Cochrane Library

Cochrane Database of Systematic Reviews

MRI versus CT for detecting cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma (Protocol)

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

Wegner I, Hooft L, Reitsma JB, Pameijer FA, de Bree R, Stegeman I.
MRI versus CT for detecting cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma.
Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD012115.
DOI: 10.1002/14651858.CD012115.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Protocol]

MRI versus CT for detecting cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma

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

Editorial group: Cochrane ENT Group.
Publication status and date: New, published in Issue 4, 2016.

Citation: Wegner I, Hooft L, Reitsma JB, Pameijer FA, de Bree R, Stegeman I. MRI versus CT for detecting cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No.: CD012115. DOI: 10.1002/14651858.CD012115.

ABSTRACT

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

To determine and compare the diagnostic accuracy of preoperative conventional MRI and conventional CT for detecting cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma, who have not previously undergone surgery, using histopathology as the reference standard.

Diagnostic accuracy may vary depending on the methodological quality of the study, study population and technical characteristics of the index test and reference test. Therefore, we will assess whether heterogeneity in sensitivity and/or specificity can be related to:

- methodological quality rating for the four QUADAS-2 domains (low risk of bias versus high risk of bias versus unclear risk of bias);
- previous radiotherapeutic treatment (yes versus no);
- sublocation distribution (glottic versus supraglottic versus subglottic versus hypopharyngeal);
- type of cartilage assessed (thyroid versus cricoid versus arytenoid);
- T-classification distribution;
- N-classification distribution;
- magnetic field strength used (1.5 Tesla versus 3.0 Tesla);
- slice thickness imaging;
- slice thickness pathology;
- diagnostic criteria used/differences in (qualitative) thresholds.

BACKGROUND

Target condition being diagnosed

The target condition being diagnosed in this review is cartilage invasion in laryngeal and hypopharyngeal squamous cell carcinoma. This protocol applies only to squamous cell carcinoma. Whenever we speak of laryngeal or hypopharyngeal cancer, we mean squamous cell carcinoma. The hypopharynx, oropharynx and nasopharynx make up the pharynx. The hypopharynx is the most caudal part of the pharynx. The larynx is subdivided into three anatomical regions for tumour staging purposes: the glottis, supraglottis and subglottis. Approximately two-thirds of laryngeal cancers originate from the glottis and one-third from the supraglottis. Subglottic laryngeal cancers are rare. Worldwide, over 150,000 new cases of laryngeal cancer are diagnosed yearly with an age-standardised rate of 2.1 per 100,000 (Cancer Research UK 2012; GLOBOCAN 2012). Reported five-year survival rates following treatment range between 61% and 72% in the United States, the United Kingdom and the Netherlands (American Cancer Society 2015; Cancer Research UK 2012; NWHHT 2010; SEER 2014). Survival is related to the stage of disease at diagnosis. Five-year survival rates are highest for patients presenting at stage I (94% to 100%) and lowest for those diagnosed with stage IV disease (40% to 65%) (Cancer Research UK 2012; NWHHT 2010).

The presence of cartilage invasion is an important prognostic factor in laryngeal and hypopharyngeal squamous cell carcinoma and one of the main determinants in the staging of laryngeal cancer according to the TNM classification (Table 1; Table 2; Table 3; Table 4; UICC 2009). Accurate assessment of cartilage invasion is essential not only in determining prognosis, but also in choosing a treatment strategy. Tumour invasion through the inner cortex of the thyroid cartilage equals a T3 tumour and tumour invasion through the outer cortex of the thyroid cartilage equals a T4 tumour. According to the Dutch national guideline on laryngeal cancer, small T3 tumours can be treated using either (accelerated) radiotherapy or larynx-preserving surgery in selected cases (NWHHT 2010). Large T3 and T4 tumours should be treated with chemoradiotherapy or total laryngectomy (NCCN 2014; NWHHT 2010). The American Society of Clinical Oncology recommends larynx preservation in T3 and T4 disease with cartilage invasion limited to the cortex (ASCO 2006). Chemoradiotherapy is the most widely applied approach. Total laryngectomy is reserved for those patients with tumour penetration through the cartilage into the surrounding soft tissues (ASCO 2006).

Index test(s)

Both magnetic resonance imaging (MRI) and computed tomography (CT) can be used to assess cartilage invasion preoperatively. CT and MRI each have their unique strengths and weaknesses. MRI provides good contrast between different soft tissues, which makes it superior in distinguishing between soft tissue and tumour. However, cartilage invasion is easily overestimated, since it is hard to distinguish from peritumoral inflammation (Maroldi 2014). MRI does not require the use of iodinated contrast and does not involve ionising radiation. CT evaluation requires less time, thereby limiting artefacts induced by movements as a result of breathing, swallowing and coughing (Castelijns 1988). CT is superior to MRI in evaluating cortical bone involvement, is less expensive than MRI and, unlike MRI, can be used in patients with ferromagnetic objects in the body.

Different criteria can be used to diagnose cartilage invasion on CT, including erosion, lysis, (asymmetric) sclerosis and extra-laryngeal spread (Appendix 1; Becker 1997). On MRI, signal intensities are used to distinguish tumour invasion from peritumoral inflammation and healthy tissue. The signal intensity of cartilage is compared to the tumour mass outside the cartilage. If the cartilage has a similar intensity to the tumour mass or has a similar enhancement following gadolinium contrast injection as the tumour mass, the cartilage is considered to be invaded by tumour (Becker 2008). If the cartilage has a higher signal intensity than the tumour mass or enhancement following gadolinium contrast is greater, peritumoral inflammation without tumour invasion is diagnosed (Becker 2008). Both CT and MRI criteria are based on a qualitative assessment. As a result, the expertise of the radiologist or clinician affects the diagnostic value of these imaging modalities.

Clinical pathway

Laryngeal and hypopharyngeal cancer are diagnosed through clinical history and physical examination using a rigid or flexible fibre-optic endoscope. If a tumour is identified, both imaging and direct laryngoscopy will be used to assess the tumour before treating it. The tumour can be primary, residual or recurrent. Patients with a residual or recurrent tumour may have undergone previous surgical treatment or chemoradiotherapy. Imaging is vitally important in staging and treatment selection. It provides assessment of both the tumour and regional lymph nodes, and occasionally unsuspected second primary tumours are detected. Both the Dutch and American guidelines recommend using one of the two (NCCN 2014; NWHHT 2010). However, given their unique strengths and weaknesses, MRI and CT can complement each other and

may be combined in the diagnostic process (Shah 2012). Direct laryngoscopy with the use of rigid telescopes and/or an operating microscope is performed following radiographic evaluation. Direct laryngoscopy ensures accurate evaluation of the primary tumour and its extension into adjacent structures. Histological evaluation is required to determine the precise nature of the lesion. Biopsy is usually performed during direct laryngoscopy. Imaging, direct laryngoscopy and histological evaluation are performed in all patients with laryngeal and hypopharyngeal cancer, regardless of TNM stage.

Rationale

The Dutch national guideline on laryngeal cancer states that CT is characterised by low sensitivity and high specificity for detecting cartilage invasion, whereas MRI is characterised by high sensitivity and low specificity (NWHHT 2010). The guideline does not advise the use of either MRI or CT, but rather leaves the choice between the two to be based on the clinician's preference and experience. These statements are based on outdated literature and this leaves room for an extensive review process on this topic (Becker 1995; Becker 1997). It is important to accurately diagnose cartilage invasion. Overestimation of cartilage invasion can result in over-staging and consequent over-treatment. This means that patients may be subjected to total laryngectomy when they could have been offered organ preservation if they were staged more accurately (Li 2011). On the other hand, underestimation may result in under-treatment, jeopardising the chance of cure.

OBJECTIVES

To determine and compare the diagnostic accuracy of preoperative conventional MRI and conventional CT for detecting cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma, who have not previously undergone surgery, using histopathology as the reference standard.

Secondary objectives

Diagnostic accuracy may vary depending on the methodological quality of the study, study population and technical characteristics of the index test and reference test. Therefore, we will assess whether heterogeneity in sensitivity and/or specificity can be related to:

- methodological quality rating for the four QUADAS-2 domains (low risk of bias versus high risk of bias versus unclear risk of bias);
- previous radiotherapeutic treatment (yes versus no);
- sublocation distribution (glottic versus supraglottic versus subglottic versus hypopharyngeal);
- type of cartilage assessed (thyroid versus cricoid versus arytenoid);
- T-classification distribution;
- N-classification distribution;
- magnetic field strength used (1.5 Tesla versus 3.0 Tesla);
- slice thickness imaging;
- slice thickness pathology;
- diagnostic criteria used/differences in (qualitative) thresholds.

METHODS

Criteria for considering studies for this review

Types of studies

We will include randomised controlled studies and cross-sectional diagnostic studies. These may be either prospective or retrospective. Patients should be included consecutively. We will exclude case-control studies. We will include studies with the following comparisons:

- studies directly comparing both CT and MRI in the same patient population;
- studies that randomise patients to one of the two index tests;
- studies evaluating one of the two index tests;
- studies in which both index tests are evaluated, but not in the same patient population.

We realise that indirect comparisons (bullet 3 and 4) can be prone to selection bias. However, a first exploration of the literature has led us to believe there are insufficient studies to limit inclusion to studies with direct comparisons (bullet 1 and 2).
Furthermore, studies must report sufficient data in order for us 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 apply any restrictions based on a minimal quality standard, minimal sample size or number of patients with the target condition (i.e. cartilage invasion).

Participants

We will include studies that include patients with histopathologically proven laryngeal or hypopharyngeal squamous cell carcinoma, or both. We will only include studies that include patients who have not previously undergone surgery, to ensure that the reference standard (definitive histopathology) can be obtained. We will include studies including patients who have previously undergone radiotherapy. Study inclusion will not be limited by age, gender, T-stage or setting.

Index tests

The index tests under review are conventional CT and conventional MRI. Conventional CT is characterised by an X-ray tube and detector that are physically rotated around a stationary object. Conventional MRI is characterised by an oscillating magnetic field that is temporarily applied to the patient and a receiving coil. Non-conventional imaging modalities include single-photon emission CT, diffusion-weighted MRI or nuclear imaging. These non-conventional imaging modalities are not eligible for inclusion. We chose conventional CT and conventional MRI because they are currently used to detect cartilage invasion/assess tumour stage. Both are widely available in radiology departments worldwide. Different criteria can be used to diagnose cartilage invasion, including erosion, lysis, (asymmetric) sclerosis and extra-laryngeal spread (Appendix 1; Becker 1997). We will not limit study inclusion based on the diagnostic criteria that were used.

We will limit study inclusion based on the quality of the CT and/or MRI. 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-weighted and T2-weighted, turbo spin echo (TSE) or short tau inversion recovery (STIR) sequences should have been used. Post-gadolinium contrast series and either a 1.5 or 3.0 Tesla magnet should have been used.

Target conditions

Cartilage invasion in patients with histopathologically proven laryngeal and hypopharyngeal squamous cell carcinoma.

Reference standards

We will only include studies that used definitive histopathology on tissue that has been removed during surgery as the reference standard. Since surgery is invasive, it is not usually performed in patients with negative index test results, which can lead to partial verification bias. Long-term clinical follow-up of at least one year is not an alternative to our primary reference standard in patients that will not undergo laryngectomy/tumour resection. All patients with laryngeal or hypopharyngeal cancer will undergo treatment (surgery, radiotherapy or chemoradiotherapy) regardless of whether they have cartilage invasion or not. A lack of cartilage invasion at one-year follow-up does not necessarily mean the patient had no cartilage invasion at the time of imaging. The absence of cartilage invasion may have been due to the patient having undergone treatment (surgery, radiotherapy or chemoradiotherapy). Furthermore, it is not possible to accurately assess cartilage invasion post-(chemo)radiotherapy using imaging without pathologic confirmation. Studies therefore will most likely only include patients with an indication for surgery. As it seems likely that in these study populations the prevalence of cartilage invasion is higher than it is in the general population of patients with laryngeal or hypopharyngeal cancer, this could lead to overestimation of the positive predictive value and underestimation of the negative predictive value.

We will restrict the studies to those that used the index and reference tests within six weeks of each other to limit bias due to changes in tumour status/cartilage invasion over time.

Search methods for identification of studies

The Cochrane ENT Trial Search Co-ordinator will conduct systematic searches for diagnostic test accuracy studies. There will be no language, publication year or publication status restrictions. We may contact original authors for clarification and further data if trial reports are unclear and 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);
- 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 2).

Searching other resources

We will scan the reference lists of identified publications for additional trials and contact trial authors if necessary. In addition, the Trials Search Co-ordinator 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 trials. 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 discrepancies (LH, JBR, FAP 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 regarding radiology and RdB regarding clinical questions. We will review the titles and abstracts and discard those articles that are obviously ineligible (see Table 5 for a full list of the inclusion and exclusion criteria). 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 conventional CT or MRI, but other imaging modalities such as diffusion-weighted MRI or nuclear imaging. We will exclude studies reviewing the use of CT and/or MRI following surgical treatment.

We will proceed by retrieving and reviewing the full-text copies. We will only include those studies evaluating the diagnostic accuracy of conventional CT and/or conventional MRI in comparison with definitive histology in patients with laryngeal and/or hypopharyngeal squamous cell carcinoma for detecting cartilage invasion. We will exclude studies evaluating the diagnostic accuracy of these imaging modalities for establishing lymph node metastases or other tumour characteristics. We will only include studies if they meet the CT and/or MRI criteria listed under Index tests. We will not include studies using reference standards other than histopathology. 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 populations, 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 shown in Appendix 3. Discrepancies will be resolved through a consensus discussion between the two review authors or a third review author will resolve discrepancies (LH, JBR, FAP 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 or RdB). We will use the QUADAS-2 tool to assess methodological quality (Whiting 2011). A full list of the criteria and their operational definitions is shown in Appendix 4. 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 with the item rated '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 and/or 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 the forest plots by imaging techniques (MRI and 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 operator characteristic (ROC) space, using different symbols for MRI and CT estimates. In case of studies providing direct evidence we will link MRI and 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 and 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 carry out the comparison of accuracy between CT and MRI by adding a covariate for the type of index test to the model to investigate whether sensitivity, specificity or both are different between the two 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).

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 (Macaskill 2010; Reitsma 2005).

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 factors mentioned under Secondary objectives. 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 (DTA) reviews and the determinants of publication bias are unknown. Therefore we elect not to investigate reporting bias.

ACKNOWLEDGEMENTS

This project was supported by the National Institute for Health Research, via Cochrane Infrastructure, Cochrane Programme Grant or Cochrane Incentive funding to Cochrane ENT. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS or the Department of Health.

REFERENCES

Additional references

American Cancer Society 2015
American Cancer Society. *Cancer Facts & Figures 2015*. Atlanta: American Cancer Society, 2015.

ASCO 2006
American Society of Clinical Oncology, Pfister DG, Laurie SA, Weinstein GS, Mendenhall WM, Adelstein DJ, Ang KK, et al. American Society of Clinical Oncology clinical practice guideline for the use of larynx-preservation strategies in the treatment of laryngeal cancer. *Journal of Clinical Oncology* 2006;**24**(22):3693–704.

Becker 1995
Becker M, Zbären P, Laeng H, Stoupis C, Porcellini B, Vock P. Neoplastic invasion of the laryngeal cartilage: comparison of MR imaging and CT with histopathologic correlation. *Radiology* 1995;**194**(3):661–9.

Becker 1997
Becker M, Zbären P, Delavelle J, Kurt AM, Egger C, Rüfenacht DA, et al. Neoplastic invasion of the laryngeal cartilage: reassessment of criteria for diagnosis at CT. *Radiology* 1997;**203**(2):521–32.

Becker 2008
Becker M, Zbären P, Casselman JW, Kohler R, Dulguerov P, Becker CD. Neoplastic invasion of laryngeal cartilage: reassessment of criteria for diagnosis at MR imaging. *Radiology* 2008;**249**:551–9.

Cancer Research UK 2012
Cancer Research UK. Laryngeal cancer statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/laryngeal-cancer (accessed June 2015).

Castelijns 1988
Castelijns JA, Gerritsen GJ, Kaiser MC, Valk J, van Zanten TE, Golding RG, et al. Invasion of laryngeal cartilage by cancer: comparison of CT and MR imaging. *Radiology* 1988;**167**(1):199–206.

GLOBOCAN 2012
Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr (accessed June 2015).

Li 2011
Li B, Bobinski M, Gandour-Edwards R, Farwell DG, Chen AM. Overstaging of cartilage invasion by multidetector CT scan for laryngeal cancer and its potential effect on the use of organ preservation with chemoradiation. *British Journal of Radiology* 2011;**84**(997):64–9.

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.

Maroldi 2014
Maroldi R, Ravanelli M, Farina D. Magnetic resonance for laryngeal cancer. *Current Opinion in Otolaryngology & Head and Neck Surgery* 2014;**22**(2):131–9.

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.

NWHHT 2010
Nederlandse Werkgroep Hoofd-Hals Tumoren NWHHT. National guideline laryngeal carcinoma, version 3 [Landelijke richtlijn larynxcarcinoom, versie 3]. Available at: http://www.oncoline.nl/larynxcarcinoom (accessed June 2015).

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.

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 2012
Shah JP, Patel SG, Singh B. Larynx and trachea. *Jatin Shah's Head and Neck Surgery and Oncology*. 4th Edition. Mosby, 2012.

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).

UICC 2009
UICC. In: Sobin LH, Gospodarowicz MK, Wittekind C editor(s). *TNM Classification of Malignant Tumours*. 7th Edition. Wiley-Blackwell, 2009.

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

* *Indicates the major publication for the study*

ADDITIONAL TABLES

Table 1. Tumour staging for glottic laryngeal cancer according to TNM classification

Tumour stage		Characteristics
T1	a	Tumour limited to one vocal cord with normal mobility; may involve anterior commissure
	b	Tumour involves both vocal cords with normal mobility; may involve anterior commissure
T2		Tumour extends to supraglottis and/or subglottis and/or impaired vocal cord mobility
T3		Tumour limited to the larynx with vocal cord fixation and/or invasion of paraglottic space and/or **inner** cortex of the thyroid cartilage
T4	a	Moderately advanced local disease: tumour invades through **outer** cortex of the thyroid cartilage and/or invades tissue beyond the larynx
	b	Very advanced local disease: tumour invades prevertebral space, encases carotid artery or invades mediastinal structures

Table 2. Tumour staging for supraglottic laryngeal cancer according to TNM classification

Tumour stage		Characteristics
T1		Tumour limited to one subsite of supraglottis with normal vocal cord mobility
T2		Tumour invades mucosa of more than one adjacent subsite of supraglottis or glottis or region outside the supraglottis without fixation of the larynx
T3		Tumour limited to larynx with vocal cord fixation and/or invades any of the following: post-cricoid area, pre-epiglottic space, paraglottic space and/or inner cortex of thyroid cartilage
T4	a	Moderately advanced local disease: tumour invades through the thyroid cartilage and/or invades tissues beyond the larynx
	b	Very advanced local disease: tumour invades prevertebral space, encases carotid artery or invades mediastinal structures

Table 3. Tumour staging for subglottic laryngeal cancer according to TNM classification

Tumour stage		Characteristics
T1		Tumour limited to subglottis
T2		Tumour extends to vocal cord(s) with normal or impaired mobility
T3		Tumour limited to larynx with vocal cord fixation
T4	a	Moderately advanced local disease: tumour invades cricoid or thyroid cartilage and/or invades tissues beyond the larynx
	b	Very advanced local disease: tumour invades prevertebral space, encases carotid artery or invades mediastinal structures

Table 4. Tumour staging for hypopharyngeal cancer according to TNM classification

Tumour stage		Characteristics
T1		Tumour limited to one subsite of hypopharynx and/or 2 cm or less in greatest dimension
T2		Tumour invades more than one subsite of hypopharynx or an adjacent site, or measures more than 2 cm but not more than 4 cm in greatest dimension without fixation of hemilarynx
T3		Tumour more than 4 cm in greatest dimension or with fixation of hemilarynx or extension to oesophagus
T4	a	Moderately advanced local disease: tumour invades thyroid/cricoid cartilage, thyroid bone, thyroid gland or central compartment soft tissue*

Table 4. Tumour staging for hypopharyngeal cancer according to TNM classification (*Continued*)

	b	Very advanced local disease: tumour invades prevertebral fascia, encases carotid artery or involves mediastinal structures

*Central compartment soft tissue includes prelaryngeal strap muscles and subcutaneous fat.

Table 5. Inclusion and exclusion criteria

Inclusion criteria	Prospective and retrospective randomised controlled studies and cross-sectional diagnostic studies
	Studies with direct comparisons and studies with indirect comparisons of CT and/or MRI
	Patients with histopathologically proven laryngeal and/or hypopharyngeal squamous cell carcinoma
	The use of conventional CT and/or MRI as the index test
	The use of definitive histopathology removed during surgery as the reference standard
	The evaluation of cartilage invasion
Exclusion criteria	Patients who have previously undergone surgical treatment of the larynx and/or hypopharynx
	CT slice thickness of more than 3 mm and/or a MRI slice thickness of more than and/or an interslice gap of more than 0.4 mm
	Lack of intravenous contrast when using CT or lack of post-gadolinium contrast series when using MRI
	Sequences other than T1-weighted and T2-weighted, turbo spin echo (TSE) or short tau inversion recovery (STIR)
	Tesla magnet other than 1.5 or 3.0
	The use of long-term clinical follow-up as a reference standard
	More than 6 weeks between the index test and the reference test
	The evaluation of lymph node metastases or tumour characteristics other than cartilage invasion
	The number of true positives, false negatives, false positives and true negatives cannot be extracted
	Case-control studies
	Systematic reviews
	Opinion papers
	Editorials

Table 5. Inclusion and exclusion criteria (*Continued*)

	Conference abstracts
	Poster presentations

CT: computed tomography
MRI: magnetic resonance imaging

APPENDICES

Appendix 1. Glossary

Term	Explanation
Erosion	A localised form of lysis limited to a sclerotic cortex
Lysis	A punched-out lesion or focal defect within sclerotic marrow
Sclerosis	An obvious thickening of the ossified inner or outer cortex or increased ossification of the medullary cavity

Appendix 2. MEDLINE search strategy

1. exp Hypopharyngeal Neoplasms/
2. exp Laryngeal Neoplasms/
3. exp Hypopharynx/
4. exp Larynx/
5. (hypopharyn* or laryngopharyn* or larynx* or laryngeal or pharyngolaryn* or laryngopharyn*).ab,ti.
6. (glott* or supraglott* or subglott*).ab,ti.
7. ((head adj3 neck) or (neck adj3 node*)).ab,ti.
8. 3 or 4 or 5 or 6 or 7
9. exp Neoplasms/
10. (carcinom* or cancer* or neoplas* or tumor* or tumour* or malignan* or SCC).ab,ti.
11. 9 or 10
12. 8 and 11
13. "Head and Neck Neoplasms"/
14. (HPC or LPC or HNSCC or LPSCC or HPSCC).ab,ti.
15. 1 or 2 or 12 or 13 or 14
16. exp Laryngeal Cartilages/
17. exp Neoplasm Staging/
18. Neoplasm Invasiveness/

19. ((Cartilage* or arytenoid* or cricoid or thyroid or epiglott* or neoplas* or tumour* or tumor* or bone or extralaryn* or locoregional) and (invasion or invad* or invasiveness or infiltrate* or destruct* or spread)).ab,ti.
20. (erosion or lysis or sclerotic or sclerosis or (punch* adj3 out adj3 lesion*) or (focal adj3 defect) or ossification or ossified).ab,ti.
21. ((over or neoplasm* or neck or tumour* or tumor* or primary) adj3 (stage or staging)).ab,ti.
22. overstaging.ab,ti.
23. 16 or 17 or 18 or 19 or 20 or 21 or 22
24. 15 and 23
25. exp Magnetic Resonance Imaging/
26. ((mr or nmr) adj3 (imag* or tomograph*)).ab,ti.
27. (mri or mris or fmri).ab,ti.
28. (magnetic adj3 resonance adj3 imag*).ab,ti.
29. (chemical adj3 shift adj3 imag*).ab,ti.
30. (proton adj3 spin adj3 tomograph*).ab,ti.
31. (Magneti#ation adj3 Transfer adj3 Contrast adj3 Imag*).ab,ti.
32. Zeugmatography.ab,ti.
33. 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
34. exp Tomography, X-Ray Computed/
35. (CT or CTs).ab,ti.
36. (comput* adj6 tomograph*).ab,ti.
37. (electron adj3 beam adj3 tomography).ab,ti.
38. Tomodensitometry.ab,ti.
39. (CAT adj3 scan*).ab,ti.
40. (mfct or mdct or msct).ab,ti.
41. 34 or 35 or 36 or 37 or 38 or 39
42. 33 and 41
43. 24 and 42

Appendix 3. 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)

(*Continued*)

	Sex distribution
	Sublocation distribution (glottic, supraglottic, subglottic, hypopharyngeal)
	TNM stage distribution
	Surgery type (partial or total laryngectomy)
	Previous treatment (surgery, radiotherapy, chemoradiotherapy)
Index test	Index test used (CT, MRI, both)
	Number of radiologists involved
	Experience level of radiologists
	Diagnostic criteria used for establishing cartilage invasion
	Blinding of radiologists for results of histopathological analysis
	Slice thickness (mm)
	Interslice gap (mm)
	Scanning time (s)
	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)
Reference test	Number of pathologists involved
	Experience level of pathologists
	Blinding of pathologists to results of CT and/or MRI
	Slice thickness (μm)
Flow and timing	Time in between index test(s) and reference test (weeks)
Study results	Type of cartilage assessed (thyroid, cricoid, arytenoid)
	True positives
	False positives

(*Continued*)

	True negatives
	False negatives
Correspondence with authors	

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

Appendix 4. Methodological quality assessment sheet

Domain	Item	Assessment
Patient selection	Signalling question 1: was a consecutive or random sample of patients enrolled?	Yes if the 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 the authors did not supply additional information
	Signalling question 3: did the study avoid inappropriate exclusions?	Yes if inappropriate exclusions were avoided. Examples of inappropriate exclusions are excluding patients who have previously undergone radiotherapy or limiting inclusion by age, gender or TNM stage. No if inappropriate exclusions were not avoided. Unclear if inclusion and exclusion were not mentioned in the article and the authors did not supply additional information
	RoB: could the selection of patients have introduced bias?	Low if at least two questions are answered 'yes', high if two 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 patients with cancers other than laryngeal and/or hypopharyngeal cancer were included or patients had undergone previous surgical treatment. No if only patients with laryngeal and/or hypopharyngeal cancer were included that had not undergone previous surgical treatment. Unclear if the subtype of head and neck cancer or previous history of patients was not mentioned in the article and the authors did not supply additional

(*Continued*)

		information
Index test	Signalling question 1: were the index test results interpreted without knowledge of the results of the reference standard?	Yes if CT and MRI were assessed before performing surgery or if radiologists were blinded for the outcomes of surgery and histopathology. No if CT and/or MRI were assessed after performing surgery and radiologists were not blinded for the outcome of surgery and histopathology. Unclear if both flow and timing and blinding of radiologists were not mentioned in the article and the authors did not supply additional information
	Signalling question 2: if a threshold was used, was it prespecified?	Yes if cartilage invasion was 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 cartilage invasion. Unclear if (the use of) diagnostic criteria was not mentioned in the article and the 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 CT and/or MRI did not fulfil the criteria listed in section Index tests. No if CT and MRI 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 the 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 was used as the reference test. No if another reference test was used, e.g. long-term follow-up. Unclear if type of reference test was not mentioned in the article and the 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 for the results of CT and MRI. No if histopathologic specimen was assessed by pathologists that were not blinded for the results of CT and/or MRI. Unclear if no information was provided in the article regarding blinding of the pathologists for one or both of the index tests and the 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 cartilage invasion as established by histopathology. No if the outcome is cartilage invasion as established by histopathology. Unclear if the outcome measure or the reference standard was not mentioned in the article and the 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 CT/MRI and histopathology was shorter than six weeks. No if the interval between CT/MRI and histopathology was longer than six weeks. Unclear if the interval between CT/MRI and histopathology was not explicitly mentioned in the article and the 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 was performed in all patients in the same standardised way: clearly defined slice thickness and staining techniques. No if histopathologic examination was not performed in all patients or the slice thickness and/or staining techniques were not the same for all patients. Unclear if it was not explicitly mentioned whether all patients received histopathologic examination, or slice thickness and/or staining techniques 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 three questions were answered 'yes', high if two or more questions were answered 'no' and the remaining combinations of answers leads to the judgement 'unclear'

Abbreviations: CT = computed tomography; MRI = magnetic resonance imaging; RoB = risk of bias.

WHAT'S NEW

Date	Event	Description
14 March 2016	Amended	Author order changed.

CONTRIBUTIONS OF AUTHORS

Inge Wegner (IW): drafting the protocol, developing the search strategy, selection of studies to include, extracting data from studies, performing quality assessment, carrying out the analyses, interpreting the analyses, drafting the final review, updating the review.

Inge Stegeman (IS): drafting the protocol, selection of studies to include, extracting data from studies, performing quality assessment, carrying out the analyses, interpreting the analyses, drafting the final review.

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

Johannes B Reitsma (JBR): drafting the protocol, selection of studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), carrying out the analyses, interpreting the analyses, drafting the final review.

Frank A Pameijer (FAP): drafting the protocol, selection of studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), interpreting the analyses, drafting the final review.

Remco de Bree (RdB): drafting the protocol, selection of studies to include (arbiter), extracting data from studies (arbiter), performing quality assessment (arbiter), interpreting the analyses, drafting the final review.

DECLARATIONS OF INTEREST

Inge Wegner: nothing to declare.

Inge Stegeman: nothing to declare.

Lotty Hooft: nothing to declare.

Johannes B Reitsma: nothing to declare.

Frank A Pameijer: nothing to declare.

Remco de Bree: 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