Diseases of the Esophagus (2021)**34**,1–12
https://doi.org/10.1093/dote/doab002

DISEASES OF THE ESOPHAGUS

ISDE The International Society for Diseases of the Esophagus

Systematic Review and Meta-analysis

Accuracy of clinical staging for T2N0 oesophageal cancer: systematic review and meta-analysis

Paul Wolfson,[1,2] Kai Man Alexander Ho,[1,2] Paul Bassett,[3] Rehan Haidry,[1,2] Alessandro Olivo,[4] Laurence Lovat,[1,2] Sarmed S. Sami[1,2]

[1]*Division of Surgery and Interventional Science, University College London, London, UK,* [2]*Department of Gastroenterology, University College Hospital NHS Foundation Trust, London, UK,* [3]*Statsconsultancy Ltd., Amersham, UK, and* [4]*Department of Medical Physics and Bioengineering, University College London, London, UK*

SUMMARY. **Oesophageal cancer is the sixth commonest cause of overall cancer mortality. Clinical staging utilizes multiple imaging modalities to guide treatment and prognostication. T2N0 oesophageal cancer is a treatment threshold for neoadjuvant therapy. Data on accuracy of current clinical staging tests for this disease subgroup are conflicting. We performed a meta-analysis of all primary studies comparing clinical staging accuracy using multiple imaging modalities (index test) to histopathological staging following oesophagectomy (reference standard) in T2N0 oesophageal cancer. Patients that underwent neoadjuvant therapy were excluded. Electronic databases (MEDLINE, Embase, Cochrane Library) were searched up to September 2019. The primary outcome was diagnostic accuracy of combined T&N clinical staging. Publication date, first recruitment date, number of centers, sample size and geographical location main histological subtype were evaluated as potential sources of heterogeneity. The search strategy identified 1,199 studies. Twenty studies containing 5,213 patients met the inclusion criteria. Combined T&N staging accuracy was 19% (95% CI, 15–24); T staging accuracy was 29% (95% CI, 24–35); percentage of patients with T downstaging was 41% (95% CI, 33–50); percentage of patients with T upstaging was 28% (95% CI, 24–32) and percentage of patients with N upstaging was 34% (95% CI, 30–39). Significant sources of heterogeneity included the number of centers, sample size and study region. T2N0 oesophageal cancer staging remains inaccurate. A significant proportion of patients were downstaged (could have received endotherapy) or upstaged (should have received neoadjuvant chemotherapy). These findings were largely unchanged over the past two decades highlighting an urgent need for more accurate staging tests for this subgroup of patients.**

KEY WORDS: **diagnosis, oesophageal cancer, staging.**

INTRODUCTION

Oesophageal cancer is the seventh most common cause of cancer worldwide and the sixth most common cause of overall cancer mortality.[1] Five-year survival has only shown modest improvement since the 1970s despite advances in diagnostic and therapeutic options.[2] The causes for this are likely to be multifactorial but include better access to patient care and advancements in medical, surgical and adjuvant therapy.[3]

Clinical staging is the most accurate reflection of cancer prognosis; it guides therapy and is a survival reference point.[4] Accurate staging has become increasingly important as the options for therapy have increased.[5]

A number of recent studies assessing oesophageal cancer therapy have shown differing results, especially with regards to the benefits of neoadjuvant chemotherapy and radiotherapy. One prospective randomized trial suggested that neoadjuvant chemoradiotherapy reduced mortality and increased disease-free survival in patients with locally advanced oesophageal cancer (T1N1M0 or T2–3 N0–1 M0).[6] However, in patients with T2N0 disease, the benefit is less clear, with a recent European multicenter retrospective study demonstrating that neoadjuvant therapy had no impact on recurrence, disease-free survival and overall survival.[7] A major limitation of those studies has been the variable accuracy of clinical staging for T2N0 disease reported in the literature. When compared with postoperative

Address correspondence to: Dr Sarmed S. Sami. Email: s.sami@ucl.ac.uk

pathological staging as the reference standard, clinical staging can be accurate in as low as 6% or as high as 42% of patients depending on which study is considered.[8–10]

Clinical staging of oesophageal cancer uses a number of modalities.[11–13] This includes any combination of endoscopic ultrasound (EUS), computerized tomography (CT), ultrasound (US), positron emission tomography (PET) and laparoscopy. A number of studies and systematic reviews have evaluated the accuracy of these tests individually.[14–16] While this information is useful, it does not accurately reflect clinical practice when modalities are used in combination.

Accurate clinical staging of oesophageal cancer is vital to ensure that appropriate therapeutic decisions are made both to direct clinical care and to enable precision research. Clinical staging of T2 oesophageal cancer (cT2) is of particular significance, because patients with cancers that are >cT2N0 undergo different treatment pathways to those with ≤cT2N0, hence this threshold has a significant clinical application. Patients with >cT2N0 are typically offered neoadjuvant chemotherapy in addition to surgery, patients with cT2N0 may be offered surgery alone and patients with <cT2N0 may be offered local therapies including surgery or endotherapy including endoscopic mucosal resection or endoscopic submucosal dissection.[11–13] Accurate staging in this setting has significant implications on mortality and morbidity outcomes.

When evaluating the accuracy of clinical staging tests, pathology is used as the reference standard. Patients with >T2N0M0 disease usually receive neoadjuvant therapy which may downstage their cancer prior to surgery; therefore, the postoperative pathological staging in those patients may not reflect the true preoperative clinical stage. On the other hand, patients with <T2N0M0 disease usually undergo endoscopic therapy, and therefore, staging in these patients is not accurate.[11–13] For these reasons, we have decided to only include patients with cT2N0M0 disease.

We believe that this is a significant clinical issue as 21–30% of patients present with stage 1 or 2 oesophageal cancer.[17] Having an accurate staging test will ensure that these patients receive appropriate evidence-based treatment.

In summary, precise data on accuracy of clinical staging remain lacking. Moreover, data on the understaging and overstaging of T2N0 oesophageal cancer are also lacking. This has major implications for treatment decisions and patient survival among other factors. We aimed to address these knowledge gaps by a systematic review and meta-analysis of the relevant literature.

MATERIALS AND METHODS

This study was conducted according to guidance provided by the Cochrane Collaboration handbook for systematic reviews[18] and followed a prespecified protocol. The study was prospectively registered on the PROSPERO international database (CRD42019157635).

Search strategy

We searched Ovid MEDLINE, Ovid EMBASE and The Cochrane Library databases for studies published from database inception to 1 June 2019 for relevant articles evaluating staging accuracy in oesophageal cancer. No restrictions were applied to the search algorithm (Supplementary Table 1).

Study selection and outcome measures

We included studies that met the following criteria:

- Performed in cohorts of adult patients defined as 18 years or older
- Provided data on T2 oesophageal cancer staging accuracy
- Used multiple imaging modalities to assess clinical staging
- Compared clinical staging to pathological staging as the reference standard test
- Provided sufficient data to allow the calculation of staging accuracy.

Accuracy was defined as the proportion of patients with correct clinical staging prior to surgery compared with postoperative pathological staging as the reference standard. The primary outcome was diagnostic accuracy of combined T and N staging. Secondary outcomes were accuracy of T stage only; percentage T downstaged; percentage T upstaged and percentage N upstaged.

We only evaluated patients with cT2N0M0 and were unable to calculate sensitivity and specificity values as these require the number of false negatives for a given test. These data were not available, because incorrectly staged patients as >T2N0 would have typically undergone neoadjuvant therapy; therefore, their postoperative pathological staging may not reflect their preoperative stage.

All titles and abstracts identified by the primary searches were screened by two reviewers (P.W. and A.H.). Full-text articles of potentially eligible studies were read and assessed for inclusion. Data were then independently extracted by P.W. and A.H. before entry into a standardized pro forma (Excel 2010; Microsoft, Redmond, Wash). Disparity between the data collected was resolved by discussion, and if no agreement was reached, a third investigator (S.S.S.) was consulted. The corresponding authors of the

primary studies were contacted to ask for any missing data.

Variables for each study were collected as follows: Year of publication, years of recruitment, number of centers, population studied (age, sex, sample size and cancer subtype), country of origin, study design and imaging modalities used.

Study quality was assessed independently by two investigators (P.W. and A.H.) using the updated version of the quality assessment of diagnostic accuracy studies (QUADAS-2) tool.[19]

Statistical analysis

All outcomes were binary in nature. Meta-analysis methods were used to pool together the results from different studies for each outcome. The Freeman–Tukey double arcsine transformation was performed before analysis. This was used to stabilize the variances when the proportions were close to zero and one, and a Normal approximation to the binomial distribution did not hold.

The DerSimonian–Laird random-effects method was used for the analysis, regardless of the degree of heterogeneity between the study results.

Heterogeneity, subgroup analyses and publication bias

The heterogeneity between studies was assessed based on the significance of the between-study heterogeneity and also on the size of the I^2 value. Low, moderate and high inconsistency were associated with I^2 values of 25%, 50 and 75, respectively.[20] Additional subgroup analyses were performed to examine if any factors could explain the heterogeneity between studies. We evaluated several factors *a priori* including: publication date (before and after January 2015), date of first recruitment (before and after 2000), number of centers (single vs. multicenter), sample size (<100 vs. ≥ 100), geographical location of study (USA, Europe or Asia) and main histological subtype included (adenocarcinoma or squamous cell carcinoma). Publication bias was assessed with the use of a Funnel plot and by the Egger test. A P value of ≤ 0.10 was considered to represent possible publication bias.

When calculating heterogeneity, the dates used were chosen to achieve a near equal split before and after the declared date.

RESULTS

Characteristics of included studies

After duplicates were removed, the search strategy identified a total of 1,199 studies which were all screened by title and abstract. Fifty-six potentially eligible articles were identified which were all read in full. Twenty studies[7–10,21–36] met the inclusion criteria ($n = 5{,}213$ patients) (Fig. 1). Of these studies, 16 ($n = 4{,}182$ patients)[8,10,21–23,25,27–36] reported the accuracy of combined T and N staging. Eighteen studies ($n = 4{,}471$ patients)[7–10,21–23,26–36] provided data on the accuracy of T stage and 18 studies ($n = 5{,}180$ patients),[7,8,10,21–25,27–36] presented data for the accuracy of N stage.

A summary of the included studies is shown in Table 1. The characteristics of the patients included in the studies are shown in Table 2. Supplementary Figure 1 displays the outcomes of the QUADAS-2 quality assessment.[19] No studies were high risk for bias.

Outcomes

T&N staging

Combined T and N staging was accurate in 19% of patients [95% confidence interval (CI), 15–24; 16 studies; $n = 4{,}182$ patients; $I^2 = 88\%$; $P < 0.01$] (Fig. 2 and Table 3).

Sources of heterogeneity in accuracy of T&N staging included: number of centers [single center (13%; 95% CI, 9–18; $I^2 = 29\%$) vs. multicenter studies (27%; 95% CI, 20–34; $I^2 = 95\%$); $P = 0.01$]; sample size [$n \leq 100$ patients (13%; 95% CI, 9–18; $I^2 = 29\%$) vs. $n \geq 100$ patients (27%; 95% CI, 20–34; $I^2 = 95\%$); $P = 0.01$]; study region [USA (17%; 95% CI, 13–23; $I^2 = 89\%$) vs. Asia (19%; 95% CI, 10–28; $n = 2$ studies) vs. Europe (38%; 95% CI, 32–46; $n = 1$ study); $P < 0.01$] and most common histological subtype included in each study [adenocarcinomas (12%; 95% CI, 7–18; $I^2 = 20\%$) vs. mixed histological types (22%; 95% CI, 15–31; $I^2 = 95\%$) vs. SCC (19%; 95% CI, 10–28; $n = 2$ studies) vs. unknown histological types (26%; 95% CI, 23–29%; $n = 2$ studies); $P < 0.01$)] (Supplementary Table 2). There was no significant evidence of publication bias ($P = 0.11$) (Supplementary Fig. 2).

T stage accuracy

T staging was accurate in 29% of patients (95% CI, 24–35; 18 studies; $n = 4{,}471$ patients; $I^2 = 91\%$; $P < 0.01$) (Fig. 3A).

Sources of heterogeneity in T stage accuracy included: number of centers [single center (23%; 95% CI, 17–28; $I^2 = 31\%$) vs. multicenter studies (36%; 95% CI, 28–44; $I^2 = 96\%$); $P = 0.01$] (Supplementary Table 3). There was no significant evidence of publication bias ($P = 0.20$) (Supplementary Fig. 3).

T downstaging

The percentage of patients downstaged after surgery was 41% (95% CI, 33–50; 18 studies; $n = 4{,}471$ patients; $I^2 = 96\%$; $P < 0.01$) (Fig. 3B).

Sources of heterogeneity in T downstaging included: number of centers [single center (53%; 95% CI, 44–62;

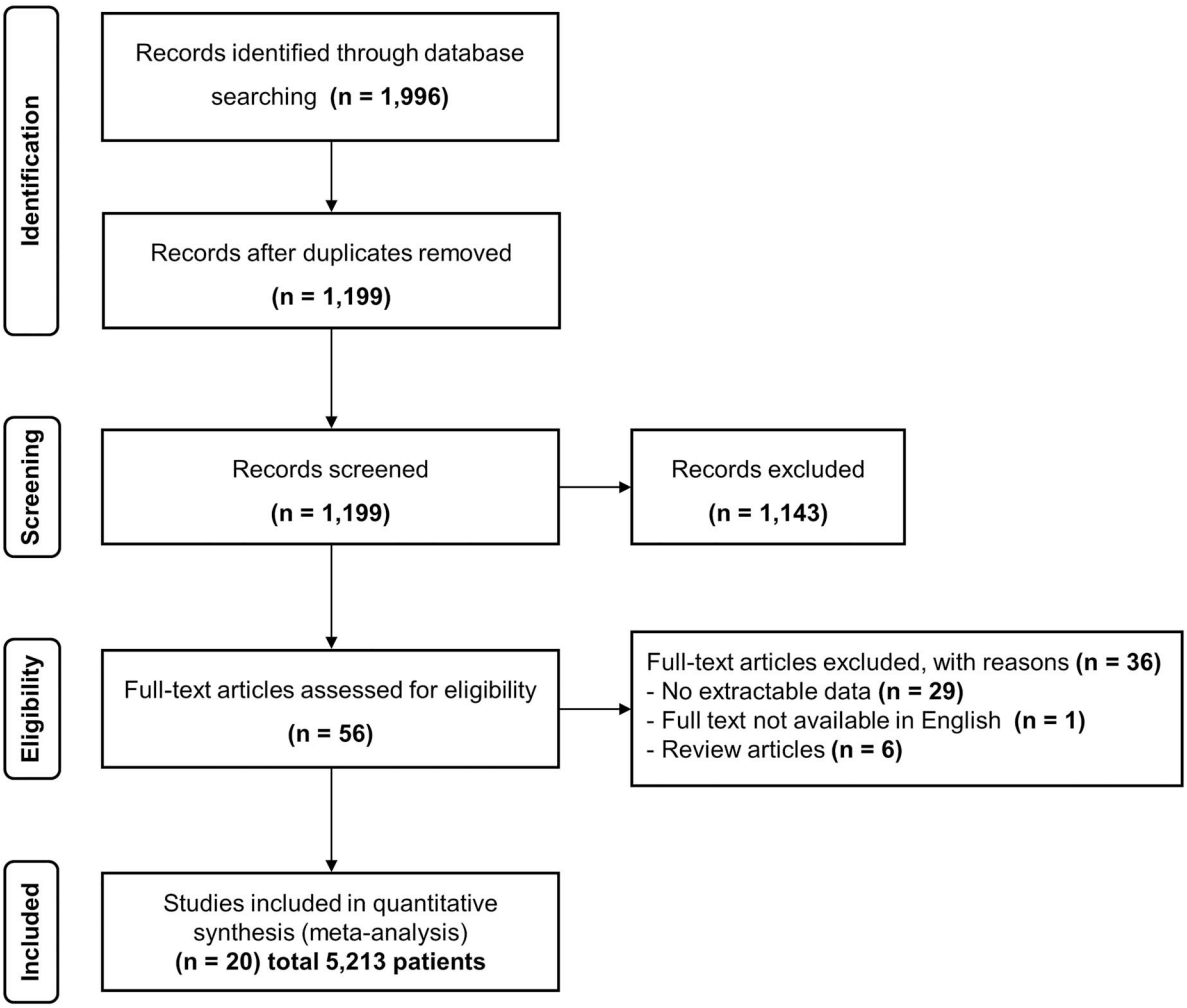

Fig. 1 Flow diagram of the search strategy and selection of studies.

$I^2 = 65\%$) vs. multicenter studies (29%; 95% CI, 19–40; $I^2 = 98\%$); $P < 0.01$)]; most common histological subtype included in each study [adenocarcinomas (55%; 95% CI, 45–65; $I^2 = 36\%$) vs. mixed histological types (33%; 95% CI, 21–46; $I^2 = 95\%$) vs. SCC (52%; 95% CI, 38–66; $n = 3$ studies) vs. unknown histological types (33%; 95% CI, 29–36%; $n = 2$ studies); $P < 0.01$)] and sample size [$n \leq 100$ patients (52%; 95% CI, 44–61; $I^2 = 62\%$) vs. $n \geq 100$ patients (27%; 95% CI, 17–38; $I^2 = 98\%$); $P < 0.01$] (Supplementary Table 4). There was evidence of publication bias as the Egger Test P-value was 0.02 (Supplementary Fig. 4).

T upstaging

The percentage of patients T upstaged after surgery was 28% (95% CI, 24–32; $n = 18$ studies; 5,184 patients; $I^2 = 86\%$, $P < 0.01$) (Fig. 4A).

Sources of heterogeneity in T upstaging included: number of centers [single center (22%; 95% CI, 16–29; $I^2 = 50\%$) vs. multicenter studies (32%; 95% CI, 27–38; $I^2 = 93\%$); $P = 0.03$)]; study region [USA (29%; 95% CI, 24–32; $I^2 = 85\%$) vs. Asia (14%; 95% CI, 8–22; $n = 3$ studies) vs. Europe (31%; 95% CI, 16–48; $n = 3$ studies); $P < 0.01$]; most common histological subtype included in each study [adenocarcinomas (22%; 95% CI, 16–28; $I^2 = 0\%$) vs. mixed histological types (32%; 95% CI, 27–37; $I^2 = 86\%$) vs. SCC (14%; 95% CI, 8–22; $n = 3$ studies) vs. unknown histological types (40%; 95% CI, 37–44%; $n = 2$ studies); $P < 0.01$)] and sample size [$n \leq 100$ patients (21%; 95% CI, 15–27; $I^2 = 51\%$) vs. $n \geq 100$ patients (34%; 95% CI, 28–39; $I^2 = 93\%$); $P = 0.01$] (Supplementary Table 5). There was no significant evidence of publication bias ($P = 0.55$) (Supplementary Fig. 5).

N upstaging

The percentage of patients N upstaged after surgery was 34% (95% CI, 30–39; 18 studies; $n = 5{,}180$ patients; $I^2 = 86\%$; $P < 0.01$) (Fig. 4B).

Sources of heterogeneity in N upstaging included: study region [USA (33%; 95% CI, 29–38; $I^2 = 88\%$) vs. Asia (29%; 95% CI, 20–40; $n = 2$ studies) vs. Europe (47%; 95% CI, 42–51; $n = 2$ studies); $P < 0.01$]

Table 1 Summary of included studies ($n = 20$)

Author and year	Country (Recruitment years)	Modalities used (study design)	No. of pts (Number of centers)	% Accuracy on T&N stage	% Accuracy on T stage	% T-downstaged	% T-upstaged	% N upstaged
Nishimaki T[9] 1999	Japan (1993–1994)	CT/EUS, abdominal US (prospective)	23 (16)	NS	43%	48%	9%	.
Rice TW[31] 2007	USA (1987–2005)	PET/EUS/CT (retrospective)	53 (1)	13%	17%	62%	21%	25%
Crabtree TD *et al.*[29] 2011	USA (2000–2008)	PET/EUS/CT (retrospective)	18 (1)	6%	28%	50%	22%	17%
Stiles BM[32] 2011	USA (1992–2005)	PET/EUS/CT (retrospective)	40 (1)	13%	30%	30%	40%	55%
Chen WH[33] 2012	Taiwan (1995–2005)	EUS/CT/bone scan/barium swallow (retrospective)	14 (1)	43%	50%	36%	14%	0%
Zhang JQ[34] 2012	USA (1989–2009)	PET/EUS/CT/Barium swallow (retrospective)	14 (1)	29%	50%	29%	21%	43%
Crabtree TD[35] 2013	USA (2002–2011)	PET/EUS/CT (retrospective)	482 (multiple)	27%	40%	34%	26%	21%
Shin S [36] 2014	Korea (2005–2010)	PET & EUS (retrospective)	66 (1)	15%	23%	61%	17%	39%
Tekola BD[21] 2014	USA (2003–2013)	EUS & PET/CT (retrospective)	30 (1)	17%	20%	67%	13%	33%
Hardacker TJ[22] 2014	USA (1990–2011)	EUS and 'cross sectional imaging' (retrospective)	35 (1)	9%	17%	43%	40%	40%
Speicher PJ[23] 2014	USA (1998–2011)	PET/EUS/CT (retrospective)	786 (multiple)	27%	27%	32%	42%	30%
Dolan JP[8] 2016	USA (1999–2011)	all had PET,CT EUS (retrospective)	16 (1)	6%	19%	50%	31%	56%
Markar SR[7] 2016	France (2000–2010)	PET/EUS/CT (retrospective)	285 (30)	NS	27%	39%	35%	48%
Samson P[24] 2016	USA (2006–2012)	PET/EUS/CT (retrospective)	713 (multiple)	NS	NS	NS	32%	34%
Luu C [25] 2017	USA (2000–2015)	PET/EUS/CT (retrospective)	29 (1)	21%	NS	NS	NS	24%
Winiker M[26] 2018	Switzerland (2000–2006)	PET/EUS/CT (retrospective)	10 (1)	NS	10%	90%	0%	NS
Goense L[10] 2018	Netherlands (2005–2014)	PET/EUS/CT/US neck (retrospective)	180 (multiple)	38%	53%	0%	47%	45%
Shridhar R[27] 2018	USA (2004–2013)	NS (retrospective)	1840 (multiple)	31%	44%	31%	26%	34%
Barbetta A *et al.*[30] 2018	USA (1997–2016)	PET & EUS (retrospective)	80 (1)	9%	18%	61%	21%	35%
Atay SM[28] 2019	USA (2002–2012)	PET/EUS/CT (retrospective)	499 (26)	14%	24%	44%	32%	39%

CT, computerized tomography; EUS, endoscopic ultrasound; NS, not stated; PET, positron emission tomography; US, ultrasound.

Table 2 Summary of characteristics of patients included in the analysis

Author and year	Country	Mean age, (years)	% Male	% Adeno	% SCC
Nishimaki T[9] 1999	Japan	62	88.0%	0.4%	96.4%
Rice TW[31] 2007	USA	65	NS	97.0%	3.0%
Crabtree TD *et al.*[29] 2011	USA	NS	NS	100.0%	0.0%
Stiles BM[32] 2011	USA	62.5	81.4%	71.6%	28.4%
Chen WH[33] 2012	Taiwan	60.9	78.6%	0.0%	100.0%
Zhang JQ[34] 2012	USA	69	85.7%	100.0%	0.0%
Crabtree TD[35] 2013	USA	63.8	83.2%	86.3%	13.7%
Shin S[36] 2014	Korea	63.1	95.0%	1.3%	97.5%
Tekola BD[21] 2014	USA	NS	89.5%	NS	NS
Hardacker TJ[22] 2014	USA	62.5	82.9%	82.9%	17.1%
Speicher PJ[23] 2014	USA	66	82.5%	N/S	N/S
Dolan JP[8] 2016	USA	68	94.0%	93.8%	6.3%
Markar SR[7] 2016	France	NS	80.7%	49.5%	50.5%
Samson P[24] 2016	USA	65.6	82.3%	69.3%	30.7%
Luu C[25] 2017	USA	64.9	82.0%	92.8%	72.0%
Winiker M[26] 2018	Switzerland	65	74.5%	59.0%	41.0%
Goense L[10] 2018	Netherlands	66	77.0%	81.0%	19.0%
Shridhar R[27] 2018	USA	67	80.4%	87.5%	12.5%
Barbetta A[30] 2018	USA	64	78.0%	100.0%	0.0%
Atay SM[28] 2019	USA	66	78.0%	88.0%	12.0%

Adeno, adenocarcinoma; NS, not stated; SCC, squamous cell carcinoma.

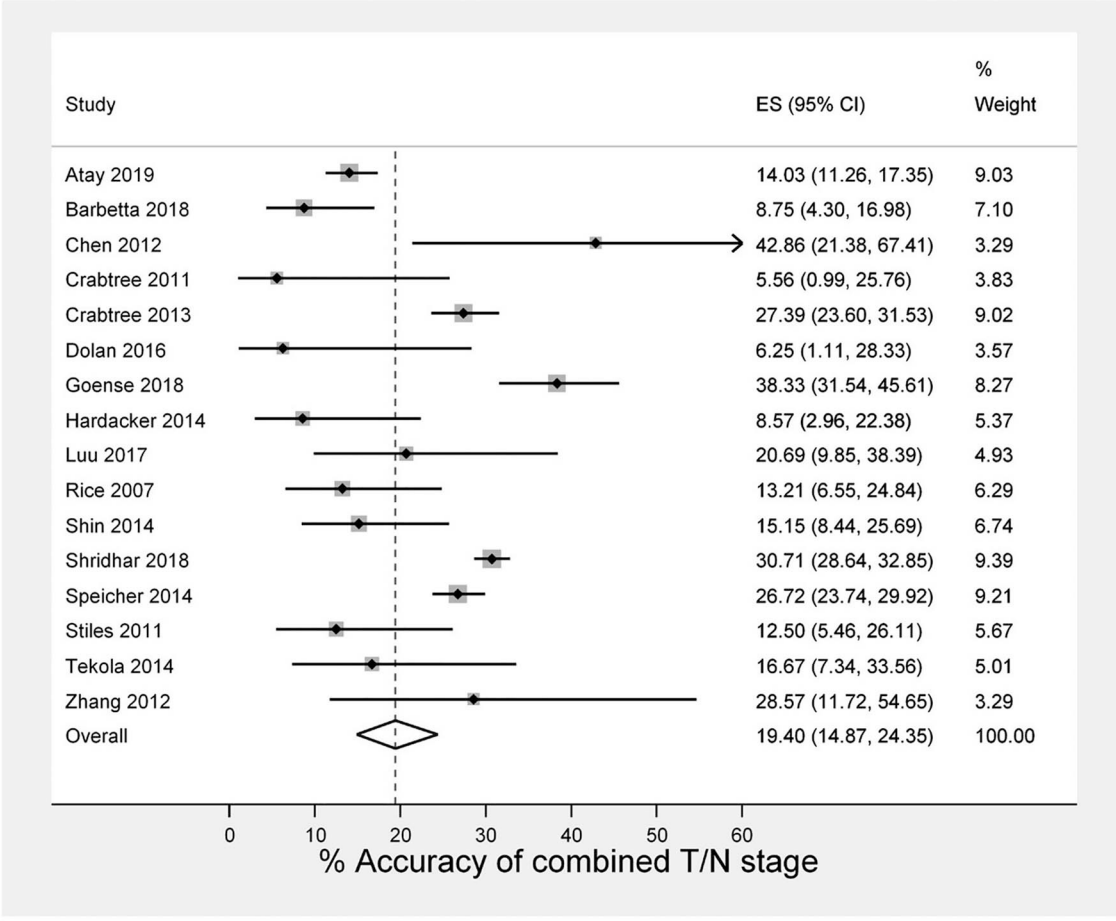

Fig. 2 Forest plot for combined T/N stage accuracy.

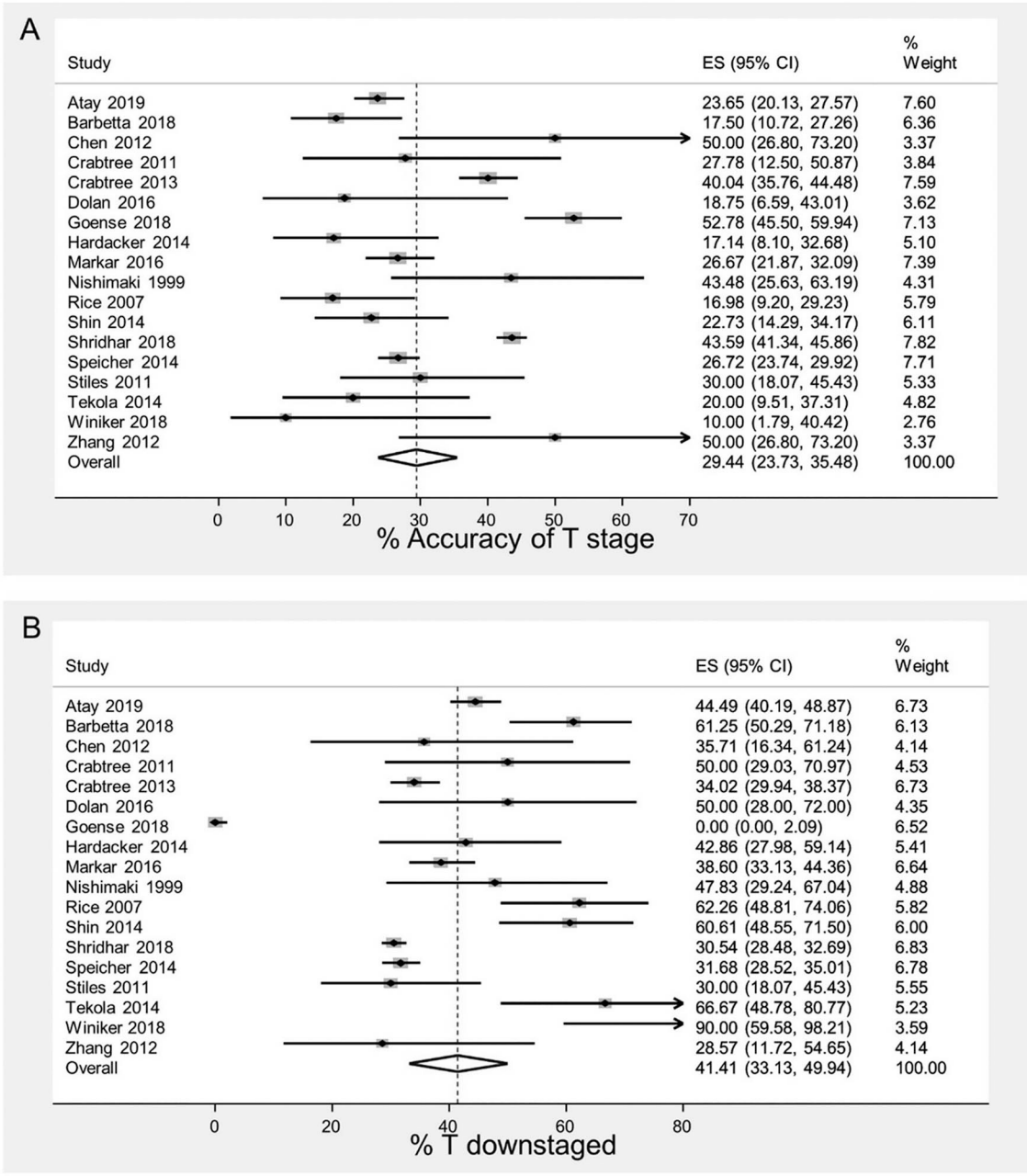

Fig. 3 Forest plots for secondary outcomes. (A) T stage accuracy. (B) T downstaging.

Table 3 Meta-analysis results for all studies combined

Outcome	Number of studies	Heterogeneity *P*-value	Pooled % I^2	(95% CI)
Combined T/N stage accuracy	16	< 001	88%	19% (15%, 24%)
T stage accuracy	18	< 001	91%	29% (24%, 35%)
T downstaged	18	< 001	96%	41% (33%, 50%)
T upstaged	19	< 001	86%	28% (24%, 32%)
N upstaged	18	< 001	86%	34% (30%, 39%)

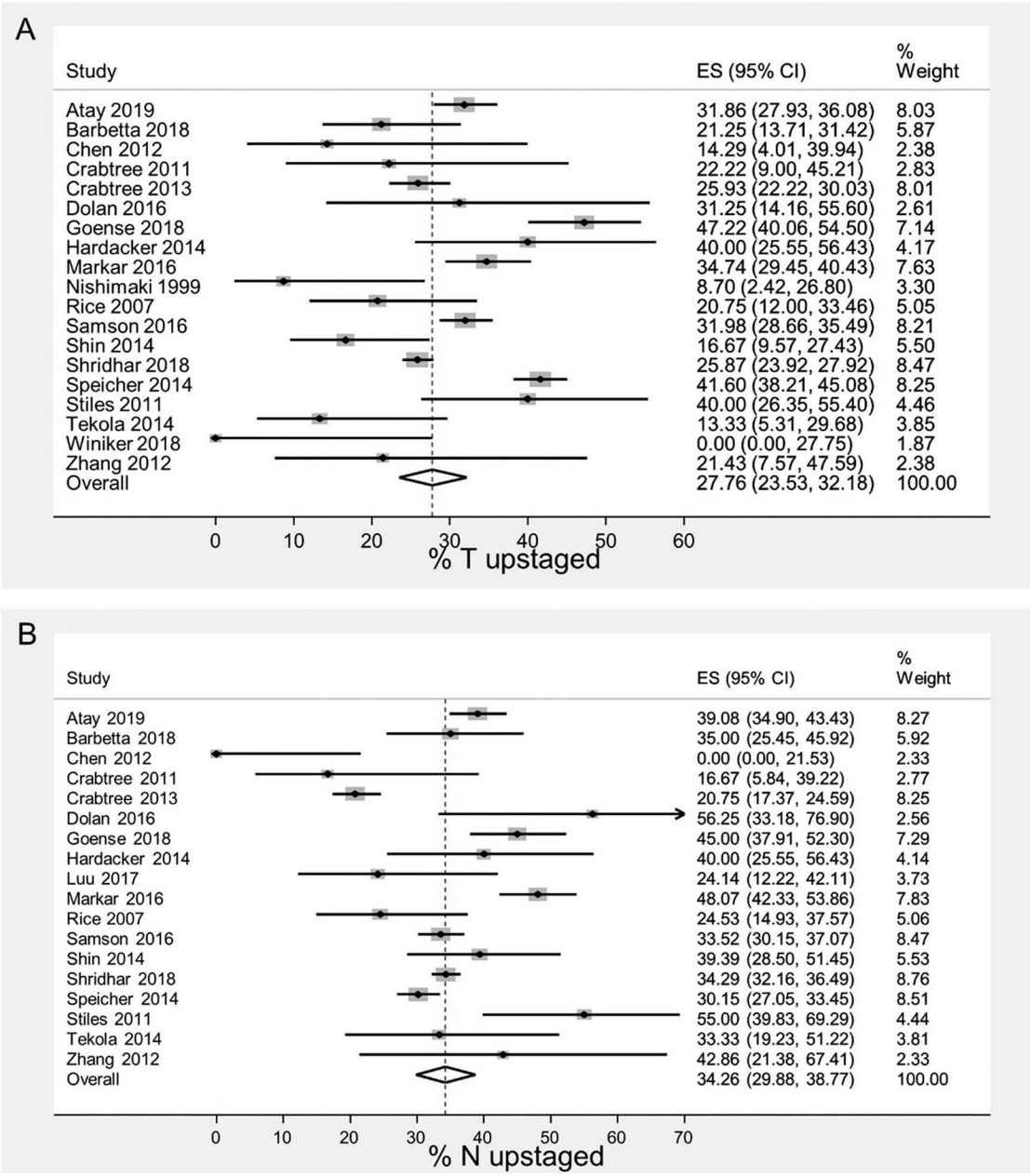

Fig. 4 Forest plots for secondary outcomes. (A) T upstaging. (B) N upstaging.

(Supplementary Table 6). There was no significant evidence of publication bias ($P = 0.63$) (Supplementary Fig. 6).

DISCUSSION

Principal findings

To our knowledge, this is the first systematic review and meta-analysis evaluating the accuracy of current clinical staging tests of T2 oesophageal cancer. We only included studies using multimodality staging in order to reflect current day-to-day clinical practice. We demonstrate that combined T/N staging was accurate in only 19% of patients. The proportion of patients with an accurate T stage was 29%. Percentage of patients who had T downstaged, T upstaged or N upstaged was 41, 28 and 34%, respectively. Small ($n < 100$ patients), single-center studies reported lower accuracy compared with larger ($n > 100$ patients), multicenter studies. The latter study design has fewer biases and is therefore more likely to be representative of the truth rather than chance findings. Of note, we found no significant improvement in accuracy in the more recent studies (published after January 2015 and recruitment started after January 2000) compared with older studies (published before January 2015 and recruitment started before January 2000). These dates were chosen to provide an equal split between earlier and later studies. There was

considerable heterogeneity between studies for all primary and secondary outcomes. This suggests that the study results should be interpreted with caution. We explored this further through subgroup analyses.

The inaccuracy of staging T2 disease is thought to stem from missed occult nodal metastases, largely at EUS, resulting in staging of the disease (N stage) prior to surgery and subsequent upstaging (N stage) following surgery.[35] However, our study shows that the accuracy for T stage only is equally poor (29%) with even higher rates of downstaging to T1 (41%) than upstaging (28%). The separation of muscle layers with EUS remains very challenging. In one study, 44% of EUS-staged cT2 OACs found to be pT1 tumors and curative endoscopic resection was achieved in more than a third of these cases.[37] Similarly, our data suggest that a significant proportion of patients are perhaps over-treated with surgery instead of being offered organ-sparing endoscopic therapy. Similarly, a significant proportion is being undertreated and potentially missing out of life prolonging neoadjuvant therapy.

Increasing the rates of EUS-guided FNA has been suggested as a way to improve node detection,[38] while the use of higher frequency probes may improve the detection of tumor depth; however, both these techniques will increase the technical complexity of the procedure and their efficacy has not been proven.[39] There are significant resources being applied to improving outcomes for oesophageal cancer patients.[40] In order for this research to be precise and for the findings from this research to be applicable to clinical practice, the clinical stage of the patients included in these studies must be accurate. EUS has been used in GI imaging since the 1980s.[41] It is used both for diagnostic and therapeutic reasons. Despite reported high sensitivity and specificity of EUS in T and N stage, limitations remain such that it does not really accurately stage the patient.[13,14,16] Accuracy is highly operator dependent and yet despite improvements in the technology continues to be poor for T2 tumors.[42] Furthermore, ~30% of oesophageal tumors are not traversable with the EUS probe at diagnosis[43] and performing FNA for a node in the vicinity of a tumor remains challenging.

New technologies to aid staging include biomarkers and magnetic resonance imaging (MRI). Despite the advances in biomarker research, none has become clinically viable for detection and staging of oesophageal cancers.[44,45] MRI has been shown to be sensitive for T stage but is subject to significant technical challenges[46] such as movement artifact, MRI bore diameters and the longer imaging times.[47] Significant improvements have been made with respiratory and cardiac gating.[48] Optical coherence tomography is an advanced imaging technique that uses the reflection of infrared light from the target tissue [KAUL]. Unfortunately, the light is only able to penetrate 3 mm, so it is unsuitable for providing accurate tumor stage.

Study strengths and limitations

Our study includes a large number of patients with T2 oesophageal cancer, which allows clinical staging accuracy to be assessed with high precision.

In our study, the majority of patients are from Western populations with adenocarcinoma. This accurately reflects the demographic of oesophageal cancer patients seen in Europe and the USA, but perhaps not in Asia.

Three studies included in the review primarily included patients with SCC histological subtype. We note that although the total number of included patients in these studies is 103, the T&N staging accuracy was similar across these three studies to studies primarily including adenocarcinomas, namely, 19% (95% CI, 10–28%) for SCC studies compared with 12% (95% CI, 7–18%) for adenocarcinoma studies ($P < 0.01$).

It should also be noted that the majority of the studies are from tertiary referral centers. This is probably an accurate representation of the clinical pathway as oesophagectomies are increasingly done in high-volume centers.[49] A large proportion of the cases reviewed came from four large USA-based studies. These studies were all retrospective and used databases to collect the published data. This may limit some of the applicability of our findings to populations outside of the USA. Unfortunately, the time lag between index test (clinical staging) and reference standard (pathological staging) was rarely published, which could lead to an incorrect assessment of clinical staging accuracy due to disease progression in the interim. While unlikely, this could lead to clinical understaging of the disease status as T and N stage could progress during the time interval between clinical staging and surgery.

Values of heterogeneity indicate an inconsistency between the different studies, which will compromise any reliable conclusions that can be drawn from this analysis. This issue is particularly evident in multicenter studies. There was a significant reduction in heterogeneity when calculations are restricted to smaller ($n < 100$) and single center studies.

Eighteen of the twenty included papers have been published in the past decade and use all imaging modalities currently available in clinical practice, which increases the clinical relevance and external validity of this work.

Unfortunately, most of the studies included did not publish some data that have been associated with the accuracy of staging such as FNA and tumor length.

Implications for clinical practice

Patient care and prognosis is fundamentally based around the staging of their cancer at diagnosis. If this staging is inaccurate, we may be offering incorrect therapies or denying patients the correct therapies. The latter may be ineffective or even harmful. For understaged patients, they may go on to have surgery with curative intent despite having more advanced disease. These patients may suffer unnecessary postoperative complications leading to significant mortality and morbidity as well as being exposed to long-lasting deterioration in the quality of life.[50,51]

Individuals who have T or N upstaging following pathological assessment would have not received neoadjuvant therapy and are therefore deprived of the additional survival benefit this therapy confers. The survival benefit of giving adjuvant chemotherapy after surgical resection remains unproven.[52–54] Similarly, individuals that have T downstaging following pathological assessment may have potentially been candidates for organ preserving endoscopic therapy rather than subjecting them to unnecessary oesophagectomy.

CONCLUSIONS

The accuracy of clinical staging for oesophageal T2 cancers remains poor and is largely unchanged in recent years.[27] Patients that were downstaged after surgery may have successfully been treated with endotherapy. The 28% of patients who were upstaged would have been offered neoadjuvant therapy according to current guidelines.[11–13]

These data reinforce the need for further research directed at improving T2 oesophageal cancer staging. In addition, this may indicate a greater need beyond T2N0 disease. Our rationale for reviewing T2N0 oesophageal cancer was to enable direct comparison between clinical staging and pathological staging in patients not undergoing neoadjuvant therapy. This is not possible for patients with clinical staging >T2N0 cancer (due to the downstaging effects of neoadjuvant therapy), and therefore, it remains a possibility that the inaccuracy we have detected may also affect patients with more advanced cancers.

It is possible that MRI may offer reliable staging in the future but the technical challenges of adapting this technology need to be resolved. Translational technologies such as X-ray phase contrast imaging which has been used in breast cancer imaging may provide more accuracy than those currently used.[55,56] Ongoing work looking at adapting current imaging techniques with postprocessing techniques such as radiomics and artificial intelligence has shown promise in enhancing CT, MRI and PET.[57]

SUPPLEMENTARY DATA

Supplementary data mentioned in the text are available to subscribers in *DOTESO* online.

ACKNOWLEDGMENTS

The authors received no financial support for the research, authorship and/or publication of this article. The authors declare that they have no conflict of interest.

AUTHORS' CONTRIBUTION

P.W., L.B.L. and S.S.S. conceived and drafted the study. P.W. and A.H. collected the data. P.W., P.B. and S.S.S. analyzed and interpreted the data. P.W., A.H., L.B.L. and S.S.S. drafted the manuscript. All authors commented on drafts of the paper. All authors have approved the final draft of the manuscript.

ABBREVIATIONS

CT, computerized tomography; EUS, endoscopic ultrasound; PET, positron emission tomography; US, ultrasound

References

1 Bray F, Ferlay J, Soerjomataram I, Siegel R L, Torre L A, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394–424.
2 Oesophageal cancer incidence statistics | Cancer Research UK. [Accessed 29 Jan 2019]. Available from URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence#ref-5
3 Njei B, Mccarty T R, Birk J W. Trends in esophageal cancer survival in United States adults from 1973 to 2009: a SEER database analysis. J Gastroenterol Hepatol 2016; 31: 1141–6.
4 Rice T W, Patil D T, Blackstone E H. 8th edition AJCC/UICC staging of cancers of the esophagus and esophagogastric junction: application to clinical practice. Ann Cardiothorac Surg 2017; 6: 119–30.
5 Mariette C, Piessen G, Briez N, Gronnier C, Triboulet J P. Oesophagogastric junction adenocarcinoma: which therapeutic approach? Lancet Oncol 2011; 12: 296–305.
6 Shapiro J, van Lanschot J J B, Hulshof M C C M *et al.* Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol 2015; 16: 1090–8.
7 Markar S R, Gronnier C, Pasquer A *et al.* Role of neoadjuvant treatment in clinical T2N0M0 oesophageal cancer: results from a retrospective multi-center European study. Eur J Cancer 2016; 56: 59–68.
8 Dolan J P, Kaur T, Diggs B S *et al.* Significant understaging is seen in clinically staged T2N0 esophageal cancer patients undergoing esophagectomy. Dis Esophagus 2016; 29: 320–5.
9 Nishimaki T, Tanaka O, Ando N *et al.* Evaluation of the accuracy of preoperative staging in thoracic esophageal cancer. Ann Thorac Surg 1999; 68: 2059–64.
10 Goense L, Visser E, Haj Mohammad N *et al.* Role of neoadjuvant chemoradiotherapy in clinical T2N0M0 esophageal cancer: a population-based cohort study. Eur J Surg Oncol 2018; 44: 620–5.

11 Ajani JA, D TA, Chair V, *et al.* [Accessed 26 July 2019]. Available from URL: NCCN guidelines version 2.2019 Esophageal and Esophagogastric junction cancers continue NCCN., 2019. https://www.nccn.org/professionals/physician_gls/pdf/esophageal.pdf
12 Kuwano H, Nishimura Y, Oyama T *et al.* Guidelines for diagnosis and treatment of carcinoma of the Esophagus April 2012 edited by the Japan Esophageal Society. Esophagus 2015; 12: 1–30.
13 Lordick F, Mariette C, Haustermans K *et al.* Oesophageal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2016; 27: v50–7.
14 Thosani N, Singh H, Kapadia A *et al.* Diagnostic accuracy of EUS in differentiating mucosal versus submucosal invasion of superficial esophageal cancers: a systematic review and meta-analysis. Gastrointest Endosc 2012; 75: 242–53.
15 Kelly S, Harris M, Berry E *et al.* A systematic review of the staging performance of endoscopic ultrasound in gastro-oesophageal carcinoma. Gut 2001; 49: 534–9.
16 Puli S R, Reddy J B, Bechtold M L *et al.* Staging accuracy of esophageal cancer by endoscopic ultrasound: a meta-analysis and systematic review ESOPHAGEAL CANCER. World J Gastroenterol 2008; 14: 1479–90.
17 Oesophageal cancer incidence statistics | Cancer Research UK. [Accessed 8 Jun 2020]. Available from URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence#heading-Three
18 Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training. [Accessed 28 Jan 2020]. Available from URL: https://training.cochrane.org/handbook/current
19 Whiting P F, Rutjes A W S, Westwood M E *et al.* Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155: 529–36.
20 Higgins J P T, Thompson S G, Deeks J J, Altman D G. Measuring inconsistency in meta-analyses. BMJ 2003; 327: 557–60.
21 Tekola B D, Sauer B G, Wang A Y, White G E, Shami V M. Accuracy of endoscopic ultrasound in the diagnosis of T2N0 Esophageal Cancer. J Gastrointest Cancer 2014; 45: 342–6.
22 Hardacker T J, Ceppa D, Okereke I *et al.* Treatment of clinical T2N0M0 Esophageal Cancer. Ann Surg Oncol 2014; 21: 3739–43.
23 Speicher P J, Ganapathi A M, Englum B R *et al.* Induction therapy does not improve survival for clinical stage T2N0 Esophageal Cancer. J Thorac Oncol 2014; 9: 1195–201.
24 Samson P, Puri V, Robinson C *et al.* Clinical T2N0 Esophageal cancer: identifying Pretreatment characteristics associated with pathologic upstaging and the potential role for induction therapy. Ann Thorac Surg 2016; 101: 2102–11.
25 Luu C, Amaral M, Klapman J *et al.* Endoscopic ultrasound staging for early esophageal cancer: are we denying patients neoadjuvant chemo-radiation? World J Gastroenterol 2017; 23: 8193–9.
26 Winiker M, Mantziari S, Figueiredo S G *et al.* Accuracy of preoperative staging for a priori resectable esophageal cancer. Dis Esophagus 2018; 31: 1–6.
27 Shridhar R, Huston J, Meredith K L. Accuracy of endoscopic ultrasound staging for T2N0 esophageal cancer: a National Cancer Database analysis. J Gastrointest Oncol 2018; 9: 887.
28 Atay S M, Correa A, Hofstetter W L *et al.* Predictors of staging accuracy, pathologic nodal involvement, and overall survival for cT2N0 carcinoma of the esophagus. J Thorac Cardiovasc Surg 2019; 157: 1264–1272.e6.
29 Crabtree T D, Yacoub W N, Puri V *et al.* Endoscopic ultrasound for early stage esophageal adenocarcinoma: implications for staging and survival. Ann Thorac Surg 2011; 91: 1509–16.
30 Barbetta A, Schlottmann F, Nobel T *et al.* Predictors of nodal metastases for clinical T2N0 Esophageal Adenocarcinoma. Ann Thorac Surg 2018; 106: 172–7.
31 Rice T W, Mason D P, Murthy S C *et al.* T2N0M0 esophageal cancer. J Thorac Cardiovasc Surg 2007; 133: 317–24.e1.
32 Stiles B M, Mirza F, Coppolino A *et al.* Clinical T2-T3N0M0 Esophageal Cancer: the risk of node positive disease. Ann Thorac Surg 2011; 92: 491–8.
33 Chen W-H, Chao Y-K, Chang H-K *et al.* Long-term outcomes following neoadjuvant chemoradiotherapy in patients with clinical T2N0 esophageal squamous cell carcinoma. Dis Esophagus 2012; 25: 250–5.
34 Zhang J Q, Hooker C M, Brock M V *et al.* Neoadjuvant chemoradiation therapy is beneficial for clinical stage T2 N0 Esophageal Cancer patients due to inaccurate preoperative staging. Ann Thorac Surg 2012; 93: 429–37.
35 Crabtree T D, Kosinski A S, Puri V *et al.* Evaluation of the reliability of clinical staging of T2 N0 esophageal cancer: a review of the Society of Thoracic Surgeons database. Ann Thorac Surg 2013; 96: 382–90.
36 Shin S, Kim H K, Choi Y S *et al.* Clinical stage T1–T2N0M0 oesophageal cancer: accuracy of clinical staging and predictive factors for lymph node metastasis. Eur J Cardio-Thoracic Surg 2014; 46: 274–9.
37 Gotink A W, Spaander M C W, Doukas M *et al.* Exploring diagnostic and therapeutic implications of endoscopic mucosal resection in EUS-staged T2 esophageal adenocarcinoma. Endoscopy 2017; 49: 941–8.
38 Vazquez-Sequeiros E, Levy M J, Clain J E *et al.* Routine vs. selective EUS-guided FNA approach for preoperative nodal staging of esophageal carcinoma. Gastrointest Endosc 2006; 63: 204–11.
39 Yoshinaga S. Endoscopic ultrasound using ultrasound probes for the diagnosis of early esophageal and gastric cancers. World J Gastrointest Endosc 2012; 4: 218.
40 Smyth E C, Lagergren J, Fitzgerald R C *et al.* Oesophageal cancer. Nat Rev Dis Primers 2017; 3: 17048.
41 Friedberg S R, Lachter J. Endoscopic ultrasound: current roles and future directions. World J Gastrointest Endosc 2017; 9: 499–505.
42 Krill T, Baliss M, Roark R *et al.* Accuracy of endoscopic ultrasound in esophageal cancer staging. J Thorac Dis 2019; 11: S1602–9.
43 Pfau P R, Ginsberg G G, Lew R J *et al.* Esophageal dilation for endosonographic evaluation of malignant esophageal strictures is safe and effective. Am J Gastroenterol 2000; 95: 2813–5.
44 Tan C, Qian X, Guan Z *et al.* Potential biomarkers for esophageal cancer. Springerplus 2016; 5: 467.
45 Lagergren J, Smyth E, Cunningham D *et al.* Oesophageal cancer. Lancet 2017; 390: 2383–96.
46 Giganti F, Ambrosi A, Petrone M C *et al.* Prospective comparison of MR with diffusion-weighted imaging, endoscopic ultrasound, MDCT and positron emission tomography-CT in the pre-operative staging of oesophageal cancer: results from a pilot study. Br J Radiol 2016; 89: 20160087.
47 Linder G, Korsavidou-Hult N, Bjerner T *et al.* 18F-FDG-PET/MRI in preoperative staging of oesophageal and gastroesophageal junctional cancer. Clin Radiol 2019; 74: 718–25.
48 van Rossum P S N, van Hillegersberg R, Lever F M *et al.* Imaging strategies in the management of oesophageal cancer: what's the role of MRI? Eur Radiol 2013; 23: 1753–65.
49 Markar SR, Low DE. The Volume-Outcome Relationship, Standardized Clinical Pathways, and Minimally Invasive Surgery for Esophagectomy. In: Minimally Invasive Foregut Surgery for Malignancy. Cham: Springer International Publishing, 2015; 25–34. [Accessed 3 Sept 2019]. http://link.springer.com/10.1007/978-3-319-09342-0_3
50 Daly J M, Fry W A, Little A G *et al.* Esophageal cancer: results of an American College of Surgeons patient care evaluation study. J Am Coll Surg 2000; 190: 562–72.
51 Jacobs M, Macefield R C, Elbers R G *et al.* Meta-analysis shows clinically relevant and long-lasting deterioration in health-related quality of life after esophageal cancer surgery. Qual Life Res 2014; 23: 1155–76.
52 Pasquali S, Yim G, Vohra R S *et al.* Survival after neoadjuvant and adjuvant treatments compared to surgery alone for resectable esophageal carcinoma. Ann Surg 2017; 265: 481–91.
53 Pouliquen X, Levard H, Hay J M *et al.* 5-fluorouracil and cisplatin therapy after palliative surgical resection of squamous cell carcinoma of the esophagus. A multicenter randomized trial. French associations for surgical research. Ann Surg 1996; 223: 127–33.

54 Saeed N A, Mellon E A, Meredith K L *et al.* Adjuvant chemotherapy and outcomes in esophageal carcinoma. J Gastrointest Oncol 2017; 8: 816–24.
55 Wolfson P, Jiang J, Wilson A *et al.* OTU-18 using X-ray phase contrast imaging to identify oesophageal pathology. Gut 2019; 68: A134–4.
56 Castelli E, Tonutti M, Arfelli F *et al.* Mammography with synchrotron radiation: first clinical experience with phase-detection technique. Radiology 2011; 259: 684–94.
57 van Rossum P S N, Xu C, Fried D V *et al.* The emerging field of radiomics in esophageal cancer: current evidence and future potential. Transl Cancer Res 2016; 5: 410–23.