Received: 13 April 2022 | Revised: 2 August 2022 | Accepted: 18 August 2022

DOI: 10.1002/hed.27183

CLINICAL REVIEW

WILEY

Correlation between radiologic and pathologic extranodal extension in HPV-associated oropharyngeal cancer: Systematic review

Tristan Morey MBBS[1] | **John-Charles Hodge FRACS, MA**[2] | **Cindy Stern PhD**[1] | **Suren Krishnan FRACS, OAM**[2] | **Andrew Foreman FRACS, PhD**[2]

[1]JBI, Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, South Australia, Australia

[2]Department of Otolaryngology - Head and Neck Surgery, Royal Adelaide Hospital, Adelaide, South Australia, Australia

Correspondence
Andrew Foreman, Department of Otolaryngology—Head and Neck Surgery, Royal Adelaide Hospital, Adelaide 5000, South Australia, Australia.
Email: aforeman@adam.com.au

Abstract

Pretreatment determination of extranodal extension (ENE) has significant clinical implications in human papillomavirus positive (HPV+) oropharyngeal squamous cell carcinoma (OPSCC). Unfortunately there is no gold-standard imaging modality for radiological assessment of ENE in HPV+ OPSCC, leading to subjective assessments and complex decision making concerning ENE. A systematic review of diagnostic test accuracy was therefore undertaken, with five databases systemically searched to evaluate the diagnostic performance of an imaging modality for detection of ENE in HPV+ OPSCC. A meta-analysis was conducted on four CT studies using a random-effects model. While a narrative synthesis was provided for the studies using PET/CT and "CT and MRI." Out of 1772 hits, six studies were included in the review. Meta-analysis on four CT studies showed CT had an overall sensitivity of 77% and specificity of 60%. PET/CT had a sensitivity of 37.5% and specificity of 97%. "CT and MRI" had a sensitivity of 62% and specificity of 78%. Further diagnostic studies involving CT, PET/CT and MRI are ultimately required.

KEYWORDS

extranodal extension, human papillomavirus, imaging modalities, oropharynx, squamous cell carcinoma

1 | INTRODUCTION

Oropharyngeal squamous cell carcinoma (OPSCC) is one of the most common types of head and neck cancer (HNC) in the western world.[1] Previously it has been well reported that sustained exposure to tobacco and alcohol increases the risk of developing OPSCC.[2] P16-positive (p16+) OPSCC has now emerged as a distinct OPSCC variant coinciding with the decline of tobacco-related OPSCC. The primary risk factor for developing p16+ OPSCC is oral human papillomavirus (HPV) infection.[3] The steep rise in the incidence of HPV-positive (HPV+)

[Correction added on 20 September 2022, after first online publication: The following section headings have been updated: "7.2" as "7.1.3" and "7.3" as "7.2".]

OPSCC has occurred on a global scale.[4,5] Interestingly, HPV+ OPSCC is a distinct entity with unique histopathological features, improved prognosis and enhanced response to treatment compared to OPSCC associated with other risk factors such as tobacco.[6–8]

Extranodal extension (ENE) is a well-known phenomenon in cancers such as head and neck, breast and bladder cancer.[9] Detection of ENE particularly in HNC is considered one of the most important adverse prognostic factors for local recurrence, distant metastasis and survival.[10,11] As such treatment is adjusted accordingly. Traditional treatment paradigms would recommend adjuvant chemoradiotherapy in the presence of pathologically confirmed ENE.[12,13] Furthermore, with a shift towards treatment de-intensification and primary non-surgical treatment, pathologic specimens may not be available to make this determination and hence there is a reliance on pre-treatment imaging to determine the presence of ENE and therefore potentially the treatment offered to the patient. Of particular importance to decision making is that HPV+ squamous cell carcinoma (SCC) has a propensity for peritumoral desmoplasia (fibrosis) which can mimic ENE making the radiological diagnosis of ENE and subsequent treatment planning in HPV+ OPSCC fraught with inaccuracies and false assumptions.[14] For the subset of patients who are presumed radiologic ENE positive and undergo primary chemoradiotherapy, the diagnosis of radiologic ENE may be a false positive and a treatment path consisting of surgery (+/− adjuvant radiotherapy) may be more appropriate. One of the main consequences of misdiagnosing radiologic ENE include an estimated 25% of patients requiring unplanned hospital admissions for the management of treatment-related toxicities from concurrent chemoradiotherapy, compared to single modality therapy.[15] Conversely, given the limitations of staging scans, patients who are incorrectly presumed ENE negative on pre-treatment imaging may undergo a surgical approach. This may lead to trimodal therapy (if pathologic ENE positive), when a bimodal approach (primary chemoradiotherapy) would likely be more appropriate and conservative.[12,13] Given the clinical implications of radiologic ENE, it is crucial that our understanding of the strengths and limitations of currently utilized imaging modalities to detect ENE are understood.

Imaging modalities including computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), combination PET/CT and ultrasound (US) can all be utilized to detect and diagnose ENE from a radiological perspective.[16] The choice of imaging modality for detecting ENE largely comes down to clinician preference as currently there is no gold-standard imaging modality for radiological assessment of ENE.

Several reviews have been published on the accuracy of imaging modalities and ENE in the broader population of HNC.[16–18] Unfortunately however there are very few studies addressing ENE in the HPV+ OPSCC population. As such, the following research is the first systematic review reporting on the correlation between radiologic and pathologic ENE exclusively in HPV+ OPSCC.

2 | MATERIALS AND METHODS

This systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.[19] The review protocol was registered in the PROSPERO International Prospective Register of Systematic Reviews with registration number CRD42021250626.

3 | LITERATURE SEARCH

Terms to help identify the inclusion criteria were identified and selected with help of a scientific librarian, and on May 18th 2021 the electronic databases MEDLINE (PubMed), Embase (Ovid), Cochrane Central Register of Controlled Trials (via Cochrane Library), Web of Science and Scopus were searched with no restriction on publication date (see Appendix I for the full search strategy). There were no language filters imposed on the search strategy however due to time restraints and available resources only studies in English were included. Title and abstract screening was performed using JBI SUMARI (System for the Unified Management of the Assessment and Review of Information), while full-text screening was performed using EndNoteX9 (Clarivate Analytics, PA, USA). The reference lists of all included studies were also screened for additional studies. Gray literature was not searched.

4 | STUDY SELECTION

4.1 | Inclusion criteria

This review recruited studies based on: (1) study type: all published studies that examined the diagnostic accuracy (including sensitivity and specificity) of a conventional imaging modality used to detect ENE in HPV+ OPSCC; (2) participants: studies using participants with a confirmed diagnosis of HPV+ OPSCC and suspected diagnosis of ENE of cervical lymph node metastases; (3) index test: studies using a conventional imaging modality to detect ENE including CT, MRI, PET, PET/CT and US; and (4) reference test: studies using histopathology (gold standard) as a

reference test for ENE detection. All of the included studies confirmed radiologic ENE assessment using histopathology.

4.2 | Exclusion criteria

This review excluded studies based on: (1) participants: those with recurrent disease, HNC other than HPV+ OPSCC and participants without nodal disease; (2) index tests: machine learning methods and studies using only indirect measures that are not routinely used in radiological assessment of ENE; (3) study type: review articles, case studies and letters to the editor.

All stages of screening were performed independently by the primary researcher (TM) and another reviewer (CN) with expertise on the topic. Any disagreements were resolved by discussion among the two reviewers. The results of the search and the study inclusion process is reported in full and presented in a PRISMA flow diagram (see Figure 1).

5 | QUALITY ASSESSMENT AND DATA EXTRACTION

5.1 | Assessment of methodological quality

Selected studies were critically appraised by two independent reviewers (TM and CS) for methodological quality using QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2). The QUADAS-2 checklist for primary diagnostic accuracy studies is structured around four domains; Patient Selection, Index Test, Reference Standard, and Flow and Timing. Signaling questions with "yes," "no" or "unclear" responses are included to help determine the risk of bias. The risk of bias is then assessed to be either "A" low risk of bias, "B" unclear risk of bias, or "C" high risk of bias based on the signaling questions for each domain.[20] Any disagreements that arose were resolved through discussion between the two reviewers. All studies were included regardless of methodology quality.

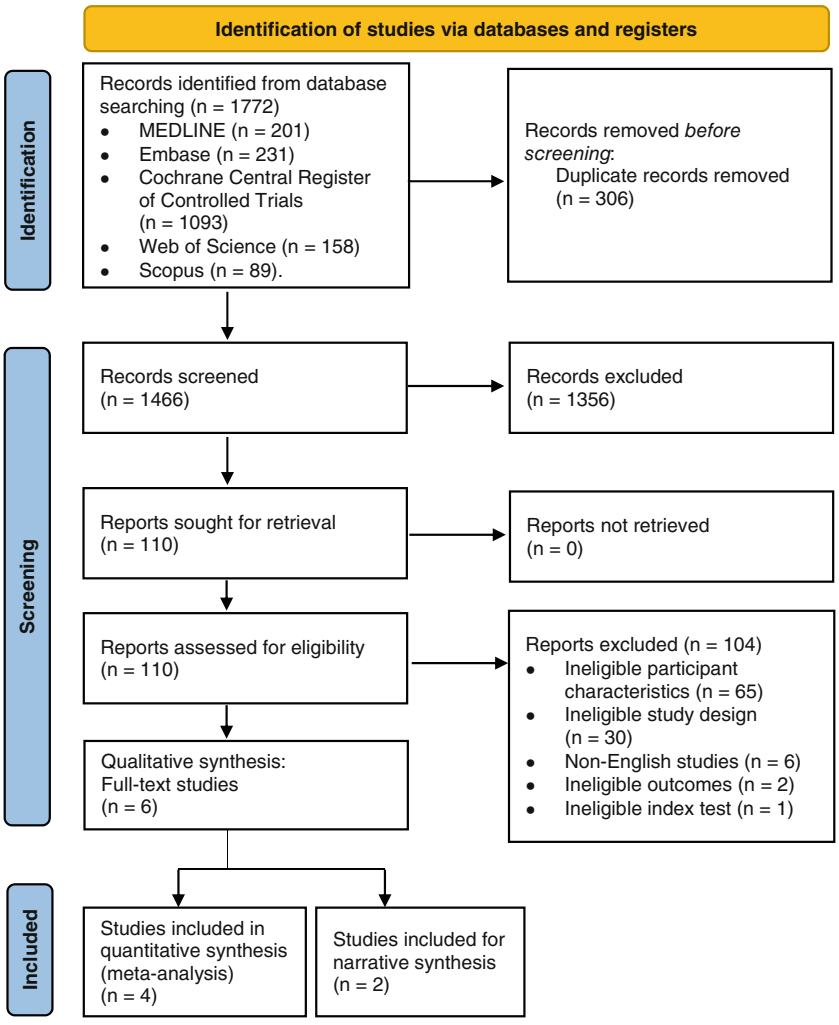

FIGURE 1 PRISMA flow diagram of study selection [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Characteristics of included studies

Study	Country	Study design	Participants/ Tumor	Tumor location	Reference test	Blinding	Imaging modality	Scoring system	Index-reference test interval	Observers	Outcomes (%)
2020 Faraji et al.[23]	United States	RS 2006–2015	$n = 73$ Median age 56.7 67 M:6 F 73 HPV+ OPSCC 32 ENE+:41 ENE−	Oropharynx - 43 Palatine tonsils - 28 BOT or lingual tonsils - 2 Unknown	Pathology	Yes	CECT	Positive/ negative	N/A	2	Absence of perinodal fat plane: *SN*: 87.1, 96.4 *SP*: 50, 34.5 *PPV*: 58.8, 52.9 *NPV*: 61.7, 63.3 Interobserver agreement: $K = 0.5$
2017 Geltzeiler et al.[24]	United States	RS 2010–2015	$n = 100$ Age not reported 80 M:20F 100 HPV+ OPSCC 39 ENE+:61 ENE−	Oropharynx - 71 Tonsil - 29 BOT	Pathology	Yes	CECT	Positive/ negative	≤6 weeks	1	*SN*: 64 *SP*: 68 *PPV*: 71 *NPV*: 61 Combination of severe irregular borders and >3 radiographically suspicious nodes: *SN*: 61 *SP*: 94 *PPV*: 92 *NPV*: 67
2019 Lee et al.[27]	Korea	RS 2006–2016	$n = 134$ Mean age 59.9 118 M:16 F 134 HPV+ OPSCC 70 ENE+:64 ENE−	Oropharynx - 115 Palatine Tonsil - 14 BOT - 1 Others - 4 Unknown	Pathology	N/A	CECT and MRI	Positive/ negative	≤8 weeks ($n = 114$) ≥8 weeks ($n = 20$)	2	For CT and MRI: *SN*: 62 *SP*: 77.8 *PPV*: 61.9 *NPV*: 77.8 Interobserver agreement: Overall: $K = 0.7$ CT: K = 0.7 MRI: K = 0.7
2019 Noor et al.[25]	Australia	RS 2010–2016	$n = 80$ Mean age 58 68 M:12F 80 HPV+ OPSCC 43 ENE+: 37 ENE–	Oropharynx - 56 Tonsil - 23 BOT - 1 Unknown	Pathology	Yes	CECT	4-point scale	≤6 weeks	2	*SN*: 56.5, 60.9 *SP*: 73.3, 66.7 *PPV*: 68.4, 65.1 *NPV*: 62.3, 62.5 Interobserver agreement: $K = 0.4$
2018 Patel et al.[26]	United States	RS 2014–2016	$n = 27$ Mean age 57 27 M:0F	Oropharynx - 14 BOT - 13 Tonsil	Pathology	Yes	CECT	Positive/ negative	N/A	2	For detecting major pENE (>2 mm extension):

TABLE 1 (Continued)

Study	Country	Study design	Participants/ Tumor	Tumor location	Reference test	Blinding	Imaging modality	Scoring system	Index-reference test interval	Observers	Outcomes (%)
			27 HPV+ OPSCC No data on ENE %								*SN*: 88, 100 *SP*: 52.6, 63.2 *PPV*: 43.8, 53.8 *NPV*: 90.9, 100 Interobserver agreement: $K = 0.7$
2021 Snyder et al.[22]	United States	RS 2011–2017	$n = 49$ Mean age 56.4 43 M:6 F 49 HPV+ OPSCC 16 ENE+:33 ENE−	Oropharynx - 27 Tonsil - 19 BOT - 3 Overlapping	Pathology	Yes	PET/CT	Positive/ negative	N/A	2	*SN*: 37.5 (mean) *SP*: 96.9 (mean) *PPV*: 85.7 (mean) *NPV*: 76.1 (mean) Interobserver agreement: $K = 0.4$

Abbreviations: BOT, base of tongue; F, female; K, Cohen's Kappa value; M, male; N/A, not available; NPV, negative predictive value; PPV, positive predictive value; RSN, radiographically suspicious nodes; RS, retrospective study; SN, sensitivity; SP, specificity.

5.2 | Data extraction

With aid from a standardized form,[21] data was extracted from papers included in the review by the primary author (TM) and reviewed by a secondary author (AF). Any disagreements that arose between the two reviewers were resolved through discussion. The data extracted includes:

1. Basic information from the study—year of publication, authors, study design, location, number of participants.
2. Participant characteristics—age, sex, tumor type, tumor location.
3. Details about the imaging modality used—CT, MRI, PET/CT.
4. Details about the index test and reference standard—diagnostic criteria, blinding, time interval.
5. Outcomes per index test (including sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV], area under the ROC curve [AUC] and inter-rater agreement [Cohen's Kappa {K}]) for studies using two observers. For studies using two observers to detect radiological ENE, the outcomes for each observer were extracted.

All included studies reported their outcomes with sensitivity, specificity, PPV and NPV values except for the PET/CT study by Snyder et al.[22] With aid from a local statistician, calculations were required to convert the Snyder et al.[22] "ENE misclassification analysis" to the review's primary outcomes and an overall diagnostic accuracy. These calculations can be found in Appendix II. The CT study by Faraji et al.[23] investigated the highest performing characteristics to aid radiological ENE detection with no overall assessment for ENE being reported. Therefore, in the following review the highest overall performing characteristic for ENE detection in the Faraji et al.[23] study ("absence of perinodal fat plane") was used in the CT meta-analysis.

5.3 | Statistical analysis

All statistical data analysis was performed using Stata version 15.1 (StataCorp LLC, Texas, USA) and MetaDiSc version 1.4 (the Unit of Clinical Biostatistics team of the Ramón y Cajal Hospital, Madrid, Spain) with assistance from a local statistician. The main outcomes were sensitivity and specificity measures with 95% confidence intervals for each imaging modality used to detect ENE in HPV+ OPSCC. Secondary outcomes included PPV, NPV, area under the ROC curve

TABLE 2 QUADAS-2 critical appraisal results

Study	Risk of bias			
	Patient selection	Index test	Reference standard	Flow and timing
Geltzeiler et al.[24]	A	A	B	B
Patel et al.[26]	A	B	A	B
Lee et al.[27]	A	B	B	B
Noor et al.[25]	A	A	B	A
Faraji et al.[23]	B	A	B	B
Snyder et al.[22]	A	A	A	B

Note: A, Low risk; B, Unclear risk; C, High risk.

(AUC) and interobserver agreements (K)(where applicable) for the different imaging modalities. If two observers were present in a study, the mean of both observers was used for analysis purposes. A meta-analysis was conducted on the four CT studies using a random-effects model.[23–26] The results for pooled CT sensitivity and specificity are displayed on paired forest plots and a summary receiver operating characteristic curve (SROC). The I^2 statistic was used to evaluate heterogeneity (with $I^2 > 50\%$ indicating significant heterogeneity) as was Cochran's Q p value (with p value <0.05 indicating significant heterogeneity). In view of the heterogeneity found for both sensitivity and specificity, a random-effects model was used throughout. Statistical tests were considered significant against the null hypotheses if p values <0.05. The remaining two studies (investigating combination PET/CT and "CT and MRI")[22,27] are described using a narrative synthesis approach, due to the heterogeneity of these index tests prohibiting pooling of data for a meta-analysis. Publication bias was unable to be assessed (as per Cochrane Guidelines) due to the limited number of included studies (<10) in the review.[28]

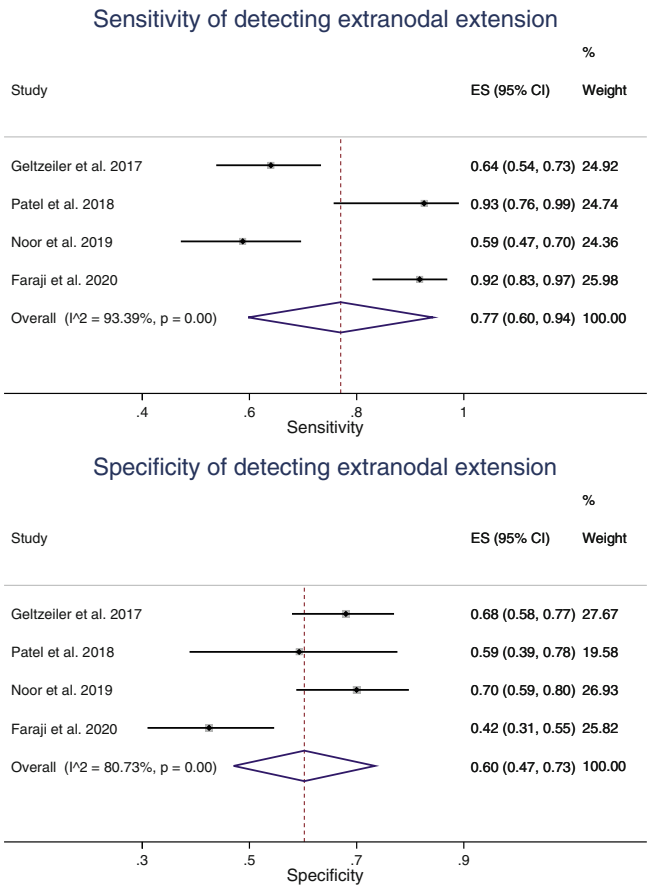

FIGURE 2 Coupled forest plot of sensitivity and specificity of CT for detecting ENE. Abbreviation: ES, effect size. [Color figure can be viewed at wileyonlinelibrary.com]

6 | RESULTS

6.1 | Study selection

A total of 1772 citations were retrieved in the initial search. Following removal of duplicates ($n = 306$), 1466 citations underwent title and abstract screening. Of the 1466, 110 papers were assessed for full-text eligibility and 104 were excluded. Six studies were included in the review, with four eligible for meta-analysis[23–26] and two undergoing a narrative synthesis.[22,27] The study selection process is shown in Figure 1.

6.2 | Study characteristics

A total of six studies with 463 participants were included in the review.[23–27] All studies were of retrospective cohort study design and conducted between 2006 and 2017. The most common site of primary tumor was tonsillar ($n = 298$) followed by base of ton-

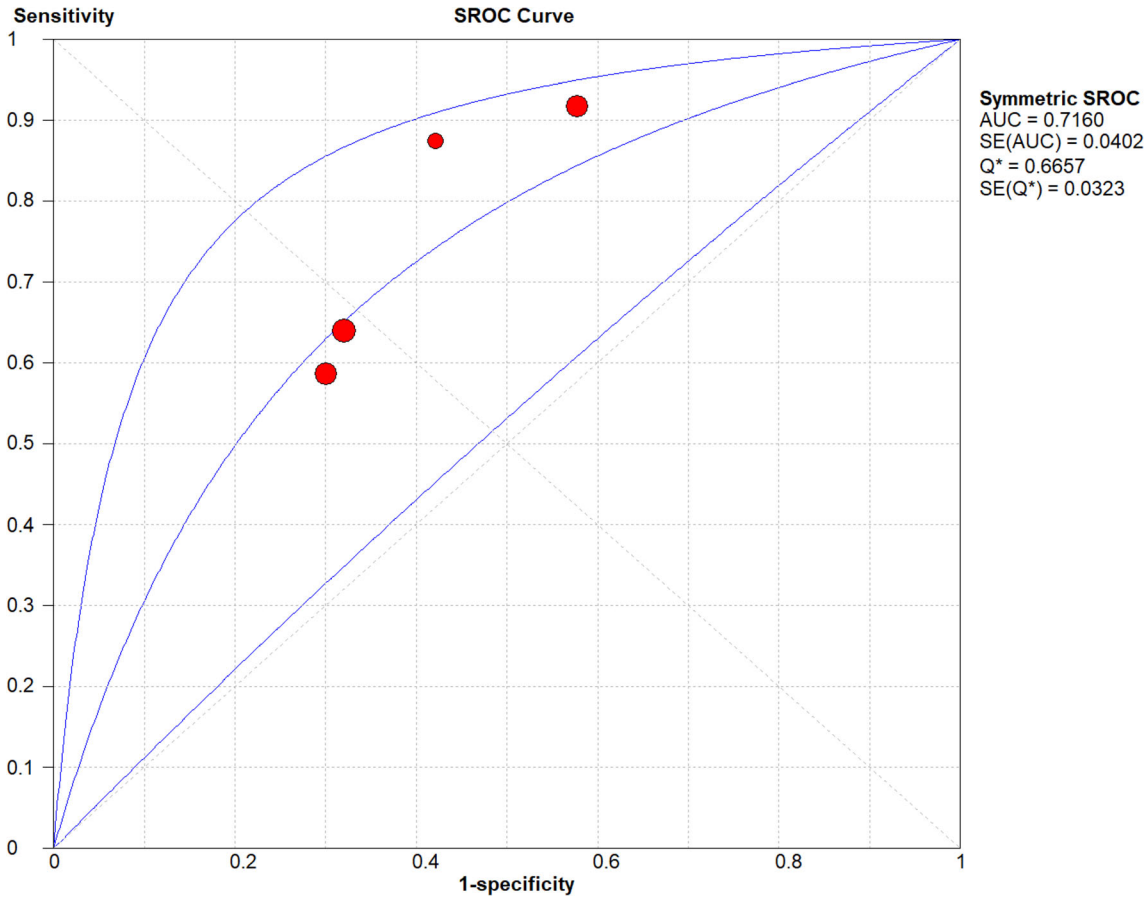

FIGURE 3 Summary receiver operating characteristic (SROC) curve of diagnostic performance of CT. The area under the SROC curve (AUC) was 0.72. Abbreviations: SE, sedimentation equilibrium; *Q**, an index defined by the point on the SROC curve where the sensitivity and specificity are equal, which is the point closest to the top-left corner of the receiver operating characteristic (ROC) space. [Color figure can be viewed at wileyonlinelibrary.com]

gue ($n = 127$). Four of the studies were conducted in the United States,[22–24,26] one in Australia[25] and one in Korea.[27] Four of the studies assessed contrast-enhanced computed tomography (ceCT),[23–26] one study assessed PET/CT,[22] and one study assessed "CT and MRI" for radiological detection of ENE.[27] No studies focusing on US or MRI met the inclusion criteria of the review. See Table 1 for a detailed description of the included studies.

6.3 | Risk of bias

None of the studies were considered to be of high risk of bias in the QUADAS-2 domains of patient selection, index test, reference standard, and flow and timing. Unclear risk of bias was found in the majority of studies in the domains of reference standard and flow and timing. This was mostly as a result of insufficient information regarding the blinding status of the pathologist to the index test, and a lack of information regarding timing between index test and reference standard. Low risk of bias was found in the majority of studies in the domains of patient selection and index test. Overall quality and applicability was not assessed per the QUADAS-2 tool and Cochrane recommendations.[20,28] A summary of results for the risk of bias assessment is presented in Table 2.

7 | REVIEW FINDINGS

7.1 | Evaluation of diagnostic performance

7.1.1 | CT

The results of the CT meta-analysis showed a pooled sensitivity of 77% (95% confidence intervals [CI] 60%–94%), and specificity of 60% (95% CI 47%–73%). A paired forest plot of the meta-analysis can be seen in Figure 2. The area under the SROC curve (AUC) was 0.72 (Figure 3). Individual sensitivity, specificity, PPV and NPV values

for each of the four CT studies can be found in Table 1. Cochran's Q test and Higgins I^2 statistics demonstrated the possibility of substantial heterogeneity in the CT meta-analysis in terms of sensitivity and specificity (I^2 93.4%, $p < 0.0001$ for Q test for sensitivity; I^2 80.7%, $p < 0.0001$ for Q test for specificity). These findings were adjusted for using a random-effects model. [Correction added on 20 September 2022, after first online publication: The in-text citation "Appendix III" has been updated as "Table 1".]

7.1.2 | PET/CT

The sole PET/CT study[22] reported their outcomes in terms of an "ENE misclassification analysis" (misclassified, under-staged, over-staged). With input from a local statistician to convert outcomes (see Appendix II), PET/CT had a calculated sensitivity of 37.5% (95% CI 15%–64%), specificity of 97% (95% CI 84%–99%), PPV of 86% and NPV of 76%. The overall diagnostic accuracy was 77.5% (95% CI 63%–88%).

7.1.3 | "CT and MRI"

The sole "CT and MRI" study[27] had a sensitivity of 62% (95% CI 53%–70%), specificity of 78% (95% CI 70%–84%), PPV of 62%, NPV of 78% and AUC of 0.70. The results were unable to be separated to individual CT and MRI outcomes.

7.2 | Interobserver agreement

Five of the studies contained assessments from two different observers to detect radiological ENE.[22,23,25–27] In these studies, inter-rater agreements were reported using the K value (Cohen's kappa coefficient). Based on Kappa statistics by Landis and Koch[29]: three studies were considered to have "moderate" inter-rater agreements (K 0.4–0.5), while the remaining two studies were considered to have "substantial" agreement between observers (K 0.7) for detecting ENE (see Table 1).

8 | DISCUSSION

By performing this review, we were able to report on the diagnostic value of CT and PET/CT for assessment of ENE in the HPV+ OPSCC population ($n = 463$). Given that HPV+ SCC has a propensity for peritumoral desmoplasia (fibrosis) which can mimic ENE, it is important guidance is available on the various imaging modalities and their accuracy for ENE in HPV+ OPSCC. Particularly, to help guide patients to appropriate treatment (i.e., whether surgery +/− adjuvant therapy vs. primary chemoradiotherapy) and to avoid trimodal therapy when there is a misdiagnosis suggesting absence of radiologic ENE.[12,13]

The pooled sensitivity, specificity and AUC values for our CT meta-analysis were 77%, 60% and 0.72, respectively. Unfortunately, due to limited numbers and the Lee et al. study design,[27] we were unable to perform a comparative meta-analysis among CT, PET/CT and MRI. However, on direct comparison, the findings from our CT meta-analysis appear to show higher sensitivity albeit lower specificity and accuracy values for ENE detection compared to the single PET/CT study[22] (86% vs. 37.5%, 76% vs. 97%, 72% vs. 77.5%, respectively). High specificity, PPV and overall diagnostic accuracy values (97%, 86%, 77.5%, respectively) possibly suggests the superior diagnostic value of PET/CT compared to CT for ENE detection in HPV+ OPSCC. In the clinical setting, CT may therefore be of greater use at helping exclude the diagnosis of ENE, while PET/CT may be of greater use in detecting the presence of ENE in HPV+ OPSCC although further studies are ultimately required. For both CT sensitivity and specificity there was substantial heterogeneity in the meta-analysis (I^2 93.4% vs. 80.7%, respectively). Due to the limited study numbers, unfortunately no meta-regression for analysis on possible causes was able to be performed.

Three previous reviews on the topic exist however all in the broader population of HNC and SCC-related HNC.[16-18] In 2015 a systematic review by Su et al.[16] investigated various imaging modalities used to detect ENE in all HNC types. The findings revealed CT to be appropriately specific (85%) however poorly sensitive (77%). MRI had potential for superior sensitivity (85%) while similar specificity (85%) to CT. US and combination PET/CT showed no evidence for their use in detecting ENE. More recently, two further reviews have been conducted by Park et al.[17] on CT and MRI in SCC-related HNC,[17] and Abdel-Halim et al.[18] on all imaging modalities in SCC-related HNC.[18] In comparison to the three previous reviews involving CT in the broader population of HNC,[16-18] the CT outcomes in our meta-analysis are inferior to all three, likely reflecting the different and more complex nature of HPV+ OPSCC. In the review by Park et al.[17] a subgroup analysis of CT in HPV+ OPSCC was performed which reported CT sensitivity to be similar to our findings (73% vs. 77%), however our pooled CT specificity was inferior (60% vs. 74%). The subgroup analysis however included the "CT and MRI" study by Lee et al.[27] which likely explains our differences in results. While comparing the single PET/CT study by Snyder et al.,[22] the sensitivity values are inferior to the previous findings in HNC by Su

et al.[16] and Abdel-Halim et al.[18] (37.5% vs. 86% vs. 80%, respectively), however the specificity values for PET/CT are superior (97% vs. 86% vs. 83%, respectively) in our review in the HPV+ OPSCC population.

In terms of diagnostic criteria for ENE, the radiological assessment of ENE remains an ongoing contentious issue. In four of the included studies,[23–26] observers utilized five or more different diagnostic features to detect radiological ENE, however with variability between studies in their choice of diagnostic features and overall approach to detecting ENE. The combination of central node necrosis, irregular nodal margins, matted nodes and perinodal stranding featured as part of ENE assessment in four of the included studies.[23–26] One study used an "overall impression" for ENE as their assessment,[22] while the other remaining study assessed for "irregularity of nodal rim" or "infiltration of adjacent soft tissues" to predict presence or absence of ENE.[27] This review alongside previous studies on ENE in HNC reflect the ongoing subjective nature of radiological ENE assessment.

Despite the exclusion of ENE in the most recent American Joint Committee on Cancer (AJCC) 8th edition staging system for HPV+ OPSCC, there is currently insufficient evidence on the prognostic role of ENE in HPV+ OPSCC to support de-intensification therapy.[30] As a result, many clinicians are continuing to utilize the AJCC 7th edition for treatment decisions concerning ENE in HPV+ OPSCC.[31] Therefore, given that the radiological or histopathological diagnosis of ENE continues to warrant treatment with chemoradiotherapy (primary or adjunct),[12,13] there is clinical concern that a certain percentage of patients may undergo unnecessary treatment with chemoradiotherapy related side-effects when the radiological diagnosis of ENE may be a false-positive. The pre-treatment diagnosis of ENE therefore plays an imperative role in the work up of patients with HPV+ OPSCC and suspected ENE. The results of our review suggest that the use CT alone may be of greater value in helping exclude rather than diagnose the presence of ENE in HPV+ OPSCC. Therefore if the use of CT suggests the absence of radiological ENE, these patients may be better served with surgery (+/− adjuvant therapy) as opposed to primary chemoradiotherapy which has been found to have a more toxic side-effect profile.[11] Conversely, the potential superior diagnostic value of PET/CT for ENE detection in HPV+ OPSCC suggests that those with radiological ENE on PET/CT may be better served with primary chemoradiotherapy compared to a surgical approach.[11]

Interestingly, new research is looking into degrees of nodal extension (microscopic versus macroscopic ENE >2 mm) and their associated prognostic implications. The ECOG-ACRIN 3311 trial suggested that only macroscopic ENE is likely to require chemotherapy and full-dose radiotherapy (66-Gy), while surgery and low-dose adjuvant radiotherapy (50 or 60-Gy) might be an appropriate treatment regimen for patients with microscopic ENE <1 mm.[32] If these findings are accepted clinically, and research continues to focus on treatment de-intensification protocols, imaging modalities will need high precision to accurately differentiate between microscopic and macroscopic ENE.

More recently, newer imaging modalities such as texture analysis and machine learning methods are currently being investigated in attempts to improve ENE detection and reduce the current subjective nature of its assessment among radiologists.[33,34] Recent findings by Kann et al.[34] in SCC-related HNC suggest their CT machine learning algorithm showed superior diagnostic advantage at detecting ENE with an achieved AUC of 0.90 (88.6% accuracy), outperforming their radiologists AUCs of 0.60 and 0.82 ($p < 0.0001$ and $p = 0.16$). The highest sensitivity and specificity values for the CT machine learning method were 82% and 91%, respectively. While the variable sensitivity and specificity values for either radiologist ranged from 24%–71% and 75%–96%, respectively, for detection of ENE. In a HPV+ SCC-related HNC subgroup analysis, the machine learning method also outperformed either radiologist with an AUC of 0.81 compared to 0.75 and 0.56. Although these findings were identified in SCC-related HNC, machine learning methods may have superior diagnostic value in the HPV+ OPSCC population and further research is required.

Despite every effort throughout the methodological steps to minimize limitations in the review, several limitations do exist. Although the search was conducted both electronically and manually with the aid of a scientific librarian, the defined exclusion criteria resulted in six non-English studies being excluded and only published studies being retrieved. It is therefore uncertain whether all relevant studies on the topic have been included in the review.

Regarding the included studies, all studies were of retrospective study design which has an impact on selection bias among the included participants. Furthermore, in the CT meta-analysis involving sensitivity and specificity substantial heterogeneity was found. Multiple factors such as patient populations, radiologist experience level, and use of different radiological features for the diagnosis of ENE could have contributed to this however no meta-regression was able to be performed to explore this further. Therefore, given the small number of included studies and findings of substantial heterogeneity, caution must be taken for widespread extrapolation of the review findings.

9 | CONCLUSION

In conclusion, pooled CT specificity values (60%) appear too low to suggest clinical value for CT in detecting ENE in

HPV+ OPSCC. The use of CT however might have clinically acceptable capabilities at helping exclude the diagnosis of radiological ENE in HPV+ OPSCC. This may be of benefit to patients who are presumed ENE negative on radiology and recommended a pathway of surgery (+/− adjuvant therapy) rather than primary chemoradiotherapy, with the attendant reduction in treatment-related toxicity. While findings from the single PET/CT study[22] suggest PET/CT may be of superior diagnostic value for ENE detection in HPV+ OPSCC however further diagnostic studies are ultimately required to allow for a true comparative analysis between CT, PET/CT and MRI. It is conceivable these will be augmented by new radiological modalities such as machine learning to improve diagnostic accuracy in the future.

AUTHOR CONTRIBUTIONS

Tristan Morey: study concepts and design, data collection, data analysis, data interpretation, manuscript writing, and critical revision. **John-Charles Hodge**: study concepts and design, critical revision, and supervision. **Cindy Stern**: data collection, data interpretation, critical revision, and supervision. **Suren Krishnan**: study concepts and design, and critical revision. **Andrew Foreman**: study concepts and design, data collection, critical revision, and supervision.

ACKNOWLEDGMENT

All the authors contributed to the preparation of the systematic review. Open access publishing facilitated by The University of Adelaide, as part of the Wiley - The University of Adelaide agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Tristan Morey https://orcid.org/0000-0002-6660-542X
Andrew Foreman https://orcid.org/0000-0002-6560-6391

REFERENCES

1. Siegel R, Miller K, Fuchs H, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin*. 2022;72(1):7-33.
2. Marur S, Forastiere A. Head and neck cancer: changing epidemiology, diagnosis, and treatment. *Mayo Clin Proc*. 2008;83(4):489-501.
3. Sturgis E, Cinciripini P. 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.
4. Chaturvedi A, Engels E, Pfeiffer R, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol*. 2011;29:4294-4301.
5. Hong A, Lee CS, Jones D, et al. Rising prevalence of human papillomavirus-related oropharyngeal cancer in Australia over the last 2 decades. *Head Neck*. 2016;38(5):743-750.
6. Jung AC, Briolat J, Millon R, et al. Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *Int J Cancer*. 2010;126:1882-1894.
7. 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(27):4142-4148.
8. Posner MR, Lorch JH, Goloubeva O, et al. Survival and human papillomavirus in oropharynx cancer in TAX324: a subset analysis from and international phase III trial. *Ann Oncol*. 2011;22(5):1071-1077.
9. Park CH, Song CM, Ji YB, et al. Significance of the extracapsular spread of metastatic lymph nodes in papillary thyroid carcinoma. *Clin Exp Otorhinolaryngol*. 2015;8(3):289-294.
10. Bernier J, Cooper JS, Pajak TF, 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):843-850.
11. Cooper JS, Zhang Q, Pajak TF, et al. Long-term follow-up of the RTOG 9501/intergroup phase III trial: postoperative concurrent radiation therapy and chemotherapy in high-risk squamous cell carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys*. 2012;84(5):1198-1205.
12. National Comprehensive Cancer Network. NCNN Guidelines for Cancer of the Oropharynx (version 1.2022). Retrieved from https://www.nccn.org/patients/guidelines/content/PDF/hn-oropharyngeal-patient.pdf.
13. Machiels JP, René Leemans C, Golusinski W, Grau C, Licitra L, Gregoire V. Squamous cell carcinoma of the oral cavity, larynx, oropharynx and hypopharynx: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2020;31(11):1462-1475.
14. Haave H, Gulati S, Brekke J, Lybak S, Vintermyr OK, Aarstad HJ. Tumor stromal desmoplasia and inflammatory response uniquely predict survival with and without stratification for HPV tumor infection in OPSCC patients. *Acta Otolaryngol*. 2018;138(11):1035-1042.
15. Moore ZR, Pham NL, Shah JL, et al. Risk of unplanned hospital encounters in patients treated with radiotherapy for head and neck squamous-cell carcinoma. *J Pain Symptom Manage*. 2019;57:738-745.e3.
16. Su Z, Duan Z, Pan W, et al. Predicting extracapsular spread of head and neck cancers using different imaging techniques: a systematic review and metaanalysis. *Int J Oral Maxillofac Surg*. 2016;45(4):413-421.
17. Park S, Guenette J, Suh C, et al. The diagnostic performance of CT and MRI for detecting extranodal extension in patients with head and neck squamous cell carcinoma: a systematic review and diagnostic meta-analysis. *Eur Radiol*. 2021;31:2048-2061.

18. Abdel-Halim C, Rosenberg T, Dyrvig A, et al. Diagnostic accuracy of imaging modalities in detection of histopathological extranodal extension: a systematic review and meta-analysis. *Oral Oncol*. 2021;114:105169.
19. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Med*. 2009;6(7): e1000097.
20. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
21. Campbell JM, Kulgar M, Ding S, et al. Diagnostic test accuracy systematic reviews. In: Aromataris E, Munn Z, eds. *JBI Manual for Evidence Synthesis*. Vol 1. JBI;2020:355. Available from: https://doi.org/10.46658/JBIMES-20-10
22. Snyder V, Goyal LK, Bowers EMR, et al. PET/CT poorly predicts AJCC 8th edition pathologic staging in HPV-related oropharyngeal cancer. *Laryngoscope*. 2021;131(7):1535-1541.
23. Faraji F, Aygun N, Coquia SF, et al. Computed tomography performance in predicting extranodal extension in HPV-positive oropharynx cancer. *Laryngoscope*. 2020;130(6):1479-1486.
24. Geltzeiler M, Clayburgh D, Gleysteen J, et al. Predictors of extracapsular extension in HPV-associated oropharyngeal cancer treated surgically. *Oral Oncol*. 2017;65:89-93.
25. Noor A, Mintz J, Patel S, et al. Predictive value of computed tomography in identifying extracapsular spread of cervical lymph node metastases in p16 positive oropharyngeal squamous cell carcinoma. *J Med Imaging Radiat Oncol*. 2019;63(4): 500-509.
26. Patel MR, Hudgins PA, Beitler JJ, et al. Radiographic imaging does not reliably predict macroscopic extranodal extension in human papilloma virus-associated oropharyngeal cancer. *ORL J Otorhinolaryngol Relat Spec*. 2018;80(2):85-95.
27. Lee B, Choi YJ, Kim SO, et al. Prognostic value of radiologic extranodal extension in human papillomavirus-related oropharyngeal squamous cell carcinoma. *Korean J Radiol*. 2019;20(8): 1266-1274.
28. Deeks JJ, Bossuyt PM, Gatsonis C. Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0. the cochrane collaboration, 2010. Available from: http://srdta.cochrane.org/
29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
30. Amin MB, Edge S, Greene F, et al. *American Joint Commission on Cancer. AJCC Cancer Staging Manual*. 8th ed. Springer International Publishing; 2017.
31. Edge S, Byrd D, Compton C, et al. *American Joint Commission on Cancer. AJCC Cancer Staging Manual*. 7th ed. Springer International Publishing; 2010.
32. Ferris R, Flamand Y, Weinstein G, et al. Phase II randomized trial of transoral surgery and low-dose intensity modulated radiation therapy in Resectable p16+ locally advanced oropharynx cancer: an ECOG ACRIN cancer research group trial (E3311). *J Clin Oncol*. 2022;40(2):138-149.
33. Frood R, Palkhi E, Barnfield M, Prestwich R, Vaidyanathan S, Scarsbrook A. Can MR textural analysis improve the prediction of extracapsular nodal spread in patients with oral cavity cancer? *Eur Radiol*. 2018;28(12):5010-5018.
34. Kann BH, Hicks DF, Payabvash S, et al. Multi-institutional validation of deep learning for pretreatment identification of extranodal extension in head and neck squamous cell carcinoma. *J Clin Oncol*. 2020;38(12):1304-1311.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Morey T, Hodge J-C, Stern C, Krishnan S, Foreman A. Correlation between radiologic and pathologic extranodal extension in HPV-associated oropharyngeal cancer: Systematic review. *Head & Neck*. 2022;44(12): 2875-2885. doi:10.1002/hed.27183