



# HHS Public Access

Author manuscript

*Oral Oncol.* Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:

*Oral Oncol.* 2015 November ; 51(11): 985–990. doi:10.1016/j.oraloncology.2015.08.003.

## Inter- and Intra-observer Variability in The Classification of Extracapsular Extension in p16 Positive Oropharyngeal Squamous Cell Carcinoma Nodal Metastases

James S Lewis Jr., MD<sup>1,2,3</sup>, Yaman Tarabishy, MD<sup>3</sup>, Jingqin Luo, PhD<sup>4</sup>, Haresh Mani, MD<sup>5</sup>, Justin A Bishop, MD<sup>6</sup>, Marino E Leon, MD<sup>7</sup>, Manju L Prasad, MD<sup>8</sup>, Haodong Xu, MD, PhD<sup>9</sup>, and Silvana Di Palma, MD<sup>10</sup>

<sup>1</sup>Vanderbilt University Medical Center, Pathology, Microbiology, and Immunology, Nashville, TN, United States

<sup>2</sup>Vanderbilt University Medical Center, Otolaryngology, Nashville, TN, United States

<sup>3</sup>Washington University, Pathology and Immunology, St. Louis, MO, United States

<sup>4</sup>Washington University, Biostatistics, St. Louis, MO

<sup>5</sup>Inova Fairfax Hospital, Pathology, Falls Church, VA, United States

<sup>6</sup>Johns Hopkins University, Pathology, Baltimore, MD, United States

<sup>7</sup>H. Lee Moffitt Cancer Center, Pathology and Cell Biology, Tampa, FL, United States

<sup>8</sup>Yale University, Pathology, New Haven, CT, United States

<sup>9</sup>UCLA, Pathology and Laboratory Medicine, Los Angeles, CA, United States

<sup>10</sup>Royal Surrey County Hospital, Histopathology, Guildford, Surrey, United Kingdom

### Abstract

**Objective**—Discern inter- and intra-observer variability in the classification of extracapsular extension (ECE) in p16+ oropharyngeal (OP) SCC comparing pathologists' own criteria versus those of a well-defined classification system.

**Methods**—Five pathologists reviewed 50 digitally scanned nodal metastasis slides in three Rounds. Round One was by their own criteria as ECE present or absent, and Rounds Two and Three were with a defined ECE system: Grade 0 (no ECE), 0c (no ECE - thick capsule; no infiltration), 1 (ECE - cells beyond capsule), and 2 (soft tissue metastasis - cells in soft tissue without residual node). Round Three assessed intra-observer variability.

---

Corresponding Author: James S. Lewis Jr., MD, Associate Professor, Pathology, Microbiology, and Immunology, Associate Professor, Otolaryngology, Vanderbilt University Medical Center, 1211 Medical Center Dr., Room 3020D – Surgical Pathology, Nashville, TN 37232-7415, ; Email: james.lewis@vanderbilt.edu, (615) 343-0233 (office), (615) 322-1831 (fax).

Conflict of Interest: None of the authors has any conflict of interest, financial or otherwise, to report relevant to this work.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Results**—In Round One, all five agreed on only 48% of cases (n=24). Fleiss's Kappa value was 0.508 (95% CI: 0.357~0.644). For Rounds Two and Three, Grade 0 and 0c and Grade 1 and 2 were separately grouped as ECE absent or present. In Round Two, all five agreed on 68% of cases (n=34). Fleiss' Kappa was 0.635 (95% CI: 0.472~0.783), indicating statistically significantly better agreement. In Round Three, all five agreed on 64% of cases (n=32) giving a Fleiss's Kappa of 0.639. Pathologists agreed with their prior reads in approximately 90% of cases (average n=45.4, range n=42~49), an average intra-observer Cohen's Kappa of 0.8 (range: 0.68~0.95). Inter- and intra-observer variability rates for classification of soft tissue metastasis (ECE2) were substantially worse.

**Conclusion**—There is substantial inter-, and modest intra-, observer variability among head and neck pathologists for ECE in p16+ OPSCC, modestly improved by a defined system.

### Keywords

Extracapsular extension; lymph node metastasis; soft tissue metastasis; interobserver variability; p16; oropharyngeal squamous cell carcinoma

### Introduction

Human papillomavirus (HPV)-related oropharyngeal squamous cell carcinoma (OPSCC) is increasing in incidence and now constitutes a large fraction of all head and neck SCC[1, 2] [3]. It has unique clinical and pathologic features including nonkeratinizing morphology in most cases, lower mutation rates, and a much better prognosis than conventional (HPV-negative) head and neck SCC, with approximately 3 to 5 fold lower risk of death from disease[4] [5]. p16 is markedly overexpressed in OPSCC with transcriptionally-active high risk HPV and has emerged as a reliable surrogate marker for the virus in these tumors and also as a remarkably strong prognostic marker[8] [5] [4] [9]. For simplicity, hereafter, HPV-related OPSCC will be referred to as p16 positive OPSCC.

The clinical presentation of p16 positive OPSCC is dominated by neck disease[4]. As many as 80 to 85% of patients have neck metastases at presentation[5], and 50% or more of patients present because of complaints in the neck [10]. Traditional studies on the treatment of head and neck SCC have shown that extracapsular extension (ECE) in cervical nodal metastases is a significant adverse prognosticator [11] [12], and it has emerged as a specific indication for the addition of adjuvant chemotherapy to postoperative radiation[13] [14, 15]. It is not clear, however, that p16 positive OPSCC patients have the same biology and clinical course due to nodal metastases with ECE as for other head and neck sites, nor is it clear that adjuvant chemotherapy for node positive patients with ECE results in improved outcomes[16].

Several recent retrospective studies have shown no clinical benefit to adjuvant chemotherapy for node positive, ECE positive patients with p16 positive OPSCC [17] [18] [19], and prospective trials to address this important question are under way. To compound this problematic area, there is little data on what constitutes bona fide, clinically significant ECE in head and neck SCC in general (from a pathologic diagnostic perspective) and particularly in p16 positive OPSCC. There is also very little data about the reproducibility of

pathologists in diagnosing ECE in cervical nodal metastases[20]. Further, the data that exists[20] does not specifically address nodal metastases in p16 positive OPSCC patients. The nodal metastases in p16 positive OPSCC are different than for typical head and neck SCC. They are larger, often are cystic[21], and develop pushing borders and thickened capsules/pseudocapsules[18] [17]. The literature and practice guidelines such as the NCCN[22] consistently ask pathologists to analyze nodal metastases in p16 positive OPSCC for ECE, but do not specifically guide pathologists on what patterns of growth in the nodal metastases actually have clinical significance and what patterns do not.

Because p16 positive OPSCC is a distinct disease with distinct patterns of neck disease, we previously developed an ECE grading system specifically tailored to these tumors. In three separate retrospective studies, we found little clinical significance to ECE, although the worst extent (termed “soft tissue metastasis” or STM) seems to associate with poorer outcomes, particularly for T3/T4 patients or those who do not receive postoperative adjuvant therapy[17, 18]. These results have largely been supported by additional studies from other institutions [19]. The current study was performed to assess the degree of inter-observer variability in the diagnosis of ECE in cervical nodal metastases in p16 positive OPSCC, to assess if a defined system could decrease variability, to assess intra-observer variability with this system, and finally, to determine variability in the diagnosis of STM, the worst extent of ECE.

## Materials and Methods

Approval was obtained from the Washington University Human Resource Protection Office prior to performance of the study which involved waiver of consent for individual patients. Power analysis was performed (see Statistics section below) based on prior ECE studies and the expected distribution of ECE positive and negative lymph nodes in order to determine the number of cases and pathologists to utilize for the study. From a well characterized study cohort of surgically treated OPSCC patients from which prior ECE studies were performed[18], 50 patients were chosen using a random number generator. Slides from their nodal metastases were digitally scanned at 400× magnification using an Aperio Scanscope XT digital scanner. A survey was created using publicly available SurveyMonkey online software (Palo Alto, CA, USA) for the 50 cases. Digitally scanned slides were shared via digital links to central slide sharing resources through Aperio ePathology Solutions using Second Slide and eSlideShare services (Leica Microsystems, Inc., Buffalo Grove, IL, USA)

Five head and neck pathologists were asked to participate in the study (JAB, MP, HM, ML, HX). All sign out on subspecialty head and neck services and come from various geographic regions across the United States including the West coast, Southeast, Northeast, and mid-Atlantic regions. The survey and digital slides were circulated for review in three different rounds (herein referred to as Rounds One, Two, and Three). In Round One, pathologists were simply instructed to read the nodal metastases for the binary presence of ECE or not in whatever manner they were using at the time in clinical practice. In Round Two, pathologists were instructed on the details of a specific grading system[17, 18] by a Figure with depictions of the Grades (Figure 1) and associated text descriptions as follows:

“Please classify the nodal metastases according to the below definitions (with associated explanatory figures shown below).

- 1) ECE absent (Grade 0 - tumor cells limited to node parenchyma without any capsular thickening)
- 2) ECE absent (Grade 0c - tumor cells limited to node but with associated thickened nodal capsule/pseudocapsule without infiltration\*)
- 3) Simple ECE present (Grade 1 - nodal tissue is still present but tumor cells infiltrate beyond the capsule into perinodal soft tissue, such as into fat or around blood vessels)
- 4) Soft tissue metastasis (Grade 2 - tumor cells in soft tissue without any residual nodal tissue or architecture remaining. There can be adjacent separate lymph nodes on the slide but if the tumor does not emanate from them, they are not considered to be “residual nodal tissue”)

\*Even if nodes seem matted together by fusion of capsules, still grade the tumor based on the above criteria for each individual lymph node.”

Metastases were thus classified as Grade 0, 0c, 1, and 2 for the study. For clarification, Grade 0c, also a pattern of no ECE, has cells limited to the node but with an associated thickened nodal capsule or pseudocapsule. In this pattern, tumor cells specifically lack infiltration into perinodal soft tissue such as adipose tissue or around nerves or blood vessels. Grade 1, a pattern considered to represent true ECE, has nodal tissue present on the slide, but has tumor cells clearly infiltrating beyond the capsule into perinodal soft tissue. Although one could try to make an allowance for the hilum of the lymph node, where the nodal capsule is not complete and in theory tumor cells could infiltrate into perinodal soft tissue without actually breaching the capsule, this system simply considers any tumor in perinodal soft tissue as ECE, regardless of location in or around the node. Grade 2, soft tissue metastasis (STM) or the worst pattern of ECE has metastatic tumor cells in soft tissue without any residual organized nodal lymphoid tissue or residual nodal architecture. These metastases typically are large, irregular or stellate, and ill-defined, with frequent invasion into large vessels and/or skeletal muscle. Any residual lymphoid tissue consisting of small, inactive, monomorphic lymphocytes with or without germinal centers, but with associated nodal capsule or subcapsular space, was considered to be residual lymph node tissue obviating this Grade. Representative histopathologic images of the Grades are shown in Figure 2.

Round Three was conducted to evaluate intra-observer variability for each of the individual reviewers and also to assess the consistency of the classification system. This third review was conducted 8 months after Round Two and consisted of pathologists simply repeating their review of the 50 cases using the same instructions for the defined ECE classification system as provided in Round Two.

## Statistics

During the study planning stage, under the assumption that ECE is present at a proportion of 63% (based on prevalence in prior studies[17, 18]), the sample size of 50 was calculated to

test a hypothesized multi-rater Kappa of 0.65 against an unfavorable Kappa of 0.4 with a 89% power at a 5% significance level [23] using the R package “kappaSize” [24]. Fleiss multi-rater Kappa was calculated to evaluate inter-observer agreement among the five pathologists while Cohen's kappa was calculated between two raters or, for intra-observer agreement, between two rounds, using the R package “irr” [25]. The 95% confidence interval on Kappa for each round or difference in Kappa coefficients between rounds or between two types of binary ECE groupings was generated as the (2.5%, 97.5%) quantile interval using 1000 nonparametric bootstrap samples. Gratings on the same cases constructed the bootstrap samples to calculate the difference in Kappa values between two rounds or ECE groupings.

## Results

Raw data for all Rounds are provided in Supplemental Table 1, and summarized results for all three Rounds in Table 1. For Round One, all five pathologists agreed on simple presence or absence of ECE in only 24 of the 50 cases (48%). For these complete agreement cases, 12 were read as ECE present and 12 as ECE absent. Fleiss's Kappa for inter-observer agreement was 0.508 (95% CI: 0.357~0.644). Comparing each of the five individual pathologists with each other in pairs, Cohen's Kappa ranged from 0.362 to 0.717. For Round Two, with results dichotomized as Grades 0 or 0c (no ECE) versus Grades 1 or 2 (ECE present), all five pathologists agreed on 34 (68%) of the 50 cases. For the complete agreement cases, 12 were read as ECE present (11 of which were graded by all pathologists as ECE present in Round One) and 22 as ECE absent. The Fleiss's Kappa for inter-observer agreement was 0.635 (95% CI: 0.472~0.783), which was higher than for Round One (Kappa 0.635 versus 0.508). The difference in Kappa values between Rounds One and Two was 0.127 (95% CI: 0.001~0.264). As the lower limit of this 95% CI is greater than 0, it indicates that the pathologists had statistically significantly better agreement in Round Two results (as binary ECE present or absent) than in Round One. For Round Two, the Cohen Kappa values for agreement between two of the five individual pathologists ranged from 0.511 to 0.754, which was also an improvement over Round One, where the Kappa values ranged from 0.362 to 0.717.

For Round Three, again with results dichotomized as Grades 0 or 0c (no ECE) versus Grades 1 or 2 (ECE present), all five pathologists agreed on 32 (64%) of the 50 cases. For these complete agreement cases, 12 were read as ECE present and 20 as ECE absent. The Fleiss's Kappa for inter-observer agreement was 0.639 (95% CI: 0.465 – 0.769), which was almost identical to the result obtained from Round Two. For Round Three, the Kappa values for agreement between two of the five individual pathologists ranged from 0.434 to 0.810.

We then analyzed inter-observer variability in the classification of STM in Rounds Two and Three, with results dichotomized as Grades 0, 0c, and 1 (not STM) versus Grade 2 (STM). All five pathologists agreed that cases were, or were not, STM in 33 (66%) and 39 (78%) of the 50 cases in Rounds Two and Three, respectively. The Fleiss's Kappa values for inter-observer variability were 0.480 (95% CI: 0.306~0.656) and 0.529 (95% CI: 0.243~0.728), respectively. The difference in Kappa values (0.049 (95% CI: - 0.269~0.324) for calling STM versus other grades was not significant between Rounds Two and Three. Both of these

were worse than for the Kappa results for review simply as either ECE absent or present, although statistically not significantly worse based on the 95% CI on the Kappa differences (Kappa difference between ECE and STM in Round Two = 0.155, -0.039~0.355 and in Round Three = 0.110, -0.116~0.372). Thus, these results indicated at least moderately less agreement for the classification of Grade 2 (STM) compared to classification of ECE as simply present or absent.

To assess for intra-observer variability (i.e. “how well do pathologists agree with themselves?”), after a “washout period” of 8 months after Round 2, pathologists were asked to re-review the same cases using the same classification system. Again, considering results as binary ECE present or absent, pathologists agreed with themselves in an average of 45.4 of the 50 cases (range 42 to 49, or 84% to 98%). Cohen's Kappa values for intra-observer agreement ranged from 0.680 to 0.955 (average 0.800). When considering results as Grade 2 (STM) versus all others, pathologists agreed with themselves in an average of 45.6 of the 50 cases (range 42 to 48, or 84% to 96%). However, when considering just the cases called Grade 2 in Round 2 (49 total “calls”), pathologists consistently called less of them Grade 2 in Round 3 (only 30 total “calls”). Kappa values for intra-observer agreement ranged from 0.296 to 0.840 (average 0.627). For example, in Round 2, for each individual pathologist, it was from four to 15 “calls” of ECE2/STM and in Round 3, only from one to 11.

## Discussion

In the 1980's and 1990's, large retrospective studies showed that the presence of ECE in cervical lymph node metastases was shown to be strongly and independently prognostically adverse in head and neck SCC [26] [13]. A number of subsequent large studies showed that adding chemotherapy to postoperative adjuvant treatment with ECE improved outcomes[13] [14, 15]. Although this is now ingrained in clinical practice and practice guidelines[22], well performed prospective trials were not performed to confirm this approach as proper management for each different head and neck anatomic subsite. Further, most of these practices were defined before the “HPV era” in head and neck cancer. ECE is important, in general, but the majority of studies did not take HPV/p16 status into account for the OPSCC cases that were included[13] [26] [14, 15]. Thus, the critical question is: do these large studies across head and neck sites and with no HPV or p16 testing of oropharyngeal patients really still truly represent the biology and clinical features of HPV-related/p16 positive tumors?

There was little data specifically addressing ECE in p16 positive OPSCC until recently where several large, retrospective studies have suggested that it has very little clinical importance when controlling for other variables, particularly T-classification and smoking status[17-19, 29, 30] [31]. These studies have also suggested that the addition of chemotherapy when ECE is diagnosed in surgical specimens does not affect disease recurrence rates or overall survival[17] [32]. Large, prospective trials that are designed to address the proper treatment for patients with p16 positive OPSCC and nodal metastases with ECE, such as ECOG 3311, ADEPT, and PATHOS, are underway. However, results are not expected for many years.

A “dirty little secret” amongst all of this discussion on ECE is that definitions of what actually constitutes ECE histologically, particularly for p16 positive OPSCC, have been lacking, and almost no one has considered if pathologists can reproducibly classify nodal metastases for ECE in a clinically meaningful way. In the literature, rates of ECE in nodal metastases in head and neck SCC have ranged widely, from ~20% to ~85% [20] [11]. The reason for the difference in rates is not clear, but may be because of differences by anatomic subsite and variation in the relative amounts of ECE cases in any given study. It also may have something to do with variation in pathologists' classification. The older studies showing clinical importance for ECE used pathologists' definitions, mostly just drawn from clinical pathology reports, but they did so amongst large numbers of patients. The current study attempts to address both the inter- and intra-observer variability for nodal metastases in p16 positive OPSCC patients, using a classification system specifically designed for such tumors and with significant prior data analyzing the clinical significance of the individual ECE “grades” [17,18,33].

The findings are of only modest consistency in the characterization of nodal metastases as having ECE versus not. However, Kappa values are difficult to translate into subjective categories of “good” or “poor” agreement, at least in isolation. They are, however, helpful in comparing relative improvements between assessments, such as comparing Rounds One, Two, and Three in this study. When the pathologists were simply told to diagnose ECE present or absent using the criteria they utilized for clinical practice, all five pathologists agreed on only 48% of the cases. The defined classification system led to a statistically significant improvement in agreement among pathologists, but the agreement was still modest. This system does, however, result in less calls of ECE overall, perhaps because Grade 0c was considered as no ECE, although it is a pattern which some pathologists may call as actual ECE. Given that in other retrospective studies, this grade/pattern was shown not to correlate with bona fide ECE based on clinical outcomes [17-19], the defined system would be an improvement in at least lowering the overall rate of ECE “calls” for such metastases. Fortunately, intra-observer agreement was shown to be better.

In several recent studies at Washington University in St. Louis, we assessed the clinical impact of ECE in p16+ OPSCC, and in different retrospective patient cohorts, found little clinical significance for ECE [17, 18, 34]. The only clinically relevant form of ECE was STM (or Grade 2), defined exactly as it is in this work. As such, we wanted to assess if pathologists could reproducibly classify nodal metastases as STM. Results of Rounds Two and Three showed that all five pathologists agreed on Grade 2/STM calls in 33 and 39 of the 50 cases (66% and 78%, respectively). However, they ranged widely from each other in the fraction of cases called Grade 2/STM because they changed reads between ECE Grade 1 and Grade 2/STM quite frequently. Again, this suggests that the pathologists could not consistently call a nodal metastasis as STM versus not.

To our knowledge, this is the only study to assess inter- or intra-observer variability in nodal metastasis evaluation in OPSCC. There are a few other studies assessing variation in assessment of ECE in cervical lymph nodes from more generalized head and neck SCC, however. The largest, by van den Brekel et al., used 41 cases of SCC positive neck lymph nodes (from primary sites across head and neck and no apparent assessment for HPV/p16)

and 10 reviewing pathologists in two rounds with ECE simply defined as “yes or no” by the pathologists' own criteria[20, 35]. They found Kappa values of 0.42 and 0.49 for inter-observer agreement and between 0.49 and 0.95 for intra-observer variability. Again, this is only modest agreement, at best[35].

It should be acknowledged that the current study was performed using digitally scanned hematoxylin and eosin slides to facilitate sharing and gathering of data. These 40× images were scanned using a high quality system, the same scanning technology and quality as used for archiving clinical consult slides and for teaching and tumor board presentations at Washington University in St. Louis. Questions have arisen about the validation of histologic review of digitally scanned slides for primary diagnosis. Data shows that digital slide review is essentially comparable to actual glass slide review[36], and the College of American Pathologists has made recommendations regarding validation of such systems for clinical work[37]. We did not, however, specifically validate pathologists on their use of digital images for diagnosis. The fact that we did attain a significant degree of inter-observer agreement in all Rounds and intra-observer agreement of binary ECE of ~90% between Rounds 2 and 3 argues that the digital system for slide sharing and review did not influence the results.

Ultimately, diagnosis and treatment come down to individual patients, not to statistics or p-values in large cohorts. If one considers an individual patient, then 50% or 60% agreement on whether or not they have ECE in their lymph node metastases is far from ideal, particularly since positive ECE is a criterion for adjuvant chemotherapy. In the current study, results indicate that head and neck service pathologists agree with each other on ECE in p16 positive OPSCC only to that modest extent, although they do a better job of agreeing with themselves. Pathologists are better at simple yes/no ECE assessment, but only with clear definitions, particularly with the category 0c specifically defined. However, they are worse at the assessment of nodal metastases as STM. This latter finding is particularly disappointing, given that STM may be the only clinically relevant form of ECE in p16 positive OPSCC. Since many retrospective studies are now suggesting that adding chemotherapy to radiation after surgery in p16 positive OPSCC patients with ECE based on the pathology reports does not improve outcomes[17, 29, 30], the lack of benefit may not only be a function of the lack of ECE as a major clinical/biological factor for patient outcomes, it may also be a function of the lack of consistency in the actual diagnosis of ECE in the nodal metastases.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank Stacey Yates, Tyler Smallman, and Jonathan Bihr from the Washington University Department of Pathology and Immunology Digital Imaging and Computer Support areas for assistance with the scanning and sharing of the digital images.



## References

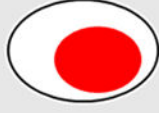

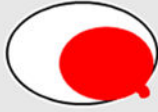

1. Chaturvedi AK, Engels EA, Anderson WF, Gillison ML. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol.* 2008; 26(4):612–9. [PubMed: 18235120]
2. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol.* 2011; 29(32):4294–301. [PubMed: 21969503]
3. 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(7):1429–35. [PubMed: 17724670]
4. 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(1):24–35. [PubMed: 20530316]
5. Ukpo OC, Flanagan JJ, Ma XJ, Luo Y, Thorstad WL, Lewis JS Jr. High-risk human papillomavirus E6/E7 mRNA detection by a novel in situ hybridization assay strongly correlates with p16 expression and patient outcomes in oropharyngeal squamous cell carcinoma. *Am J Surg Pathol.* 2011; 35(9):1343–50. [PubMed: 21836494]
6. Shi W, Kato H, Perez-Ordóñez B, Pintilie M, Huang S, Hui A, et al. Comparative prognostic value of HPV16 E6 mRNA compared with in situ hybridization for human oropharyngeal squamous carcinoma. *J Clin Oncol.* 2009; 27(36):6213–21. [PubMed: 19884544]
7. Mehrad M, Zhao H, Gao G, Wang X, Lewis JS Jr. Transcriptionally-active human papillomavirus is consistently retained in the distant metastases of primary oropharyngeal carcinomas. *Head Neck Pathol.* 2014; 8(2):157–63. [PubMed: 24271958]
8. Mellin Dahlstrand H, Lindquist D, Bjornestal L, Ohlsson A, Dalianis T, Munck-Wikland E, et al. P16(INK4a) correlates to human papillomavirus presence, response to radiotherapy and clinical outcome in tonsillar carcinoma. *Anticancer Res.* 2005; 25(6C):4375–83. [PubMed: 16334111]
9. Bishop JA, Ma XJ, Wang H, Luo Y, Illei PB, Begum S, et al. Detection of transcriptionally active high-risk HPV in patients with head and neck squamous cell carcinoma as visualized by a novel E6/E7 mRNA in situ hybridization method. *Am J Surg Pathol.* 2012; 36(12):1874–82. [PubMed: 23060353]
10. McIlwain WR, Sood AJ, Nguyen SA, Day TA. Initial symptoms in patients with HPV-positive and HPV-negative oropharyngeal cancer. *JAMA Otolaryngol Head Neck Surg.* 2014; 140(5):441–7. [PubMed: 24652023]
11. Myers JN, Greenberg JS, Mo V, Roberts D. Extracapsular spread. A significant predictor of treatment failure in patients with squamous cell carcinoma of the tongue. *Cancer.* 2001; 92(12):3030–6. [PubMed: 11753980]
12. Shaw RJ, Lowe D, Woolgar JA, Brown JS, Vaughan ED, Evans C, et al. Extracapsular spread in oral squamous cell carcinoma. *Head Neck.* 2010; 32(6):714–22. [PubMed: 19827119]
13. Bernier J, Cooper JS, Pajak TF, van Glabbeke M, Bourhis J, Forastiere A, et al. Defining risk levels in locally advanced head and neck cancers: a comparative analysis of concurrent postoperative radiation plus chemotherapy trials of the EORTC (#22931) and RTOG (# 9501). *Head Neck.* 2005; 27(10):843–50. [PubMed: 16161069]
14. Bernier J, Dommé C, Ozsahin M, Matuszewska K, Lefebvre JL, Greiner RH, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Engl J Med.* 2004; 350(19):1945–52. [PubMed: 15128894]
15. Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. *N Engl J Med.* 2004; 350(19):1937–44. [PubMed: 15128893]
16. Haughey BH, Sinha P. Prognostic factors and survival unique to surgically treated p16+ oropharyngeal cancer. *Laryngoscope.* 2013; 122(2):S13–33. [PubMed: 22926949]
17. Sinha P, Lewis JS Jr, Piccirillo JF, Kallogjeri D, Haughey BH. Extracapsular spread and adjuvant therapy in human papillomavirus-related, p16-positive oropharyngeal carcinoma. *Cancer.* 2012; 118(14):3519–30. [PubMed: 22086669]

18. Lewis JS Jr, Carpenter DH, Thorstad WL, Zhang Q, Haughey BH. Extracapsular extension is a poor predictor of disease recurrence in surgically treated oropharyngeal squamous cell carcinoma. *Mod Pathol*. 2011; 24(11):1413–20. [PubMed: 21701534]
19. Maxwell JH, Ferris RL, Gooding W, Cunningham D, Mehta V, Kim S, et al. Extracapsular spread in head and neck carcinoma: impact of site and human papillomavirus status. *Cancer*. 2013; 119(18):3302–8. [PubMed: 23797868]
20. Lodder WL, van den Brekel MW. Re: Extracapsular tumor extension in cervical lymph nodes: reconciling the literature and seer data. *Head Neck*. 2011; 33(12):1809. [PubMed: 22006529]
21. Goldenberg D, Begum S, Westra WH, Khan Z, Sciubba J, Pai SI, et al. Cystic lymph node metastasis in patients with head and neck cancer: An HPV-associated phenomenon. *Head Neck*. 2008; 30(7):898–903. [PubMed: 18383529]
22. Pfister DG, Ang KK, Brizel DM, Burtness BA, Busse PM, Caudell JJ, et al. Head and neck cancers, version 2.2013. Featured updates to the NCCN guidelines. *J Natl Compr Canc Netw*. 2013; 11(8):917–23. [PubMed: 23946171]
23. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med*. 1987; 6(4):441–8. [PubMed: 3629046]
24. Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. *J Clin Epidemiol*. 2012; 65(7):778–84. [PubMed: 22560852]
25. Gamer, M.; Lemon, J.; Fellows, I.; Singh, P. irr: Various coefficients of interrater reliability and agreement. R package. 2012. [version 0.84]. Available from: <http://CRAN.R-project.org/package=irr>
26. Johnson JT, Myers EN, Bedetti CD, Barnes EL, Schramm VL Jr, Thearle PB. Cervical lymph node metastases. Incidence and implications of extracapsular carcinoma. *Arch Otolaryngol*. 1985; 111(8):534–7. [PubMed: 4026664]
27. Greenberg JS, Fowler R, Gomez J, Mo V, Roberts D, El Naggar AK, et al. Extent of extracapsular spread: a critical prognosticator in oral tongue cancer. *Cancer*. 2003; 97(6):1464–70. [PubMed: 12627511]
28. Ferlito A, Rinaldo A, Devaney KO, MacLennan K, Myers JN, Petruzzelli GJ, et al. Prognostic significance of microscopic and macroscopic extracapsular spread from metastatic tumor in the cervical lymph nodes. *Oral Oncol*. 2002; 38(8):747–51. [PubMed: 12570052]
29. Geiger JL, Lazim AF, Walsh FJ, Foote RL, Moore EJ, Okuno SH, et al. Adjuvant chemoradiation therapy with high-dose versus weekly cisplatin for resected, locally-advanced HPV/p16-positive and negative head and neck squamous cell carcinoma. *Oral Oncol*. 2014; 50(4):311–8. [PubMed: 24467937]
30. Rahmati R, Dogan S, Pyke O, Palmer F, Awad M, Lee N, et al. Squamous cell carcinoma of the tonsil managed by conventional surgery and postoperative radiation. *Head Neck*. 2015; 37(6):800–7. [PubMed: 24616252]
31. Klozar J, Kratochvil V, Salakova M, Smahelova J, Vesela E, Hamsikova E, et al. HPV status and regional metastasis in the prognosis of oral and oropharyngeal cancer. *Eur Arch Otorhinolaryngol*. 2008; 265(1):S75–82. [PubMed: 18094985]
32. Sinha P, Piccirillo JF, Kallogjeri D, Spitznagel EL, Haughey BH. The role of postoperative chemoradiation for oropharynx carcinoma: A critical appraisal of the published literature and National Comprehensive Cancer Network guidelines. *Cancer*. 2015; 121(11):1747–54. [PubMed: 25588360]
33. Sinha P, Lewis JS Jr, Kallogjeri D, Nussenbaum B, Haughey BH. Soft tissue metastasis in p16+ oropharynx carcinoma: Prevalence and association with distant metastasis. *Oral Oncol*. 2015 epub ahead of print.
34. Sinha P, Kallogjeri D, Gay H, Thorstad WL, Lewis JS Jr, Chernock R, et al. High metastatic node number, not extracapsular spread or N-classification is a node-related prognosticator in transorally-resected, neck-dissected p16-positive oropharynx cancer. *Oral Oncol*. 2015; 51(5):514–20. [PubMed: 25771076]

35. van den Brekel MW, Lodder WL, Stel HV, Bloemena E, Leemans CR, van der Waal I. Observer variation in the histopathologic assessment of extranodal tumor spread in lymph node metastases in the neck. *Head Neck*. 2012; 34(6):840–5. [PubMed: 22095886]
36. Bauer TW, Schoenfield L, Slaw RJ, Yerian L, Sun Z, Henricks WH. Validation of whole slide imaging for primary diagnosis in surgical pathology. *Arch Pathol Lab Med*. 2013; 137(4):518–24. [PubMed: 23323732]
37. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, Contis L, et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med*. 2013; 137(12): 1710–22. [PubMed: 23634907]

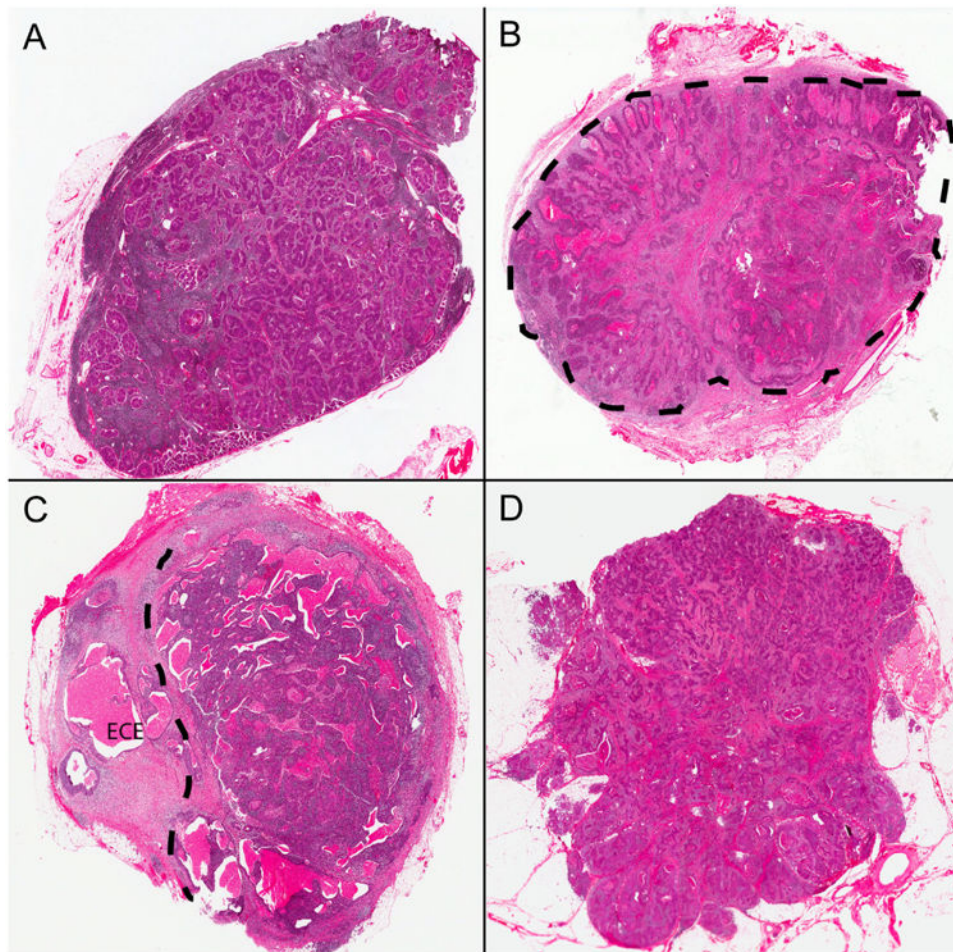
### Highlights

1. Inter-observer variability for ECE in p16+ oropharynx cancer is substantial.
2. Pathologists agree significantly better with a well-defined classification system.
3. Agreement on soft tissue metastasis is significantly worse than for simple ECE.
4. Pathologists, after a washout period, generally agreed with themselves (~90%)
5. The results, considered for each individual patient, show unacceptable variation.

ECE grade	0	0(c)	1	2
Image				
	Tumor confined to the substance of the lymph node (surrounded by lymphoid tissue)	Tumor expanding to the capsule of the lymph node (no intervening lymphoid tissue) with thickening of overlying capsule.	Tumor invading beyond lymph node capsule into perinodal tissue.	<b>Soft tissue metastasis.</b> Tumor mass without residual nodal tissue or architecture (no remaining LN germinal centers or subcapsular sinus).

**Figure 1.**

Defined classification system for extracapsular extension in nodal metastases. Grade 0 and 0c are considered to represent no extracapsular extension, while Grade 1 represents simple extracapsular extension with tumor cells invading perinodal soft tissue in the background of an at least partially preserved lymph node and Grade 2 represents nodal metastases where tumor cells have obliterated all nodal tissue and are just growing in irregular collections (so-called “soft tissue metastasis”).



**Figure 2.** Histopathologic examples of extracapsular extension Grades from the defined classification system. A) Tumor cells limited to lymph node parenchyma without alteration of the capsule or nodal architecture (Grade 0). B) Tumor cells limited to lymph node but with expansion of the node and development of a thickened capsule/pseudocapsule (designated by dashed line) around them (Grade 0c). C) Tumor cells invading into perinodal soft tissue (beyond dashed line and indicated by “ECE”) but with at least partial preservation of the lymph node (Grade 1). D) Tumor cells growing as irregular collections in the neck soft tissues without any histologic evidence of residual nodal parenchyma (Grade 2 or “soft tissue metastasis”) (all images hematoxylin and eosin stained; 15× magnification).  
ECE = extracapsular extension

**Table 1**

Summarized results of reviewers' classifications from all three Rounds.

	All 5 Reviewers Agree	At Least 4 of 5 Reviewers Agree	At Least One ECE Call	Overall Number of ECE Calls
Round 1 (ECE Y/N)	24/50 (48%)	40/50 (80%)	38/50 (76%)	114
Round 2 *	34/50 (68%)	39/50 (78%)	28/50 (56%)	95
Round 3 *	32/50 (64%)	44/50 (88%)	30/50 (60%)	92

\* Dichotomized as Grades 0 or 0c (no ECE) versus Grades 1 or 2 (ECE present)

ECE = Extracapsular extension

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript