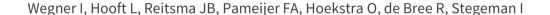![Cochrane Library logo]

Cochrane Database of Systematic Reviews

MRI versus CT versus 18F-FDG PET-CT for detecting lymph node metastases in patients with head and neck squamous cell carcinoma (Protocol)

Wegner I, Hooft L, Reitsma JB, Pameijer FA, Hoekstra O, de Bree R, Stegeman I

www.cochranelibrary.com

WILEY

T A B L E   O F   C O N T E N T S

[Diagnostic Test Accuracy Protocol]

MRI versus CT versus 18F-FDG PET-CT for detecting lymph node metastases in patients with head and neck squamous cell carcinoma

Inge Wegner[1], Lotty Hooft[2], Johannes B Reitsma[3], Frank A Pameijer[4], Otto Hoekstra[5], Remco de Bree[6], Inge Stegeman[1]

[1]Department of Otorhinolaryngology & Head and Neck Surgery, University Medical Center Utrecht, Utrecht, Netherlands. [2]Dutch Cochrane Centre, Julius Center for Health Sciences and Primary Care / University Medical Center Utrecht, Utrecht, Netherlands. [3]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands. [4]Department of Radiology, University Medical Center Utrecht, Utrecht, Netherlands. [5]Department of Radiology & Nuclear Medicine, VU Medical Center, Amsterdam, Netherlands. [6]Department of Head and Neck Surgical Oncology, UMC Utrecht Cancer Center, University Medical Center Utrecht, Utrecht, Netherlands

Contact address: Inge Wegner, Department of Otorhinolaryngology & Head and Neck Surgery, University Medical Center Utrecht, Utrecht, Netherlands. i.wegner@umcutrecht.nl.

A B S T R A C T

This is the protocol for a review and there is no abstract. The objectives are as follows:

To determine the diagnostic accuracy of MRI, CT and PET-CT for detecting lymph node metastases in patients with head and neck squamous cell carcinoma.

B A C K G R O U N D

Head and neck cancer encompasses malignant tumours of the upper aerodigestive tract including the pharynx, larynx, oral cavity, nasal cavity, paranasal sinuses and salivary glands. Approximately 550,000 new cases are diagnosed each year throughout the world (Jemal 2011), the majority of which are mucosal squamous cell carcinoma (NCCN 2014). Tobacco and alcohol abuse are two important risk factors associated with oral, pharyngeal and laryngeal cancer (Blot 1988; Decker 1982). Viral infection also plays a role, with associations between the human papilloma virus and oropharyngeal cancer (Gillison 2000), as well as the Epstein-Barr virus and nasopharyngeal carcinoma (Sankaranarayanan 1998).

An individual patient's prognosis is determined by the type and extent of their cancer, established during staging. Head and neck cancer staging takes into consideration anatomic subsite, tumour size, cervical lymph node involvement and the presence of distant metastasis (AJCC 2010). Up to 40% of patients have early stage I and II cancer when they first present (NCCN 2014).

Treatment options include surgery, radiation and chemotherapy. The majority of early stage I and II patients can be treated with single modality therapy using either surgery or radiation alone, and survival rates are similar for both types of treatment (Gregoire

2010; Higgins 2009; NCCN 2014). In contrast, when advanced stage III and IV cancer is treated with the aim of curing the patient, this requires multimodality therapy to include surgery with adjuvant radiotherapy or organ preservation chemoradiation. Adjuvant chemotherapy has proven beneficial for some patients with advanced disease (Forastiere 2003). Ultimately, head and neck cancer treatment is individualised to the patient and based not only on the stage of the cancer and the likely prognosis associated with that stage, but also the patient's comorbidities and wishes. Sometimes treatment is palliative and not intended to try and elicit a cure.

Target condition being diagnosed

The target condition is lymph node metastasis in head and neck squamous cell carcinoma. More than 45% of patients with squamous cell carcinomas of the oral cavity and pharynx and more than 20% of patients with laryngeal squamous cell carcinoma present with lymph node metastasis at the time of initial diagnosis (SEER 2014). The presence of lymph node metastasis is an important prognostic factor in head and neck squamous cell carcinoma (Shah 1990). The presence of nodal metastasis reduces survival by nearly 25% to 30% (SEER 2014).

Accurate assessment of lymph node status is essential not only in determining prognosis, but also in choosing a treatment strategy (Snow 1992). The clinically node-positive neck (cN+) is treated surgically with or without adjuvant (chemo)radiation or with primary (chemo)radiation (de Bree 2014). Management of the clinically node-negative neck (cN0) is controversial. Elective treatment of the neck is indicated in case of a high chance of occult lymph node metastases, when the neck needs to be entered anyway to resect the primary tumour or reconstruct a surgical defect, or when regular follow-up is not an option (de Bree 2014). If occult metastases are not treated electively, patients will ultimately develop clinically manifest disease, some of which is extensive or even inoperable. Delayed treatment of the neck is more extensive and associated with higher morbidity than elective neck dissection. On the other hand, elective treatment is also associated with morbidity, costs and altered routes of cancer spread in case of recurrence or second primary tumours. A recent randomised clinical trial showed that after limited diagnostic work-up, elective neck dissection results in better overall survival as compared to follow-up with watchful waiting and therapeutic neck dissection when lymph node metastases become clinically manifest (D'Cruz 2015). Improved diagnostic work-up and follow-up using ultrasound-guided fine needle aspiration cytology showed, in experienced hands, a similar disease-free survival and overall survival for elective neck dissection and watchful waiting patients (Flach 2013). The reliable identification of lymph node metastasis is thus very important in choosing a treatment strategy.

Index test(s)

Magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET)-CT can be used to assess lymph node status. Criteria for suspicious lymph nodes on MRI and CT are based on the size of the node, contrast enhancement and the presence of central necrosis or extranodal extension (de Bree 2014). Since both MRI and CT rely on morphological and size-related criteria, small metastases or micrometastatic nodes are easily missed (de Bondt 2007; de Bree 2009a). PET imaging with 18F-fluorodeoxyglucose (18F-FDG) is a functional imaging modality that is increasingly used for staging of head and neck squamous cell carcinoma. However, 18F-FDG accumulates in tumours as well as inflammatory and infectious diseases (de Bree 2009b). As a result, the use of 18F-FDG PET-CT in the detection of cancer (metastases) may be associated with higher rates of false positives. Whenever we speak of PET-CT in this protocol, we mean the use of 18F-FDG PET-CT.

Clinical pathway

All patients with a diagnosis of head and neck primary cancer are routinely screened for lymph node metastases. They undergo a thorough physical examination, including palpation of the neck. However, physical examination cannot accurately detect metastasis, not even in the superficial cervical region (Ali 1985; van den Brekel 1991). Size of the lymph nodes, location in the neck, body habitus of the patient and prior surgery or radiotherapy make palpation of early disease difficult. Furthermore, mediastinal, parapharyngeal and retropharyngeal lymph nodes are not accessible to clinical examination. Therefore, imaging is usually performed to assess lymph node status preoperatively in both cN0 and cN+. CT, MRI and/or PET-CT are frequently used for evaluating lymph node status. Clinically or radiographically detected lymph nodes are biopsied under ultrasound or, less frequently, CT guidance.

Rationale

There is no general agreement on whether to use MRI or CT in assessing lymph node status and the choice between the two is largely influenced by availability, local expertise and experience, and local guidelines (NCCN 2014). PET-CT is increasingly used for staging of head and neck squamous cell carcinoma. Combining PET and CT unites both anatomic and functional imaging, and thus might increase diagnostic accuracy, particularly if the CT is of diagnostic quality and not only used for anatomical localisation of hot spots and attenuation correction.

O B J E C T I V E S

To determine the diagnostic accuracy of MRI, CT and PET-CT for detecting lymph node metastases in patients with head and neck squamous cell carcinoma.

METHODS

Criteria for considering studies for this review

Types of studies

We will include randomised controlled studies, cross-sectional diagnostic studies and cohort studies. These may be either prospective or retrospective. Patients should be included consecutively. We will exclude case-control studies. We will include studies with a direct comparison of two or three imaging modalities in the same patient population. We will also include studies in which only one of the three index tests and studies in which multiple index tests are evaluated, but not in the same patient population. We realise that ideally all three index tests are used in the same patient population. However, a first exploration of the literature has led us to believe there are insufficient studies to limit inclusion to studies with such a design.

Furthermore, studies must report sufficient data in order to extract the following:

- the number of true positives: patients categorised as diseased by both the index and reference test;
- the number of false negatives: patients categorised as diseased by the reference test, but as non-diseased by the index test;
- the number of true negatives: patients categorised as non-diseased by both the index and reference test;
- the number of false positives: patients categorised as non-diseased by the reference test, but as diseased by the index test.

We will not use any restrictions based on a minimal quality standard, minimal sample size or number of patients with the target condition (i.e. lymph node metastases).

Participants

We will include studies that include patients with histopathologically proven head and neck squamous cell carcinoma. We will exclude studies that predominantly include patients with nasopharyngeal, sinonasal, thyroid, salivary gland, skin and unknown primary cancer. These tumours are characterised by a different pathophysiology. We will not exclude studies with a mixed study population, including up to 10% of patients with any of these cancer types. We will only include studies including patients without previous neck dissection, to ensure that the reference standard (definitive histology) can be obtained. We will not exclude studies including patients that have undergone previous (chemo)radiation.

Study inclusion will not be limited by age, gender or T-classification. All included patients should have received an MRI, CT and/or PET-CT for lymph node staging purposes. Study inclusion will not be limited by clinical N-classification; we will include studies including both cN0 and cN+ patients.

Index tests

The index tests under review are MRI, CT and PET-CT. These tests were chosen because they are currently used to detect lymph node metastases in head and neck cancer patients. Different criteria can be used to diagnose lymph node metastases, including morphological criteria and size-related criteria for MRI and CT, and quantitative metrics for PET. We will not limit study inclusion based on the diagnostic criteria that were used. It needs to be noted though that differences in diagnostic criteria between MRI, CT and PET-CT may introduce variation in the percentage of cN0 and cN+ patients.

We will limit study inclusion based on the quality of the MRI, CT and/or PET-CT. Regarding CT, we will exclude studies using a CT slice thickness of more than 3 mm. Intravenous contrast should have been used in the studies in order to be included. Regarding MRI, we will exclude studies using a MRI slice thickness of more than 4 mm and/or an interslice gap of more than 0.4 mm. At least T1- and T2-weighted, turbo spin echo (TSE) or short tau inversion recovery (STIR) sequences should have been used, as well as a post-Gadolinium contrast series. Either a 1.5 or 3.0 Tesla magnet should have been used. Regarding PET-CT, at least 18F-FDG should have been used. Even though we will limit exclusion based on the quality of the MRI, CT and/or PET-CT, small variations in these imaging techniques may introduce bias.

Target conditions

Lymph node metastases in patients with histologically proven head and neck squamous cell carcinoma.

Reference standards

The primary reference standard is definitive histopathology following surgery, preferably in combination with one-year clinical follow-up during which patients will not receive any further treatment. Since the former is invasive, it is not usually performed on patients with negative index test results, which can lead to verification bias. Long-term clinical follow-up of at least one year is the primary differential reference standard in patients that will not undergo neck dissection. If no obvious lymph node metastases are present at one-year clinical follow-up, these patients are considered negative. However, studies often only include patients with an indication for surgery and exclude patients that will not undergo treatment or radio(chemo)therapy. As it seems likely that in these study populations the reported prevalence of lymph node metastases is higher than it is in the general population of patients

with head and neck squamous cell carcinoma, this could lead to overestimation of the positive predictive value and underestimation of the negative predictive value. We will not pool the results for these two reference standards (histopathology and clinical follow-up). We will perform a subgroup analysis instead.

It is important to note that different histopathological techniques can be used to assess lymph nodes. Routine histopathological examination can miss minimal lymph node metastases. Step-serial sectioning and immunohistochemistry can increase the yield of diseased lymph nodes by up to 15% (Rinaldo 2004). The use of routine histopathological examination likely leads to an overestimation of the number of true negatives. Furthermore, the reported pre-test probability or prevalence of disease is influenced by the amount of lymph nodes harvested. Increased lymph node harvesting lowers the risk of missing micrometastases (van den Brekel 1996). The lymph node ratio (ratio of the number of positive lymph nodes to total number of lymph nodes in the surgical specimen) is not only a strong prognostic factor in head and neck cancer, it is also influenced by the amount of harvested lymph nodes (Marres 2014). We will be using the total number of harvested lymph nodes to assess the quality of the neck dissections performed.

Search methods for identification of studies

The Cochrane ENT Information Specialist will conduct systematic searches for diagnostic test accuracy studies (Appendix 1). There will be no language, publication year or publication status restrictions. We may contact original authors for clarification and further data if study reports are unclear. If data from the data extraction sheets (Appendix 2) or methodological quality assessment criteria (Appendix 3) are missing, we will contact the original authors. We will arrange translations of papers where necessary.

Electronic searches

Published, unpublished and ongoing studies will be identified by searching the following databases from their inception:
- Ovid MEDLINE (1946 to date);
- Ovid MEDLINE (In-Process & Other Non-Indexed Citations);
- PubMed (as a top up to searches in Ovid MEDLINE);
- Ovid EMBASE (1974 to date);
- EBSCO CINAHL (1982 to date);
- Web of Knowledge, Web of Science (1945 to date);
- ClinicalTrials.gov, www.clinicaltrials.gov (search via the Cochrane Register of Studies to date);
- World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (search to date);
- ISRCTN, www.isrctn.com (search to date);
- Google Scholar (search to date);
- Google (search to date).

The subject strategies for databases will be modelled on the search strategy designed for MEDLINE (Appendix 1).

Searching other resources

We will scan the reference lists of identified publications for additional studies and contact study authors if necessary. In addition, the Information Specialist will search PubMed, TRIPdatabase and Google to retrieve existing systematic reviews relevant to this systematic review, so that we can scan their reference lists for additional studies. We will search for conference abstracts using the Cochrane ENT Trials Register and EMBASE.

Data collection and analysis

Selection of studies

We will import all references identified by the electronic searches into RefWorks and remove duplicates. Two review authors (IW and IS) will independently screen the retrieved records. Discrepancies will be resolved through a consensus discussion between the two review authors. If disagreements persist, a third review author will resolve the discrepancies (LH, JBR, FAP, OSH or RdB). The third review author will be chosen based on his or her expertise. LH and JBR will resolve discrepancies regarding epidemiology/ methodology, FAP and OSH regarding radiology, and RdB regarding clinical questions. We will review titles and abstracts and exclude those articles that are obviously ineligible. We will exclude systematic reviews, opinion papers, editorials, conference abstracts and poster presentations at this stage. We will exclude studies that clearly do not evaluate the diagnostic accuracy of MRI, CT or PET-CT, but other imaging modalities such as diffusion-weighted MRI. We will exclude studies reviewing the use of MRI, CT and/ or PET-CT following surgical treatment, as well as studies evaluating treatment response.

We will proceed by retrieving and reviewing the full-text copies. We will only include those studies evaluating the diagnostic accuracy of MRI, CT and/or PET-CT in comparison with definitive histology or clinical follow-up in patients with head and neck squamous cell carcinoma for detecting lymph node metastases. We will exclude studies evaluating the diagnostic accuracy of these imaging modalities for establishing tumour characteristics other than lymph node metastases. We will only include studies if they meet the MRI, CT and/or PET-CT criteria listed under Index tests. Studies will not be limited by language, location or setting at any stage of the selection procedure.

Data extraction and management

Two review authors (IW and IS) will independently extract key data regarding the study population, the index test, the reference

test and the diagnostic test results from the included studies using data extraction sheets. A complete list of characteristics and results that we will be extracting is enclosed in Appendix 2. Discrepancies will be resolved through a consensus discussion between the two review authors or a third review author will resolve the discrepancies (LH, JBR, FAP, OSH or RdB).

The two review authors (IW and IS) are capable of assessing and evaluating articles written in English, Dutch, German and French. Data from articles written in other languages will be extracted by radiologists or ENT surgeons with knowledge of the language, correspondence with the study authors, or by a translator working in conjunction with two review authors (IW and IS).

Assessment of methodological quality

Two review authors (IW and IS) will independently assess the methodological quality of the included studies using predefined criteria. Discrepancies will be resolved through a consensus discussion between the two review authors or a third review author will resolve discrepancies (LH, JBR, FAP, OSH or RdB). We will use the QUADAS-2 tool to assess the methodological quality (Whiting 2011). A full list of the criteria and their operational definitions is enclosed in Appendix 3. We will include a filled out QUADAS-2 form for each of the included studies, as well as a methodological quality summary figure and a methodological quality graph. The methodological quality summary figure presents for each included study the 'yes', 'no' and 'unclear' judgements for each quality assessment item in graphical form. The methodological quality graph presents for each quality assessment item the percentage of the included studies that rate the item 'yes', 'no' and 'unclear' in a stacked bar chart.

We will include studies regardless of their quality of evidence. Instead, we will explore the impact of (poor) methodological quality using meta-regression analyses. The methods are described in the section Investigations of heterogeneity.

Statistical analysis and data synthesis

Disease status and index test results are both binary, leading to the extraction of true-positive, false-positive, true-negative and false-negative test results for MRI, CT and/or PET-CT for each included study. We will construct forest plots showing estimates of sensitivity and specificity and their corresponding 95% confidence intervals for each study. We will sort forest plots by imaging technique (MRI, CT and PET-CT). We will calculate confidence intervals around these proportions according to the method of Wilson (Newcombe 1998). We will also plot sensitivities and specificities in receiver operating characteristic (ROC) space, using different symbols for MRI, CT and PET-CT estimates. For all direct comparisons, we will link MRI, CT and/or PET-CT points in ROC space derived from the same study. We will use random-effects bivariate regression models to meta-analyse the logit transformed sensitivity and specificity of MRI, CT and PET-CT to

obtain pooled estimates and 95% confidence intervals of these parameters (Macaskill 2010). We will use the exact binomial distribution to model the within-study variance. Therefore, we will use non-linear mixed models to estimate the parameters of interest. We will compare the accuracy between MRI, CT and PET-CT by adding a covariate for the type of index test to the model to investigate whether sensitivity, specificity, or both, are different between the three imaging techniques. If sufficient studies providing direct evidence are available (n = 5 or higher), we will perform a sensitivity analysis focusing on these studies alone, taking the paired nature into account (Trikalinos 2013). In case of paired diagnostic studies (study participants receive more than one index test under investigation) there is additional information available if such studies report the fully cross-classified data. It means a two-by-two table of one index test against the other index test among the patients with the target condition and another two-by-two table among patients without target condition. Unfortunately, paired diagnostic accuracy studies hardly report their data in such format, and this is also our experience within this field. Therefore, we will perform a sensitivity analysis of the paired studies in which we will focus on the difference in logit sensitivity between the index tests in each study and then summarise these differences across studies. We will do the same for differences in logit specificity. When the number of studies is either small or the correlation is low, we will perform separate meta-analysis, otherwise we will extend the model by incorporating the correlation between logit difference sensitivity and logit difference specificity over studies. We will compare the results of the paired studies with the results of the analysis including all studies.

We will present the results of the bivariate regression models in ROC space, showing pooled estimates of sensitivity and specificity together with 95% confidence intervals around these pooled estimates (Reitsma 2005).

We will perform analyses using SAS and R software.

Investigations of heterogeneity

We will add covariates to the bivariate regression models to assess whether heterogeneity in sensitivity and/or specificity can be related to the following factors:

- methodological quality rating for the four QUADAS-2 domains (low risk of bias versus high risk of bias versus unclear risk of bias);
- previous (chemo)radiotherapeutic treatment of the neck region (yes versus no);
- subsite distribution (oral versus oropharyngeal versus hypopharyngeal versus laryngeal versus other);
- N-classification distribution (cN0 versus cN+);
- magnetic field strength used (1.5 Tesla versus 3.0 Tesla);
- slice thickness imaging;
- slice thickness pathology;

- reference test (histopathology with clinical follow-up versus histopathology versus clinical follow-up);
  - use of immunohistochemistry;
  - use of step serial sectioning;
  - number of harvested lymph nodes.

We will include one covariate in the analyses at a time.

Sensitivity analyses

We will perform two sensitivity analyses:

- a separate analysis of studies with an overall low risk of bias, characterised by a low risk of bias in three out of four QUADAS-2 domains (patient selection, index test, reference standard, and flow and timing) compared to the analysis of all studies combined regardless of their methodological quality;

- a separate analysis of studies with direct comparisons compared to the analysis of studies with both direct and indirect comparisons.

Assessment of reporting bias

There is a lack of sensitive tests for use in diagnostic test accuracy reviews and the determinants of publication bias are unknown. Therefore we elect not to investigate reporting bias.

REFERENCES

Additional references

AJCC 2010
Edge S, Byrd DR, Compton CC, Fritz, AG, Greene FL, et al. *AJCC Cancer Staging Manual*. 7th Edition. New York, NY: Springer, 2010.

Ali 1985
Ali S, Tiwari RM, Snow GB. False positive and false negative neck nodes. *Head and Neck Surgery* 1985;**8**(2):78–82.

Blot 1988
Blot WJ, McLaughlin JK, Winn DM, Austin DF, Greenberg RS, Preston-Martin S, et al. Smoking and drinking in relation to oral and pharyngeal cancer. *Cancer Research* 1988;**48**(11):3282.

D'Cruz 2015
D'Cruz AK, Vaish R, Kapre N, Dandekar M, Gupta S, Hawaldar R, et al. Head and Neck Disease Management Group. Elective versus therapeutic neck dissection in node-negative oral cancer. New England Journal of Medicine 2015 May 31 [Epub ahead of print].

de Bondt 2007
de Bondt RB, Nelemans PJ, Hofman PA, Casselman JW, Kremer B, van Engelshoven JM, et al. Detection of lymph node metastases in head and neck cancer: a meta-analysis comparing US, USg FNAC, CT and MR imaging. *European Journal of Radiology* 2007;**64**(2):266-72.

de Bree 2009a
de Bree R, Castelijns JA, Hoekstra OS, Leemans CR. Advances in imaging in the work-up of head and neck cancer patients. *Oral Oncology* 2009;**45**:930–5.

de Bree 2009b
de Bree R, van der Putten L, Brouwer J, Castelijns JA, Hoekstra OS, Leemans CR. Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging. *Oral Oncology* 2009;**45**(4-5): 386–93.

de Bree 2014
de Bree R, Takes RP, Castelijns JA, Medina JE, Stoeckli SJ, Mancuso AA, et al. Advances in diagnostic modalities to detect occult lymph node metastases in head and neck squamous cell carcinoma. Head and Neck 2014 Jun 21 [Epub ahead of print].

Decker 1982
Decker J, Goldstein JC. Risk factors in head and neck cancer. *New England Journal of Medicine* 1982;**306**(19): 1151–5.

Flach 2013
Flach GB, Tenhagen M, de Bree R, Brakenhoff RH, van der Waal I, Bloemena E, et al. Outcome of patients with early stage oral cancer managed by an observation strategy towards the N0 neck using ultrasound guided fine needle aspiration cytology: no survival difference as compared to elective neck dissection. *Oral Oncology* 2013;**49**(2):157–64.

Forastiere 2003
Forastiere AA, Goepfert H, Maor M, Pajak TF, Weber R, Morrison W, et al. Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. *New England Journal of Medicine* 2003;**349**:2091–8.

Gillison 2000

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. *Journal of the National Cancer Institute* 2000;**92**: 709–20.

Gregoire 2010

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. *Annals of Oncology* 2010;**21**(Suppl 5):184–6.

Higgins 2009

Higgins KM, Shah MD, Ogaick MJ, Enepekides D. Treatment of early-stage glottic cancer: meta-analysis comparison of laser excision versus radiotherapy. *Journal of Otolaryngology - Head & Neck Surgery* 2009;**38**(6):603–12.

Jemal 2011

Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: A Cancer Journal for Clinicians* 2011;**61**(2):69.

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. Available from: http://srdta.cochrane.org/. The Cochrane Collaboration, 2010.

Marres 2014

Marres CCM, de Ridder M, Hegger I, van Velthuysen ML, Hauptmann M, Navran A, et al. The influence of nodal yield in neck dissections on lymph node ratio in head and neck cancer. *Oral Oncology* 2014;**50**(1):59–64.

NCCN 2014

National Comprehensive Cancer Network. Head and Neck Cancers (Version 2.2014). http://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf (accessed June 2015).

Newcombe 1998

Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine* 1998;**17**(8):857-72.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982-90.

Rinaldo 2004

Rinaldo A, Devaney KO, Ferlito A. Immunohistochemical studies in the identification of lymph node micrometastases in patients with squamous cell carcinoma of the head and neck. *ORL; Journal of Oto-Rhino-Laryngology and Its Related Specialties* 2004;**66**(1):38-41.

Sankaranarayanan 1998

Sankaranarayanan R, Masuyer E, Swaminathan R, Ferlay J, Whelan S. Head and neck cancer: a global perspective on epidemiology and prognosis. *Anticancer Research* 1998;**18** (6B):4779.

SEER 2014

Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, et al (editors). SEER Cancer Statistics Review, 1975-2012, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2012/, based on November 2014 SEER data submission, posted to the SEER website, April 2015.

Shah 1990

Shah J. Cervical lymph node metastases: diagnostic, therapeutic, and prognostic implications. *Oncology* 1990;**4** (10):61–9.

Snow 1992

Snow GB, Patel P, Leemans CR, Tiwari R. Management of cervical lymph nodes in patients with head and neck cancer. *European Archives of Oto-Rhino-Laryngology* 1992;**249**(4): 187–94.

Trikalinos 2013

Trikalinos TA, Hoaglin DC, Small KM, Schmid CH. Evaluating Practices and Developing Tools for Comparative Effectiveness Reviews of Diagnostic Test Accuracy: Methods for the Joint Meta-Analysis of Multiple Tests. Methods Research Report. AHRQ Publication No. 12(13)-EHC151-EF. January 2013. www.effectivehealthcare.ahrq.gov.

van den Brekel 1991

van den Brekel MWM, Castelijns JA, Croll GA, Stel HV, Valk J, van der Wall I, et al. Magnetic resonance imaging versus palpation of cervical lymph node metastasis. *Archives of Otolaryngology -- Head and Neck Surgery* 1991;**117**(6): 666–73.

van den Brekel 1996

van den Brekel MW, van der Waal I, Meijer CJ, Freeman JL, Castelijns JA, Snow GB. The incidence of micrometastases in neck dissection specimens obtained from elective neck dissections. *Laryngoscope* 1996;**106**(6):987-91.

Whiting 2011

Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36.

* *Indicates the major publication for the study*

APPENDICES

Appendix 1. MEDLINE search strategy

1. "Head and Neck Neoplasms"/
2. mouth neoplasms/ or gingival neoplasms/ or lip neoplasms/ or palatal neoplasms/ or tongue neoplasms/ or otorhinolaryngologic neoplasms/ or laryngeal neoplasms/ or pharyngeal neoplasms/ or hypopharyngeal neoplasms/ or exp oropharyngeal neoplasms/
3. exp larynx/ or pharynx/ or exp hypopharynx/ or exp oropharynx/
4. exp Mouth/
5. exp tongue/
6. exp Palate/
7. 3 or 4 or 5 or 6
8. exp Neoplasms/
9. 7 and 8
10. ((mouth or gingival or lip* or palat* or tongue or Laryn* or pharyn* or hypopharyn* or oropharyn* or tonsil* or otorhinolaryngologic or oral) adj6 (cancer* or carcinoma* or neoplas* or tumor* or tumour* or malignan* or SCC)).ab,ti.
11. (head adj3 neck adj6 (cancer* or carcinoma* or neoplas* or tumor* or tumour* or malignan* or SCC)).ab,ti.
12. (H&N adj6 (cancer* or carcinoma* or neoplas* or tumor* or tumour* or malignan* or SCC)).ab,ti.
13. (HNSCC or SCCHN or OP-SCC or OPSCC or "HN SCC" or OP-SCC or "OP SCC" or SCC-HN or "SCC HN")).ab,ti.
14. 1 or 2 or 9 or 10 or 11 or 12 or 13
15. exp Lymphatic Metastasis/
16. exp lymph nodes/
17. exp Neoplasm Metastasis/
18. exp Neoplasm Staging/
19. 17 or 18
20. 16 and 19
21. ((Lymph or neck) adj3 (node* or nodal) adj6 (metastasis or staging or status or Metastases or metastatic or positive or negative or "cN+" or "cN0" or "N0" or "N+")).ab,ti.
22. ((lymphatic or occult) adj6 (metastasis or staging or status or Metastases or metastatic or "cN+" or "cN0" or "N0" or "N+")).ab,ti.
23. MNN.ab,ti.
24. 15 or 20 or 21 or 22 or 23
25. 14 and 24
26. exp Magnetic Resonance Imaging/
27. exp Tomography, X-Ray Computed/
28. exp Positron-Emission Tomography/
29. ((mr or nmr) adj3 (imag* or tomograph*)).ab,ti.
30. (comput* adj6 tomograph*).ab,ti.
31. (magnetic adj3 resonance adj3 imag*).ab,ti.
32. (chemical adj3 shift adj3 imag*).ab,ti.
33. (proton adj3 spin adj3 tomograph*).ab,ti.
34. (Magneti#ation adj3 Transfer adj3 Contrast adj3 Imag*).ab,ti.
35. (electron adj3 beam adj3 tomography).ab,ti.
36. (Zeugmatography or Tomodensitometry).ab.
37. (CAT adj3 scan*).ab,ti.
38. (Positron adj3 Emission adj3 Tomograph*).ab,ti.
39. exp Fluorodeoxyglucose F18/du [Diagnostic Use]
40. (Fluorodeoxyglucose or "Fluorine 18" or (fluoro adj3 deoxy*)).ab,ti.
41. (mri or mris or fmr or CT or CTs or mfct or mdct or msct or PET or FDG or 18F or F18 or 18FDG).ab,ti.
42. 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41
43. 25 and 42
44. exp *"Head and Neck Neoplasms"/di, ra [Diagnosis, Radiography]
45. exp Lymphatic Metastasis/di, ra [Diagnosis, Radiography]

46. exp Lymph Nodes/di, ra [Diagnosis, Radiography]
47. 43 or 46
48. 44 and 47
49. 43 or 48

Appendix 2. Data extraction sheet

Domain	Items
Study characteristics	Author, publication year and journal
	Language
	Study design
	Start and finish dates of study
	Study location (country)
	Setting (community, university, tertiary)
	Number of participating centres
Study population	Population source (clinic, surgical records)
	Age distribution (mean/median, standard deviation, range)
	Sex distribution
	Sublocation distribution (glottic, supraglottic, subglottic, hypopharyngeal)
	TNM-stage distribution
	Surgery type (selective or radical lymph node dissection)
	Previous treatment (surgery, radiotherapy, chemoradiotherapy)
Index test	Index test used (MRI, CT and/or PET-CT)
	Number of radiologists involved
	Experience level of radiologists
	Diagnostic criteria used for establishing lymph node metastasis
	Blinding of radiologists for results of histopathological analysis

	Slice thickness (mm)	
	Interslice gap (mm)	
	Scanning time (seconds)	
	Intravenous contrast use (yes or no)	
	MRI sequence used (T1- and T2-weighted, TSE, STIR)	
	Strength of Tesla magnet used (1.5 or 3.0)	
	Patient preparation before PET-CT (hours of fasting, use of beta-blockers and use of benzodiazepines)	
	Interval between administration of 18F-FDG and start of acquisition (minutes)	
	Use of time-of-flight scanner (yes or no)	
	Image reconstruction algorithm for PET-CT	
	Reconstruction parameters for PET-CT (method, matrix and post-reconstruction resolution)	
	Method of generating volume of interest on PET	
	Method of normalising PET image data into standardised uptake values	
Reference test	Number of pathologists involved	
	Experience level of pathologists	
	Blinding of pathologists for results of MRI, CT and/or PET-CT	
	Slice thickness (μm)	
	Use of immunohistochemistry (yes or no)	
	Use of step-serial sectioning (yes or no)	
	Total number of harvested lymph nodes	
Flow and timing	Time in between index test(s) and reference test (weeks)	
Study results	True positives	
	False positives	
	True negatives	

	False negatives	
Correspondence with authors		

Abbreviations: MRI = magnetic resonance imaging; CT = computed tomography; PET = positron emission tomography.

Appendix 3. Methodological quality assessment sheet

Domain	Item	Assessment
Patient selection	Signalling question 1: was a consecutive or random sample of patients enrolled?	Yes if authors explicitly mention that patients were consecutively enrolled or a random sample of patients was enrolled. No if patients were not consecutively enrolled or patients were not enrolled randomly. Unclear if inclusion and exclusion were not mentioned and it was not explicitly mentioned whether patients were consecutively enrolled
	Signalling question 2: was a case-control design avoided?	Yes if the study design was not case-control. No if the study design was case-control. Unclear if no information was provided in the article regarding study design and authors did not supply additional information
	Signalling question 3: did the study avoid inappropriate exclusions?	Yes if inappropriate exclusions were avoided. No if inappropriate exclusions were not avoided. Unclear if inclusion and exclusion were not mentioned in the article and authors did not supply additional information
	RoB: could the selection of patients have introduced bias?	Low if at least 2 questions are answered 'yes', high if 2 or more questions are answered 'no', and the remaining combinations of answers leads to the judgement 'unclear'
	Applicability: is there concern that the included patients do not match the review question?	Yes if > 10% of patients were diagnosed with nasopharyngeal, sinonasal, thyroid, salivary gland, skin and/or unknown primary cancer or patients had undergone neck dissection previously. No if only patients with oral, oropharyngeal, hypopharyngeal or laryngeal squamous cell carcinoma were included or no more than 10% of patients were diagnosed with cancer in other subsites, and patients were not previously treated with neck dissection. Unclear if the subtype of head and neck cancer or previous history of patients was not mentioned in the article and authors did not supply additional information

Index test	Signalling question 1: were the index test results interpreted without knowledge of the results of the reference standard?	Yes if MRI, CT and PET-CT were assessed before performing surgery or if radiologists were blinded to the outcomes of surgery and histopathology and/or clinical follow-up. No if MRI, CT and/or PET-CT were assessed after performing surgery and/or clinical follow-up and radiologists were not blinded to the outcome of surgery and histopathology and/or clinical follow-up. Unclear if both flow and timing and blinding of radiologists were not mentioned in the article and authors did not supply additional information
	Signalling question 2: if a threshold was used, was it pre-specified?	Yes if lymph node metastases were established or excluded based on predefined diagnostic criteria. No if diagnostic criteria were not predefined or no standardised diagnostic criteria were used for establishing presence or absence of lymph node metastases. Unclear if (the use of) diagnostic criteria were not mentioned in the article and authors did not supply additional information
	RoB: could the conduct or interpretation of the index test have introduced bias?	Low if both questions are answered 'yes', high if either question is answered 'no', and the remaining combinations of answers leads to the judgement 'unclear'
	Applicability: is there concern that the index test, its conduct, or interpretation differ from the review question?	Yes if MRI, CT and/or PET-CT did not fulfil the criteria listed in section Index tests. No if MRI, CT and PET-CT fulfilled all of the criteria listed in section Index tests. Unclear if no information was provided in the article regarding any of the criteria listed in the section Index tests and authors did not supply additional information.
Reference test	Signalling question 1: is the reference standard likely to correctly classify the target condition?	Yes if histopathology of surgical specimens and/or clinical follow-up were used as the reference test. No if another reference test was used. Unclear if type of reference test was not mentioned in the article and authors did not supply additional information
	Signalling question 2: were the reference standard results interpreted without knowledge of the results of the index test?	Yes if histopathologic specimen was assessed by pathologists that were blinded to the results of MRI, CT and PET-CT. No if histopathologic specimen was assessed by pathologists that were not blinded to the results of MRI, CT and/or PET-CT. Unclear if no information was provided in the article regarding blinding of the pathologists to one or both of the index test(s) and authors did not supply additional information
	RoB: could the reference standard, its conduct, or its interpretation have introduced bias?	Low if both questions are answered 'yes', high if either question is answered 'no', and the remaining combinations of answers leads to the judgement 'unclear'

(Continued)

	Applicability: is there concern that the target condition as defined by the reference standard does not match the review question?	Yes if the outcome is any other than lymph node metastases as established by histopathology and/or clinical follow-up. No if the outcome is lymph node metastases as established by histopathology and/or clinical follow-up. Unclear if the outcome measure or the reference standard was not mentioned in the article and authors did not supply additional information
Flow and timing	Signalling question 1: was there an appropriate interval between the index test(s) and reference standard?	Yes if the interval between MRI/CT/PET-CT and histopathology was shorter than 6 weeks or the interval between the index test(s) and clinical follow-up was at least 1 year. No if the interval between MRI/CT/PET-CT and histopathology was longer than 6 weeks or the interval between the index test(s) and clinical follow-up was shorter than 1 year. Unclear if the interval between MRI/CT/PET-CT and histopathology was not explicitly mentioned in the article and authors did not supply additional information
	Signalling question 2: did all patients receive a reference standard?	Yes if all patients received a reference test, regardless of what type of reference test. No if any of the included patients did not receive a reference test. Unclear if it was not explicitly mentioned whether all patients received a reference test
	Signalling question 3: did patients receive the same reference standard?	Yes if histopathologic examination and/or clinical follow-up was performed in all patients in the same standardised way: clearly defined slice thickness and staining techniques and predefined (imaging) protocol for clinical follow-up. No if histopathologic examination and/or clinical follow-up was not performed in all patients or the slice thickness, staining techniques and/or diagnostic follow-up were not the same for all patients. Unclear if it was not explicitly mentioned whether all patients received histopathologic examination, or slice thickness, staining techniques and/or protocol for follow-up were not mentioned
	Signalling question 4: were all patients included in the analysis?	Yes if at least 90% of the patients that were initially recruited were included in the analyses. No if more than 10% of the initially recruited patients were not included in the analyses. Unclear if it was not mentioned how many patients were initially included and how many were eventually included for analyses, or how many patients were excluded from the final analyses
	RoB: could the patient flow have introduced bias?	Low if at least 3 questions were answered 'yes', high if 2 or more questions were answered 'no', and the remaining combinations of answers leads to the judgement "un-

(Continued)

		clear"

Abbreviations: CT = computed tomography; MRI = magnetic resonance imaging; PET = positron emission tomography; RoB: risk of bias

CONTRIBUTIONS OF AUTHORS

IW: drafting protocol, developing search strategy, selecting studies to include, extracting data from studies, performing quality assessment, carrying out analyses, interpreting analyses, drafting final review, updating review.

IS: drafting protocol, selecting studies to include, extracting data from studies, performing quality assessment, carrying out analyses, interpreting analyses, drafting final review.

LH: drafting protocol, developing search strategy, selecting studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), carrying out analyses, interpreting analyses, drafting final review.

JBR: drafting protocol, selecting studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), carrying out analyses, interpreting analyses, drafting final review.

FAP: drafting protocol, selecting studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), interpreting analyses, drafting final review.

OSH: drafting protocol, selecting studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), interpreting analyses, drafting final review.

RdB: drafting protocol, selecting studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), interpreting analyses, drafting final review.

DECLARATIONS OF INTEREST

IW: nothing to declare.

IS: nothing to declare.

LH: nothing to declare.

JBR: nothing to declare.

FAP: nothing to declare.

OSH: nothing to declare.

RdB: nothing to declare.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- National Institute for Health Research, UK.
Infrastructure funding for Cochrane ENT