EA Jr, Yueh B. Influences and predictors of long-term quality of life in head and neck cancer survivors. *Arch Otolaryngol Head Neck Surg*. 2009;135(4):380-384.

14. Li Y, Taylor JM, Ten Haken RK, Eisbruch A. The impact of dose on parotid salivary recovery in head and neck cancer patients treated with radiation therapy. *Int J Radiat Oncol Biol Phys*. 2007;67(3):660-669.

15. Logemann JA, Pauloski BR, Rademaker AW, et al. Swallowing disorders in the first year after radiation and chemoradiation. *Head Neck*. 2008;30(2):148-158.

16. Sinclair CF, McColloch NL, Carroll WR, Rosenthal EL, Desmond RA, Magnuson JS. Patient-perceived and objective functional outcomes following transoral robotic surgery for early oropharyngeal carcinoma. *Arch Otolaryngol Head Neck Surg*. 2011;137(11):1112-1116.

17. Tschudi D, Stoeckli S, Schmid S. Quality of life after different treatment modalities for carcinoma of the oropharynx. *Laryngoscope*. 2003;113(11):1949-1954.

18. Dowthwaite SA, Franklin JH, Palma DA, Fung K, Yoo J, Nichols AC. The role of transoral robotic surgery in the management of oropharyngeal cancer: a review of the literature. *ISRN Oncol*. 2012;2012:945162.

19. Calais G, Alfonsi M, Bardet E, et al. Randomized trial of radiation therapy versus concomitant chemotherapy and radiation therapy for advanced-stage oropharynx carcinoma. *J Natl Cancer Inst*. 1999;91(24):2081-2086.

ORIGINAL ARTICLE

Surgical management of oropharyngeal squamous cell carcinoma: Survival and functional outcomes

Bhavna Kumar, MS, Michael J. Cipolla, MD, Matthew O. Old, MD, Nicole V. Brown, MS, Stephen Y. Kang, MD, Peter T. Dziegielewski, MD, Kasim Durmus, MD, Enver Ozer, MD, Amit Agrawal, MD, Ricardo L. Carrau, MD, David E. Schuller, MD, Marino E. Leon, MD, Quintin Pan, PhD, Pawan Kumar, PhD, Valerie Wood, MD, Jessica Burgers, MD, Paul E. Wakely Jr, MD, Theodoros N. Teknos, MD*

Department of Otolaryngology – Head and Neck Surgery, The James Cancer Hospital and Solove Research Institute, The Ohio State University Wexner Medical Center, Columbus, Ohio.

Accepted 19 September 2015
Published online 23 December 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/hed.24319

ABSTRACT: *Background.* The purpose of this study was to further define the impact of primary surgery in the management of oropharyngeal squamous cell carcinoma (SCC).

Methods. Two hundred ninety-six patients with oropharyngeal SCC treated with primary surgery were included. Multivariable analysis and recursive partitioning analysis (RPA) identified predictors of survival and gastrostomy tube presence.

Results. Multivariable analysis identified that HPV negativity ($p = .0002$), presence of extranodal extension ($p = .0025$), and advanced T classification ($p = .0081$) were independent predictors of survival. For HPV-positive patients, surgical approach ($p = .0111$) and margin status ($p = .0287$) were significant predictors of survival. For HPV-negative patients, extranodal extension ($p = .0021$) and advanced T classification ($p = .0342$) were significant predictors of survival. Smoking status and advanced neck disease did not impact survival, and the addition of adjuvant chemotherapy did not confer survival benefit in HPV-positive or HPV-negative subgroups.

Conclusion. Independent predictors of survival are unique in patients with oropharyngeal SCC treated with primary surgery. © 2015 Wiley Periodicals, Inc. *Head Neck* **38**: E1794–E1802, 2016

KEY WORDS: surgery oropharynx, oropharyngeal cancer, human papillomavirus, squamous cell carcinoma, transoral surgery

INTRODUCTION

The worldwide incidence of oropharyngeal squamous cell carcinoma (SCC) is rising at an alarming rate.[1,2] Once a rare disease, oropharyngeal SCC is now the most common malignancy encountered by the head and neck oncologist.[1,3] This dramatic shift in tumor incidence has been linked to increasing rates of infection with the carcinogenic strains of human papillomavirus (HPV).[4] Traditionally, head and neck malignancies, including oropharyngeal SCC, have been treated with open surgical resection, reconstruction, and postoperative radiotherapy.[5,6] However, after the publication of the Veterans Affairs Laryngeal Cancer Study Group trial in 1991, there has been an increased emphasis on nonsurgical approaches to therapy.[7,8] Specifically with regard to oropharyngeal SCC, a meta-analysis by Parsons et al[5] noted similar survival outcomes in patients treated with surgery followed by radiotherapy as those treated with primary radiotherapy and surgical salvage. Furthermore, because of functional and cosmetic morbidity associated with conventional open en bloc resections in oropharyngeal SCC, "organ preservation" approaches began to be explored.[9–14] In time, radiotherapy alone was supplanted by concurrent chemoradiotherapy because of improved primary tumor control.[10,15] Novel and ever-intensifying chemotherapeutic approaches were also investigated in oropharyngeal SCC.[16–20] However, with the proliferation of "organ preservation" approaches to oropharyngeal SCC, dramatic increases in the rates of treatment-related toxicities have been documented.[20–23] There have been notable increases in the rates of xerostomia (33%), gastrostomy tube dependence (12%), cervical stricture (6%), and osteoradionecrosis, even with the use of the latest radiation techniques.[24] Published gastrostomy tube rates have ranged from 7% to as high as 31% at 1 year after chemoradiotherapy.[25,26]

In a landmark publication, Ang et al[15] retrospectively reviewed patients with oropharyngeal SCC enrolled in Radiation Therapy Oncology Group (RTOG) 0129, comparing high-dose cisplatin given concurrently with either

***Corresponding author:** T. N. Teknos, Department of Otolaryngology – Head and Neck Surgery, The James Cancer Hospital and Solove Research Institute, The Ohio State University Wexner Medical Center, 915 Olentangy River Road, Suite 4000, Columbus OH 43212. E-mail: ted.teknos@osumc.edu

Contract grant sponsor: This work was supported by The Ohio State University Comprehensive Cancer Center funds.

This work was presented as an Abstract at the American Head and Neck Society Annual Meeting, Orlando, Florida, April 10–11, 2013; and it was also presented in part at the 2014 Multidisciplinary Head and Neck Cancer Symposium, American Society of Clinical Oncology (ASCO)/American Society for Radiation Oncology (ASTRO)/American Head and Neck Society (AHNS), Scottsdale, Arizona, February 20–22, 2014.

standard fraction or accelerated fraction radiotherapy. The authors provided strong evidence that HPV-related oropharyngeal SCC is a unique disease entity with improved survival outcomes. They also classified patients with oropharyngeal SCC as having either a low, intermediate, or high-risk of death dependent upon HPV status, smoking history, neck disease, and primary tumor classification (all of which were independent predictors of survival). The 3-year survival rates were 93.0%, 70.8%, and 46.2%, respectively. The data from RTOG 0129 suggests that less intense therapy may be warranted for the low-risk group of patients and more intense therapy may be needed for the high-risk group. The intermediate group of patients, which includes 36% of all HPV-positive patients, should not be deintensified but clearly have poorer overall survival outcomes than the low-risk group because of their smoking status and extensive neck disease.

Nonsurgically treated patient survival and functional outcomes for oropharyngeal SCC have been well-documented; however, outcomes data for patients treated with primary surgery is sparse. Since the reports by Parsons et al,[5,6] surgical and reconstructive technology has dramatically advanced and improved functional outcomes. The use of transoral laser microsurgery and transoral robotic surgery (TORS) have allowed surgeons to access tumors without disrupting normal anatomy, while providing superior visualization of tumor margins.[27–32] Furthermore, when open procedures are necessary, free flap surgeons are able to provide superior cosmetic and functional outcomes. In addition, surgical resection may be a way to deintensify therapy for patients in the lowest or intermediate risk categories by obviating the need for concurrent chemoradiotherapy.

The purpose of this study was to determine the impact of primary surgery in the treatment of oropharyngeal SCC. Predictors of survival will be determined and functional outcomes will be reported. In addition, survival and functional outcomes will be compared between open surgery and transoral surgical approaches for tumor extirpation.

MATERIALS AND METHODS

After institutional review board approval, a prospective database of patients with head and neck cancer treated with primary surgery was assembled and continually maintained. This study was retrospective in nature and exempt from consent. The database was searched for patients with oropharyngeal SCC treated from January 1, 2002, to August 31, 2012. Patients who were treated with primary surgery for histologically confirmed SCC were assessed for eligibility. Patients who were previously untreated, had tissue available for analysis, and had available clinical follow-up data were included in the study. For each patient, demographic data, complete medical history, pathology, and follow-up were recorded and verified in real time. Survival data was ascertained through medical record review and confirmed through tumor registry files and the Social Security Death Index data. The type of surgery performed was recorded and classified as: (a) TORS; (b) transoral nonrobotic; (c) open transcervical (ie, suprahyoid pharyngotomy, lateral pharyngotomy); (d) mandibulotomy; or (e) composite resection (ie, mandibulectomy with pharyngectomy and/or base of tongue removal). For analysis purposes, this was condensed into groups of transoral (a or b) and open (c, d, or e) surgical approaches. All patients underwent neck dissections at the time of their primary resections, according to therapeutic guidelines.[33] Most patients were treated with adjuvant radiotherapy or adjuvant chemoradiotherapy based on standard National Comprehensive Cancer Network guidelines, which notably changed over time. After the publication by Bernier et al,[34] most high-risk patients were treated with postoperative concurrent chemoradiotherapy, whereas patients before 2004 were treated with radiation alone.

Multiple studies have shown gastrostomy tube dependence to be a major negative predictor of quality of life in the head and neck cancer population.[35–37] In this study, gastrostomy tube presence was assessed and defined as the presence of a gastrostomy tube that was used for at least a portion of the diet. Gastrostomy tube presence was assessed at 0, 6, and 12 months postsurgery, and at last follow-up visit.

A high density tissue microarray was created with representative samples from patients in the study.[38] Tumor p16 expression was evaluated by means of immunohistochemical staining using a mouse monoclonal antibody (MTM Laboratories CINTEC, Westborough, MA) and visualized with a Ventana XT autostainer (Ventana Medical Systems, Tucson, AZ). Positive p16 expression was defined as diffuse nuclear and cytoplasmic staining in 50% or more of the tumor cells. All samples were further evaluated for HPV positivity via in situ hybridization for HPV16 (GenPoint HPV DNA Probe) or for high-risk HPV genotypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 66 (INFORM HPV III Family 16 Probe (B); Ventana Medical Systems). Any definitive nuclear staining in the tumor cells was considered positive. Low-, intermediate-, and high-risk patients were defined as in the article by Ang et al.[15]

Statistical methods

Overall survival was defined as the time from the date of surgery to the date of death, with patients alive at the date of the last observation censored. Cox proportional hazards models were used to assess univariate associations of biomarkers as predictors for death. Unadjusted hazard ratios (HRs) and 95% confidence intervals (CIs) are reported. Predictors of both clinical and statistical significance were evaluated in multivariable models for the entire cohort and within the HPV-negative and HPV-positive subgroups. These predictors included: surgical approach (transoral vs open), HPV status (positive vs negative), extranodal extension (no vs yes), mucosal margins (free of carcinoma vs positive), perineural invasion (no vs yes), smoking status ($\leq$10 pack-years vs $>$10 pack-years), and tumor classification (T1/T2 vs T3/T4).

To profile the risk of death, a recursive partitioning analysis (RPA) was used as an exploratory analysis. The "rpart" library in the R package was used to fit a regression tree with the overall survival data.[39,40] In evaluating prognostic factors for overall survival, the predictors used in the multivariable Cox proportional hazards model were included in the building of the tree. All analyses were

TABLE 1. Patient characteristics.

Characteristics	No. of patients (%)
Age, y, mean (SD)	57.8 (9.2)
Marital status	
Single/divorced/widowed	115 (41.8)
Married	160 (58.2)
Race	
African American/black	12 (4.1)
White	281 (95.9)
Sex	
Female	61 (20.6)
Male	235 (79.4)
Smoking status	
≤10 pack-years	82 (29.0)
>10 pack-years	201 (71.0)
Approach	
TORS	84 (28.4)
Transoral, nonrobotic	90 (30.4)
Transcervical (eg, suprahyoid)	37 (12.5)
Mandibulotomy	50 (16.9)
Composite resection	35 (11.8)
ECS	
No	175 (61.2)
Yes	111 (38.8)
HPV status	
Negative	117 (40.5)
Positive	172 (59.5)
Mucosal margins	
Free of carcinoma	239 (81.6)
Positive margins	54 (18.4)
N classification	
N0	38 (12.9)
N1	57 (19.4)
N2	186 (63.3)
N3	13 (4.4)
p16 status	
Negative	66 (22.9)
Positive	222 (77.1)
Perineural invasion	
No	209 (71.8)
Yes	82 (28.2)
Primary site	
Base of tongue	74 (25.1)
Other	23 (7.8)
Tonsil	198 (67.1)
Rick group, HPV	
Low	106 (38.4)
Intermediate	69 (25.0)
High	101 (36.6)
Rick group, p16	
Low	129 (46.9)
Intermediate	85 (30.9)
High	61 (22.8)
TNM stage	
I	11 (3.7)
II	12 (4.1)
III	62 (21.1)
IV	209 (71.1)
T classification	
T1	77 (26.0)
T2	126 (42.6)
T3	48 (16.2)
T4	45 (15.2)

Abbreviations: TORS, transoral robotic surgery; ECS, extracapsular extension; HPV, human papillomavirus.

conducted in SAS version 9.3 (SAS Institute, Cary, NC) or the R language environment for statistical computing (R version 3.1.0).

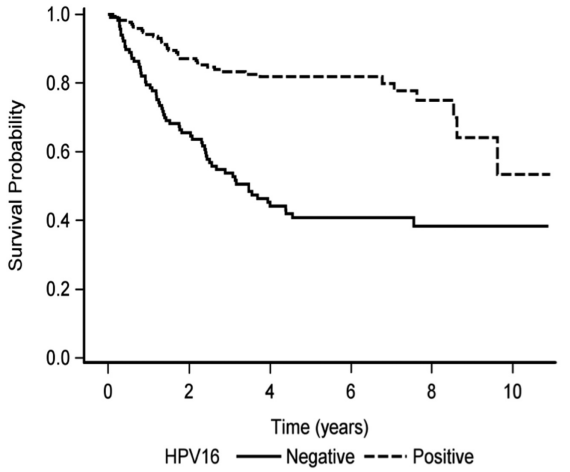
RESULTS

From January 1, 2002, to August 31, 2012, a total of 296 previously untreated patients with oropharyngeal SCC underwent surgical resection for curative intent at the Arthur G. James Cancer Hospital and Richard J. Solove Research Institute. Patient demographic and clinical characteristics are summarized in Table 1.

The vast majority of patients were p16-positive (77.1%) and 168 patients (58.5%) were confirmed to be both p16-positive and HPV-positive. Fifty-three patients were p16-positive/HPV-negative by in situ hybridization and 3 were p16-negative/HPV-positive. With regard to surgical approach, 58.8% underwent transoral and 41.2% had open surgical approaches for tumor extirpation.

Overall survival

Overall 1-, 3-, and 5-year survival rates for all patients, regardless of HPV status, were 88.5%, 71.3%, and 65.1%, respectively. HPV-positive patients had better overall survival than HPV-negative patients (log-rank p value $<$ .0001; see Figure 1). HPV-positive patients had overall survival rates of 94.2%, 83.3%, and 81.8% at 1, 3, and 5 years compared to 79.5%, 53.3%, and 40.3% for HPV-negative patients.



	Survival Rates			
	1 year (%)	3 year (%)	5 year (%)	10 year (%)
Overall	88.5	71.3	65.1	52.2
HPV Negative	79.5	53.3	40.3	37.9
HPV Positive	94.2	83.3	81.8	58.3

Abbreviations: HPV, human papillomavirus.

FIGURE 1. Kaplan–Meier estimates of survival according to human papillomavirus (HPV) status. Patients with HPV-positive tumors had significantly better overall survival than patients with HPV-negative tumors ($p < .0001$).

TABLE 2. Univariate analysis of outcome predictors.

	Overall				HPV negative				HPV positive			
Variables	HR	95% CI	*p* value	No. of patients	HR	95% CI	*p* value	No. of patients	HR	95% CI	*p* value	No. of patients
Age, y	1.027	1.006–1.05	.0129	295	1.023	0.997–1.050	.0871	117	1.011	0.974–1.050	.5602	171
Approach												
Transoral	Ref				Ref				Ref			
Open	2.732	1.822–4.096	< .0001	296	1.449	0.865–2.426	.1589	117	3.899	1.964–7.742	.0001	172
Extranodal extension												
No	Ref				Ref				Ref			
Yes	2.364	1.582–3.532	< .0001	286	2.396	1.432–4.010	.0009	111	1.921	0.985–3.744	.0553	169
HPV												
Positive	Ref											
Negative	3.393	2.255–5.106	< .0001	289	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Marital status												
Married	Ref				Ref				Ref			
Single/divorced/widowed	1.625	1.064–2.481	.0247	275	1.671	0.963–2.898	.0680	104	1.203	0.595–2.433	.6067	164
Mucosal margins												
Free of carcinoma	Ref				Ref				Ref			
Positive	1.591	1.005–2.517	.0473	293	1.317	0.727–2.387	.3640	116	1.775	0.828–3.806	.1401	170
N classification												
N0/N1	Ref				Ref				Ref			
N2/N3	1.059	0.701–1.600	.7849	294	1.385	0.836–2.294	.2064	117	1.057	0.508–2.200	.8817	171
p16 status												
Positive	Ref											
Negative	2.983	1.999–4.453	< .0001	288	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Perineural invasion												
No	Ref				Ref				Ref			
Yes	2.453	1.653–3.642	< .0001	291	1.806	1.102–2.958	.0190	116	1.884	0.920–3.858	.0831	168
Primary site												
BOT	Ref				Ref				Ref			
Other	2.417	1.233–4.739	.0102		2.836	1.323–6.080	.0074		0.561	0.072–4.403	.5828	
Tonsil	0.963	0.605–1.533	.8739	295	1.798	0.987–3.274	.0550	117	0.584	0.279–1.221	.1529	171
Race												
White	Ref											
African American/black	2.556	1.238–5.278	.0112	293	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HPV risk group												
Low	Ref											
Intermediate	1.818	0.962–3.438	.0659		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
High	4.801	2.831–8.141	< .0001	276	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
p16 risk group												
Low	Ref											
Intermediate	2.322	1.397–3.861	.0012		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
High	4.284	2.590–7.086	< .0001	275	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Sex												
Female	Ref				Ref				Ref			
Male	1.223	0.743–2.013	.4292	296	0.949	0.507–1.776	.8691	117	1.268	0.553–2.909	.5752	172

TABLE 2. *Continued*

Variables	Overall				HPV negative				HPV positive			
	HR	95% CI	*p* value	No. of patients	HR	95% CI	*p* value	No. of patients	HR	95% CI	*p* value	No. of patients
Smoking status												
≤10 pack-years	Ref				Ref				Ref			
>10 pack-years	2.006	1.216–3.308	.0064	283	1.707	0.778–3.745	.1819	116	1.226	0.606–2.482	.5706	160
TNM classification												
I/II	Ref				Ref				Ref			
III/IV	1.025	0.498–2.111	.9461	294	1.918	0.828–4.445	.1287	117	0.739	0.175–3.114	.6799	171
T classification												
T1/T2	Ref				Ref				Ref			
T3/T4	2.946	1.999–4.341	< .0001	296	2.158	1.321–3.525	.0021	117	3.192	1.643–6.201	.0006	172

Abbreviations: HPV, human papillomavirus; HR, hazard ratio; 95% CI, 95% confidence interval; BOT, base of tongue.

Predictors of survival

For all patients with oropharyngeal SCC in this cohort, the following factors were found to be statistically significant predictors of survival in the univariate models (Table 2): age, surgical approach, extranodal extension, HPV status, marital status, mucosal margins, p16 status, perineural invasion, primary tumor site, race, smoking status, T classification, and low/intermediate/high-risk stratification (by HPV or p16).

For HPV-negative patients, the following factors were significant predictors of survival in the univariate models: extranodal extension, perineural invasion, primary tumor site, and T classification (Table 2). In the HPV-positive cohort, the following factors were significantly associated with survival (Table 2): surgical approach and T classification. The presence of extranodal extension nearly reached significance ($p = .0553$).

In the multivariable analysis ($n = 260$), after adjustment for all covariates, HPV status, extranodal extension, and T classification were independent predictors of survival (Table 3). Of note, aside from HPV status, several of the factors that were predictive of survival in the RTOG 0129 (smoking status and neck disease) failed to reach significance in this patient cohort after adjustment in the multivariable model.

For HPV-negative patients (multivariable model $n = 108$), extranodal extension (HR of 2.322 for those with extracapsular spread [ECS]; 95% CI = 1.359–3.968; $p = .0021$) and T classification (HR of 2.029 for those with T3/T4 disease; 95% CI = 1.054–3.906; $p = .0342$) were the only independent predictors of survival. For HPV-positive patients (multivariable model $n = 152$), the surgical approach and mucosal margins were the only 2 factors predictive of survival (Table 3). Within the HPV-positive subgroup, those with open surgical approaches had over 3 times the hazard of death than those with transoral surgical approaches (HR = 3.09; 95% CI = 1.293–7.385; $p = .0111$). Those with positive margins had a HR of 2.519 (95% CI = 1.101–5.766; $p = .0287$). There were no differences in positive margin rates between the transoral and open surgical approaches ($p = .2868$).

Recursive partitioning analysis

RPA of this patient cohort revealed that HPV status was the most important predictor of overall survival (see Figure 2). For HPV-positive patients, the best outcomes were achieved in those who underwent transoral surgery and had no evidence of perineural invasion at the primary site (9 deaths/104 patients; 91.3% survival). In addition, for those HPV-positive patients who were treated with an open surgical approach, the margin status was the next most important predictor of survival, with 17 deaths out of 46 patients (63.0% survival) at last follow-up with negative margins as opposed to 6 deaths out of 8 patients (25.0% survival) if the margins were positive. For HPV-negative patients, the most important predictor of survival was the presence of ECS (58.2% vs 27.5% survival). Survival rates were worst for T3/T4 tumors with ECS (13.8%).

TABLE 3. Multivariable analysis of outcome predictors.

	Analysis								
	Overall, $n = 260$			HPV negative, $n = 108$			HPV positive, $n = 152$		
Variables	HR	95% CI	*p* value	HR	95% CI	*p* value	HR	95% CI	*p* value
Approach									
Transoral	Ref			Ref			Ref		
Open	1.343	0.788–2.288	.2784	0.79	0.4–1.561	.4984	3.09	1.293–7.385	.0111
Extranodal extension									
No	Ref			Ref			Ref		
Yes	1.938	1.262–2.976	.0025	2.322	1.359–3.968	.0021	1.119	0.536–2.335	.7644
HPV									
Positive	Ref								
Negative	2.362	1.496–3.731	.0002	N/A	N/A	N/A	N/A	N/A	N/A
Mucosal margins									
Free of carcinoma	Ref			Ref			Ref		
Positive	1.621	0.969–2.711	.0656	1.213	0.614–2.395	.5787	2.519	1.101–5.766	.0287
Perineural invasion									
No	Ref			Ref			Ref		
Yes	1.398	0.901–2.17	.1347	1.407	0.82–2.414	.2153	1.14	0.514–2.529	.7466
Smoking status									
$\leq$10 pack-years	Ref			Ref			Ref		
$>$10 pack-years	1.316	0.763–2.27	.3235	1.452	0.637–3.311	.3747	1.302	0.616–2.751	.4894
T classification									
T1/T2	Ref			Ref			Ref		
T3/T4	1.937	1.187–3.16	.0081	2.029	1.054–3.906	.0342	2.125	0.983–4.595	.0555

Abbreviations: HPV, human papillomavirus; HR, hazard ratio; 95% CI, 95% confidence interval.

Effect of adjuvant therapy on survival

In this cohort, patients either underwent surgery alone ($n = 26$), surgery with postoperative radiotherapy ($n = 89$), or surgery with postoperative chemoradiotherapy ($n = 143$). In the remaining patients ($n = 38$), the data were incomplete with regard to adjuvant therapy, and these patients with missing data were excluded from the multivariable analysis.

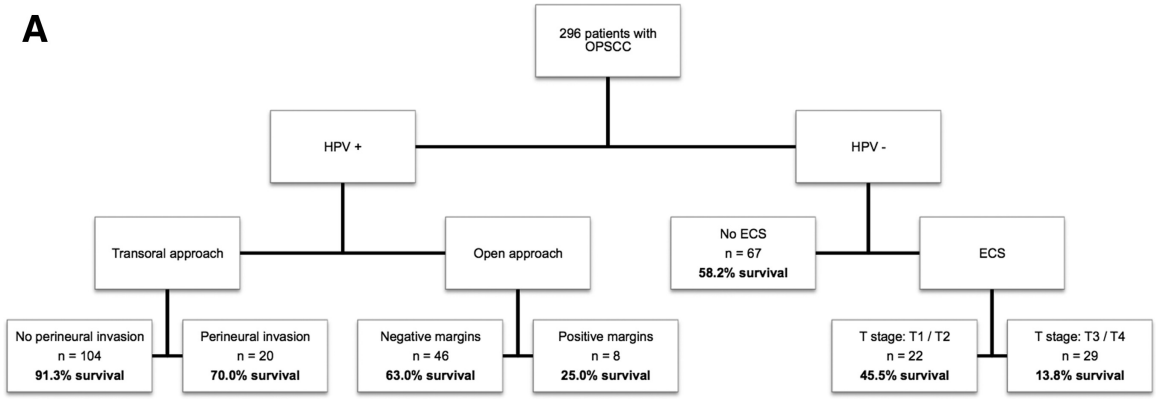
Multivariable analysis showed that there was no significant difference in survival in patients treated with surgery alone, surgery with adjuvant radiation, or surgery with adjuvant chemoradiation. The small number of patients treated with surgery alone made statistical comparisons between this group of patients and those receiving adjuvant radiation or chemoradiation treatment underpowered. However, the robust number of patients receiving adjuvant radiation treatment ($n = 89$) and adjuvant chemoradiation treatment ($n = 143$) allowed us to draw meaningful statistical comparisons between these groups. After controlling for all other variables, no difference in survival was seen between patients treated with surgery with adjuvant radiation versus surgery with adjuvant chemoradiation ($p = .6306$). In the HPV-positive subgroup, no difference was seen in patients treated with surgery with adjuvant radiation versus surgery with adjuvant chemoradiation ($p = .4707$). Similarly, in the HPV-negative subgroup, no difference was seen in patients treated with surgery with adjuvant radiation versus surgery with adjuvant chemoradiation ($p = .8493$).

Gastrostomy tube outcomes

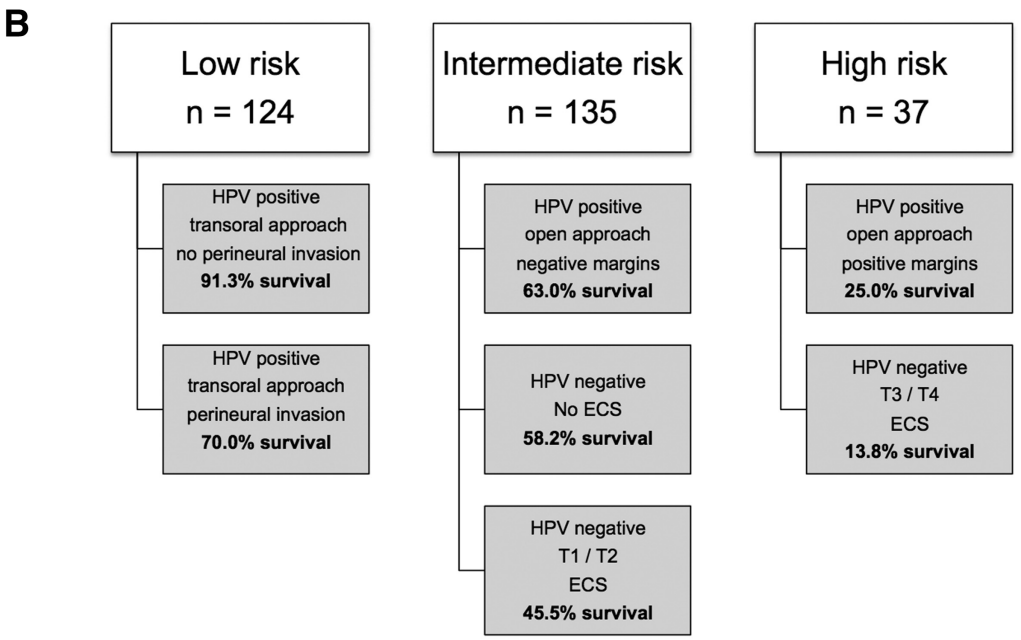
In an effort to determine functional outcomes in this cohort, the presence of a gastrostomy tube was documented. Patients were included based on the presence of a gastrostomy tube, rather than gastrostomy tube dependence, and this group included patients who were using the gastrostomy tube for at least a portion of their diet. Gastrostomy tubes were present in 12.8%, 23.3%, and 32.1% of patients at 1, 3, and 5 years, respectively. Gastrostomy tube dependence rates were impacted by surgical approach, with the open approach cohort accounting for the majority of patients who were gastrostomy tube dependent. Gastrostomy tube presence was lowest in patients with T1/T2 tumors undergoing transoral resection, whereas the rate of tube dependence was 7.84% for T1/T2 classification and 9.52% in patients with T3/T4 classification (Table 4).

DISCUSSION

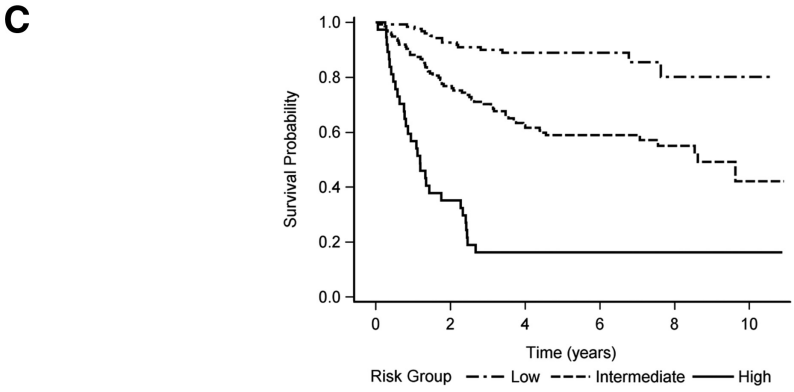
This study provides strong evidence that HPV status is the most important predictor of overall survival in a large, surgically treated oropharyngeal SCC cohort. For the entire cohort, 3 and 5 year survival rates were 71.3% and 65.1%, respectively. However, for HPV-positive patients, survival rates at 3 and 5 years were 83.3% and 81.8%, respectively, compared to 53.3% and 40.3% for HPV-negative patients. These outcomes are comparable to chemoradiotherapy trials for the same disease site.[13] When adjusting for other covariates (surgical approach, mucosal margins, perineural invasion, and smoking status), HPV status, ECS, and tumor classification were significantly associated with overall survival. Unlike previously reported primary chemoradiation studies, smoking status, stratified risk levels, and neck disease did not impact survival in surgically treated patients.[15]



Abbreviations: OPSCC, oropharyngeal squamous cell carcinoma; HPV, human papillomavirus; ECS, extracapsular spread



Abbreviations: HPV, human papillomavirus; ECS, extracapsular spread.



Risk Group	Survival Rates			
	1 year (%)	3 year (%)	5 year (%)	10 year (%)
Low	98.4	89.9	88.8	80.5
Intermediate	88.1	70.1	58.7	43.4
High	56.8	16.2	16.2	16.2

FIGURE 2. (A) Recursive partitioning analysis (RPA) showed that human papillomavirus (HPV) status was the most important predictor of overall survival. For HPV-positive tumors, the best outcomes were achieved in patients undergoing resection via the transoral approach with no perineural invasion. For HPV-positive tumors resected with an open approach, margin status was a critical determinant of survival. In HPV-negative tumors, extracapsular spread (ECS) was the next most important determinant of survival. Patients with HPV-negative tumors, ECS, and advanced T classification had the worst overall survival. (B) RPA allowed classification of patients into categories of low, intermediate, and high-risk of death. The low-risk group comprised 41.9% of the entire cohort and the intermediate-risk group comprised 45.6% of the entire cohort, whereas the high-risk group represented 12.5% of the entire cohort. (C) Kaplan–Meier estimates of survival according to low, intermediate, and high-risk of death determined by RPA.

TABLE 4. Percentage of patients with gastrostomy tube present at 12 months postsurgery.

Surgical approach	T1/T2	T3/T4
Transoral	7.84%	9.52%
Open	34.0%	33.33%

HPV-negative and HPV-positive patients had distinct independent predictors of survival. The most important predictor of survival in HPV-negative patients was the presence of ECS. On the other hand, survival in HPV-positive patients was associated with the surgical approach (transoral vs open), and whether or not negative margins were achieved. Specifically, in the HPV-positive cohort, those with transoral resection tended to have more favorable outcomes; patients resected via the open approach were over 3 times more likely to die than those treated transorally. When controlling for all factors, including T classification, smoking status, etc., patients who underwent transoral resection had improved survival. Our multivariable analysis confirmed that the surgical approach was a significant independent predictor of overall survival and not simply a surrogate marker for advanced disease. This finding may reflect the greater morbidity and swallowing dysfunction associated with open approaches, placing these patients at greater risk of postoperative aspiration pneumonia. Without randomization to surgical approach, however, potential unidentified confounders cannot be ruled out.

RPA of the entire patient population revealed that HPV status was the major determinant of overall survival. In HPV-positive patients, the next most important determinant of survival was the surgical approach utilized followed by the pathologic factors of margin status and perineural invasion. In the HPV-negative patient population, the surgical approach was not a significant predictor of outcome, but rather the presence of ECS, followed by the T classification of the primary tumor. If the RPA trees are pruned further, 3 survival outcome groups emerge that may be deemed: low-, intermediate-, and high-risk (see Figure 2). HPV-positive patients who are resected transorally have the lowest risk of death (15 deaths out of 124 patients; 87.9% survival). For HPV-positive patients undergoing transoral resection, the presence of perineural invasion was a significant prognostic factor, as shown in Figure 2A, which is contrary to the study by Haughey and Sinha,[27] who did not find perineural invasion to be a significant prognostic factor in surgically treated p16-positive patients. The intermediate-risk group consists of those patients who are HPV-positive and resected with an open approach and negative margins (60.3% survival), HPV-negative patients with no ECS (58.2% survival), or HPV-negative T1/T2 tumors with ECS (45.5% survival). Finally, the high-risk group consists of HPV-positive tumors resected with an open approach and positive margins (25.0% survival) and HPV-negative T3/T4 tumors with ECS (13.8% survival).

To complement the survival data, functional outcomes were also investigated. As shown in Table 4, patients undergoing an open approach had much higher rates of gastrostomy tube dependence compared with patients undergoing a transoral approach. Patients with T1 and T2 tumors who underwent a transoral resection had a gastrostomy tube present at 1 year in 7.84% of the cases, regardless of HPV status. These numbers are remarkably similar to other surgical trials and reinforce that higher rates of gastrostomy tube presence are primarily seen with T3 and T4 tumors (9.52% for the transoral approach and 33.33% for the open approach). This number compares favorably to gastrostomy tube and dysphagia rates in chemoradiation trials. Best et al[23] reported a 19% rate of stricture and Shiley et al[26] reported that 47% of patients continue to require gastrostomy tube feedings even 1 year after chemoradiotherapy. Even in studies evaluating the use of intensity-modulated radiotherapy, sparing pharyngeal constrictors, 4 of 73 patients (5.6%) report significant change in diet and 1 of 73 patients (1.3%) was exclusively gastrostomy tube dependent. In quality of life surveys, a sharp deterioration of swallowing is seen postchemoradiotherapy treatment and this improves slightly between 3 and 12 months posttherapy. Only 15.6% of patients reported a normal diet at 1 year postchemoradiotherapy, 57% have objective swallowing impairment, and 23% exhibit silent aspiration on modified barium swallowing studies.[22]

In this patient cohort, postoperative concurrent chemoradiotherapy was delivered for "high-risk patients" as defined by the paired *New England Journal of Medicine* manuscripts published in 2004.[34,41] Based on our analysis, when controlling for other variables, there was no significant difference in survival between patients treated with postoperative radiation versus those treated with concurrent chemoradiation. With the recent emphasis on treatment deintensification for HPV-positive patients, transoral surgery with postoperative radiotherapy alone may be an effective strategy to pursue based on these results. On the other hand, patients with HPV-negative tumors, T3/T4 primary, and ECS have unusually poor outcomes from both a survival and functional perspective. Intensification of nonsurgical therapy may be the best treatment options to consider in this group of patients.

The retrospective nature of this study could lend to selection biases, such as changing treatment patterns and techniques. Another weakness of this study was the lack of data on disease-specific, progression-free, and disease-free survival. However, this study is the largest analysis of primary surgical therapy for oropharyngeal SCC. We were able to control for many factors to arrive at the significant results of this study, demonstrating excellent survival and functional outcomes for selected populations and treatment modalities of oropharyngeal SCC. This analysis further supports the future use and study of primary surgical therapy for certain cohorts of oropharyngeal SCC, particularly in our attempts at deintensifying therapy for HPV-positive patients.

REFERENCES

1. Chaturvedi AK, Anderson WF, Lortet–Tieulent J, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *J Clin Oncol* 2013;31:4550–4559.
2. Gillison ML, D'Souza G, Westra W, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst* 2008;100:407–420.

3. Hafkamp HC, Manni JJ, Haesevoets A, et al. Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. *Int J Cancer* 2008;122:2656–2664.
4. Sturgis EM, Cinciripini PM. Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers? *Cancer* 2007;110:1429–1435.
5. Parsons JT, Mendenhall WM, Stringer SP, et al. Squamous cell carcinoma of the oropharynx: surgery, radiation therapy, or both. *Cancer* 2002;94: 2967–2980.
6. Parsons JT, Mendenhall WM, Million RR, Stringer SP, Cassisi NJ. The management of primary cancers of the oropharynx: combined treatment or irradiation alone? *Semin Radiat Oncol* 1992;2:142–148.
7. [No authors listed]. Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. The Department of Veterans Affairs Laryngeal Cancer Study Group. *N Engl J Med* 1991;324:1685–1690.
8. Spaulding MB, Fischer SG, Wolf GT. Tumor response, toxicity, and survival after neoadjuvant organ-preserving chemotherapy for advanced laryngeal carcinoma. The Department of Veterans Affairs Cooperative Laryngeal Cancer Study Group. *J Clin Oncol* 1994;12:1592–1599.
9. Calais G, Alfonsi M, Bardet E, et al. Randomized trial of radiation therapy versus concomitant chemotherapy and radiation therapy for advanced-stage oropharynx carcinoma. *J Natl Cancer Inst* 1999;91:2081–2086.
10. Denis F, Garaud P, Bardet E, et al. Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma. *J Clin Oncol* 2004;22:69–76.
11. Machtay M, Moughan J, Trotti A, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. *J Clin Oncol* 2008;26:3582–3589.
12. Huang K, Xia P, Chuang C, et al. Intensity-modulated chemoradiation for treatment of stage III and IV oropharyngeal carcinoma: the University of California-San Francisco experience. *Cancer* 2008;113:497–507.
13. Worden FP, Kumar B, Lee JS, et al. Chemoselection as a strategy for organ preservation in advanced oropharynx cancer: response and survival positively associated with HPV16 copy number. *J Clin Oncol* 2008;26:3138–3146.
14. Rischin D, Young RJ, Fisher R, et al. Prognostic significance of p16INK4A and human papillomavirus in patients with oropharyngeal cancer treated on TROG 02.02 phase III trial. *J Clin Oncol* 2010;28:4142–4148.
15. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010;363:24–35.
16. May JT, Rao N, Sabater RD, et al. Intensity-modulated radiation therapy as primary treatment for oropharyngeal squamous cell carcinoma. *Head Neck* 2013;35:1796–1800.
17. Posner MR, Lorch JH, Goloubeva O, et al. Survival and human papillomavirus in oropharynx cancer in TAX 324: a subset analysis from an international phase III trial. *Ann Oncol* 2011;22:1071–1077.
18. Garden AS, Harris J, Trotti A, et al. Long-term results of concomitant boost radiation plus concurrent cisplatin for advanced head and neck carcinomas: a phase II trial of the radiation therapy oncology group (RTOG 99-14). *Int J Radiat Oncol Biol Phys* 2008;71:1351–1355.
19. Setton J, Caria N, Romanyshyn J, et al. Intensity-modulated radiotherapy in the treatment of oropharyngeal cancer: an update of the Memorial Sloan–Kettering Cancer Center experience. *Int J Radiat Oncol Biol Phys* 2012;82:291–298.
20. Prestwich RJ, Kancherla K, Oksuz DC, et al. A single centre experience with sequential and concomitant chemoradiotherapy in locally advanced stage IV tonsillar cancer. *Radiat Oncol* 2010;5:121.
21. Greven KM, White DR, Browne JD, Williams DW III, McGuirt WF Sr, D'Agostino RB Jr. Swallowing dysfunction is a common sequelae after chemoradiation for oropharynx carcinoma. *Am J Clin Oncol* 2008;31:209–212.
22. Wilson JA, Carding PN, Patterson JM. Dysphagia after nonsurgical head and neck cancer treatment: patients' perspectives. *Otolaryngol Head Neck Surg* 2011;145:767–771.
23. 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;33:1727–1734.
24. Feng FY, Kim HM, Lyden TH, et al. Intensity-modulated chemoradiotherapy aiming to reduce dysphagia in patients with oropharyngeal cancer: clinical and functional results. *J Clin Oncol* 2010;28:2732–2738.
25. Setton J, Lee NY, Riaz N, et al. A multi-institution pooled analysis of gastrostomy tube dependence in patients with oropharyngeal cancer treated with definitive intensity-modulated radiotherapy. *Cancer* 2015;121:294–301.
26. Shiley SG, Hargunani CA, Skoner JM, Holland JM, Wax MK. Swallowing function after chemoradiation for advanced stage oropharyngeal cancer. *Otolaryngol Head Neck Surg* 2006;134:455–459.
27. Haughey BH, Sinha P. Prognostic factors and survival unique to surgically treated p16+ oropharyngeal cancer. *Laryngoscope* 2012;122 Suppl 2:S13–S33.
28. Licitra L, Perrone F, Bossi P, et al. High-risk human papillomavirus affects prognosis in patients with surgically treated oropharyngeal squamous cell carcinoma. *J Clin Oncol* 2006;24:5630–5636.
29. Moore EJ, Henstrom DK, Olsen KD, Kasperbauer JL, McGree ME. Transoral resection of tonsillar squamous cell carcinoma. *Laryngoscope* 2009; 119:508–515.
30. Hurtuk A, Agrawal A, Old M, Teknos TN, Ozer E. Outcomes of transoral robotic surgery: a preliminary clinical experience. *Otolaryngol Head Neck Surg* 2011;145:248–253.
31. Hurtuk AM, Marcinow A, Agrawal A, Old M, Teknos TN, Ozer E. Quality-of-life outcomes in transoral robotic surgery. *Otolaryngol Head Neck Surg* 2012;146:68–73.
32. Weinstein GS, O'Malley BW Jr, Magnuson JS, et al. Transoral robotic surgery: a multicenter study to assess feasibility, safety, and surgical margins. *Laryngoscope* 2012;122:1701–1707.
33. NCCN Clinical Practice Guidelines in Oncology. Head and Neck Cancers version 2.2014. Available at: http://www.nccn.org/professionals/physician_gls/f_guidelines.asp. Accessed May 1, 2015.
34. Bernier J, Domenge C, Ozsahin M, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Engl J Med* 2004;350:1945–1952.
35. Terrell JE, Ronis DL, Fowler KE, et al. Clinical predictors of quality of life in patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2004;130:401–408.
36. Padilla GV, Grant MM. Psychosocial aspects of artificial feeding. *Cancer* 1985;55(1 Suppl):301–304.
37. Roberge C, Tran M, Massoud C, et al. Quality of life and home enteral tube feeding: a French prospective study in patients with head and neck or oesophageal cancer. *Br J Cancer* 2000;82:263–269.
38. Radhakrishnan R, Solomon M, Satyamoorthy K, Martin LE, Lingen MW. Tissue microarray – a high-throughput molecular analysis in head and neck cancer. *J Oral Pathol Med* 2008;37:166–176.
39. The R project for statistical computing. R: a language and environment for statistical computing. Available at: http://www.R-project.org/. Accessed May 1, 2015.
40. Therneau T, Atkinson B, Ripley B. Rpart: recursive partitioning and regression trees. R package version 4.1-8. Available at: https://cran.r-project.org/web/packages/rpart/rpart.pdf. Accessed May 1, 2015.
41. Cooper JS, Pajak TF, Forastiere AA, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. *N Engl J Med* 2004;350:1937–1944.

Reprinted by permission of Head Neck. 2016; 38(12):1752-1758.

ORIGINAL ARTICLE

Cutaneous head and neck melanoma in OPTiM, a randomized phase 3 trial of talimogene laherparepvec versus granulocyte-macrophage colony-stimulating factor for the treatment of unresected stage IIIB/IIIC/IV melanoma

Robert H. I. Andtbacka, MD, CM[1]* Sanjiv S. Agarwala, MD[2] David W. Ollila, MD[3] Sigrun Hallmeyer, MD[4] Mohammed Milhem, MD[5] Thomas Amatruda, MD[6] John J. Nemunaitis, MD[7] Kevin J. Harrington, PhD[8] Lisa Chen, PhD[9] Mark Shilkrut, MD, PhD[9] Merrick Ross, MD[10] Howard L. Kaufman, MD[11]

[1]University of Utah Huntsman Cancer Institute, Salt Lake City, Utah, [2]St. Luke's University Hospital and Temple University, Philadelphia, Pennsylvania, [3]University of North Carolina, Chapel Hill, North Carolina, [4]Advocate Lutheran General Hospital, Park Ridge, Illinois, [5]University of Iowa Hospitals and Clinics, Iowa City, Iowa, [6]Minnesota Oncology, Fridley, Minnesota, [7]Mary Crowley Cancer Research Center, Dallas, Texas, [8]The Institute of Cancer Research/The Royal Marsden Hospital, London, UK, [9]Amgen, Inc., Thousand Oaks, California, [10]The University of Texas MD Anderson Cancer Center, Houston, Texas, [11]Rutgers Cancer Institute of New Jersey, Rutgers, New Jersey.

Accepted 16 May 2016
Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/hed.24522

ABSTRACT: *Background.* Cutaneous head and neck melanoma has poor outcomes and limited treatment options. In OPTiM, a phase 3 study in patients with unresectable stage IIIB/IIIC/IV melanoma, intralesional administration of the oncolytic virus talimogene laherparepvec improved durable response rate (DRR; continuous response $\geq$6 months) compared with subcutaneous granulocyte-macrophage colony-stimulating factor (GM-CSF).

Methods. Retrospective review of OPTiM identified patients with cutaneous head and neck melanoma given talimogene laherparepvec ($n = 61$) or GM-CSF ($n = 26$). Outcomes were compared between talimogene laherparepvec and GM-CSF treated patients with cutaneous head and neck melanoma.

Results. DRR was higher for talimogene laherparepvec–treated patients than for GM-CSF treated patients (36.1% vs 3.8%; $p = .001$). A total of 29.5% of patients had a complete response with talimogene laherparepvec versus 0% with GM-CSF. Among talimogene laherparepvec–treated patients with a response, the probability of still being in response after 12 months was 73%. Median overall survival (OS) was 25.2 months for GM-CSF and had not been reached with talimogene laherparepvec.

Conclusion. Treatment with talimogene laherparepvec was associated with improved response and survival compared with GM-CSF in patients with cutaneous head and neck melanoma. *Head Neck* **00:** 000–000, 2016

KEY WORDS: cutaneous head and neck melanoma, talimogene laherparepvec, oncolytic virus, cancer immunotherapy

***Corresponding author:** R. Andtbacka, MD, CM, FRCSC, Department of Surgery, University of Utah, The Huntsman Cancer Institute, 2000 Circle of Hope Drive, Salt Lake City, UT 84112-5550. E-mail: Robert.Andtbacka@hci.utah.edu

This research was supported by Amgen Inc.

Conflict of interest statement: R.H.I.A. has received honoraria from Amgen Inc. S.H. has been a consultant to Bristol–Myers Squibb, Cardinal Health, and iCAN, and served on speaker's bureau for Bristol–Myers Squibb. M.M. has been an advisory board member for Amgen Inc., Genentech, EMD Serono, and Novartis. K.J.H. discloses payments by Amgen Inc. for consultant work and honoraria for symposia. M.R. has been an advisory board member for Merck, GSK, and Amgen Inc., has received travel expenses and honoraria from Merck, GSK, and Amgen Inc., and has been a publications steering committee member for GSK. L.C. and M.S. are employees of and stockholders in Amgen Inc. H.L.K. has been a consultant/advisor for BioVex and Amgen Inc. and receives research funding from Amgen Inc. and Viralytics. H.L.K. has also served as a consultant for Alkermes, EMD Serono, Merck, Prometheus, and Sanofi and receives research funding from Amgen Inc., Bristol Myers Squibb, EMD Serono, Merck, Prometheus, and Viralytics. H.L.K. has served on a speaker's bureau for Merck and returns all honoraria to Rutgers University. S.S.A., D.W.O., T.A., and J.J.N. have no conflicts to declare.

Contract grant sponsor: Amgen Inc.

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

INTRODUCTION

Overall, 15% to 20% of cutaneous melanomas arise from head and neck locations despite this region representing <10% of total body surface area.[1–3] Outcomes associated with cutaneous head and neck melanoma are poorer when compared with all other body sites, with a higher rate of recurrence and shorter disease-free and overall survival (OS).[1] Surgical treatment of cutaneous head and neck melanoma is technically challenging, owing to the difficulty in achieving appropriate margins in this cosmetically sensitive region.[4–6] Because of the increased risk of recurrence and regional and systemic spread and recurrence with this location of melanoma, adjuvant therapy (including radiation therapy) is often used after surgical resection.[7–9] For patients with unresectable head and neck disease, treatment options have been even more limited, with radiation therapy frequently used for locoregional disease control and palliation. Therefore, new treatment strategies are of high priority.

Oncolytic viruses are novel cancer treatments that mediate antitumor activity by selectively replicating in tumors and lysing tumor cells, subsequently releasing tumor-derived antigens to promote antitumor immunity.[10]

TABLE 1. Baseline demographics and clinical characteristics.

	Talimogene laherparepvec	GM-CSF
No. of patients	$N = 61$	$N = 26$
Median (IQR) age, y	70 (61–79)	66 (58–75)
Men, no. (%)	51 (84)	17 (65)
ECOG PS, no. (%)		
0	43 (70)	20 (77)
1	18 (30)	6 (23)
Disease stage at screening,* no. (%)		
IIIB	9 (15)	5 (19)
IIIC	17 (28)	6 (23)
IVM1a	11 (18)	6 (23)
IVM1b	15 (25)	4 (15)
IVM1c	9 (15)	5 (19)
Elevated LDH, no. (%)	2 (3)	1 (4)
BRAF status,† no. (%)		
Mutant	10 (16)	6 (23)
Wild-type	6 (10)	4 (15)
Unknown/missing	45 (74)	16 (62)
Location of first recurrence,‡ no. (%)		
Surgical scar (local)	17 (28)	4 (15)
In-transit/satellitosis	21 (34)	7 (27)
Regional lymph node(s)	16 (26)	3 (12)
Distant skin site	7 (11)	6 (23)
Distant lymph node(s)	0	1 (4)
Visceral	3 (5)	2 (8)
Other	4 (7)	4 (15)
Missing	3 (5)	2 (8)
Median (IQR) time from initial diagnosis to first recurrence, y	0.6 (0.3–1.2)	0.5 (0.3–1.6)
Line of therapy, no. (%)		
First line	37 (61)	15 (58)
Second line or greater	24 (39)	11 (42)
HSV-1 status, no. (%)		
Seropositive	38 (62)	13 (50)
Seronegative	18 (30)	13 (50)
Unknown	5 (8)	0

Abbreviations: GM-CSF, granulocyte-macrophage colony-stimulating factor; IQR, interquartile range; ECOG PS, Eastern Cooperative Oncology Group performance status; LDH, lactate dehydrogenase; HSV-1, herpes simplex virus type 1.

*Per case report form at screening.

†Because tissue was not collected prospectively, *BRAF* mutation analysis was reported by investigators and not evaluated centrally.

‡Patients may have had more than one site of first recurrence. Site of first recurrence was evaluated at screening.

Oncolytic viruses can be modified to express genes that further augment the antitumor immune response.[11] Talimogene laherparepvec is a modified herpes simplex virus (HSV) type-1 designed to specifically replicate in and lyse tumor cells.[12] In addition to modifications designed to attenuate viral pathogenicity in normal tissues and to restore antigen presentation by HSV-infected cells, talimogene laherparepvec is engineered to express the gene encoding human granulocyte-macrophage colony-stimulating factor (GM-CSF).[12] GM-CSF can act to recruit and activate antigen-presenting cells to process and present tumor-derived antigens to help promote tumor specific T-cell responses.[13] Release of immune-stimulatory viral proteins may further enhance the antitumor immune response.[11] Responses in uninjected tumors, including visceral metastases, have been seen in patients treated with talimogene laherparepvec (in the OPTiM study responses to talimogene laherparepvec were observed in 34% of evaluable uninjected nonvisceral and 15% of evaluable visceral lesions),[14–17] indicating that an effective systemic antitumor response can be achieved.

In the randomized phase 3 OPTiM study, intralesional talimogene laherparepvec improved the primary endpoint of durable response rate (DRR; defined as complete response [CR] or partial response [PR] lasting continuously for $\geq$6 months) from 2% to 16% ($p < .0001$), compared to subcutaneous GM-CSF in patients with stage IIIB/IIIC/IV melanoma that was not surgically resectable. The overall response rate (ORR), as evaluated by an independent Endpoint Assessment Committee, was also improved from 6% with GM-CSF to 26% with talimogene laherparepvec ($p < .0001$, descriptive). Similarly, 11% of patients had a CR in the talimogene laherparepvec arm versus <1% in the GM-CSF arm. Median OS with talimogene laherparepvec treatment was 23.3 months compared with 18.9 months with GM-CSF treatment (hazard ratio [HR] = 0.79; 95% confidence interval [CI] = 0.62–1.00; $p = .051$).[16] At the final planned analysis of OS, median OS was 23.3 months in the talimogene laherparepvec arm and 18.9 months in the GM-CSF arm (HR = 0.79; 95% CI = 0.62–1.00; $p = .049$, descriptive]).[18] This article describes a retrospective analysis of the subgroup of patients from the phase 3 OPTiM study who had cutaneous head and neck melanoma. DRR, ORR, time to treatment failure (TTF), and OS are reported to describe clinical outcomes with talimogene laherparepvec treatment in this melanoma subtype.

PATIENTS AND METHODS

Study design, patients, and treatment

Eligibility criteria and study design for the randomized, phase 3, open-label multicenter OPTiM study are summarized in Supplementary Figure S1, online only, and have been reported in detail previously.[16] Briefly, eligible patients were $\geq$18 years old with histologically confirmed cutaneous injectable and unresectable stage IIIB/IIIC/IV melanoma. Patients were excluded from the study if they had 3 or more visceral metastases, except lung metastases or nodal metastases associated with visceral organs, or visceral metastases >3 cm. This subgroup analysis included patients enrolled in the study who, at initial diagnosis, had melanoma located in the head and neck region (ie, scalp, face, and neck) as determined by the investigator. Patients were randomly assigned 2:1 to receive intralesional talimogene laherparepvec ($\leq$4 mL initially at 10^6 pfu/mL, then after 3 weeks 10^8 pfu/mL once every 2 weeks) or subcutaneous GM-CSF (125 $\mu g/m^2$ daily for 14 days in 28-day cycles). Discontinuation of study treatment because of disease progression was not required before 24 weeks unless alternate therapy was required or intolerance to treatment developed. All patients provided written informed consent, and all study procedures were approved by institutional review boards or ethics committees. The trial was registered with ClinicalTrials.gov (identifier NCT00769704).

DRR was the primary endpoint (defined as the rate of CR or PR lasting $\geq$6 months continuously and beginning within the first 12 months of treatment). Key secondary endpoints included OS (time from randomization to death), ORR, onset

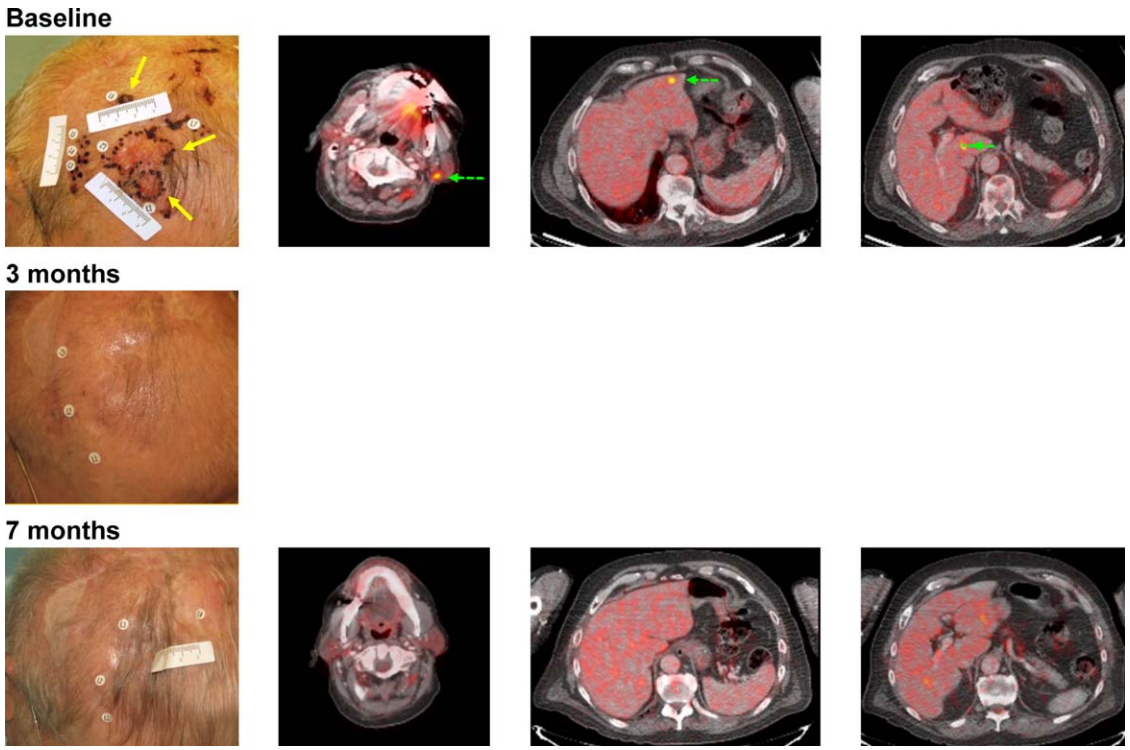


FIGURE 1. (A) Representative images from a patient with melanoma of the scalp with metastasis to cervical lymph nodes and liver (stage IVM1c). The patient was diagnosed 2 years before enrollment in OPTiM and had 2 surgeries: one at diagnosis, and another 1 year after recurrence. Top row: injection sites shown in yellow arrows at baseline (left panel). Uninjected sites are shown with green dashed arrows. Black dots mark tumor lesions. Sites included 1 fluorodeoxyglucose (FDG)-avid left upper level V cervical lymph node (center left panel) and 2 FDG-avid liver lesions (center right and right panels). Middle row: injections were stopped after complete resolution of scalp lesions after cycle 2 (1 cycle = 2 injections of talimogene laherparepvec). Bottom row: Complete resolution of cervical and liver tumors was documented by FDG-PET CT at cycle 7. Patient was in complete response until the end of the trial, duration of response (complete response) was approximately 17 months.

and duration of response, TTF (time from date of randomization to the date of the first clinically relevant progressive disease not followed by response or until death), and safety. Patients were evaluated clinically every treatment cycle (4 or 5 weeks) and/or radiographically every 12 weeks. DRR and ORR were determined using modified World Health Organization Criteria for Tumor Response Evaluation.[16,19] Patients with a best response of CR or PR per investigator assessment or who had received study treatment for $\geq$9 months were evaluated by an independent blinded endpoint assessment committee (EAC).

Statistical analysis

Efficacy analyses were done for all patients with cutaneous head and neck melanoma who met the criteria for inclusion in this subgroup analysis and received at least 1 dose of study medication (see Patients above). All analyses were exploratory. The Fisher exact test was used to compare DRR and ORR between treatment arms. Time-to-event endpoints were evaluated using Cox proportional hazard models and unadjusted log-rank tests. DRR and ORR were based on data from the primary DRR analysis; data cutoff for this analysis was December 21, 2012. OS and TTF analyses were based on data from the primary OS analysis, which was done after 290 survival events had occurred in the overall study population; the data cutoff date for this analysis was March 31, 2014. Multivariate analysis was conducted to adjust for imbalances in baseline prognostic factors. Statistical significance was interpreted at a two-sided 5% confidence level without multiplicity adjustment.

RESULTS

Patient characteristics, disposition, and treatment

Of the 436 patients enrolled in the OPTiM study, retrospective review identified 87 patients (20%) with cutaneous head and neck melanoma (treated with talimogene laherparepvec, n = 61 [21%]; treated with GM-CSF, n = 26 [18%]). The baseline clinical characteristics of these patients are shown in Table 1. Baseline demographics and characteristics for the intent-to-treat population are shown in Supplementary Table S1, online only. The median duration of follow-up at the primary analysis of OS was 35 months (interquartile range [IQR], 13–43 months) for the talimogene laherparepvec group and 25 months (IQR, 13–39 months) for the GM-CSF group.

Durable and overall response

DRR per EAC was 9.5-times higher in the talimogene laherparepvec arm (36.1%; 95% CI = 24.2% to 49.4%) compared to the GM-CSF arm (3.8%; 95% CI = 0.1% to 19.6%; $p = .001$). ORR was higher in the talimogene laherparepvec arm (47.5%; 95% CI = 34.6% to 60.7%)

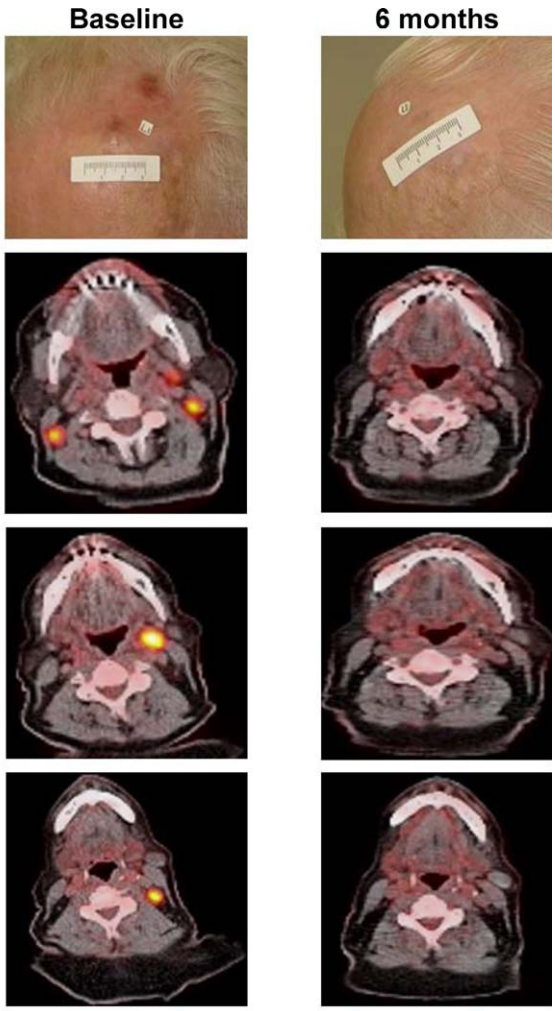


FIGURE 2. Representative images from a patient with stage IIIC disease randomized to talimogene laherparepvec who had a complete response. The patient was enrolled in the study with desmoplastic melanoma of the forehead with bilateral cervical fluorodeoxyglucose-avid lymph nodes (left panel). Talimogene laherparepvec was injected only into the cutaneous lesion marked by the label (top row). At month 4, a partial response was reported and injection of talimogene laherparepvec was stopped. At cycle 6, a complete remission was reported that continued until the end of the study. Duration of response was 15.5 months. The patient was disease-free at last follow-up contact approximately 3 years after enrollment.

than in the GM-CSF arm (7.7%; 95% CI = 1.0% to 25.1%; $p = .0004$). Eighteen patients (29.5%) in the talimogene laherparepvec arm had a CR, whereas no patient in the GM-CSF arm had a CR. Eleven patients (18.0%) in the talimogene laherparepvec arm had a PR, compared with 2 patients (7.7%) in the GM-CSF arm. DRRs and ORRs were more common among patients with disease stages IIIB, IIIC, and IVM1a (Supplementary Table S2, online only). Although ORR was numerically greater among patients with HSV-seropositive disease (55.3%; 95% CI = 38.3–71.4) than patients with HSV-seronegative disease (27.8%; 95% CI = 9.7–53.5), the difference between the 2 groups was not statistically significant ($p = .14$). Similarly, the DRR in patients with HSV-seropositive disease (29.4%; 95% CI = 17.5–43.8) was numerically greater but not significantly different from that in patients with HSV-seronegative disease (16.1%; 95% CI = 5.5–33.7; $p = .20$).

In the talimogene laherparepvec arm, responses were identified in 63.8% of injected lesions, 7.9% of uninjected nonvisceral lesions, and 10.8% of visceral lesions. Among 341 responding injected lesions, 311 (91.2%) were cutaneous or subcutaneous, and 29 (8.5%) were nodal; among 88 responding uninjected nonvisceral lesions, 65 (73.9%) were cutaneous or subcutaneous, and 6 (6.8%) were nodal.

Photographs and radiographic images from representative patients with cutaneous head and neck melanoma who received treatment with talimogene laherparepvec are shown in Figures 1 and 2.

Duration of response and probability of responders remaining in response at landmark time points are shown in Figure 3. Among patients in the talimogene laherparepvec arm with a response ($n = 29$), the estimated probability of being in response after 9 months was 73% (95% CI = 56% to 90%); this remained unchanged at the 12-month and 15-month time points.

Time to treatment failure

Median TTF was significantly prolonged for patients in the talimogene laherparepvec group (18.3 months [IQR, 8.6–not estimable]) compared with patients in the GM-CSF group (4.1 months [IQR, 2.8–7.4]; HR = 0.32; 95% CI = 0.17–0.61; $p = .0002$). Kaplan–Meier curves for TTF are shown in Figure 4A.

Overall survival and multivariate analysis

Kaplan–Meier curves for primary OS are shown in Figure 4B. Median OS was not estimable in the talimogene laherparepvec group (IQR, 29.7 months–not estimable) and was 25.2 months (IQR, 12.8–37.4 months) in the GM-CSF group. The unadjusted HR for OS was 0.57 (95% CI = 0.32–1.03) favoring the talimogene laherparepvec group (unadjusted $p = .062$). At 24 and 48 months, estimated survival was 67.2% and 52.9%, respectively, in patients in the talimogene laherparepvec group and 50.0% and 29.6%, respectively, in patients in the GM-CSF group. To adjust for potential clinically meaningful imbalances in prognostic factors of sex, disease stage, and Eastern Cooperative Oncology Group (ECOG) performance status, a multivariate sensitivity analysis was conducted. In this analysis, talimogene laherparepvec treatment was associated with improved OS compared to GM-CSF (HR = 0.38; 95% CI = 0.20–0.72; $p = .003$; Table 2).

DISCUSSION

OPTiM was the first randomized, controlled, phase 3 study with an oncolytic virus to show therapeutic benefit in melanoma. The study met its primary endpoint, with the results indicating intralesional talimogene laherparepvec treatment improved DRR compared to subcutaneous GM-CSF.[16] This retrospective analysis of the OPTiM study evaluated clinical outcomes in the patients with cutaneous head and neck melanoma cohort and

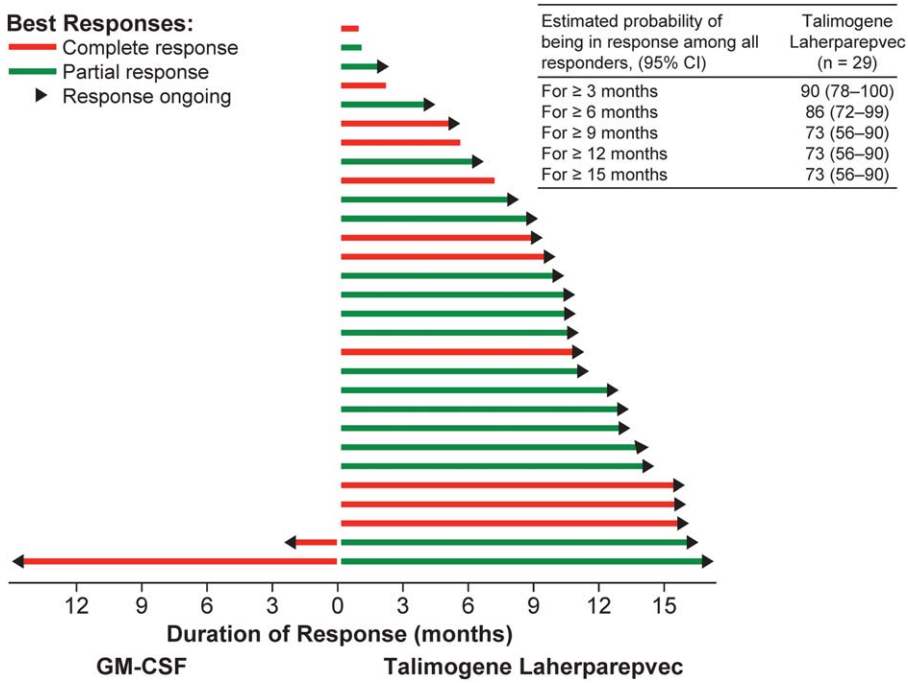


FIGURE 3. Duration of response for all patients with a response per endpoint assessment committee (EAC) was censored (marked by arrow) if at the last tumor assessment there was no evidence (per EAC) that the response had ended. Probability of being in response was estimated using the Kaplan–Meier method. Because only 1 patient in the granulocyte-macrophage colony-stimulating factor (GM-CSF) group had a response lasting >3 months, probability of being in response was not calculated for this group.

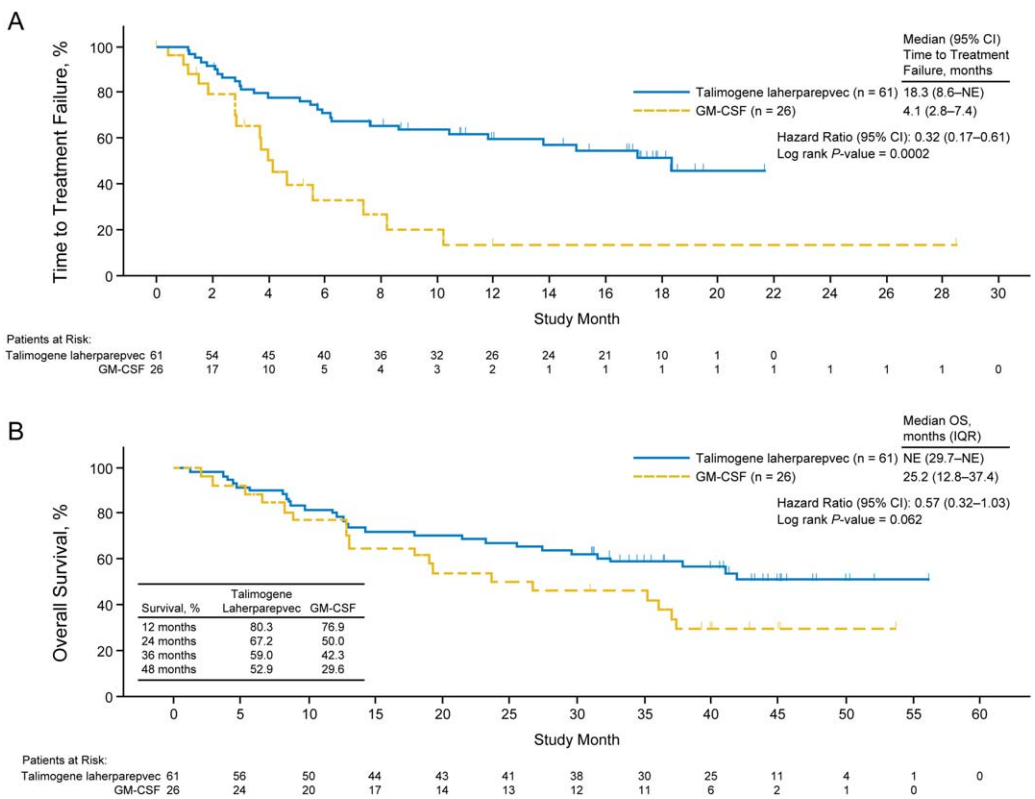


FIGURE 4. (A) Time to treatment failure per investigator assessment. (B) Overall survival. CI, confidence interval; GM-CSF, granulocyte-macrophage colony-stimulating factor; IQR, interquartile range; NE, not estimable; OS, overall survival.

TABLE 2. Multivariate analysis of the effect of talimogene laherparepvec on overall survival.

Covariate*	HR (95% CI)	*p* value
Sex		
Female vs male	0.40 (0.18–0.89)	.025
ECOG PS		
0 vs 1	0.27 (0.14–0.53)	< .001
Disease stage		
IIIC vs IIIB	0.15 (0.04–0.55)	< .001
IV M1a vs IIIB	0.91 (0.35–2.41)	
IV M1b vs IIIB	2.07 (0.83–5.19)	
IV M1c vs IIIB	1.05 (0.39–2.87)	
Treatment		
Talimogene laherparepvec vs GM-CSF	0.38 (0.20–0.72)	.003

Abbreviations: HR, hazard ratio; CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group performance status; GM-CSF, granulocyte-macrophage colony-stimulating factor.
* Multivariate analysis includes prognostic covariates with imbalances at baseline.

showed that talimogene laherparepvec demonstrated clinical benefit across different outcome measures in this difficult-to-treat subgroup.

Administration of talimogene laherparepvec was associated with higher DRR compared to GM-CSF (36.1% vs 3.8%; $p < .0001$). In addition, responding patients had an estimated 73% probability of being in response 15 months or longer. As shown in the representative images (see Figure 1), some patients receiving talimogene laherparepvec had resolution of all lesions. The rate of CR (30%) was noteworthy. Achievement of CR is a particularly important consideration in patients with cutaneous head and neck melanoma because resection of these often cosmetically disfiguring lesions can be challenging, and some effective regional treatment options, such as isolated infusion/perfusion with antitumor agents, are not feasible for this anatomic site.[20]

Because retrospective comparisons in general can be flawed, particularly when comparing groups of patients that were not prospectively stratified, a multivariate sensitivity analysis that adjusted for imbalances in clinically important prognostic factors between the treatment arms in the cutaneous head and neck melanoma subgroup was performed. This analysis demonstrated a 62% lower risk of death in patients treated with talimogene laherparepvec compared with the GM-CSF group (HR = 0.38; 95% CI = 0.20–0.72; $p = .003$). The median OS times in this retrospective analysis of the cutaneous head and neck melanoma subgroup are notable, and stand in contrast to previous reports that have noted poorer survival outcomes in patients with cutaneous head and neck melanoma.[1] Importantly, treatment with talimogene laherparepvec has been associated with responses at uninjected tumor sites, including lesions in visceral organs,[14,16] indicating that a systemic antitumor response was initiated.

The better outcomes for patients with cutaneous head and neck melanoma compared with the overall study population are notable. One potential explanation for the better outcomes observed with talimogene laherparepvec in patients with cutaneous head and neck melanoma may be the higher proportion of patients with stage IIIB/IIIC disease than the overall study population (43% vs 30%). In an exploratory analysis of OPTiM, patients with stages IIIB/IIIC/IVM1a melanoma benefited the most from talimogene laherparepvec, with DRR as high as 33% for stages IIIB/IIIC and 16% for stage IVM1a, and median OS that was 41.1 months for patients with stage IIIB/IIIC/IVM1a disease in the talimogene laherparepvec arm compared to 21.5 months in the GM-CSF arm (HR = 0.57; 95% CI = 0.40–0.80; $p < .001$ descriptive).[16]

Recently, a number of new immunotherapy and targeted therapy agents[21–27] have been shown to be effective in patients with advanced melanoma but it is unclear what proportion of patients receiving these new therapies in these studies had cutaneous head and neck melanoma. Given its activity in patients with unresectable melanoma, its intralesional mode of administration, its ability to induce durable PRs and CRs, and responses at distant uninjected sites coupled with the prolonged TTF and OS, talimogene laherparepvec may represent a potential treatment option for patients with unresectable cutaneous head and neck melanoma. Notably, talimogene laherparepvec demonstrated a tolerable safety profile with most adverse events being within a spectrum of flu-like symptoms, and generally transient and mild to moderate in severity.[16]

The key limitation of this study was its retrospective nature, which did not allow for control of clinical features across the treatment groups. As noted above, there were imbalances in duration of median follow-up (1.4-fold longer for patients treated with talimogene laherparepvec) and in baseline prognostic factors between arms that may have influenced the assessment of OS. It is also important to note that randomization of patients to treatment was not stratified by tumor location and that, although randomization in the overall population was 2:1 (talimogene laherparepvec:GM-CSF), fewer patients with cutaneous head and neck melanoma were randomized to the GM-CSF arm; the ratio in this analysis was 2.35:1. The influence on outcomes of this imbalance in randomization is uncertain.

In conclusion, in this retrospective analysis of the OPTiM study, administration of talimogene laherparepvec was associated with improved ORR, DRR, and OS compared to GM-CSF in patients with cutaneous head and neck melanoma, consistent with results seen in the intent-to-treat population of the primary study.[16] Talimogene laherparepvec is a potential novel treatment option for patients with regionally and distantly metastatic unresectable cutaneous head and neck melanoma.

Acknowledgments

The authors thank Peng He, PhD, Amgen Inc., for statistical support, and Ali Hassan, PhD, and Meghan Johnson, PhD (Complete Healthcare Communications, LLC, Chadds Ford, PA) for medical writing assistance in the preparation of this article. Their work was funded by Amgen Inc. K.J.H. acknowledges support from the RM/ICR NIHR Biomedical Research Centre.

REFERENCES

1. Fadaki N, Li R, Parrett B, et al. Is head and neck melanoma different from trunk and extremity melanomas with respect to sentinel lymph node status and clinical outcome? *Ann Surg Oncol* 2013;20:3089–3097.

2. Garbe C, Büttner P, Bertz J, et al. Primary cutaneous melanoma. Identification of prognostic groups and estimation of individual prognosis for 5093 patients. *Cancer* 1995;75:2484–2491.
3. Golger A, Young DS, Ghazarian D, Neligan PC. Epidemiological features and prognostic factors of cutaneous head and neck melanoma: a population-based study. *Arch Otolaryngol Head Neck Surg* 2007;133:442–447.
4. Shashanka R, Smitha BR. Head and neck melanoma. *ISRN Surg* 2012; 2012:948302.
5. Kienstra MA, Padhya TA. Head and neck melanoma. *Cancer Control* 2005;12:242–247.
6. Ballantyne AJ. Malignant melanoma of the skin of the head and neck. An analysis of 405 cases. *Am J Surg* 1970;120:425–431.
7. Evans MS, Drabick JJ. Systemic treatment in the management of head and neck cutaneous malignancies. *Oper Tech Otolaryngol Head Neck Surg* 2013;24:63–68.
8. Mocellin S, Pasquali S, Rossi CR, Nitti D. Interferon alpha adjuvant therapy in patients with high-risk melanoma: a systematic review and meta-analysis. *J Natl Cancer Inst* 2010;102:493–501.
9. Patel JN, Walko CM. Sylatron: a pegylated interferon for use in melanoma. *Ann Pharmacother* 2012;46:830–838.
10. Guo ZS, Liu Z, Bartlett DL. Oncolytic immunotherapy: dying the right way is a key to eliciting potent antitumor immunity. *Front Oncol* 2014;4:74.
11. Forbes NE, Krishnan R, Diallo JS. Pharmacological modulation of anti-tumor immunity induced by oncolytic viruses. *Front Oncol* 2014;4:191.
12. Hu JC, Coffin RS, Davis CJ, et al. A phase I study of OncoVEXGM-CSF, a second-generation oncolytic herpes simplex virus expressing granulocyte macrophage colony-stimulating factor. *Clin Cancer Res* 2006;12:6737–6747.
13. Kaufman HL, Ruby CE, Hughes T, Slingluff CL Jr. Current status of granulocyte-macrophage colony-stimulating factor in the immunotherapy of melanoma. *J Immunother Cancer* 2014;2:11.
14. Senzer NN, Kaufman HL, Amatruda T, et al. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. *J Clin Oncol* 2009;27:5763–5771.
15. Ross MI, Andtbacka RHI, Puzanov I, et al. Patterns of durable response with intralesional talimogene laherparepvec (T-VEC): results from a phase III trial in patients with stage IIIb-IV melanoma. *J Clin Oncol* 2014; 32(suppl):abstract 9026.
16. Andtbacka RH, Kaufman HL, Collichio F, et al. Talimogene laherparepvec improves durable response rate in patients with advanced melanoma. *J Clin Oncol* 2015;33:2780–2788.
17. Andtbacka RH, Ross MI, Delman K, et al. Responses of injected and uninjected lesions to intralesional talimogene laherparepvec (T-VEC) in the OPTiM study and the contribution of surgery to response. Ann Surg Oncol 2014;21(suppl 1):abstract 52.
18. Andtbacka RHI, Collichio FA, Amatruda T, et al. Final planned overall survival (OS) from OPTiM, a randomized phase III trial of talimogene laherparepvec (T-VEC) versus GM-CSF for the treatment of unresected stage IIIB/C/IV melanoma (NCT00769704). *J Immunother Cancer* 2014;2(Suppl 3):P263.
19. World Health Organization. WHO handbook for reporting results of cancer treatment. Geneva, Switzerland: World Health Organization; 1979.
20. Hayes AJ, Clark MA, Harries M, Thomas JM. Management of in-transit metastases from cutaneous malignant melanoma. *Br J Surg* 2004;91: 673–682.
21. Chapman PB, Hauschild A, Robert C, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 2011;364: 2507–2516.
22. Hauschild A, Grob JJ, Demidov LV, et al. Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* 2012;380:358–365.
23. Falchook GS, Lewis KD, Infante JR, et al. Activity of the oral MEK inhibitor trametinib in patients with advanced melanoma: a phase 1 dose-escalation trial. *Lancet Oncol* 2012;13:782–789.
24. Larkin J, Ascierto PA, Dréno B, et al. Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. *N Engl J Med* 2014;371:1867–1876.
25. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010;363: 711–723.
26. Ribas A, Hodi FS, Kefford R, et al. Efficacy and safety of the anti-PD-1 monoclonal antibody MK-3475 in 411 patients (pts) with melanoma (MEL). *J Clin Oncol* 2014;32:abstract LBA9000.
27. Weber JS, Minor DR, D'Angelo SP, et al. A phase 3 randomized, open-label study of nivolumab (anti-PD-1; BMS-936558; ONO-4538) versus investigator's choice chemotherapy (ICC) in patients with advanced melanoma after prior anti-CTLA-4 therapy. *Ann Oncol* 2014;25(suppl 4): LBA3_PR.

OncoTargets and Therapy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Efficacy and safety of vascular endothelial growth factor receptor tyrosine kinase inhibitors in the treatment of advanced thyroid cancer: a meta-analysis of randomized controlled trials

This article was published in the following Dove Press journal:
OncoTargets and Therapy
7 March 2016
Number of times this article has been viewed

Wufuer Yimaer*
Aizizi Abudouyimu*
Ye Tian
Sailike Magaoweiya
Duman Bagedati
Hao Wen

Department of Vascular Thyroid Surgery, Gastrointestinal Vascular Center, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang Province, People's Republic of China

*These authors contributed equally to this work

Background: We performed a systematic review and meta-analysis to determine the efficacy and safety of the US Food and Drug Administration approved vascular endothelial growth factor receptor tyrosine kinase inhibitors (VEGFR-TKIs) in the treatment of advanced thyroid cancer.

Patients and methods: We included prospective randomized controlled trials that compared VEGFR-TKIs with placebo for advanced thyroid cancer. The endpoints included safety (fatal adverse events [FAEs], treatment discontinuation, and any severe [grade 3 or 4] adverse events [AEs]) and efficacy (objective response rate, progression-free survival, and overall survival). The pooled relative risk (RR) or hazard ratio (HR) was calculated by using either random-effects or fixed-effects models according to the heterogeneity of included studies.

Results: A total of 1,614 advanced thyroid cancer patients from five randomized controlled trials were identified for analysis. Compared with placebo alone, VEGFR-TKIs significantly increased the risk of treatment discontinuation (RR: 3.80, 95% confidence interval [CI]: 2.56–5.65, $P<0.001$) and any severe AEs (RR: 2.63, 95% CI: 1.72–4.03, $P<0.001$), but not of FAEs (RR: 1.24, 95% CI: 0.65–2.39, $P=0.52$). The use of VEGFR-TKIs in advanced thyroid cancer was associated with a significant improvement in objective response rate (RR: 8.73, 95% CI: 1.72–44.4, $P=0.009$) and progression-free survival (HR: 0.41, 95% CI: 0.27–0.61, $P<0.001$), with a tendency to improve overall survival (HR: 0.83, 95% CI: 0.68–1.01, $P=0.06$).

Conclusion: The use of small-molecule VEGFR-TKIs in advanced thyroid cancer did significantly increase the risk of treatment discontinuation and any severe AEs, but not of FAEs, compared with placebo alone. It is important for physicians to weigh the risk of toxicities as well as the potential survival benefits associated with VEGFR-TKI treatment in advanced thyroid cancer patients.

Keywords: angiogenesis inhibitors, toxicity, clinical trials, thyroid cancer, meta-analysis

Introduction

Thyroid cancer is the most common neoplasm of the endocrine system with incidence rates steadily increasing over the past 10 years.[1] In 2014, ~62,980 new cases of thyroid cancer were diagnosed and ~1,890 cancer deaths occurred from the disease in USA.[2] Although the prognosis is excellent for the majority of patients treated by surgery, thyroid-stimulating hormone-suppressive therapy, and radioiodine ablation, with an overall survival rate of 97.7% at 5 years,[3] local recurrence occurs in up to 20% of patients and distant metastases in ~10% at 10 years.[4] Until now, the medical approach

Correspondence: Hao Wen
Gastrointestinal Vascular Center, The First Affiliated Hospital of Xinjiang Medical University, No 137, South Liyushan road, Urumqi, Xinjiang province 830054, People's Republic of China
Tel +86 991 436 2974
Fax +86 991 436 4780
Email haowen20151212@sina.com

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/OTT.S102265

for the treatment of advanced or metastatic thyroid cancer refractory to conventional treatment is considered particularly challenging and few therapeutic options are available for these patients. Historically, the role of cytotoxic chemotherapy has been quite limited in these patients due to low efficacy and unfavorable toxicity profile when used.[5]

In the past decades, a better understanding of the molecular events involved in the tumorigenesis of thyroid cancers has led to development of new targeted agents for the management of advanced and refractory disease. Previous research has shown that vascular endothelial growth factor (VEGF) is overexpressed and its main receptor VEGFR-2 is upregulated in many thyroid cancers, which is associated with neoplastic progression and aggressiveness.[6] The VEGF and its receptors are, therefore, regarded as attractive therapeutic targets in the treatment of thyroid cancers.[7] Since 2011, four tyrosine kinase inhibitors (TKIs) targeting vascular endothelial growth factor receptor (VEGFR) have been approved by the US Food and Drug Administration for thyroid cancer: cabozantinib and vandetanib for medullary thyroid cancer and sorafenib and lenvatinib for differentiated thyroid cancer. All of the four drugs are multikinase inhibitors that act on multiple molecular pathways involved in growth, angiogenesis, and local and distant spread of thyroid cancer.[8] Sorafenib is a multitargeted TKI with inhibitory activity against VEGFR-2 and -3, c-Kit, platelet-derived growth factor receptor (PDGFR), rearranged during transfection (RET)/papillary thyroid carcinoma, and Raf kinases, and the Raf/Mek/Erk pathway (MAPK pathway).[9] Vandetanib has a low molecular weight and a good inhibitory activity against VEGFR-2, and targets VEGFR-3, EGFR, and RET kinases.[10] Sunitinib (SU011248) is a selective inhibitor of VEGFR-1, -2, and -3, PDGFR, c-Kit, and RET/papillary thyroid carcinoma subtypes 1 and 3.[11] Lenvatinib is an oral, multitargeted TKI of VEGFR-1, -2, and -3, fibroblast growth factor receptor-1, -2, -3, and -4, PDGFR-α, RET, and KIT.[12] To our best knowledge, there is no meta-analysis to assess the overall efficacy and toxicities of these four approved VEGFR-TKIs in advanced thyroid cancer. We, therefore, conducted this comprehensive meta-analysis to assess the efficacy and toxicities of approved VEGFR-TKIs in advanced thyroid cancer.

Methods

Data sources

Selection of studies

The Cochrane Central Register of Controlled Trials, PubMed (up to October 2015), and Web of Science (up to October 2015) databases were searched for articles. The search was extended to abstracts from oncology meetings containing the same terms ("VEGFR-TKIs", "vandetanib", "sorafenib", "lenvatinib", "cabozantinib", "advanced thyroid cancer", "metastatic thyroid cancer", "randomized controlled trial", and "humans"). Using the same search terms, we also searched abstracts and virtual meeting presentations from the American Society of Clinical Oncology conferences held up to October 2015 in order to identify relevant trials. An independent search of the Web of Science, Embase, and Cochrane electronic databases was also performed to ensure that no additional clinical trials were overlooked.

Data extraction and clinical end points

Data extraction and analysis were conducted independently by two independent investigators and any discrepancy was resolved by consensus according to the Quality of Reporting of Meta-Analyses guidelines.[13]

Clinical trials that met the following criteria were included: 1) Phase II and III trials in patients with advanced thyroid cancer; 2) random assignment of participants to treatment with VEGFR-TKIs or placebo alone; and 3) reporting data for at least one of the safety or efficacy outcomes. Independent reviewers screened reports that included the key terms by their titles and abstracts for relevance. Then, full texts of the relevant articles were retrieved to assess eligibility.

For each study, the following information was extracted: year of publication; first author; number of enrolled subjects; number of patients in each arm; median age; doses of VEGFR-TKIs administered; combination drug; median progression-free survival (PFS) (time to progression if not available), median overall survival (OS), objective response rate (ORR), fatal adverse events (FAEs), hazard ratios (HRs) for PFS and OS, treatment discontinuation related to adverse events (AEs), and any severe AE. The quality of included trials was rated using the five-point Jadad scale, which was based on the reporting of randomization method, blinding method, and withdrawals and dropouts.[14]

Statistical analysis

Incidence, relative risk (RR), and corresponding 95% confidence intervals (CIs) were the summary measures of ORR, FAEs, treatment discontinuation related to AEs, and any severe (grade 3 or 4) AE. We calculated the RRs and CIs, comparing the incidence of each AE in patients assigned to VEGFR-TKIs with those assigned to placebo alone in the same trial. For one study that reported zero events in the treatment or control arm, we applied the classic half-integer correction to calculate the RR and variance.[15] The summary

measures of PFS and OS were HR and the corresponding 95% CIs, which were extracted from each randomized controlled trial (RCT). For each meta-analysis, the Cochran Q statistic and I^2 score were first calculated to determine heterogeneity among the proportions of the included trials.[16,17] For $P<0.10$ values of the Cochran Q statistic, the assumption of homogeneity was deemed invalid and a random-effects model was reported.[18] Otherwise, results from the fixed-effects model were reported. Finally, potential publication biases were evaluated for severe AEs using Begg's and Egger's tests.[19] A two-tailed P-value of <0.05 without adjustment for multiplicity was considered statistically significant. The results of the meta-analysis were reported as classic forest plots. All statistical analyses were performed by using Version 2 of the Comprehensive MetaAnalysis program (Biostat, Englewood, NJ, USA).

Results

Search results

A total of 146 studies were identified from the database search, of which 141 reports were retrieved for full-text evaluation. Five trials met the inclusion criteria and were included in this systematic review (Figure 1).[20–33] Table 1 shows the characteristics of the included studies. Overall, a total of 1,614 patients were included for analysis. According to the inclusion criteria of each trial, patients were required to have an adequate renal, hepatic, and hematologic function. In all trials, randomization was between doublet combination group and single agent group. The quality of each included study was roughly assessed according to Jadad score, and all of these trials were double-blind, placebo-controlled trials and thus had a Jadad score of 5.

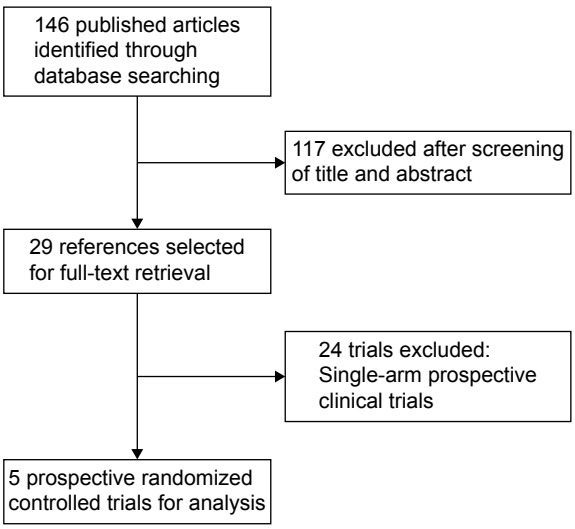


Figure 1 Studies eligible for inclusion in the meta-analysis.

Safety of VEGFR-TKIs versus placebo

Fatal adverse events

FAEs were diagnosed in 43 patients: 31 (2.7%, 95% CI: 1.2%–6.3%) in VEGFR-TKI arms and 12 (1.7%, 95% CI: 0.5%–5.8%) in placebo arms. The RR obtained for the studies ranged from 1.01 to 6.55. Overall, no increased risk was observed for the studies (RR=1.24; 95% CI: 0.65–2.39; P=0.52) (Figure 2A) using a fixed-effects model (I^2=0, P=0.81).

Any severe AEs

The incidence of any severe AE related to VEGFR-TKIs and placebo alone was, respectively, 52.2% (95% CI: 43.3%–60.8%) and 46.6% (95% CI: 32.9%–60.9%) by using the random-effects model. The use of VEGFR-TKIs significantly increased the risk of any severe AEs, when compared to placebo (RR=2.63, 95% CI: 1.72–4.03, $P<0.001$) (Figure 2B) using a random-effects model (I^2=79.7, P=0.001).

Treatment discontinuation

The incidence of treatment discontinuation due to VEGFR-TKIs and placebo alone was, respectively, 17.7% (95% CI: 13.0%–23.8%) and 4.6% (95% CI: 2.9%–7.2%) by using the random-effects model. The risk of discontinuing treatment because of AEs was higher with the use of VEGFR-TKIs compared with the controls (RR: 3.80, 95% CI: 2.56–5.65, $P<0.001$) (Figure 2C). The test for heterogeneity was nonsignificant and a fixed-effects model was used (I^2=25.6, P=0.25).

Efficacy of VEGFR-TKIs versus placebo

Overall survival

The pooled HR for OS did not show significant difference between VEGFR-TKIs and placebo alone (HR: 0.83, 95% CI: 0.68–1.01, P=0.06) (Figure 3A). The fixed-effects model was used because there was no significant heterogeneity (P=0.90, I^2=0).

Progression-free survival

In comparison with placebo alone, VEGFR-TKIs significantly improved PFS (HR: 0.41, 95% CI: 0.27–0.61, $P<0.001$) (Figure 3B). The test for heterogeneity was significant and a random-effects model was used ($P<0.001$, I^2=89.3).

Objective response rate

In comparison with placebo, the use of VEGFR-TKIs significantly improved ORR (RR: 8.73, 95% CI: 1.72–44.4, P=0.009) (Figure 3C). The test for heterogeneity was

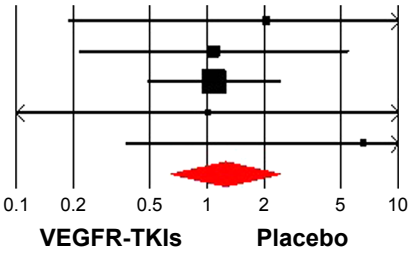
Table 1 Baseline characteristics of the five trials included for analysis

Author (year)	Phase	Total patients	Age (years)	Treatment regimens	No for analysis	FAEs	Median PFS (months)	Jadad score
Leboulleux et al[32] (2012)	II	145	63	Vandetanib 300 mg qd po	72	2	11.1	5
			64	Placebo	73	1	5.9	
Wells et al[30] (2012)	III	331	50.7	Vandetanib 300 mg qd po	231	5	30.5	5
			53.4	Placebo	100	2	19.3	
Elisei et al[29] (2013)	III	330	55	Cabozantinib 140 mg qd po	214	17	11.4	5
			55	Placebo	109	8	4	
Brose et al[31] (2014)	III	416	63	Sorafenib 400 mg bid po	207	12	10.8	5
			63	Placebo	209	6	5.8	
Schlumberger et al[33] (2015)	III	392	64	Lenvatinib 24 mg qd po	261	6	18.3	5
			61	Placebo	131	0	3.6	

Abbreviations: bid, twice daily; FAEs, fatal adverse events; PFS, progression-free survival.

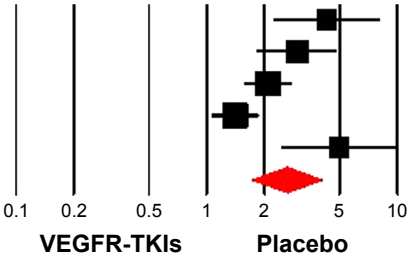
A

Study name	Statistics for each study					Events/total		Risk ratio and 95% CI
	Risk ratio	Lower limit	Upper limit	*Z*-value	*P*-value	Group-A	Group-B	
Leboulleux et al[32]	2.028	0.188	21.872	0.583	0.560	2/72	1/73	
Wells et al[30]	1.082	0.214	5.485	0.095	0.924	5/231	2/100	
Elisei et al[29]	1.082	0.482	2.428	0.192	0.848	17/214	8/109	
Brose et al[31]	1.010	0.064	16.035	0.007	0.995	1/207	1/209	
Schlumberger et al[33]	6.550	0.372	115.380	1.284	0.199	6/261	0/131	
Pooled results	1.241	0.646	2.386	0.648	0.517	31/985	12/622	



B

Study name	Statistics for each study					Events/total		Risk ratio and 95% CI
	Risk ratio	Lower limit	Upper limit	*Z*-value	*P*-value	Group-A	Group-B	
Leboulleux et al[32]	4.281	2.236	8.197	4.388	0.000	38/72	9/73	
Wells et al[30]	2.973	1.824	4.844	4.374	0.000	103/231	15/100	
Elisei et al[29]	2.094	1.580	2.776	5.138	0.000	148/214	36/109	
Brose et al[31]	1.414	1.060	1.885	2.357	0.018	77/207	55/209	
Schlumberger et al[33]	4.956	2.471	9.944	4.506	0.000	79/261	8/131	
Pooled results	2.631	1.720	4.025	4.459	0.000	445/985	123/622	



C

Study name	Statistics for each study					Events/total		Risk ratio and 95% CI
	Risk ratio	Lower limit	Upper limit	*Z*-value	*P*-value	Group-A	Group-B	
Leboulleux et al[32]	6.083	2.222	16.656	3.514	0.000	24/72	4/73	
Wells et al[30]	4.040	1.257	12.983	2.345	0.019	28/231	3/100	
Elisei et al[29]	1.981	0.988	3.970	1.927	0.054	35/214	9/109	
Brose et al[31]	4.922	2.358	10.276	4.244	0.000	39/207	8/209	
Schlumberger et al[33]	6.190	1.945	19.701	3.086	0.002	37/261	3/131	
Pooled results	3.802	2.559	5.648	6.611	0.000	163/985	27/622	

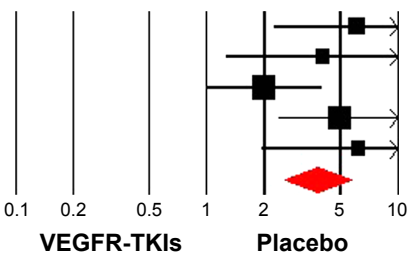


Figure 2 Risk of severe adverse outcomes associated with VEGFR-TKIs treatment compared with placebo treatment: (**A**) FAEs, (**B**) any severe adverse events, and (**C**) treatment discontinuation.
Notes: Group A: VEGFR-TKIs group; Group-B: placebo group.
Abbreviations: CI, confidence interval; FAEs, fatal adverse events; VEGFR-TKIs, vascular endothelial growth factor receptor tyrosine kinase inhibitors.

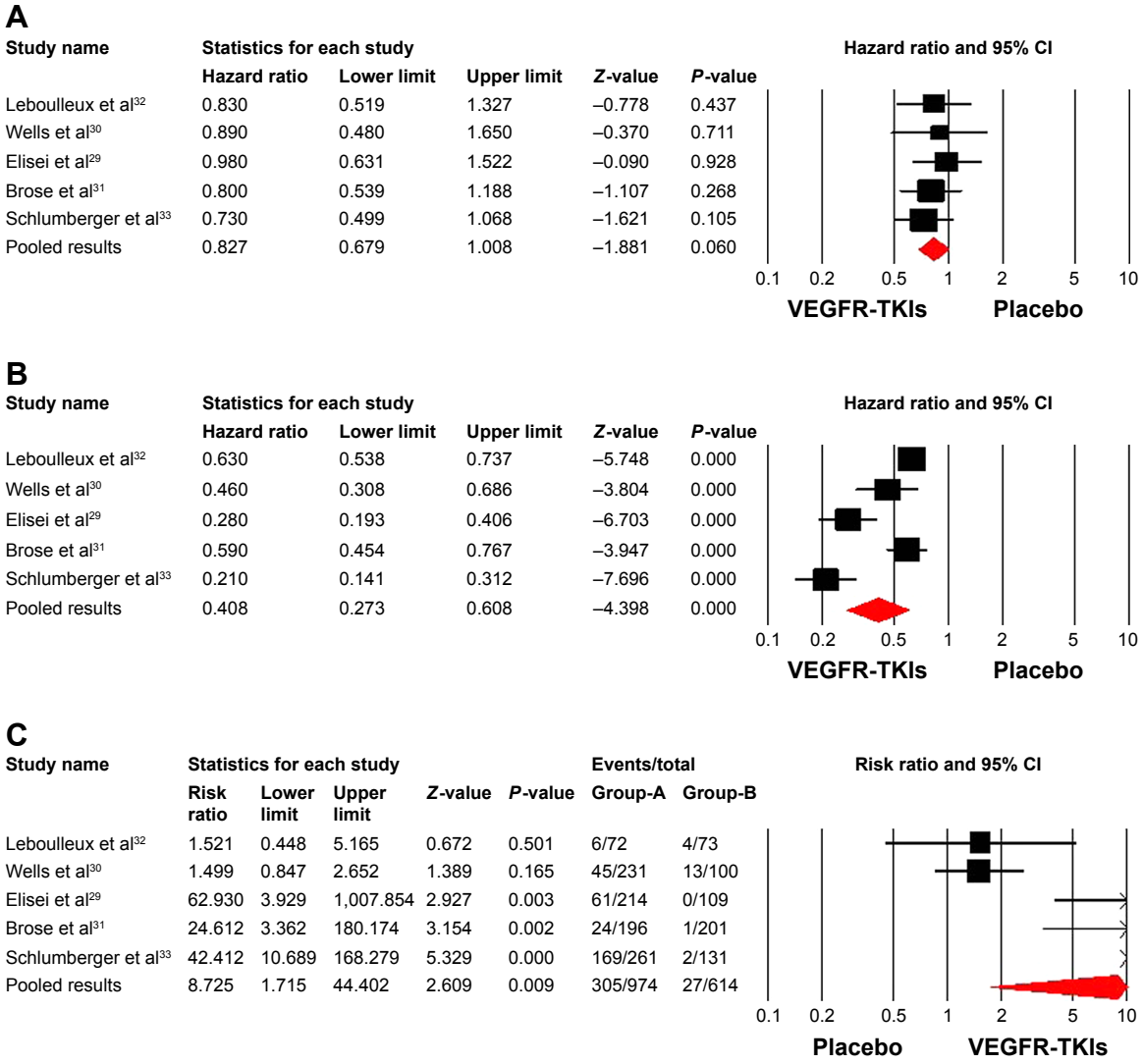


Figure 3 Efficacy associated with VEGFR-TKIs treatment compared with placebo treatment: (**A**) OS, (**B**) PFS, (**C**) ORR.
Abbreviations: CI, confidence interval; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; VEGFR-TKIs, vascular endothelial growth factor receptor tyrosine kinase inhibitors.

significant and a random-effects model was used (I^2=86.6, P<0.001).

Publication bias

No publication bias was detected with Begg's or Egger's test for the efficacy and AEs studied, except for any severe AEs (Begg's test, P=0.05; Egger's test, P=0.04) (Table 2).

Discussion

Increased vascularity has been reported in thyroid cancer. Angiogenesis, especially VEGF signal pathway, plays a pivotal role in tumor growth, progression, and metastasis.[34,35] Previous research had demonstrated that thyroid cancer cell lines were characterized by high expression of both VEGF and its receptors.[36] Thus, the VEGF signal pathway has been targeted as a therapeutic option for thyroid cancer. In fact, four VEGFR-TKIs including vandetanib, sorafenib, lenvatinib, and cabozantinib have been approved by the US Food and Drug Administration for use in radioiodine-refractory differentiated thyroid cancer or medullary thyroid cancer;[37–39] thus, it is anticipated that the use of VEGFR-TKIs would be increasing

Table 2 Publication bias by Begg's and Egger's tests (*P*-value)

	Begg's test	Egger's test
Overall survival	0.14	0.37
Progression-free survival	0.14	0.08
Objective response rate	0.22	0.10
Fatal adverse event	0.08	0.23
Treatment discontinue	0.62	0.27
Any severe adverse events	0.05	0.04

in the near future. In contrast with traditional chemotherapy agents, VEGFR-TKIs present an anti-VEGF toxicity profile, such as hypertension,[40–42] proteinuria,[43] thrombosis,[44,45] and hemorrhage.[46] However, the toxicities associated with VEGFR-TKIs in advanced thyroid cancer remains unknown. Moreover, the overall efficacy of VEGFR-TKIs in these patients has not been comprehensively assessed.

Our study, which included 1,614 patients from five RCTs, demonstrates that the use of VEGFR-TKIs in advanced thyroid cancer significantly improves ORR and PFS, and there is also a tendency to improve OS in comparison with the placebo groups. Safety of systematic treatments is of particular importance in palliative setting in advanced thyroid cancer patients, given the potential negative impact on benefit ratio and quality of life. As for toxicities, a previous meta-analysis conducted by Hong et al[47] reported that the use of VEGFR-TKIs significantly increased the risk of FAEs when compared with controls (odds ratio: 1.85, 95% CI: 1.33–2.58, $P<0.01$), while subgroup analysis according to tumor types showed that the use of VEGFR-TKIs did not significantly increase the risk of FAEs (odds ratio: 2.25, 95% CI: 0.61–8.30, $P=0.22$). Findings of our study indicate that the use of VEGFR-TKIs significantly increased the risk of treatment discontinuation and any severe AEs, but not of FAEs, which is consistent with the findings of a previous study. Based on our results, we conclude that VEGFR-TKIs could be recommended for use in advanced thyroid cancer due to their potential survival benefits, although the use of these drugs would increase the risk of developing treatment discontinuation and any severe AEs, but not of FAEs. Long-term follow-up studies for OS of advanced thyroid cancer patients receiving these VEGFR-TKIs are still needed because survival data in these published studies are immature at the time of analysis.

Our study has several limitations. First, this meta-analysis only considers published literature, and lack of individual patient data prevents us from adjusting the treatment effect according to disease and patient variables. Second, toxicity data in RCTs have been reported to be suboptimal and variable as toxicity is usually not the primary outcome measure. Furthermore, there is some degree of subjectivity in the process by which investigators in trials adjudicate whether a patient's death was the result of an AE, cancer progression, or other unrelated causes. Third, these studies exclude patients with poor renal, hematological, and hepatic functions, and are performed mostly at major academic centers and research institutions; the analysis of these studies may not apply to patients with organ dysfunctions and in the community. Finally, as in all meta-analyses, our results may be biased as a result of potential publication bias. However, a funnel plot evaluation for AEs and efficacy does not indicate publication bias except for any severe AEs.

Conclusion

In conclusion, the use of small-molecule VEGFR-TKIs in advanced thyroid cancer does significantly increase the risk of developing treatment discontinuation and any severe AEs, but not of FAEs, compared with placebo alone. Additionally, the use of VEGFR-TKIs in advanced thyroid cancer significantly improves ORR and PFS, and has a tendency to improve OS. These observations may aid medical oncologists in weighing up the risks and benefits associated with VEGFR-TKIs in treating patients with advanced thyroid cancer.

Acknowledgment

No funding has been received for this study.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. *J Cancer Epidemiol*. 2013;2013:965212.
2. Tuttle RM, Haddad RI, Ball DW, et al. Thyroid carcinoma, version 2. 2014. *J Natl Compr Canc Netw*. 2014;12(12):1671–1680; quiz 1680.
3. Stjepanovic N, Capdevila J. Multikinase inhibitors in the treatment of thyroid cancer: specific role of lenvatinib. *Biologics*. 2014;8:129–139.
4. Eustatia-Rutten CF, Corssmit EP, Biermasz NR, Pereira AM, Romijn JA, Smit JW. Survival and death causes in differentiated thyroid carcinoma. *J Clin Endocrinol Metab*. 2006;91(1):313–319.
5. Sherman SI. Cytotoxic chemotherapy for differentiated thyroid carcinoma. *Clin Oncol (R Coll Radiol)*. 2010;22(6):464–468.
6. Minucci S, Pelicci PG. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer*. 2006;6(1):38–51.
7. Fallahi P, Mazzi V, Vita R, et al. New therapies for dedifferentiated papillary thyroid cancer. *Int J Mol Sci*. 2015;16(3):6153–6182.
8. Ferrari SM, Fallahi P, Politti U, et al. Molecular targeted therapies of aggressive thyroid cancer. *Front Endocrinol*. 2015;6:176.
9. Fallahi P, Ferrari SM, Santini F, et al. Sorafenib and thyroid cancer. *BioDrugs*. 2013;27(6):615–628.
10. Carlomagno F, Vitagliano D, Guida T, et al. ZD6474, an orally available inhibitor of KDR tyrosine kinase activity, efficiently blocks oncogenic RET kinases. *Cancer Res*. 2002;62(24):7284–7290.
11. Rini BI. Sunitinib. *Expert Opin Pharmacother*. 2007;8(14):2359–2369.
12. Matsui J, Yamamoto Y, Funahashi Y, et al. E7080, a novel inhibitor that targets multiple kinases, has potent antitumor activities against stem cell factor producing human small cell lung cancer H146, based on angiogenesis inhibition. *Int J Cancer*. 2008;122(3):664–671.
13. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet*. 1999;354(9193):1896–1900.
14. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17(1):1–12.

15. Seng S, Liu Z, Chiu SK, et al. Risk of venous thromboembolism in patients with cancer treated with cisplatin: a systematic review and meta-analysis. *J Clin Oncol*. 2012;30(35):4416–4426.
16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
17. Zintzaras E, Ioannidis JP. Heterogeneity testing in meta-analysis of genome searches. *Genet Epidemiol*. 2005;28(2):123–137.
18. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. *Contemp Clin Trials*. 2015;45:139–145.
19. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–634.
20. Zinner RG, Obasaju CK, Spigel DR, et al. PRONOUNCE: randomized, open-label, phase III study of first-line pemetrexed + carboplatin followed by maintenance pemetrexed versus paclitaxel + carboplatin + bevacizumab followed by maintenance bevacizumab in patients ith advanced nonsquamous non-small-cell lung cancer. *J Thorac Oncol*. 2015;10(1):134–142.
21. Zhou C, Wu YL, Chen G, et al. BEYOND: a randomized, double-blind, placebo-controlled, multicenter, phase III study of first-line carboplatin/paclitaxel plus bevacizumab or placebo in Chinese patients with advanced or recurrent nonsquamous non-small-cell lung cancer. *J Clin Oncol*. 2015;33(19):2197–2204.
22. Galetta D, Cinieri S, Pisconti S, et al. Cisplatin/pemetrexed followed by maintenance pemetrexed versus carboplatin/paclitaxel/bevacizumab followed by maintenance bevacizumab in advanced nonsquamous lung cancer: The GOIM (Gruppo Oncologico Italia Meridionale) ERACLE phase III randomized trial. *Clin Lung Cancer*. 2015;16(4):262–273.
23. Seto T, Kato T, Nishio M, et al. Erlotinib alone or with bevacizumab as first-line therapy in patients with advanced non-squamous non-small-cell lung cancer harbouring EGFR mutations (JO25567): an open-label, randomised, multicentre, phase 2 study. *Lancet Oncol*. 2014;15(11):1236–1244.
24. Niho S, Kunitoh H, Nokihara H, et al. Randomized phase II study of first-line carboplatin-paclitaxel with or without bevacizumab in Japanese patients with advanced non-squamous non-small-cell lung cancer. *Lung Cancer*. 2012;76(3):362–367.
25. Herbst RS, Ansari R, Bustin F, et al. Efficacy of bevacizumab plus erlotinib versus erlotinib alone in advanced non-small-cell lung cancer after failure of standard first-line chemotherapy (BeTa): a double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2011;377(9780):1846–1854.
26. Reck M, von Pawel J, Zatloukal P, et al. Phase III trial of cisplatin plus gemcitabine with either placebo or bevacizumab as first-line therapy for nonsquamous non-small-cell lung cancer: AVAil. *J Clin Oncol*. 2009;27(8):1227–1234.
27. Sandler A, Gray R, Perry MC, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *N Engl J Med*. 2006;355(24):2542–2550.
28. Johnson DH, Fehrenbacher L, Novotny WF, et al. Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small-cell lung cancer. *J Clin Oncol*. 2004;22(11):2184–2191.
29. Elisei R, Schlumberger MJ, Muller SP, et al. Cabozantinib in progressive medullary thyroid cancer. *J Clin Oncol*. 2013;31(29):3639–3646.
30. Wells SA Jr, Robinson BG, Gagel RF, et al. Vandetanib in patients with locally advanced or metastatic medullary thyroid cancer: a randomized, double-blind phase III trial. *J Clin Oncol*. 2012;30(2):134–141.
31. Brose MS, Nutting CM, Jarzab B, et al. Sorafenib in radioactive iodine-refractory, locally advanced or metastatic differentiated thyroid cancer: a randomised, double-blind, phase 3 trial. *Lancet*. 2014;384(9940):319–328.
32. Leboulleux S, Bastholt L, Krause T, et al. Vandetanib in locally advanced or metastatic differentiated thyroid cancer: a randomised, double-blind, phase 2 trial. *Lancet Oncol*. 2012;13(9):897–905.
33. Schlumberger M, Tahara M, Wirth LJ, et al. Lenvatinib versus placebo in radioiodine-refractory thyroid cancer. *N Engl J Med*. 2015;372(7):621–630.
34. Folkman J. Tumor angiogenesis: therapeutic implications. *N Engl J Med*. 1971;285(21):1182–1186.
35. Folkman J. Anti-angiogenesis: new concept for therapy of solid tumors. *Ann Surg*. 1972;175(3):409–416.
36. Lin JD, Chao TC. Vascular endothelial growth factor in thyroid cancers. *Cancer Biother Radiopharm*. 2005;20(6):648–661.
37. Weitzman SP, Cabanillas ME. The treatment landscape in thyroid cancer: a focus on cabozantinib. *Cancer Manag Res*. 2015;7:265–278.
38. Yeung KT, Cohen EE. Lenvatinib in advanced, radioactive iodine-refractory, differentiated thyroid carcinoma. *Clin Cancer Res*. 2015;21(24):5420–5426.
39. Ferrari SM, Politti U, Spisni R, et al. Sorafenib in the treatment of thyroid cancer. *Expert Rev Anticancer Ther*. 2015;15(8):863–874.
40. Qi WX, Shen Z, Lin F, et al. Incidence and risk of hypertension with vandetanib in cancer patients: a systematic review and meta-analysis of clinical trials. *Br J Clin Pharmacol*. 2013;75(4):919–930.
41. Qi WX, Lin F, Sun YJ, et al. Incidence and risk of hypertension with pazopanib in patients with cancer: a meta-analysis. *Cancer Chemother Pharmacol*. 2013;71(2):431–439.
42. Wu S, Chen JJ, Kudelka A, Lu J, Zhu X. Incidence and risk of hypertension with sorafenib in patients with cancer: a systematic review and meta-analysis. *Lancet Oncol*. 2008;9(2):117–123.
43. Zhang ZF, Wang T, Liu LH, Guo HQ. Risks of proteinuria associated with vascular endothelial growth factor receptor tyrosine kinase inhibitors in cancer patients: a systematic review and meta-analysis. *PLoS One*. 2014;9(3):e90135.
44. Sonpavde G, Je Y, Schutz F, et al. Venous thromboembolic events with vascular endothelial growth factor receptor tyrosine kinase inhibitors: a systematic review and meta-analysis of randomized clinical trials. *Crit Rev Oncol Hematol*. 2013;87(1):80–89.
45. Qi WX, Min DL, Shen Z, et al. Risk of venous thromboembolic events associated with VEGFR-TKIs: A systematic review and meta-analysis. *Int J Cancer*. 2013;132(12):2967–2974.
46. Qi WX, Tang LN, Sun YJ, et al. Incidence and risk of hemorrhagic events with vascular endothelial growth factor receptor tyrosine-kinase inhibitors: an up-to-date meta-analysis of 27 randomized controlled trials. *Ann Oncol*. 2013;24(12):2943–2952.
47. Hong S, Fang W, Liang W, et al. Risk of treatment-related deaths with vascular endothelial growth factor receptor tyrosine kinase inhibitors: a meta-analysis of 41 randomized controlled trials. *Onco Targets Ther*. 2014;7:1851–1867.

The Laryngoscope

Gastroesophageal Reflux Disease and Odds of Head and Neck Squamous Cell Carcinoma in North Carolina

Evan L. Busch, PhD; Jose P. Zevallos, MD, MPH; Andrew F. Olshan, PhD

Objectives/Hypothesis: Exposure to excess gastric acid resulting from gastroesophageal reflux disease, also known as acid reflux or heartburn, might contribute to initiation of head and neck squamous cell carcinoma, particularly laryngeal cancer. Prior epidemiologic studies have reported inconsistent results. We sought to clarify this relationship using an observational study with a larger available sample size and better-characterized exposure information than most prior studies.

Study Design: A population-based case-control study of head and neck cancer in North Carolina with 1,340 newly diagnosed cases and 1,378 controls matched on age, race, and sex.

Methods: We used unconditional logistic regression to examine associations between self-reported heartburn and development of overall head and neck cancer as well as development of cancer at specific tumor sites. Subgroup analysis by smoking and alcoholic drinking status was used to make comparisons with a previous study that used a similar study design.

Results: Overall, an increased odds of head and neck cancer was not associated with either self-reported history of heartburn symptoms (odds ratio = 0.85; 95% confidence interval 0.68, 1.06) or self-reported medical diagnosis of GERD (OR = 0.89; 95% CI 0.71, 1.11). These patterns held for specific tumor sites. For laryngopharyngeal cancer, we did not detect any associations regardless of joint smoking and alcoholic drinking status.

Conclusion: Gastroesophageal reflux does not appear to play a role in development of head and neck cancer.

Key Words: Gastroesophageal reflux disease, head and neck squamous cell carcinoma, self-reported measures, epidemiology, population-based studies.

Level of Evidence: 3b.

***Laryngoscope*, 126:1091–1096, 2016**

INTRODUCTION

Gastroesophageal reflux disease (GERD), also called acid reflux or heartburn, has been linked to increased risk of multiple complications such as esophageal stricture, coughing, and esophageal ulcers.[1,2] It consists of excess acid from the stomach passing up through the esophagus and into the upper aerodigestive tract. This acid exposure has been associated with carcinogenesis, most notably in relation to the development of Barrett's esophagus and subsequently to esophageal cancer.[3,4]

Thus, it is possible that GERD could contribute to the development of head and neck squamous cell carcinoma (HNSCC).[5,6] More specifically, reflux of gastric acid is known to affect the larynx and cause laryngopharyngeal reflux.[7] A large cohort study found that, when compared to the general population, GERD patients had greater incidence of oropharyngeal and hypopharyngeal cancers.[8] Additionally, cell-line studies have shown that gastric acid is carcinogenic for both laryngeal[9] and hypopharyngeal cells.[9,10] Unlike the esophagus, the larynx lacks protective mechanisms against acid such as mucus, peristalsis, and carbonic anhydrase enzyme.[7] Due to its proximity to the upper esophagus, it has been suggested that the larynx could be at higher risk for GERD-based carcinogenesis compared to the oropharynx or oral cavity.

To further address these questions, we examined the associations between GERD and the development of HNSCC in a large population-based case-control study of HNSCC. Relationships between GERD and both overall HNSCC as well as specific tumor sites within the head and neck were evaluated. We hypothesized that a history of having GERD would be associated with greater odds of developing HNSCC, especially laryngeal cancer.

MATERIALS AND METHODS

Study Population

Subjects were drawn from the Carolina Head and Neck Cancer Epidemiology (CHANCE) study, a population-based case-control study that enrolled 1,368 incident cases of HNSCC aged 20 to 80 in a 46-county region of North Carolina during 2002 to 2006.[11,12] Cases were identified by a rapid-case

From the Department of Epidemiology (E.L.B., J.P.Z., A.F.O.); the Department of Otolaryngology/Head and Neck Surgery (J.P.Z., A.F.O.), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; the Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School (E.L.B.); and the Department of Epidemiology, Harvard T.H. Chan School of Public Health (E.L.B.), Boston, Massachusetts, U.S.A.

Editor's Note: This Manuscript was accepted for publication September 8, 2015.

Financial Support: This study was supported in part by a grant from the National Cancer Institute (R01-CA90731). E.L.B. was supported in part by a grant from the National Cancer Institute (5T32CA009001). The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Jose P. Zevallos, MD, MPH, FACS, Departments of Otolaryngology/Head and Neck Surgery and Epidemiology, University of North Carolina at Chapel Hill, Physicians Office Building, CB #7070, Chapel Hill, NC 27599. Email: jose_zevallos@med.unc.edu

DOI: 10.1002/lary.25716

ascertainment system through the North Carolina Central Cancer Registry and by contacting cancer registrars at 54 hospitals in the 46 counties during the study period. To be eligible as cases, subjects had to have received a diagnosis of first primary invasive squamous cell carcinoma of the larynx (International Classification of Disease for Oncology, 3rd Edition, topography codes C32.0–C32.9) or oral cavity or pharynx (codes C0.00–C14.8). The study enrolled 1,396 controls who were frequency-matched to cases on age, race, and sex using stratified random sampling. Controls were identified through the North Carolina Department of Motor Vehicle records as residents of the study region aged 20 to 80 years old who had never received a diagnosis of HNSCC. The study collected questionnaire data. Due to the sparse numbers, the present analysis excluded 28 cases and 18 controls whose race was not white or black.

The institutional review board at the University of North Carolina at Chapel Hill approved the protocol. All subjects provided informed consent.

Gastroesophageal Reflux Disease Measures

Gastroesophageal reflux disease exposure was assessed via two different questionnaire items administered in-person by a nurse-interviewer within 2 months of diagnosis. The first question, considered a measure of self-reported GERD symptoms, was: "Were you ever bothered by frequent heartburn?" The second question, considered a measure of medical diagnosis, was: "Did your doctor ever tell you that you had GERD?" Both items were answered "yes," "no," "refused," "don't know," or were recorded as missing. For purposes of analysis, we recoded responses of "refused" or "don't know" as missing.

Covariates

Variables that were considered to be common causes of GERD and HNSCC incidence were selected a priori to include as confounders in multivariable models. All covariates were measured at baseline interview. These included age (categorized as 20–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–80), sex, race (white or black), years smoked cigarettes (never smoker, 1–19 years, 20–39 years, 40–49 years, and 50+ years), lifetime alcohol consumption as described previously (never had alcohol; <11,232 mL; 11,232-<204,469 mL; 204,469-<927,946 mL; 927,946+ mL),[12] body mass index (< 18.5, 18.5– < 25.0, 25.0– < 30.0, and 30.0+), and education (less than high school, high school graduate/vocational training/technical training, and at least some college).

To assign alcohol consumption status for the subgroup analysis by joint alcohol consumption and smoking history status, alcohol consumption was measured in terms of 12-ounce beers, 5-ounce wines, and 1.5-ounce hard liquors per week to more closely approximate the definition of Langevin et al.[6]

Outcomes

Case-control status was the outcome variable. Some analyses examined the associations with overall case-control status (any HNSCC case vs. controls), whereas others examined the associations with specific HNSCC tumor sites (laryngeal, hypopharyngeal, oropharyngeal, or oral cavity) compared to controls. Further analyses combined hypopharyngeal and oropharyngeal cases into overall pharyngeal cases and compared them to controls. Subgroup analyses by joint alcohol consumption and smoking status compared combined laryngeal and pharyngeal cases (i.e., laryngopharyngeal cases) to controls.

The CHANCE enrolled 251 cases designated as not otherwise specified (NOS), that is, those whose tumors could not be assigned to a particular tumor site. Of these, 247 cases were eligible for inclusion in the present analysis. We included NOS cases in the overall case-control status variable but excluded them from tumor site-specific analyses.

Statistical Analysis

Distributions of all variables included in statistical models were computed as frequencies and percentages for overall cases, tumor site-specific cases, and controls. The covariate distributions of overall HNSCC cases and controls were compared using chi-square tests.

To evaluate associations between GERD and overall case-control status, we used standard unconditional logistic regression for a dichotomous outcome to estimate odds ratios (OR) and 95% confidence interval (CI). For analyses of relationships between GERD and specific tumor sites, we used polytomous logistic regression to compare each of laryngeal, hypopharyngeal, oropharyngeal, overall pharyngeal, or oral cavity cases, respectively, to controls. Different multilevel tumor site variables were constructed to include, on the one hand, hypopharyngeal and oropharyngeal cases as separate categories, and on the other hand, overall pharyngeal cases.

To allow comparison with the study by Langevin et al.,[6] we conducted an analysis of GERD with joint stratification by alcohol consumption and smoking history comparing laryngeal and pharyngeal cases combined to controls. Similar to the Langevin study, heavy drinkers were defined as those consuming more than 14 alcoholic drinks per week. One alcoholic drink was defined as, equivalently, 12 ounces of beer, 5 ounces of wine, or 1.5 ounces of hard liquor. Also per Langevin et al., heavy smokers were defined as subjects with more than 18.3 pack-years of cigarette use.

Every model adjusted for all of the confounders described above. In addition, to account for the CHANCE frequency matching, each model adjusted for 2-way and 3-way interaction terms between the matching factors of age, sex, and race. Each model excluded subjects with incomplete information.

P values less than 0.05 were considered statistically significant. All analyses were performed using SAS 9.3 (SAS Institute, Cary, NC).

RESULTS

The study population included 1,340 head and neck cancer cases and 1,378 controls. The site distribution for cases was as follows: 473 larynx, 361 oropharynx, 192 oral cavity, 67 hypopharynx, and 247 not-otherwise-specified site. Table I presents descriptive statistics for subject characteristics. Relative to controls, HNSCC cases smoked for a greater number of years, had greater lifetime alcohol consumption, and were less likely to have attended college. In this univariate analysis, we found no differences between cases and controls in terms of whether they self-reported having had GERD symptoms or received a medical diagnosis of GERD.

Using multivariable modeling, we found no associations between self-reported history of GERD symptoms and case-control status, either for overall case-control status or for specific tumor sites (Table II). Most of the ORs showed that cases had moderately decreased odds of exposure compared to controls. The OR for hypopharyngeal cancer showed an almost 50% increase in odds.

TABLE I.
Subject Characteristics.

Variable	Controls (N = 1,378)	All Cases (N = 1,340)	*P* Value*	Hypopharynx Cases (N = 67)	Larynx Cases (N = 473)	NOS Cases (N = 247)	Oral Cavity Cases (N = 192)	Oropharynx Cases (N = 361)
Age (years)			< 0.0001					
20–49	156 (11%)	254 (19%)		9 (13%)	64 (14%)	58 (23%)	37 (19%)	86 (24%)
50–54	161 (12%)	210 (16%)		16 (24%)	50 (11%)	40 (16%)	20 (10%)	84 (23%)
55–59	207 (15%)	222 (17%)		13 (19%)	71 (15%)	38 (15%)	31 (16%)	69 (19%)
60–64	205 (15%)	229 (17%)		10 (15%)	95 (20%)	41 (17%)	30 (16%)	53 (15%)
65–69	247 (18%)	178 (13%)		7 (10%)	81 (17%)	30 (12%)	25 (13%)	35 (10%)
70–74	231 (17%)	152 (11%)		3 (4%)	73 (15%)	22 (9%)	29 (15%)	25 (7%)
75–80	171 (12%)	95 (7%)		9 (13%)	39 (8%)	18 (7%)	20 (10%)	9 (2%)
Sex			0.0001					
Male	960 (70%)	1,021 (76%)		56 (84%)	372 (79%)	171 (69%)	123 (64%)	299 (83%)
Female	418 (30%)	319 (24%)		11 (16%)	101 (21%)	76 (31%)	69 (36%)	62 (17%)
Race			< 0.0001					
White	1,114 (81%)	989 (74%)		39 (58%)	345 (73%)	188 (76%)	140 (73%)	277 (77%)
Black	264 (19%)	351 (26%)		28 (42%)	128 (27%)	59 (24%)	52 (27%)	84 (23%)
Years Smoked Cigarettes			< 0.0001					
Never smoker	525 (38%)	173 (13%)		5 (8%)	19 (4%)	57 (23%)	21 (11%)	71 (20%)
1–19	293 (21%)	118 (9%)		3 (5%)	26 (6%)	30 (12%)	12 (6%)	47 (13%)
20–39	334 (24%)	499 (38%)		26 (41%)	180 (38%)	68 (28%)	79 (42%)	146 (41%)
40–49	142 (10%)	344 (26%)		19 (30%)	150 (32%)	65 (27%)	47 (25%)	63 (18%)
50+	78 (6%)	194 (15%)		10 (16%)	97 (21%)	25 (10%)	31 (16%)	31 (9%)
Lifetime Alcohol Consumption (mL)			< 0.0001					
Never had alcohol	296 (22%)	125 (10%)		1 (2%)	45 (10%)	27 (12%)	23 (13%)	29 (9%)
<11,232	161 (12%)	58 (5%)		3 (5%)	21 (5%)	15 (7%)	6 (3%)	13 (4%)
11,232-<204,469	406 (31%)	234 (19%)		3 (5%)	77 (18%)	46 (20%)	23 (13%)	85 (25%)
204,469-<927,946	321 (24%)	319 (26%)		12 (20%)	122 (28%)	53 (23%)	42 (23%)	90 (27%)
927,946+	144 (11%)	497 (40%)		40 (68%)	169 (39%)	85 (38%)	86 (48%)	117 (35%)
Body Mass Index			<0.0001					
Underweight (<18.5)	30 (2%)	100 (7%)		14 (21%)	25 (5%)	20 (8%)	20 (10%)	21 (6%)
Normal (18.5- <25.0)	405 (29%)	482 (36%)		26 (39%)	161 (34%)	93 (38%)	88 (46%)	114 (32%)
Overweight (25.0- <30.0)	551 (40%)	434 (32%)		19 (28%)	159 (34%)	71 (29%)	55 (29%)	130 (36%)
Obese (30.0+)	392 (28%)	324 (24%)		8 (12%)	128 (27%)	63 (26%)	29 (15%)	96 (27%)
Education			< 0.0001					
Less than high school	217 (16%)	458 (34%)		32 (48%)	197 (42%)	64 (26%)	67 (35%)	98 (27%)
High school/vocational/tech	490 (36%)	492 (37%)		18 (27%)	173 (37%)	89 (36%)	73 (38%)	139 (39%)
At least some college	671 (49%)	390 (29%)		17 (25%)	103 (22%)	94 (38%)	52 (27%)	124 (34%)
Ever had frequent heartburn			0.8					
No	1,007 (76%)	989 (77%)		43 (69%)	344 (75%)	194 (81%)	145 (81%)	263 (74%)
Yes	315 (24%)	303 (23%)		19 (31%)	112 (25%)	47 (20%)	34 (19%)	91 (26%)
Ever diagnosed with GERD			0.1					
No	994 (77%)	1,008 (79%)		53 (84%)	329 (73%)	200 (84%)	147 (83%)	279 (80%)
Yes	303 (23%)	266 (21%)		10 (16%)	120 (27%)	37 (16%)	31 (17%)	68 (20%)

*Chi-square comparisons between controls and overall cases.
NOS = Not otherwise specified, GERD = gastroesophageal reflux disease

We also found no association between self-reported medical diagnosis of GERD and the odds of overall HNSCC (Table II). Most ORs were again in an inverse direction, but laryngeal cases had a slightly greater odds of having been diagnosed with GERD compared to controls.

In analyses of combined laryngeal and pharyngeal cases, among those who were neither heavy smokers nor heavy drinkers we detected no association between GERD and the development of laryngopharyngeal cancer for either self-reported history of GERD symptoms or medical diagnosis of GERD (Table III). Likewise, no associations between GERD and laryngopharyngeal cancer were detected among subjects who were heavy smokers and/or heavy drinkers (Table III).

TABLE II.
Effects of Self-Reported Heartburn Symptoms and Medical Diagnosis of Gastroesophageal Reflux Disease on Odds of Developing Overall or Tumor Site-Specific Head and Neck Squamous Cell Carcinoma.

	Self-Reported History of Frequent Heartburn*				GERD Diagnosis*			
Cases	Exposed Cases (%)†	Exposed Controls (%)†	OR‡	95% CI	Exposed Cases (%)†	Exposed Controls (%)†	OR‡	95% CI
Overall	303 (23%)	315 (24%)	0.85	0.68, 1.06	266 (21%)	303 (23%)	0.89	0.71, 1.11
Hypopharynx	19 (31%)	315 (24%)	1.49	0.80, 2.79	10 (16%)	303 (23%)	0.74	0.34, 1.64
Larynx	112 (25%)	315 (24%)	0.88	0.65, 1.19	120 (27%)	303 (23%)	1.27	0.94, 1.70
Oral cavity	34 (19%)	315 (24%)	0.72	0.46, 1.11	31 (17%)	303 (23%)	0.85	0.54, 1.32
Oropharynx	91 (26%)	315 (24%)	0.92	0.68, 1.26	68 (20%)	303 (23%)	0.84	0.61, 1.18
Pharynx	110 (26%)	315 (24%)	0.99	0.73, 1.32	78 (19%)	303 (23%)	0.83	0.60, 1.14

*Recorded as dichotomous ever/never.
†Percentages exclude subjects with missing data.
‡Reference group is controls. Estimates adjusted for age, sex, race, years smoked cigarettes, lifetime alcohol consumption, body mass index, education, and 2-way and 3-way interaction terms between age/sex/race.
CI = confidence interval, GERD = gastroesophageal reflux disease, OR = odds ratio.

DISCUSSION

We assessed associations between GERD exposure and the odds of developing HNSCC in a large, population-based case-control study for both overall HNSCC and specific head and neck tumor sites. We did not detect any strong positive associations between GERD and either development of overall HNSCC or development of cancer at any particular head and neck tumor site.

Although none of our associations was statistically significant, the magnitude of some of the point estimates was notable. The point estimate for the association between self-reported history of GERD symptoms and overall HNSCC was 0.85, and the point estimates for most specific tumor sites were clustered near to that value. However, the point estimate for hypopharyngeal cancer was elevated (1.49), suggesting that GERD could be associated with a greater odds of developing hypopharyngeal cancer relative to the other tumor sites that were examined.

When the exposure was medical diagnosis of GERD rather than self-reported history of GERD symptoms, the point estimate for the association of diagnosed GERD with overall HNSCC (0.89) was close to what it had been for self-reported history of GERD symptoms. Again, most of the point estimates for specific tumor sites were clustered around the null value. There were exceptions, however, with laryngeal cancer having an OR of 1.27 and hypopharyngeal cancer having an OR of 0.74.

Our findings for subgroup analyses by joint alcohol consumption and smoking status were not consistent with previous research. A study of 631 cases of laryngopharyngeal cancer conducted in the Boston area with a similar design to our North Carolina study found that, among subjects who were neither heavy drinkers nor heavy smokers, reporting a history of frequent heartburn was associated with a greater odds of developing laryngopharyngeal cancer (OR = 1.78; 95% CI 1.00, 3.16).[6] In our analysis, no association between heartburn and laryngopharyngeal cancer was found despite using similar definitions of heavy drinking and heavy smoking. Among subjects who were heavy drinkers and/or heavy smokers, both studies found no association between heartburn and the development of laryngopharyngeal cancer.

TABLE III.
Odds of Laryngopharyngeal Cancer Associated With Self-Reported History of Heartburn and Formal Diagnosis of GERD Stratified by Heavy Smoking and/or Heavy Drinking Status.

	Self-Reported History of Heartburn*				GERD Diagnosis*			
Subjects	Cases	Controls	OR§	95% CI	Cases	Controls	OR§	95% CI
Neither a heavy smoker nor a heavy drinker†,‡								
Never had heartburn/GERD	103	543	1.00	–	107	541	1.00	–
Ever had heartburn/GERD	26	152	0.91	0.54, 1.54	23	146	0.87	0.51, 1.48
Heavy smoker and/or heavy drinker†,‡								
Never had heartburn/GERD	497	424	1.00	–	501	416	1.00	–
Ever had heartburn/GERD	175	149	0.96	0.72, 1.28	158	142	1.05	0.79, 1.41

*Recorded as dichotomous ever/never.
†Heavy smoking was defined as more than 18.3 pack-years.
‡Heavy drinking was defined as consumption of more than 14 alcoholic drinks per week.
§Estimates adjusted for age, sex, race, years smoked cigarettes, lifetime alcohol consumption, body mass index, education, and 2-way and 3-way interaction terms between age/sex/race.
CI = confidence interval, GERD = gastroesophageal reflux disease, OR = odds ratio.

Likewise, other prior research has reported conflicting results. A European case-control study of 1,774 HNSCC cases found no association between heartburn and specific HNSCC tumor sites, except for an inverse association with hypopharyngeal cancer (OR = 0.64; 95% CI 0.44, 0.93).[13] Another epidemiologic study (1,303 cases) reported that ever-smoker/ever-drinker HNSCC subjects were not more likely to have had a history of GERD than never-smoker/never-drinker HNSCC patients.[14] A small case-control study (120 cases) examined associations between *H. pylori* infection, a cause of GERD, and odds of laryngopharyngeal cancer but found no association (OR = 1.53; 95% CI 0.69, 3.41).[15]

Studies of the etiologic role of GERD in laryngeal cancer have arrived at different conclusions. A meta-analysis concluded that GERD was associated with an increased odds of laryngeal cancer (OR = 2.21; 95% CI 1.53, 3.19) but not pharyngeal cancer, although the pharynx OR was extremely imprecise.[16] A literature review that was not a meta-analysis could not conclude that GERD caused laryngeal cancer, but noted that confounding by alcohol and tobacco consumption were inadequately controlled.[17] A large case-control study conducted in the Veterans Health Administration system (14,449 cases) found no association between GERD and laryngeal cancer (OR = 1.01; 95% CI 0.92, 1.12).[18]

An important strength of our study was examination of two different measures of GERD exposure: 1) self-reported history of symptoms and medical diagnosis and 2) development of HNSCC. A medical diagnosis of GERD is more likely to indicate substantial GERD morbidity than a self-reported history of having had frequent heartburn, resulting in less misclassification. Second, our analysis was based on a large, population-based case-control study, making it representative of a well-defined source population and increasing the precision of the effect estimates. Third, CHANCE has detailed information on alcohol and tobacco consumption, important causes of HNSCC not well measured in a number of previous studies that examined relationships between GERD and HNSCC.[17] This enabled us to appropriately control for the effects of tobacco and alcohol.

In terms of limitations, our assessment of GERD by self-report was not as accurate as an objective measurement such as pH monitoring would be,[5] even when the self-reported measure was medical diagnosis of GERD rather than self-assessment of GERD symptoms. Medical diagnosis of GERD was ascertained by asking subjects whether they were ever diagnosed with GERD by a doctor rather than abstracting the information from medical records. Furthermore, even among self-reported measures of GERD, our simple one-question assessment might not be as accurate or reliable as validated multi-question instruments such as the Reflux Symptom Index.[19] However, alternative measures of self-reported GERD were not available in CHANCE. There is the possibility of misclassification of the history of GERD, especially if subjects are not aware of the criteria for frequency and severity of symptoms used to diagnose GERD.[20] For example, subjects who are not aware of frequency criteria and assume that their symptoms do not occur frequently enough to warrant being considered a medical diagnosis could falsely report not having had GERD exposure, thereby possibly attenuating estimates of an association between GERD and HNSCC.

Future research on GERD and HNSCC must consider the differing study designs and inconsistent findings of results reported to date. A larger study may be beneficial to further elucidate this association. Such a study would need to provide adequate control for tobacco and alcohol consumption as well as obesity, as was done here.[5] It would be informative to compare the effect of GERD when measured by self-report, medical diagnosis as ascertained by medical records, and by pH monitoring or another objective measurement. Multiple measures of self-reported GERD could be used for purposes of comparison, including questionnaires such as the Reflux Symptom Index.[19] Because a few of our site-specific associations suggested greater risk, estimates of the association of GERD with HNSCC should be conducted for both overall HNSCC as well as individual tumor sites, as was done here.

CONCLUSION

In summary, we find no general pattern of association between GERD and the development of HNSCC. Subgroup analysis of subjects who were neither heavy drinkers nor heavy smokers does not show an association between GERD and the development of laryngopharyngeal cancer, a finding that conflicts with prior research. However, whereas none of our associations is statistically significant, the patterns of some point estimates, such as the larynx result, are suggestive and should be further investigated in future larger studies. Such additional work would help to resolve the inconsistencies observed in this literature.

BIBLIOGRAPHY

1. Spechler SJ. Clinical manifestations and esophageal complications of GERD. *Am J Med Sci* 2003;326:279–284.
2. Williams JL. Gastroesophageal reflux disease: clinical manifestations. *Gastroenterol Nurs* 2003;26:195–200.
3. Kapoor H, Agrawal DK, Mittal SK. Barrett's esophagus: recent insights into pathogenesis and cellular ontogeny. *Transl Res* 2015;166:28–40. doi: 10.1016/j.trsl.2015.01.009.
4. Wang RH. From reflux esophagitis to Barrett's esophagus and esophageal adenocarcinoma. *World J Gastroenterol* 2015;21:5210–5219.
5. Herbella FA, Neto SP, Santoro IL, Figueiredo LC. Gastroesophageal reflux disease and nonesophageal cancer. *World J Gastroenterol* 2015;21:815–819.
6. Langevin SM, Michaud DS, Marsit CJ, et al. Gastric reflux is an independent risk factor for laryngopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev* 2013;22:1061–1068.
7. Johnston N, Dettmar PW, Strugala V, Allen JE, Chan WW. Laryngopharyngeal reflux and GERD. *Ann N Y Acad Sci* 2013;1300:71–79.
8. Kuo CL, Chen YT, Shiao AS, Lien CF, Wang SJ. Acid reflux and head and neck cancer risk: A nationwide registry over 13 years. *Auris Nasus Larynx* 2015;42:401–405. doi: 10.1016/j.anl.2015.03.008.
9. Johnston N, Yan JC, Hoekzema CR, et al. Pepsin promotes proliferation of laryngeal and pharyngeal epithelial cells. *Laryngoscope* 2012;122:1317–1325.
10. Kelly EA, Samuels TL, Johnston N. Chronic pepsin exposure promotes anchorage-independent growth and migration of a hypopharyngeal squamous cell line. *Otolaryngol Head Neck Surg* 2014;150:618–624.
11. Divaris K, Olshan AF, Smith J, et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. *Cancer Causes Control* 2010;21:567–575.
12. Stingone JA, Funkhouser WK, Weissler MC, Bell ME, Olshan AF. Racial differences in the relationship between tobacco, alcohol, and squamous

cell carcinoma of the head and neck. *Cancer Causes Control* 2013;24: 649–664.
13. Macfarlane TV, Macfarlane GJ, Thakker NS, et al. Role of medical history and medication use in the aetiology of upper aerodigestive tract cancers in Europe: the ARCAGE study. *Ann Oncol* 2012;23:1053–1060.
14. Dahlstrom KR, Little JA, Zafereo ME, Lung M, Wei Q, Sturgis EM. Squamous cell carcinoma of the head and neck in never smoker-never drinkers: a descriptive epidemiologic study. *Head Neck* 2008; 30:75–84.
15. Nurgalieva ZZ, Graham DY, Dahlstrom KR, Wei Q, Sturgis EM. A pilot study of Helicobacter pylori infection and risk of laryngopharyngeal cancer. *Head Neck* 2005;27:22–27.
16. Zhang D, Zhou J, Chen B, Zhou L, Tao L. Gastroesophageal reflux and carcinoma of larynx or pharynx: a meta-analysis. *Acta Otolaryngol* 2014;134:982–989.
17. Coca-Pelaz A, Rodrigo JP, Takes RP, et al. Relationship between reflux and laryngeal cancer. *Head Neck* 2013;35:1814–1818.
18. Francis DO, Maynard C, Weymuller EA, Reiber G, Merati AL, Yueh B. Reevaluation of gastroesophageal reflux disease as a risk factor for laryngeal cancer. *Laryngoscope* 2011;121:102–105.
19. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). *J Voice* 2002;16:274–277.
20. Rasmussen S, Jensen TH, Henriksen SL, et al. Overlap of symptoms of gastroesophageal reflux disease, dyspepsia and irritable bowel syndrome in the general population. *Scand J Gastroenterol* 2015;50:162–169.

Reprinted by permission of Cancer Prev Res (Phila). 2014; 7(12):1258-1269.

Research Article

Cancer Prevention Research

Proton Pump Inhibitors and Histamine 2 Blockers Are Associated with Improved Overall Survival in Patients with Head and Neck Squamous Carcinoma

Silvana Papagerakis[1,2], Emily Bellile[3], Lisa A. Peterson[1], Maria Pliakas[1], Katherine Balaskas[1], Sara Selman[1], David Hanauer[4,5], Jeremy M.G. Taylor[3,6], Sonia Duffy[1,7,8,9], and Gregory Wolf[1]

Abstract

It has been postulated that gastroesophageal reflux plays a role in the etiology of head and neck squamous cell carcinomas (HNSCC) and contributes to complications after surgery or during radiotherapy. Antacid medications are commonly used in patients with HNSCC for the management of acid reflux; however, their relationship with outcomes has not been well studied. Associations between histamine receptor-2 antagonists (H2RA) and proton pump inhibitors (PPI) use and treatment outcomes were determined in 596 patients with previously untreated HNSCC enrolled in our SPORE epidemiology program from 2003 to 2008 (median follow-up 55 months). Comprehensive clinical information was entered prospectively in our database. Risk strata were created on the basis of possible confounding prognostic variables (age, demographics, socioeconomics, tumor stage, primary site, smoking status, HPV16 status, and treatment modality); correlations within risk strata were analyzed in a multivariable model. Patients taking antacid medications had significantly better overall survival (OS; PPI alone: $P < 0.001$; H2RA alone, $P = 0.0479$; both PPI + H2RA, $P = 0.0133$). Using multivariable Cox models and adjusting for significant prognostic covariates, both PPIs and H2RAs used were significant prognostic factors for OS, but only H2RAs use for recurrence-free survival in HPV16-positive oropharyngeal patients. We found significant associations between the use of H2RAs and PPIs, alone or in combination, and various clinical characteristics. The findings in this large cohort study indicate that routine use of antacid medications may have significant therapeutic benefit in patients with HNSCC. The reasons for this association remain an active area of investigation and could lead to identification of new treatment and prevention approaches with agents that have minimal toxicities. *Cancer Prev Res; 7(12); 1258–69. ©2014 AACR.*

Introduction

Pathologic gastroesophageal reflux is a common condition in patients with head and neck cancer (1–4). There is evidence that acid reflux may play a role in the etiology of head and neck squamous cell cancer (HNSCC) and contribute to complications after surgery or during radiation and chemotherapy (2, 5–9); acid reflux has been recently reported as an independent risk factor for squamous cancers of the pharynx and larynx (10). Histamine receptor-2 antagonists (H2RA) and proton pump inhibitors (PPI) are distinct groups of medications known for their similar ability to decrease and/or inhibit gastric acid production. At the University of Michigan (Ann Arbor, MI), these medications are commonly and regularly administered in patients with HNSCC as part of their cancer treatment for the management of acid reflux and complications from conventional therapies. It is unknown whether preventing acid reflux might prevent tumor recurrences and improve clinical outcome in patients with HNSCC.

The objective of this study was to determine whether clinical use of antacid drugs is associated with better clinical outcomes in a large retrospective cohort of 596 previously untreated patients enrolled in our Head and Neck Cancer Specialized Program of Research Excellence (SPORE) epidemiology program from 2003 to 2008. This is the first study to identify an association of the PPI and H2RA class of drugs with treatment outcomes and survival in patients with HNSCC. Elucidation of antacid drugs biologic effects on

[1]Department of Otolaryngology-Head and Neck Surgery University of Michigan Medical School, Ann Arbor, Michigan. [2]Department of Periodontics-Oral Medicine, School of Dentistry, University of Michigan, Ann Arbor, Michigan. [3]Center for Cancer Biostatistics, University of Michigan, Ann Arbor, Michigan. [4]Clinical Informatics, Comprehensive Cancer Center Bioinformatics Core, University of Michigan, Ann Arbor, Michigan. [5]Department of Pediatrics and Communicable Diseases, University of Michigan Medical School, Ann Arbor, Michigan. [6]Department of Biostatistics, University of Michigan Medical School, Ann Arbor, Michigan. [7]School of Nursing, University of Michigan, Ann Arbor, Michigan. [8]Department of Psychiatry, University of Michigan Medical School, Ann Arbor, Michigan. [9]VA Ann Arbor Healthcare System, University of Michigan, Ann Arbor, Michigan.

Corresponding Author: Silvana Papagerakis, Department of Otolaryngology-Head and Neck Surgery, University of Michigan Medical School, 1150 W. Medical Center Drive, Room 5434 Med Sci I, Ann Arbor, MI 48109-5616. Phone: 734-615-7085; Fax: 734-764-0014; E-mail: silvanap@umich.edu

doi: 10.1158/1940-6207.CAPR-14-0002

tumor progression could lead to new strategies for cancer prevention and treatment.

Materials and Methods

Patient population

Permission from the Institutional Review Board (IRB) for Human studies was granted to retrospectively analyze the patients that presented to the Department of Otolaryngology between January29, 2003 and November 7, 2008 with HNSCC who were enrolled in our prospective Head and Neck SPORE epidemiology program. IRB approval was also granted for use of existing clinical health data regarding medication use from the medical records of the patients. All patients included provided informed and signed consent form.

The initial cohort of 884 unselected subjects prospectively completed longitudinal health surveys which collected health behaviors (tobacco and alcohol usage), quality of life measures, patient demographics (age, gender, race, marital status, US Armed Forces veteran status), and socioeconomic status (education level and median income from Census tract). The clinical and treatment outcome data were collected through SPORE data collection forms and health surveys. The investigators collected clinical and histopathologic information (primary tumor site, TNM stage, HPV16 status for oropharyngeal primaries), and follow-up information (type of treatment, duration of follow-up in months, incidence of recurrences, patterns of relapse, overall, and cause-specific survival). Patient drug use was identified by retrospective chart review and data abstraction from patient electronic health records CareWeb using the University of Michigan's EMERSE (Electronic Medical Record Search Engine) software. Using this custom designed software, we were able to create complex yet precise search queries to identify drugs taken and in which time periods (pre- or post-treatment), baseline demographics, clinical and histopathologic data in this cohort. Data were independently collected by three investigators to minimize errors.

Computerized database (BioDBx)

The collected data was transferred to a clinical database (BioDBx) for analysis. Our Head and Neck SPORE has developed and instituted this powerful integrated database with an outstanding record of data collection, management, and analysis. BioBDx runs on a dedicated server, is firewall protected, and supported by the University of Michigan Medical Center Information Technology department and Center of Advancement of Clinical Research. It is linked to the Health System clinical database (Careweb) for automatic download of clinical and demographic data and tracking of patient visits. Each patient entered in this database had identity protection through assignment of a unique identifying number. Categories of data entry included patient demographics, tumor site, tumor staging characteristics, health habits: tobacco use (cigarette smoking with average pack years: current, former (quit within 1 month vs. > 1 month) and never; alcohol use (AUDIT score), and HPV16 status for oropharyngeal primaries), treatment and detailed clinical follow-up. Our SPORE Program Tissue Core uses this same data management system for specimen tracking.

Data collection on various medications use

We searched for usage of all known members of each antacid class under their various generic and propriety names. Only usage documented after diagnosis date was counted. Within H2RA: cimetidine (Tagamet), ranitide (Zinetac, Zantac), famotidine (Pepcidine, Pepcid), and nizatidine were included. Within PPIs: omeprazole (Prilosec, Zegerid, Losec), pantoprazole (Protonix, Somac, Zurcal), esomeprazole (Nexium, Esotrex), lansoprazole (Prevacid, Zoton, Levant), rabeprazole (Zechin, Rabecid, AcipHex), and dexlansoprazole (Kapidex, Dexilant) were included.

Statistical analysis

We performed general survival analyses using Cox proportional hazards models to investigate which clinical factors and health behaviors measured by our SPORE Epidemiology project were associated with overall survival (OS), disease-specific survival (DSS), time-to-recurrence, and patterns of relapse that included local recurrence, regional, or distant metastasis in these patients with HNSCC. The development of second primary cancers was also assessed. These patients were censored at time of diagnosis of second primary in analyses of disease-specific survival, time-to-recurrence, and patterns of relapse. We created multivariable models using available covariates such as age, clinical stage, primary disease site, treatment modality, smoking status, etc. We tested whether PPI and/or H2RA usage adds to the prognostic ability of our time-to-event models using a likelihood ratio test. HRs and their 95% confidence intervals (CI) were estimated to quantify the magnitude and direction of any associations.

Pairwise comparisons between PPI and H2RA use and other characteristics were explored. The following variables were analyzed for association with medication usage: gender, age, race, marital status, education, income, tumor site, stage, smoking and drinking history, and primary treatment. Pearson χ^2 was used for categorical data and student t test for continuous data. All P values reported correspond to two-sided comparisons.

Cox proportional hazard models were used for survival outcomes (including time to recurrences). Multivariable models using all covariates and also parsimonious analysis using only covariates which displayed significant relationships in bivariate analysis or were *a priori* determined to be scientifically important were performed. A subset analysis of PPI/H2RA use and outcomes according to HPV status was performed among patients with oropharyngeal cancers that had available tissues for HPV16 testing. Survival time was defined as the time from diagnosis to death or last follow-up. Death from any cause was defined as an event for OS, only death from cancer was defined as an event for DSS. A recurrence event in the time-to-recurrence analysis was

defined as any recurrence (local, regional, and/or distant). All statistical analyses were done in SAS version 9.2 (SAS Institute). A two-tailed P value ≤ 0.05 was considered statistically significant.

Results

Cohort characteristics

From an initial 884 cases enrolled in our Head and Neck SPORE epidemiology project, 706 were treated at University of Michigan Hospital and were eligible for this study of medication usage. After further review of the medical record, other reasons for exclusion included: withdrawn of consent ($n = 1$), nonsquamous cell cancer ($n = 2$), unknown primary or nasal cavity primary ($n = 2$), unresectable or palliation ($n = 25$), incomplete clinical information ($n = 65$), treatment for HNSCC before enrollment ($n = 5$), cancer *in situ* ($n = 8$), multiple primaries ($n = 2$). Thus, our analyses for association between clinical data and use of various antacid medications was performed on a total of 596 previously untreated patients, diagnosed and treated at the University of Michigan for HNSCC between January 29, 2003 and November 7, 2008. The sociodemographics and clinicopathologic characteristics of this cohort are summarized in Table 1. The majority of cases were patients with advanced stage disease (stage III or IV cases = 482, 81%); 244 cases (41%) were stage T0, T1, or T2; 305 cases (51.7%) T3 or T4; and no T staging was possible in 44 cases (7.4%). The male/female ratio was 3:1 (448 males, 75% versus 148 females, 25%), average age: 58 years (range 21–92); average age by gender: 59.4 (females) versus 59.7 (males) years. By primary tumor sites: 150 cases (25%) of oral carcinomas, 251 cases (42%) of oropharyngeal carcinomas, 135 cases (23%) of hypopharynx and laryngeal carcinomas, and 58 cases (10%) in other head and neck sites (e.g., sinus, nasopharynx). The majority of patients had higher education (56%, with some college or more), 91% lived in counties with median income over 30,000 per year. There were 170 tumor recurrences and 222 deaths observed during follow-up; 28 patients presented with a second primary during follow-up (typically we consider a cancer a second primary if it is >2 cm from the original primary or it has been at least 3 years since the original primary was diagnosed). The Kaplan–Meier estimate for OS was 73% at 2 years and 59% at 5 years. Median follow-up time for OS was 55 months with a 95% CI of 50–60 months. HNSCC conventional treatment was categorized according with standard treatment modalities: surgery-only 68 cases (11%), radiation-only 31 cases (5%), surgery + radiation 75 cases (13%), radiation + chemotherapy 246 cases (41%), radiation + chemotherapy + surgery 176 cases (30%); there were no cases treated by chemotherapy alone, nor by a combination of surgery + chemotherapy.

Antacids usage and its impact on the clinical outcome of HNSCC patients

We defined users of antacid drugs in our association analyses as only those patients who had antacid usage documented after diagnosis date. Out of the 596 patients, 191 cases (32%) used only PPIs after diagnosis, 83 cases (14%) used only H2RAs, and 136 cases (23%) used both (H2RA + PPI) sometime after diagnosis (Table 2A). We also collected data on drug class use before diagnosis (recorded as "prior use"). Most patients with prior use continued to use PPIs after diagnosis but a small proportion of patients with prior use had no records of use after diagnosis date. Ten of 16 patients with records of prior H2RA use did not have records of H2RA use within 2 years after diagnosis and consequently were categorized as nonusers for analysis. "Late-post use" was recorded when the first record of antacid use dated more than 2 years after diagnosis and these patients were not included as PPI or H2RA users in our analysis. Frequencies of "prior" and "late-post" users of antacid drug classes are summarized in Table 2B.

The analyses were done initially using any H2RA use and any PPI use separately as predictors. We then created a categorical variable combining the information from both drug classes into 4 categories: PPI use only, H2RA use only, PPI and H2RA use, and no antacid use. The bivariate demographic information of our cohort by these categories are summarized in Table 3.

Clinical significance of H2RA usage

Our analysis of H2RA usage and its potential therapeutic benefit identified 219 patients (37%) who received H2RAs within 2 years of diagnosis with HNSCC. These patients received cimetidine ($n = 16$), ranitidine ($n = 215$), famotidine ($n = 37$; note that we did not find any nizatidine usage).

Bivariate demographic. Our analysis indicated a statistically significant association ($P < 0.05$) between H2RA usage and primary HNSCC tumor site, treatment modality, and patient education (Table 3).We observed higher H2RA use in patients with primary disease site in the oral cavity among all HNSCC sites, with higher education, and among those with trimodal (surgery, radiation and, chemotherapy) treatment. H2RA usage was lowest among those treated with radiation only. We also observed more frequent H2RA usage in patients with higher T stage (48% in T3, 4 vs. 31% in T0, T1, T2). Patients on H2RAs had a lower average age at diagnosis (57 vs. 59 years), but the distribution of ages across both groups was not notably different after closer look.

Patient survival and H2RA intake. In univariate analysis, we observed that patients taking H2RA had significantly better OS ($P = 0.0479$; Fig. 1A); when we considered drugs individually (cimetidine, ranitide, famotidine), this association was not maintained for any one particular drug. The statistical significance of the association with OS proved stronger in multivariable analysis after controlling for potential confounding variables such as age, gender, tumor site, stage, smoking, socioeconomic status, and treatment ($P = 0.02$; HR (95% CI): 0.67 (0.47–0.95); Table 4). In addition, when a backward selection algorithm was used to choose a best multivariable prediction model, H2RA usage was consistently chosen as a significant predictor of survival along with age, primary tumor site, and smoking

Table 1. Sociodemographic and clinicopathologic characteristics of the HNSCC cohort

Numerical measure	Mean (SD), median	Range
Age, y	57.9 years (11.2), 57 years	21–92
Categorical measures	***n* (%)**	
Gender		
Male	448 (75%)	
Female	148 (25%)	
Primary tumor subsite		
OC	150 (25%)	
OP	251 (42%)	
LA, HP	135 (23%)	
Other	58 (10%)	
Stage		
Early (CIS, I, II)	110 (19%)	
Late (III, IV)	482 (81%)	
T stage		
0,1,2	244 (41%)	
3,4	305 (52%)	
X,x	44 (7%)	
Smoking		
Never	145 (24%)	
Former (quit >1 month)	226 (38%)	
Current (quit within 1 month)	223 (38%)	
Race		
European American/white	560 (94%)	
Non-white	34 (6%)	
Married, Yes/No		
Married	369 (62%)	
Not married	223 (38%)	
Education		
HS or less	236 (44%)	
Some college or more	305 (56%)	
Treatment		
Surgery-only	68 (11%)	
Radiation-only	31 (5%)	
Surgery + radiation	75 (13%)	
Radiation + chemotherapy	246 (41%)	
Radiation, chemotherapy, and surgery	176 (30%)	

NOTE: The study included 596 previously untreated patients with HNSCC that were enrolled in the epidemiology program of the University of Michigan Head and Neck Cancer Specialized Program of Excellence in Research (SPORE) from 2003–2008. The International Classification of Diseases for Oncology (ICD-9 codes) based on the Union for International Cancer Control (UICC) standard classification criteria for head and neck tumors were used. Pct may not add to 100% due to rounding.

Abbreviations: CIS: carcinoma *in situ*; HP, hypopharynx; HS: high school; LA, larynx; NP: nasopharynx; OC, oral cavity; OP, oropharynx; X, unknown.

status. In the whole cohort of patients, we did not find evidence of a benefit of H2RA use for recurrence-free survival.

Interestingly, subset analysis of the patients with oropharyngeal carcinomas and available HPV16 status indicated H2RA usage as prognostic for better recurrence-free survival in multivariate analysis after controlling for HPV16 [$P = 0.03$; HR (95% CI) = 0.34 (0.12–0.92)].

Clinical significance of PPI usage

Our analysis of PPI usage identified 327 patients who received PPI within 2 years of diagnosis of HNSCC (55% of

Table 2. Antacid drug usage in 596 patients with HNSCC

A: Drug usage documented after diagnosis date in this cohort of previously untreated patients with HNSCC

Family of drugs	*N* % (out of 596)
PPI alone	191 (32%)
H2RA alone	83 (14%)
PPI and H2RA	136 (23%)
No record of usage	186 (31%)
Total	596 (100%)

B: Prior- and late-post drug usage in this cohort of previously untreated patients with HNSCC

Family of drugs	Prior use	Prior use with no post use	Late-post use
PPI	40	4	42
H2RA	16	10	26
Combination of both	13	1	8

NOTE: The data collection on the administration of the drugs of interest was conducted independently by three investigators. Drug usage of all known members of each antacid class under their various generic and propriety names was identified using a custom designed software program EMERSE (Electronic Medical Record Search Engine) and users of antacid drugs in our association analyses were defined as only those patients who had antacid usage documented after diagnosis date.

the total 596 patients). These patients received omeprazole ($n = 179$, 30%), lansoprazole ($n = 115$, 19.3%), esoprazole ($n = 104$, 17.45%), pantoprazole ($n = 127$, 21.3%), and rabeprazole ($n = 10$, 1.7%). Note that we did not find any dexlansoprazole usage.

Bivariate demographic. Our analysis indicated statistically significant associations between PPI usage and primary HNSCC tumor site and marital status (Table 3). We observed higher PPI usage in patients with primary disease site in the oropharynx and in those who were married.

Patient survival and PPI intake. We observed in univariate analysis that patients taking PPI had significantly better OS ($P < 0.0001$; Fig. 1B); this also was observed in multivariate analysis [$P < 0.0001$; HR (95% CI) = 0.55 (0.40–0.74); Table 4]. The statistical significance of the association proved stronger after controlling for potential confounding variables. Interestingly, when we considered drugs individually, this association with OS was maintained for omeprazole ($P = 0.0008$) and esomeprazole ($P = 0.001$); only a trend was noted for lansoprazole ($P = 0.06$) while pantoprazole did not demonstrate a significant association ($P = 0.67$). Univariate analysis failed to demonstrate an association or a trend between PPI use and unadjusted recurrence-free survival [$P = 0.39$; HR (95% CI) = 0.83 (0.60–1.14); Table 4]. However, there was a trend for better recurrence-free survival in PPI users in multivariate analysis after controlling for potential confounding variables such as age, gender, tumor site, stage, smoking, socioeconomic status, and treatment [$P = 0.06$; HR (95% CI) = 0.71 (0.50–1.01); Table 4]. In addition, when a backward selection algorithm (with stay criteria $\alpha = 0.10$) was used to choose a best multivariable prediction model, PPI usage was consistently chosen as a significant predictor of recurrence-free survival, along with age, smoking status, and treatment.

Clinical significance of H2RA ± PPI usage

Our analysis identified 136 patients who received both PPI and H2RA within 2 years of diagnosis of HNSCC (23% of the total 596 patients).

Bivariate demographic. Our analysis indicated a statistically significant association between H2RA + PPI usage and age, smoking, and treatment modality. Higher incidence of combined H2RA + PPI was observed in those that quit within 1 month and those who received trimodal therapy. Only a trend was noted in relation with primary HNSCC tumor site ($P = 0.08$) and median income level ($P = 0.06$).

Patient survival and H2RA+PPI intake. We observed that patients taking H2RA + PPI had significantly better OS than patients taking no antacid at all ($P < 0.0001$; Fig. 1C), and than those taking H2RA alone ($P = 0.05$); we failed to find evidence that the combination was better than PPI alone ($P = 0.88$) in univariate analysis. We did not find evidence of better recurrence-free survival in patients taking H2RA + PPI.

Discussion

To our knowledge, this is the first epidemiologic study that indicates therapeutic benefit of common antacid medication intake in patients with head and neck cancer. Our findings in this large epidemiologic cohort study indicate that clinical usage of the two classes of antacids (PPIs and H2RAs) after diagnosis with HNSCC may have significant benefit by enhancing patient survival. It is known that antacid medications have the ability to decrease and/or

Table 3. Bivariate demographic information by patients' intake of PPI and H2RA, alone and in combination

Characteristic		Overall ($N = 596$)	PPI alone ($N = 191$)	H2RA alone ($N = 83$)	PPI and H2RA ($N = 136$)	None ($N = 186$)
Age, y	Mean (SD), median, range among users	58.3 (11.3), 57, 21–92	59.2 (10.4), 58, 33—86	57.2 (12.2), 55, 21–86	56.3 (11.2), 55, 22–85	58.2 (11.4), 57, 27–92
	P		**0.06**	0.56	**0.05**	0.72
		Overall	**PPI alone *N* (% users)**	**H2RA alone, *N* (% users)**	**PPI+H2RA *N* (% users)**	**None *N* (% users)**
Gender	Male	448	152 (34%)	60 (13%)	98 (22%)	138 (31%)
	Female	148	39 (26%)	23 (16%)	38 (26%)	48 (32%)
	P		0.09	0.51	0.34	0.71
Primary tumor site	OC	150	32 (21%)	43 (29%)	43 (29%)	32 (21%)
	OP	251	67 (27%)	29 (12%)	60 (24%)	95 (38%)
	HP, LAR	135	65 (48%)	8 (6%)	23 (17%)	39 (29%)
	NP, other, unknown	58	27 (47%)	2 (3%)	10 (17%)	19 (33%)
	P		<0.0001	<0.0001	0.08	0.01
Stage	Early	110	42 (38%)	16 (15%)	19 (17%)	33 (30%)
	Late	482	148 (31%)	66 (14%)	117 (24%)	151 (31%)
	Missing	4				
	P		0.13	0.82	0.12	0.79
T stage	0,1,2	244	78 (32%)	31 (13%)	51 (21%)	84 (34%)
	3,4	305	96 (31%)	48 (16%)	78 (26%)	83 (27%)
	X,x	44	17 (39%)	3 (7%)	7 (16%)	17 (39%)
	Missing	3				
	P		0.63	0.22	0.22	0.10
Smoking	Never	145	39 (27%)	17 (12%)	45 (31%)	44 (30%)
	Former	226	77 (34%)	33 (15%)	45 (20%)	71 (31%)
	Current-quit within 1 month	223	75 (34%)	32 (14%)	46 (21%)	70 (31%)
	Missing	2				
	P		0.30	0.70	0.03	0.97

(Continued on the following page)

Table 3. Bivariate demographic information by patients' intake of PPI and H2RA, alone and in combination (Cont'd)

		Overall	PPI alone *N* (% users)	H2RA alone, *N* (% users)	PPI+H2RA *N* (% users)	None *N* (% users)
Race	White	560	178 (32%)	79 (14%)	131 (23%)	172 (31%)
	Non-White	34	13 (38%)	3 (9%)	5 (15%)	13 (38%)
	Missing	2				
	P		0.43	0.39	0.24	0.36
Married (Yes/No)	Married	369	138 (37%)	49 (13%)	81 (22%)	101 (27%)
	Not married	223	53 (24%)	33 (15%)	54 (24%)	83 (37%)
	Missing	4				
	P		0.0006	0.60	0.52	0.01
Education some college	HS or less	236	74 (31%)	42 (18%)	50 (21%)	70 (30%)
	Some college or more	305	102 (33%)	34 (11%)	74 (24%)	95 (31%)
	Missing	55				
	P		0.61	0.03	0.40	0.71
County median income from census	30,000 or below	55	16 (29%)	8 (15%)	7 (13%)	24 (44%)
	Above 30,000	541	175 (32%)	75 (14%)	129 (24%)	162 (30%)
	P		0.62	0.89	0.06	0.04
Treatment	Surgery only	68	25 (37%)	18 (26%)	9 (13%)	16 (24%)
	Radiation only	31	15 (48%)	1 (3%)	3 (10%)	12 (39%)
	Surgery + radiation	75	24 (32%)	13 (17%)	16 (21%)	22 (29%)
	Radiation + chemotherapy	246	79 (32%)	20 (8%)	50 (20%)	97 (39%)
	Radiation + chemotherapy + surgery	176	48 (27%)	31 (18%)	58 (33%)	39 (22%)
	P		0.18	<0.0003	0.001	0.002

NOTE: The patients in PPI + H2RA treatment group have been excluded from PPI alone and H2RA alone groups. All statistical analyses were done in SAS version 9.2 (SAS Institute). A two-tailed $P \leq 0.05$ was considered statistically significant.

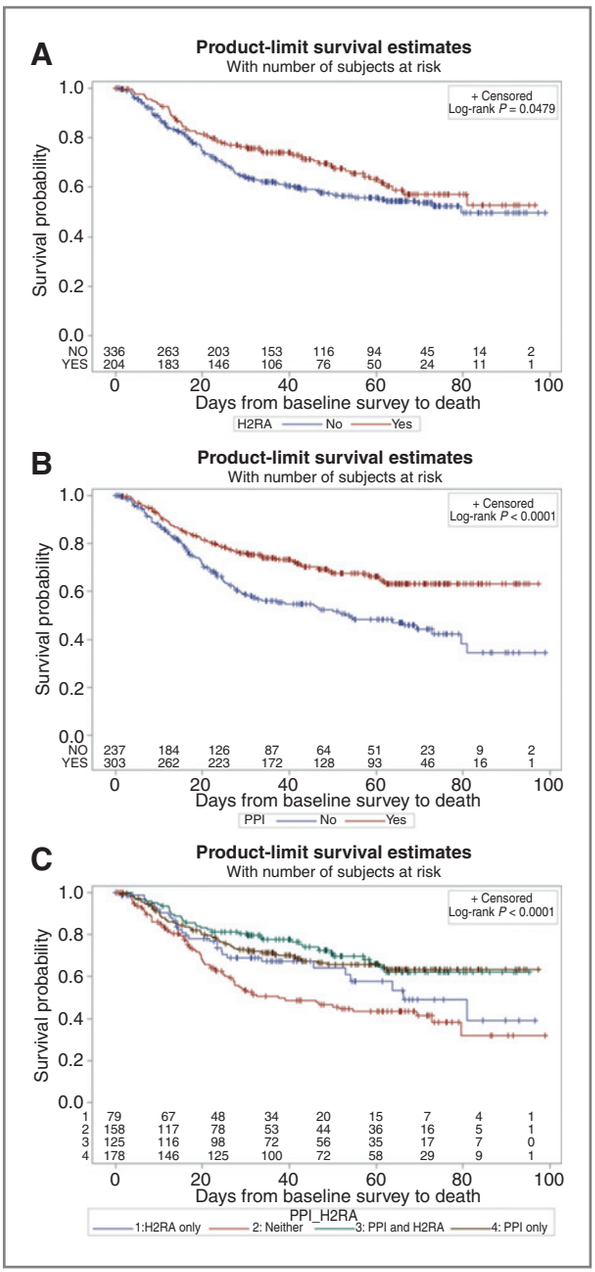


Figure 1. Survival benefits according with intake of antacids in patients with HNSCC. A, unadjusted OS in relation with usage of H2RA (A) and PPIs (B). C, each antacid class alone and in combination versus nonusers. Median follow-up = 55 months; 95% CI, 50–60 months.

inhibit the production of gastric acid and are commonly and chronically used in patients with HNSCC for the management of their gastroesophageal reflux disease. However, the potential effects of antacid medications and any potential mechanisms for altering HNSCC progression and outcome are unknown. Identifying molecular mechanisms associated with HNSCC progression and metastasis is key to improving clinical outcomes.

HNSCC are marked by their aggressiveness and invasiveness (5). HNSCC are known for poor clinical outcomes with mortality among the highest of all carcinomas mainly due to the development of metastatic disease (11, 12). The ability for cancer to metastasize seems to associate with the expression of endothelial adhesion molecules ligands by circulating tumor cells that allow them to bind to the endothelium lining the vasculature initiating extravasation (13, 14). Sialyl Lewis X (sLeX) is an endothelial adhesion molecule known to play the key role in the initiation of the metastatic spread in gastrointestinal cancers by initiating dissemination through direct interaction with E-selectin expressing endothelium (15). In agreement with findings from other types of human cancer (e.g. gastric, breast, colon; refs. 15–18), our previous studies have shown that cimetidine, the prototypical drug of the H2RAs, may have an effect on E-selectin, a molecule with critical roles in cancer dissemination (19). In addition, cimetidine seems to affect other players with important roles in tumor growth and progression (e.g. epithelial growth factor signaling pathway), and to prevent metastasis (20, 21, 22). Our *in vitro* analysis of a well-characterized set of human cell lines derived from the most common locations of the HNSCC indicates that oral squamous cell carcinomas expressed higher sLeX, which increases with advanced stage (23). Our current study has identified the highest H2RA usage in patients with oral carcinomas. It is interesting to note that in contrast to cimetidine, the most frequently prescribed H2RA drug in our cohort, ranitidine, has not proven to have similar effects as cimetidine (22); it is also known that the two also differ in molecular structure. In our patient cohort, cimetidine alone was used by only a few patients (16/596) compared with ranitidine (215/596). When analyzed per individual drug, despite the significant number of ranitidine users, our analysis failed to demonstrate the same benefit on patient survival as the entire H2RA class. Therefore, we postulate that H2RA drugs may differ in their mechanisms of action and may alter expression of other factors besides key endothelial adhesion molecules that could explain their clinical benefits in patients with HNSCC.

Remarkably, our analysis identified H2RA class usage as significant prognostic factor for recurrence-free survival only in patients with oropharyngeal tumors positive for HPV16. HPV has recently emerged as the primary etiologic factor for patients with tumors in the oropharynx that are also associated with younger age at diagnosis; 65% to 85% of the oropharyngeal cancers diagnosed this year in the United States are HPV-related with 3-year failure rates of 30% to 36% (24–31). Consequently, unique pathologic profiles have emerged that are consistent with the changing incidence of HNSCC (32–34). Patients with HPV^+head and neck cancer have a distinct risk profile, associated with a less remarkable history of tobacco and alcohol use (35, 36), a more beneficial micronutrient profile (37), improved cellular immunity (38), and improved survival compared to those with HPV^- tumors (39–42). Notably, a significant subset (20% 30%) of HPV^+ tumors fails to respond to therapy and recur principally as distant metastases. Studies

Table 4. PPI and H2RA usage and patient OS and time to recurrence

		Characteristics			
		OS univariate	**OS Multivariable**	**Recur univariate**	**Recur multivariable**
PPI Usage	Yes	0.55 (0.42–0.73)	0.55 (0.40–0.74)	0.83 (0.60–1.14)	0.71 (0.50–1.01)
	No	Ref.	Ref.	Ref.	Ref.
	P	<0.0001	<0.0001	0.39	0.06
H2RA Usage	Yes	0.74 (0.55–1.00)	0.67 (0.47–0.95)	1.02 (0.73–1.42)	0.90 (0.61–1.34)
	No	Ref.	Ref.	Ref.	Ref.
	P	0.0479	0.02	0.92	0.61
Gender	Male	1.14 (0.82–1.59)	1.30 (0.90–1.88)	1.09 (0.74–1.61)	0.94 (0.62–1.42)
	Female	Ref.	Ref.	Ref.	Ref.
	P	0.42	0.16	0.68	0.75
Primary tumor site	OC	1.43 (1.02–2.00)	2.44 (1.50–3.96)	1.26 (0.84–1.87)	1.90 (1.12–3.23)
	HP, LAR	1.43 (0.99–2.06)	1.43 (0.94–2.18)	1.35 (0.90–2.02)	1.45 (0.92–2.29)
	NP, other, unknown	0.86 (0.50–1.47)	1.11 (0.62–2.00)	0.73 (0.38–1.43)	1.03 (0.50–2.13)
	OP	Ref.	Ref.	Ref.	Ref.
	P	0.05	0.004	0.22	0.08
Stage	Late	1.65 (1.09–2.49)	1.59 (0.97–2.60)	1.79 (1.10–2.94)	1.14 (0.65–1.98)
	Early	Ref.	Ref.	Ref.	Ref.
	P	0.02	0.07	0.02	0.65
Smoking status	Never	0.44 (0.30–0.66)	0.51 (0.33–0.79)	0.42 (0.26–0.67)	0.48 (0.29–0.79)
	Former	0.72 (0.53–0.98)	0.62 (0.44–0.88)	0.80 (0.57–1.14)	0.73 (0.49–1.07)
	Current-quit within 1 month	Ref.	Ref.	Ref.	Ref.
	P	0.0002	0.003	0.002	0.01
Education some college	HS or less	Ref.	Ref.	Ref.	Ref.
	Some college or more	0.72 (0.54–0.96)	0.83 (0.61–1.13)	0.78 (0.56–1.09)	0.92 (0.64–1.31)
	P	0.02	0.24	0.14	0.64
County median income from census					
	30,000 or below	1.61 (1.05–2.47)	1.16 (0.71–1.90)	1.70 (1.05–2.75)	1.36 (0.79–2.33)
	Above 30,000	Ref.	Ref.	Ref.	Ref.
	P	0.03	0.55	0.03	0.27
Treatment	Surgery only	0. 56 (0.33–0.96)	0.39 (0.20–0.78)	0.16 (0.06–0.44)	0.12 (0.04–0.35)
	Radiation only	0.26 (0.08–0.83)	0.22 (0.06–0.74)	0.08 (0.01–0.55)	0.07 (0.01–0.51)
	Surgery + radiation	1.13 (0.74–1.73)	0.56 (0.34–0.93)	0.67 (0.40–1.25)	0.44 (0.24–0.80)
	Radiation + chemotherapy	1.05 (0.76–1.45)	0.96 (0.66–1.39)	0.71 (0.50–1.01)	0.75 (0.50–1.12)
	Radiation + chemotherapy+ surgery	Ref.	Ref.	Ref.	Ref.
	P	0.03	0.009	0.0004	<0.0001
Age, y	10-year increase	1.43 (1.26–1.62)	1.60 (1.38–1.86)	1.02 (1.01–1.04)	1.03 (1.01–1.05)
	P	<0.0001	<0.0001	0.006	0.0004

NOTE: OS and time to recurrence were analyzed using log-rank for univariate analysis and Cox proportional hazards models for multivariate analysis. All statistical analyses were done in SAS version 9.2 (SAS Institute). A two-tailed $P \leq 0.05$ or less was considered statistically significant. HR followed by the 95% CI values between parentheses are provided. Abbreviations: HP, hypopharynx; HS, high school; LA, larynx; NP, nasopharynx; OC, oral cavity, OP, oropharynx; Recur, time to recurrence.

conducted at the University of Michigan have made significant contributions to the understanding of the impact of HPV infection on the pathobiology of HNSCC and response to therapy (40–41). Our current clinical findings have prompted laboratory studies to explore potential mechanisms of the correlations observed clinically using the HPV^+ versus HPV^- carcinoma–derived cell lines from our large SPORE collection.

The major challenge in the management of patients with HNSCC today is the development of evasive resistance to conventional therapies. Our recent evidence demonstrates that cancer stem cells (CSC) play a critical role in the development of metastases in HNSCC and that sLex can help identify the metastatic CSC subset (23). Malignant progression in cancer requires populations of CSCs endowed with unlimited self-renewal, survival under stress and low pH, and establishment of distant metastases. It is also known that increasing tumor mass leads to an acidic tumor microenvironment, while acidity contributes to both tumor progression and resistance to chemotherapy (42, 44). Tumor cells are capable of maintaining a fine state of homeostasis with normal intracellular pH despite the acidic extracellular milieu because of proton pumps expressed in their plasma membranes. A key mechanism to counteract the cytosolic acidification is active proton extrusion by proton pumps. This causes intracellular alkalinization and extracellular acidification, which creates a pH gradient. Low pH of the extracellular microenvironment promotes the secretion and activation of proteolytic enzymes, and release of proangiogenic factors contributing to neovessel formation, cancer invasion, and metastasis (45, 46). This pH gradient also has been associated with multidrug resistance, likely from drug sequestration and neutralization in the acidic organelles or in the acidic extracellular environment (47, 48). Although several pH regulatory mechanisms are operating in tumor cells (Na^+/H^+ exchangers, carbonic anhydrases, bicarbonate transporters, H^+-linked monocarboxylate transporters), the major mechanism is represented by the proton pumps such the vacuolar ATPase (V-ATPase) that are ubiquitously expressed on the plasma membrane of the tumor cells. Highly metastatic cells preferentially use V-ATPases, suggesting that the proton pumps are critical for acquisition of a more metastatic and invasive phenotype (48, 49). Therefore, disruption of this pH gradient with PPIs may be an important antimetastatic mechanism.

Although the specific targets of PPIs are H^+-ATPases contained within the lumen of gastric parietal cells, PPIs also inhibit the activity of V-ATPases, thus broadly blocking proton transport across membranes through the entire body. Our study identified that patients with HNSCC take PPIs, more often alone rather than in combination with H2RA, to treat symptoms that accompany conventional therapeutic regimens, and that their usage may lead to a better patient overall and recurrence-free survival with a higher ratio than with the H2RA use alone or of the combination of both. Interestingly, among the various class members, individual drug usage of only omeprazole and esomeprazole maintained the same survival benefit. At this time we do not fully understand the complex biologic mechanisms by which antacid medications may influence patient outcome. Death from other causes and comorbidities is a major contributor to poor OS rates in patients with head and neck cancer, thus it is possible that PPIs and H2RAs influence deaths from other causes. Studies are currently underway in our laboratory to seek biologic evidences (e.g., potential effects on tumor cells and stroma, modulation of microenvironment, effects on immunity, etc.) in support of the significant association with improved patient outcome observed in the clinical settings.

Elucidation of the novel link between the pathobiology of HNSCC and antacid medication use could lead to important new chemopreventive strategies for patients with HNSCC, for whom the current preventive armamentarium is still limited. HNSCCs are an ideal model for the study of chemoprevention because they follow a histopathologic progression from normal tissue to hyperplasia to severe dysplasia to carcinoma *in situ* to invasive and metastatic carcinomas. Moreover, the phenomenon of field cancerization is well understood in HNSCC, having been characterized first in oral cancer (50). Because of this retained risk for cancer development in the epithelium adjacent to primary disease, second primary tumors act as a possible target for secondary chemoprevention in patients previously diagnosed and treated for HNSCC; furthermore, oral premalignant lesions could also serve as prime targets for chemopreventive agents.

This is the first study to report an association of the PPI and H2RA class of drugs with treatment outcomes and survival in patients with HNSCC. Despite the limitations of the current study (absence of randomization), the intriguing associations observed in our cohort will deserve further validation in randomized prospective trials to provide comprehensive support for a novel therapeutic approach that could be readily translated into clinical benefit. Further elucidation of the mechanisms of action is necessary to determine whether the beneficial effects might be extrapolated to other types of cancer. A series of focused clinical trials will be necessary to further evaluate the antacids anticancer potential in clinical settings, with the ultimate goal of improving the outcome of patients afflicted with HNSCC. If confirmed in prospective studies, new chemopreventive approaches may be possible with drugs that have a favorable therapeutic ratio and are readily available in the clinical settings.

Disclosure of Potential Conflicts of Interest

G.T. Wolf is a consultant/advisory board member for IRX Therapeutics. No potential conflicts of interest were disclosed by the other authors.

Disclaimer

None of the funding sources had any role in the design, conduct, or interpretation of the experiments.

Authors' Contributions

Conception and design: S. Papagerakis, G.T. Wolf
Development of methodology: S. Papagerakis, E. Bellile, K. Balaskas, S. Selman

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S. Papagerakis, L.A. Peterson, M. Pliakas, S.A. Duffy, G.T. Wolf
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S. Papagerakis, E. Bellile, S. Selman, J.M. G. Taylor, G.T. Wolf
Writing, review, and/or revision of the manuscript: S. Papagerakis, E. Bellile, L.A. Peterson, K. Balaskas, D. Hanauer, J.M.G. Taylor, S.A. Duffy, G.T. Wolf
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S. Papagerakis, E. Bellile, L.A. Peterson, G.T. Wolf
Study supervision: S. Papagerakis, G.T. Wolf
Other (informatics support: developed and supported medical record search engine used in study and trained data abstractors in its use for this specific project; worked with study team to refine searches to obtain comprehensive and accurate data from the clinical record): D. Hanauer

Grant Support

This study was supported by the NCI/NIDCR P50 CA097248 [University of Michigan Head and Neck Cancer Specialized Program of Research Excellence (SPORE, Principal Investigator (PI): Gregory Wolf), the Research Scholar Grant RSG-13-103-01–CCE from the American Cancer Society (PI: Silvana Papagerakis), and Undergraduate Research Opportunity Program at the University of Michigan (Ann Arbor, MI).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 9, 2014; revised September 15, 2014; accepted September 28, 2014; published online December 2, 2014.

References

1. Copper MP, Smit CF, Stanojcic LD, Devriese PP, Schouwenburg PF, Mathus-Vliegen LM. High incidence of laryngopharyngeal reflux in patients with head and neck cancer. Laryngoscope 2000;110: 1007–11.
2. Sato K, Umeno H, Chitose S, Nakashima T. Patterns of laryngopharyngeal and gastroesophageal reflux. J Laryngol Otol Supplement 2009; Suppl 31:42–7.
3. Toohill RJ, Kuhn JC. Role of refluxed acid in pathogenesis of laryngeal disorders. Amer J Med 1997;103:100S–6S.
4. Ulualp SO, Roland PS, Toohill RJ, Shaker R. Prevalence of gastroesophagopharyngeal acid reflux events: an evidence-based systematic review. Amer J Otolaryngol 2005;26:239–44.
5. Fennerty MB. The continuum of GERD complications. Cleve Clin J Med 2003;70:S33–50.
6. Turcotte S, Duranceau A. Gastroesophageal reflux and cancer. Thorac Surg Clin 2005;15:341–52.
7. Pondugula K, Wani S, Sharma P. Barrett's esophagus and esophageal adenocarcinoma in adults: long-term GERD or something else? Curr Gastroenterol Rep 2007;9:468–74.
8. Gilbert EW, Luna RA, Harrison VL, Hunter JG. Barrett's esophagus: a review of the literature. J Gastrointest Surg 2011;15:708–18.
9. Tae K, Jin BJ, Ji YB, Jeong JH, Cho SH, Lee SH. The role of laryngopharyngeal reflux as a risk factor in laryngeal cancer: a preliminary report. Clin Exp Otorhinolaryngol 2011;4:101–4.
10. Langevin SM, Michaud DS, Marsit CJ, Nelson HH, Birnbaum AE, Eliot M, et al. Gastric reflux is an independent risk factor for laryngopharyngeal carcinoma. Cancer Epidemiol, Biomarkers Prev 2013;22: 1061–8.
11. American Cancer Society. Global cancer facts & figures; 2011.
12. Shah JP, Johnson NW, Batsakis JK. Oral cancer. London Martin Dunitz 2003;367–72.
13. Haier J, Nicolson GL. The role of tumor cell adhesion as an important factor in formation of distant colorectal metastasis. Dis Colon Rectum 2001;44:876–84.
14. Orr FW, Wang HH, Lafrenie RM, Scherbarth S, Nance DM. Interactions between cancer cells and the endothelium in metastasis. J Pathol 2000;190:310–29.
15. Matsumoto S, Imaeda Y, Umemoto S, Kobayashi K, Suzuki H, Okamoto T. Cimetidine increases survival of colorectal cancer patients with high levels of sialyl Lewis-X and sialyl Lewis-A epitope expression on tumour cells. Br J Cancer 2002;86:161–7.
16. Kawase J, Ozawa S, Kobayashi K, Imaeda Y, Umemoto S, Matsumoto S, et al. Increase in E-selectin expression in umbilical vein endothelial cells by anticancer drugs and inhibition by cimetidine. Oncol Rep 2009; 22:1293–7.
17. Liu F-R, Jiang C-G, Li Y-S, Li J-B, Li F. Cimetidine inhibits the adhesion of gastric cancer cells expressing high levels of sialyl Lewis x in human vascular endothelial cells by blocking E-selectin expression. Int J Mol Med 2011;27:537–44.
18. Tang N-H, Chen Y-L, Wang X-Q, Li X-J, Yin F-Z, Wang X-Z. Cooperative inhibitory effects of antisense oligonucleotide of cell adhesion molecules and cimetidine on cancer cell adhesion. World J Gastroenterol 2004;10:62–6.
19. Papagerakis S, Thornhill M. Therapeutic targets in oral cancer. Toxicol Pathol 2006;34:1009–1009.
20. Fujikawa T, Shiraha H, Nakanishi Y, Takaoka N, Ueda N, Suzuki M, et al. Cimetidine inhibits epidermal growth factor-induced cell signaling. J Gastroenterol Hepatol 2007;22:436–43.
21. Kubecova M, Kolostova K, Pinterova D, Kacprzak G, Bobek V. Cimetidine: an anticancer drug? Eur J Pharm Sci 2011;42:439–44.
22. Kobayashi K, Matsumoto S, Morishima T, Kawabe T, Okamoto T. Cimetidine inhibits cancer cell adhesion to endothelial cells and prevents metastasis by blocking E-selectin expression. Cancer Res 2000;60:3978–84.
23. Czerwinski MJ, Desiderio V, Shkeir O, Papagerakis P, Lapadatescu MC, Owen JF , et al. *In vitro* evaluation of sialyl Lewis X relationship with head and neck cancer stem cells. Otolaryngol Head Neck Surg 2013;149:97–104.
24. Agrawal Y, Koch WM, Xiao W, Westra WH, Trivett AL, Symer DE, et al. Oral human papillomavirus infection before and after treatment for human papillomavirus 16-positive and human papillomavirus 16-negative head and neck squamous cell carcinoma. Clin Cancer Res 2008;14:7143–50.
25. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med 2010;363:24–35.
26. Chaturvedi AK, Engels EA, Anderson WF, Gillison ML. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. J Clin Oncol 2008;26:612–9.
27. D'Souza G, Agrawal Y, Halpern J, Bodison S, Gillison ML. Oral sexual behaviors associated with prevalent oral human papillomavirus infection. J Infect Dis 2009;199:1263–9.
28. Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst 2000; 92:709–20.
29. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. Lancet Oncol 2010;11:781–9.
30. Pannone G, Santoro A, Papagerakis S, Lo Muzio L, De Rosa G, Bufo P. The role of human papillomavirus in the pathogenesis of head and neck squamous cell carcinoma: an overview. Infect Agent Cancer 2011;6:4.
31. Hausen H. Infections causing human cancer. Wiley-VCH Verlag, Weinheim, Germany; 2006.
32. Chenevert J, Seethala RR, Barnes EL, Chiosea SI. Squamous cell carcinoma metastatic to neck from an unknown primary: the potential impact of modern pathologic evaluation on perceived incidence of

human papillomavirus-positive oropharyngeal carcinoma prior to 1970. Laryngoscope 2012;122:793–6.
33. Vidal L, Gillison ML. Human papillomavirus in HNSCC: recognition of a distinct disease type. Hematol Oncol Clin North Am 2008;22: 1125–42.
34. Tang A, Owen JH, Hauff S, Park J, Papagerakis S, Bradford C, et al. Head and neck cancer stem cells: the effect of HPV, an *in vitro* and mouse study. Otolaryngol Head Neck Surg 2013;149:252–60.
35. Applebaum KM, Furniss CS, Zeka A, Posner MR, Smith JF, Bryan J, et al. Lack of association of alcohol and tobacco with HPV16-associated head and neck cancer. J Nat Cancer Inst 2007;99:1801–10.
36. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. J Natl Cancer Inst 2008;100:407–20.
37. Arthur AE, Duffy SA, Sanchez GI, Gruber SB, Terrell JE, Hebert JR, et al. Higher micronutrient intake is associated with human papillomavirus-positive head and neck cancer: a case-only analysis. Nutr Cancer 2011;63:734–42.
38. Wansom D, Light E, Thomas D, Worden F, Prince M, Urba S, et al. Infiltrating lymphocytes and human papillomavirus-16–associated oropharyngeal cancer. Laryngoscope 2012;122:121–7.
39. Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 2008;100:261–9.
40. Maxwell JH, Kumar B, Feng FY, Worden FP, Lee JS, Eisbruch A, et al. Tobacco use in human papillomavirus-positive advanced oropharynx cancer patients related to increased risk of distant metastases and tumor recurrence. Clin Cancer Res 2010;16:1226–35.
41. Worden FP, Kumar B, Lee JS, Wolf GT, Cordell KG, Taylor JM, et al. Chemoselection as a strategy for organ preservation in advanced oropharynx cancer: response and survival positively associated with HPV16 copy number. J Clin Oncol 2008;26:3138–146.
42. Cardonne RA, Casavola V, Reshkin SJ. The role of disturbed pH dynamics and the Na+/H +exchanger in metastasis. Nat Rev Cancer 2005;5:786–95.
43. De Milito A, Marino ML, Fais S. Rationale for the use of proton pump inhibitors as antineoplastic agents. Curr Pharm Des 2012;18: 1395–406.
44. Martinez-Zaguilan R, Seftor EA, Seftor RE, Chu YW, Gillies RJ, Hendrix MJ. Acidic pH enhances the invasive behavior of human melanoma cells. Clin Exp Metastasis 1996;14:176–86.
45. Rofstad EK, Mathiesen B, Kindem K, Galappathi K. Acidic extracellular pH promotes experimental metastasis of human melanoma cells in athymic nude mice. Cancer Res 2006;66:6699–707.
46. Raghunand N, Martínez-Zaguilán R, Wright SH, Gillies RJ. pH and drug resistance. Turnover of acidic vesicles and resistance to weakly basic chemotherapeutic drugs. Biochem Pharmacol 1999;57: 1047–58.
47. Tannock IF, Rotin D. Acid pH in tumors and its potential for therapeutic exploitation. Cancer Res 1989;49:4373–84.
48. Martinez-Zaguilan R, Raghunand N, Lynch RM. pH and drug resistance. Functional expression of plasmalemmal Vtype HATPase in drug resistant human breast carcinoma cell lines. Biochem Pharmacol 1999;57:1037–46.
49. Sennoune SR, Bakunts K, Martínez GM, Chua-Tuan JL, Kebir Y, Attaya MN, et al. Vacuolar H+-ATPase in human breast cancer cells with distinct metastatic potential: distribution and functional activity. Amer J Physiol Cell Physiol 2004;286:1443–52.
50. Braakhuis BJ, Tabor MP, Leemans Cr, van der Waal I, Snow GB, Brakenhoff RH. Second primary tumors and field cancerization in oral and oropharyngeal cancer: molecular techniques provide new insights and definitions. Head Neck 2002;24:198–206.

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