Diagnostic accuracy of EUS in differentiating mucosal versus submucosal invasion of superficial esophageal cancers: a systematic review and meta-analysis

Nirav Thosani, MD, MHA,[1] Harvinder Singh, MD,[2] Asha Kapadia, PhD,[5] Nobuo Ochi, MD, PhD,[2] Jeffrey H. Lee, MD,[2] Jaffer Ajani, MD,[3] Stephen G. Swisher, MD,[4] Wayne L. Hofstetter, MD,[4] Sushovan Guha, MD, PhD,[2] Manoop S. Bhutani, MD[2]

Houston, Texas, USA

Background: The prognosis of esophageal cancer (EC) depends on the depth of tumor invasion and lymph node metastasis. EC limited to the mucosa (T1a) can be treated effectively with minimally invasive endoscopic therapy, whereas submucosal (T1b) EC carries relatively high risk of lymph node metastasis and requires surgical resection.

Objective: To determine the diagnostic accuracy of EUS in differentiating T1a EC from T1b EC.

Design: We performed a comprehensive search of MEDLINE, SCOPUS, Cochrane, and CINAHL Plus databases to identify studies in which results of EUS-based staging of EC were compared with the results of histopathology of EMR or surgically resected esophageal lesions. DerSimonian-Laird random-effects model was used to estimate the pooled sensitivity, specificity, and likelihood ratio, and a summary receiver operating characteristic (SROC) curve was created.

Setting: Meta-analysis of 19 international studies.

Patients: Total of 1019 patients with superficial EC (SEC).

Interventions: EUS and EMR or surgical resection of SEC.

Main Outcome Measurements: Sensitivity and specificity of EUS in accurately staging SEC.

Results: The pooled sensitivity, specificity, and positive and negative likelihood ratio of EUS for T1a staging were 0.85 (95% CI, 0.82-0.88), 0.87 (95% CI, 0.84-0.90), 6.62 (95% CI, 3.61-12.12), and 0.20 (95% CI, 0.14-0.30), respectively. For T1b staging, these results were 0.86 (95% CI, 0.82-0.89), 0.86 (95% CI, 0.83-0.89), 5.13 (95% CI, 3.36-7.82), and 0.17 (95% CI, 0.09-0.30), respectively. The area under the curve was at least 0.93 for both mucosal and submucosal lesions.

Limitations: Heterogeneity was present among the studies.

Conclusion: Overall EUS has good accuracy (area under the curve $\geq$0.93) in staging SECs. Heterogeneity among the included studies suggests that multiple factors including the location and type of lesion, method and frequency of EUS probe, and the experience of the endosonographer can affect the diagnostic accuracy of EUS. (Gastrointest Endosc 2012;75:242-53.)

Abbreviations: AUC, area under the curve; CI, confidence interval; DOR, diagnostic odds ratio; EC, esophageal cancer; ESD, endoscopic submucosal dissection; NLR, negative likelihood ratio; PLR, positive likelihood ratio; SCC, squamous cell carcinoma; SEC, superficial esophageal cancer; SROC, summary receiver operating characteristic.

DISCLOSURE: *The authors disclosed no financial relationships relevant to this publication.*

0016-5107/$36.00
doi:10.1016/j.gie.2011.09.016

Received June 30, 2010. Accepted September 6, 2010.

Current affiliations: Department of Gastroenterology, Hepatology and Nutrition (1), The University of Texas Medical School, Houston, Texas, Departments of Gastroenterology, Hepatology, and Nutrition (2), GI Medical Oncology (3), and Thoracic and Cardiovascular Surgery (4), The University of Texas MD Anderson Cancer Center, Houston, Texas, School of Public Health (5), The University of Texas, Houston, Texas, USA.

Reprint requests: Manoop S. Bhutani, MD, Department of Gastroenterology, Hepatology, and Nutrition, Unit 1466, 1515 Holcombe Boulevard, The University of Texas MD Anderson Cancer Center, Houston, TX 77030.

If you would like to chat with an author of this article, you may contact Dr Bhutani at Manoop. Bhutani@mdanderson.org.

Esophageal cancer (EC) is a devastating disease with a significant impact on patients' lives and health-care systems worldwide. From 1973 to 2002, incidence of esophageal adenocarcinoma has increased fourfold.[1] EC with infiltration limited to the mucosa or submucosa is defined as superficial EC (SEC). With a widespread increase in the surveillance of patients with Barrett's esophagus and improvement in the endoscopic technologies in the form of narrow-band imaging, chromoendoscopy, and high-magnification and confocal endoscopy, SEC is being diagnosed with increasing frequency. Prognosis of EC depends on the depth of tumor invasion and lymph node metastasis.[2] There is a strong relationship between the depth of tumor invasion and lymph node metastasis in SEC. SEC involving only mucosa (T1a) has less than a 5% to 9% chance of metastasis compared with a 19% to 44% chance of lymph node metastasis with SEC invading the submucosa (T1b).[3-5]

Preoperative correct staging is extremely important because T1a SEC can be effectively treated with minimally invasive endoscopic treatment as opposed to T1b SEC that is treated as advanced EC.[6,7] Because the noninvasive imaging modalities like CT and magnetic resonance imaging lack the ability to clearly differentiate the layers of the esophageal wall, EUS has become the investigation of choice for staging ECs. Recently, the accuracy of the EUS has been questioned by several authors, and EMR has been suggested as the investigation of choice as a first step in the diagnosis and staging of the SEC.[8-10] The aim of the current study was to evaluate the accuracy of EUS in differentiating mucosal and submucosal SEC.

Take-home Message

- The results of this meta-analysis support the use of EUS for staging of suspected superficial esophageal cancer, which can further direct the therapeutic management of these patients. However, clinicians must be aware of the performance limitation of this tool based on different cancer types and the location of the cancer.

METHODS

Search strategy

The systemic review was performed by using the developed guidelines for conducting a systematic review.[11] We searched MEDLINE (PubMed and Ovid from 1980 to June 2010), SCOPUS (Consisting of MEDLINE and Embase databases), Cochrane Database of Systemic Reviews, Google scholar, and CINAHL Plus databases. A systemic literature search was performed by using several search terms: (A) *esophageal cancer*, *EUS*; (B) *esophageal cancer*, *EUS*, *T1*; (C) *EUS*, *superficial esophageal cancer*; (D) *esophagus*, *cancer*, *ultrasound*, *probe*. We also screened the reference list of all of the selected articles for any potential related articles that were not identified by the initial search. Our search was restricted to human subjects. The studies in Japanese language were translated into English by an independent translator. Two reviewers (N.T. and H.S.) independently screened the titles and abstracts of all of the articles according to predefined inclusion and exclusion criteria. We resolved differences by discussion with 2 reviewers (M.S.B and S.G.). The final complete report of all selected articles was then retrieved and reviewed by the same 2 reviewers (N.T. and H.S.).

Inclusion criteria

The study population consisted of patients with esophageal lesions suspicious for EC or confirmed EC based on endoscopic biopsy and imaging studies like EUS, CT scan, and MRI. The intervention was EUS for EC.

Acceptable criterion standards included final pathologic staging per histologic evaluation of EMR or surgically resected specimen.

The acceptable study designs were either retrospective or prospective studies in which staging results of individual esophageal lesions by EUS were compared with the results of a criterion standard as defined above.

The outcome consisted of the reporting of results in sufficient detail to allow reconstruction of a diagnostic 2×2 table (true positive, false positive, true negative, and false negative).

Exclusion criteria

We excluded case reports and case series, studies that did not provide sufficient data for reconstruction of a diagnostic 2×2 table, and studies that included m3-sm1 staging because we were not able to separate them in the mucosal or submucosal category. We also excluded studies in which the total number patients or lesions was fewer than 15 that had both EUS and EMR/surgery available for comparison. All SECs in which lymph node suspicious for metastasis (>1 cm in size) detected during EUS were also excluded from the study.

Data abstraction

From the selected studies, 2 independent reviewers (N.T. and H.S) extracted the following data and placed them on standardized data forms (in Microsoft Excel, Microsoft Corporation, Redmond, Wash):

- Study characteristics: design, country, year of publication, setting, sample size, clinical context, and criterion standard
- Demographic characteristics: mean age, proportion of male and female patients
- Interventions: manufacture and operating frequencies of endoscope and/or probe, number of EUS procedures done, number of EMRs done, number of surgeries done
- Outcomes: number of true-positive, true-negative, false-positive, and false-negative values for T1a and T1b staging

Quality criteria

Current quality assessment guidelines focus on randomization, selection bias of the arms in the study, concealment of allocation, and blinding of outcome to evaluate the quality of the clinical trials with the control arm.[12] There is no consensus or criteria to evaluate the quality of the studies without a control arm.[12] Almost all of the studies focusing on the accuracy of EUS in differentiating mucosal from submucosal esophageal lesions were either retrospective or prospective studies without a control arm. Therefore, for this systematic review and meta-analysis, we selected studies based on our predefined inclusion and exclusion criteria and completeness of data reporting in the studies.

Statistical Analysis

Based on comparison of EUS diagnosis with final histopathological diagnosis by EMR or surgical resection, we constructed 2 × 2 statistical tables for both mucosal and submucosal lesions for each study. Where 0 counts occurred in at least 1 cell of study data, a continuity correction of 0.5 was added to every value for that study to make the calculation of sensitivity and specificity. Based on the 2 × 2 statistical tables, we calculated the true-positive, false-positive, true-negative, and false-negative values for T1a and T1b staging by EUS. Meta-Disc version 1.4 statistical software (Meta-Disc, Unit of Clinical Biostatistics Team of the Roman y Cajal Hospital, Madrid, Spain) was used to calculate the sensitivity, specificity, positive likelihood ratio (PLR), negative LR (NLR), diagnostic accuracy, and diagnostic odds ratio (DOR) (PLR/NLR) for T1a and T1b staging for each study.[13] We used the DerSimonian-Laird random-effects model to pool together final sensitivity, specificity, PLR, NLR, and DOR.[14] Forest plots were drawn to show the point estimates in each study in relation to the summary pooled estimates. A summary receiver operating characteristic curve (SROC) was constructed based on the Moses-Shapiro-Littenberg method.[15] An SROC is similar in principle to a standard receiver operating characteristic curve for a single study, except that the data points for the SROC curve are obtained from the values of sensitivity and specificity in the individual studies in the meta-analysis. As with the area under the curve (AUC) of a receiver operating characteristic curve, the AUC of an SROC is a measure of the overall performance of a diagnostic test to accurately differentiate those with and those without the condition of interest.[15] A preferred test has an AUC close to 1, and a poor test has an AUC close to 0.5.[16] Q* index was calculated as per Moses-Shapiro-Littenberg method.[15] Q* index is defined by the point where sensitivity and specificity are equal, which is the point closest to the ideal top-left corner of the SROC space.[15] The Fisher exact test was used to test the significance between the underdiagnosis and overdiagnosis rates of EUS for squamous cell carcinoma (SCC) and adenocarcinoma. Heterogeneity was assessed by using χ^2 statistics, I2 measure of inconsistency, and Cochran's Q test.[17-19] The χ^2 test, with a df = number of studies − 1, was used to assess the observed differences in study results that were compatible with chance alone or not. A P value <.05 (or a large χ^2 statistics relative to the df) was considered evidence of heterogeneity rather than chance. I^2 index describes the percentage of total variations across studies that are attributed to heterogeneity rather than chance. Generally, an I2 index of 25%, 50%, and 75% represents low, moderate, and high heterogeneity, respectively.[17] The homogeneity of the likelihood ratio and DOR were tested by using Cochran's Q test based on inverse variance weights.[18] We did 7 subgroup analyses: (1) studies from Japan, (2) studies outside Japan, (3) studies focusing on SCC, (4) studies focusing on adenocarcinoma, (5) studies using only radial scanning EUS, (6) studies using a high-frequency probe (≥15 MHz), and (7) studies from Japan using a high-frequency miniprobe to further explore the heterogeneity. We also performed a meta-regression analysis by adding covariates to the SROC model per the Moses-Shapiro-Littenberg method and calculated the relative DOR of the corresponding covariables.[15]

The robustness of the meta-analysis to the publication bias was assessed by various bias indicators, including the Egger and Fail-safe N tests, and the trim-and-fill method.[20,21] Funnel plots were constructed to evaluate the publication bias by using the standard error and DOR.[22,23] Analysis was done by using the comprehensive Meta-analysis version 2.0 (Biostat, Englewood, NJ). For all statistical methods used in the meta-analysis, $P <$ .05 was regarded as significant. There were no corrections made to P values even though there were multiple testings of meta-analytic outcome data arising from individual studies because the purpose of the research was to estimate key meta-analytic statistics and to highlight any potential differences. Also, some statistical tests were performed on data aggregated from individual studies, not by using meta-analytic methods, and these findings should also be taken as descriptive.

RESULTS

Literature search

A total of 1057 titles and/or abstracts were initially identified by using the search strategy. Of them, 951 were excluded by 2 independent reviewers (N.T. and H.S.) after preliminary review of titles and abstracts, which left 106 articles for detailed evaluation. Of these, 87 articles/studies failed to meet the predefined inclusion criteria: 38 focused on T1 to T4 staging and lymph node metastasis, 19 were case reports and case series, 12 were review articles, 8 articles had insufficient details to construct 2 × 2 tables, 6 studies compared m3-sm1, 1 study was in vitro study, and in 3 studies the total number of patients was fewer than 15. In total, 19 studies were selected for this meta-analysis.[24-42]

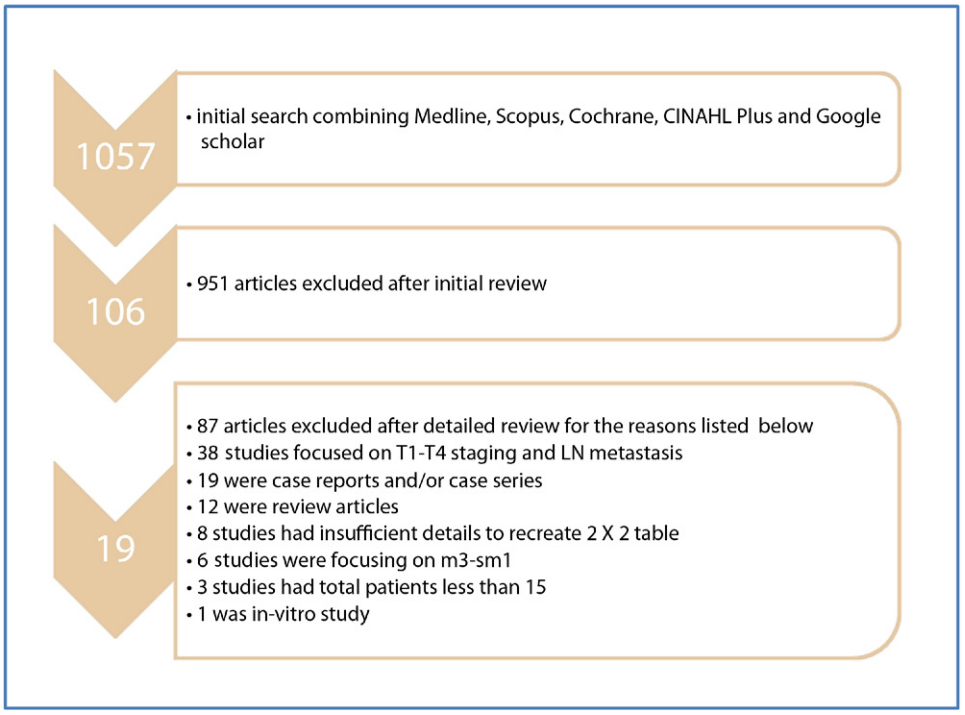

Figure 1. Flow diagram of the study selection process.

The study selection process is shown in Figure 1. There were 12 studies from Japan,[24-32,34,35,37], 2 studies from the United States,[33,39] 2 studies from Germany,[38,40] and 1 study each from Italy,[41] France,[42] and the Netherlands.[36] There were 12 prospective studies[24,28-33,35,38-41] and 7 retrospective studies.[25-27,34,36,37,42] In total, 1019 patients had diagnostic EUS and also had a final diagnosis in terms of mucosal or submucosal invasion by histopathology of either an EMR or surgical specimen. The study characteristics of the included studies are shown in Table 1.

EUS method

Diagnostic EUS was performed in most of the studies with a radial scanning echoendoscope, including GF-UM130, GF-UM2O GF-UM2, GF-UM3, GFUM200, and GIF-2T100 (Olympus America, Inc, Melville, NY). (Probes were used in some studies, including Sp101 [linear], SP-501 [linear and radial], SP-501 [Fuji Photo Optical Co, Ltd, Omiya, Japan]; US probes MP-PN15–08M, UM-3R-2 [Olympus America]; Sp701 [Fuji Photo Optical Co, Ltd]; and mini-probes UM-BS20-26R, UM-S30-20R, UM-3R [Olympus].)

Meta-analysis

Diagnostic accuracy. Figure 2 shows the Forest plots of sensitivity, specificity, PLR, and NLR of EUS for T1a staging. Similarly, Figure 3 shows the Forest plots of sensitivity, specificity, PLR, and NLR of EUS for T1b staging. Point estimates were plotted with 95% confidence intervals (CIs) for each cohort. The pooled sensitivity and specificity of EUS for T1a staging were 0.85 (95% CI, 0.82-0.88) and 0.87 (95% CI, 0.84-0.90), respectively. Similarly, the PLR and NLR were 6.62 (95% CI, 3.61-12.12) and 0.20 (95% CI, 0.14-0.30), respectively. For T1b staging, EUS had a sensitivity of 0.86 (95% CI, 0.82-0.89), a specificity of 0.86 (95% CI, 0.83-0.89), a PLR of 5.13 (95% CI, 3.36-7.82), and an NLR of 0.17 (95% CI, 0.09-0.30). The P value for χ^2 heterogeneity for all pooled estimates was $<.05$.

Results for the subgroup analysis for (1) Japanese studies, (2) studies outside Japan, (3) studies focusing of SCC, (4) studies focusing on adenocarcinoma, (5) studies using only radial scanning EUS, (6) studies using a high-frequency mini-probe, and (7) studies from Japan using a high-frequency mini-probe are shown in detail in Table 2. For both mucosal and submucosal staging, the studies from Japan had a DOR almost 7.3 times higher than that of the studies outside Japan. Also the CIs for DORs for both subgroups did not overlap, suggesting significant difference. We also performed meta-regression analysis to simultaneously evaluate multiple covariates in the same analysis. The outcomes of the regression analysis as the relative DOR are shown in Table 3. In summary, the country of the study, Japan versus studies outside Japan, was associated with higher relative DOR. The disease type and the EUS method were not statistically significant in the regression model. We also pooled the results of studies focusing on SCC and adenocarcinoma to evaluate for understaging and overstaging rates of EUS for SCC and adenocarcinoma. Overall, EUS accurately staged 143 of 170 lesions (84%) with adenocarcinoma and 75 of 93 lesions

TABLE 1. Characteristics of included studies

Index	Authors/y/country	Study type	EUS frequencies, MHz	EUS method	Disease type	Sample size	Confirmatory study
1	Murata et al/1988/Japan	Retrospective	7.5, 10	Radial	NS	52	Surgery
2	Kouzu et al/1992/Japan	Retrospective	7.5, 20	Radial and mini-probe	NS	101	Surgery
3	Toh et al/1993/Japan	Retrospective	7.5, 12	Radial	NS	26	Surgery
4	Yoshikane et al/1994/Japan	Prospective	7.5, 12	Radial	NS	26	Surgery and EMR
5	Simizu et al/1995/Japan	Prospective	7.5, 12, 20	Radial and/or mini-probe	NS	40	Surgery
6	Murata et al/1996/Japan	Prospective	15, 20	Mini-probe	NS	49	Surgery and EMR
7	Yanai et al/1996/Japan	Prospective	20	Mini probe	SCC	17	Surgery and EMR
8	Shinkai et al/2000/Japan	Prospective	7.5, 12, 15, 20	Radial and/or mini-probe	SCC	50	Surgery and EMR
9	Fukuda et al/2000/Japan	Prospective	12, 20	Mini-probe	NS	25	Surgery
10	Scotiniotis et al/2001/USA	Prospective	7.5-12	Radial	AC	22	Surgery
11	Kawano et al/2003/Japan	Retrospective	20	Mini-probe	NS	96	Surgery and EMR
12	Yanai et al/2003/Japan	Prospective	12, 20	Mini-probe	SCC	26	Surgery and EMR
13	Buskens et al/2004/the Netherlands	Retrospective	7.5, 12, 20, 30	Radial and/or mini-probe	AC	45	Surgery
14	Arima et al/2004/Japan	Retrospective	20	Mini-probe	NS	91	Surgery
15	May et al/2004/Germany	Prospective	20	Mini-probe	SCC and AC	93	Surgery and EMR
16	Larghi et al/2005/USA	Prospective	7.5, 20		AC	48	Surgery and EMR
17	Pech et al/2006/Germany	Prospective	7.5, 12.5, 20	Radial and mini-probe	AC	55	Surgery and EMR
18	Rampado et al/2008/Italy	Prospective	20	Mini-probe	SCC and AC	55	Surgery and EMR
19	Chemaly et al/2008/France	Retrospective	20, 30	Mini-probe	SCC and AC	102	Surgery and EMR

NS, Not specified; *SCC*, squamous cell carcinoma; *AC*, adenocarcinoma.

(81%) with SCC. For adenocarcinoma, EUS understaged 12 of 170 (7%) and overstaged 15 of 170 lesions (9%) while evaluating for submucosal invasion. For SCC, EUS understaged 14 of 93 (15%) and overstaged 4 of 93 lesions (4%) while evaluating for submucosal invasion. When comparing adenocarcinoma and SCC by using the Fisher exact test, the 2-sided *P* values for understaging and overstaging rates were 0.038 and 0.175, respectively. The overall accuracy of EUS was further explored by drawing SROC curves and finding the AUC (Fig. 4). For T1a and T1b staging, EUS had an AUC of 0.93 for all 19 studies and an AUC of 0.97 for all Japanese studies. For SCC and Barrett's esophagus–related early adenocarcinoma for T1a and T1b staging, the AUC values were 0.96 and 0.94, respectively.

The funnel plots for publication bias are shown in Figure 5. The Egger test for publication bias was statistically significant for both T1a ($P = .001$) and T1b ($P = .001$). The fail-safe N test indicated that for the combined 2-tailed *P* value to no longer be significant ($P > .05$), it would take an additional 1113 studies for T1a and 1122 studies for T1b with no significant findings. By using the random-effects model, the DORs and 95% CI for the combined studies for T1a and T1b were 40.64 (95% CI, 18.55-89.04) and 39.62 (95% CI, 18.38-85.42), respectively. After adjusting for the publication bias with the trim-and-fill method, the imputed DORs for T1a and T1b were 13.49 (95% CI, 5.85-31.09) and 13.46 (95% CI, 5.93-30.58), respectively.

DISCUSSION

In EC, the survival is greatly dependent on the stage at diagnosis, with a 5-year survival rate of greater than 90% with stage 1 disease, but only around 10% for stage 3 disease.[43] Surgery with curative intent is considered the criterion standard treatment for resectable EC.[44] The current European Society of Medical Oncology guideline recommends surgery as the treatment of choice in

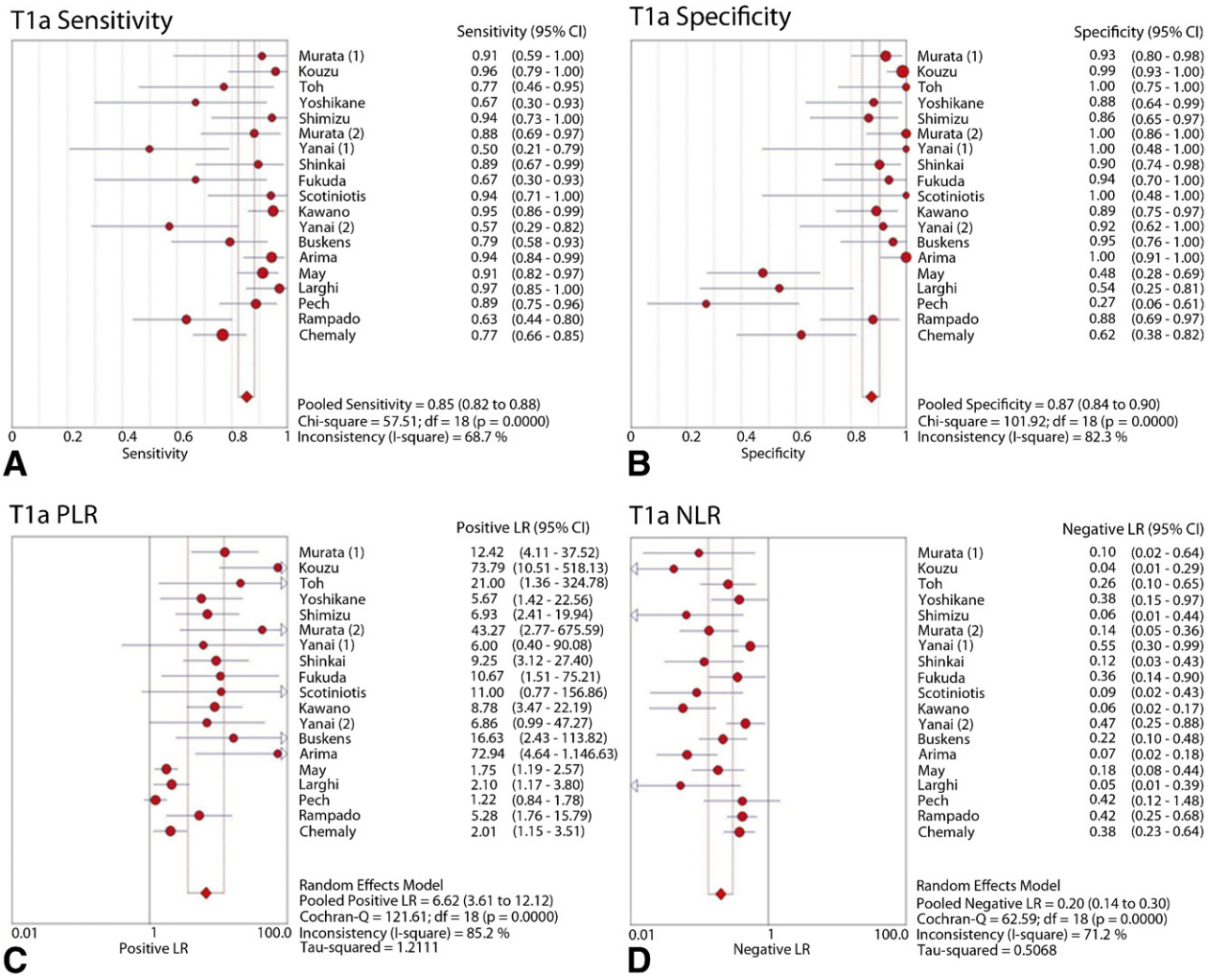

Figure 2. Pooled sensitivity (**A**), specificity (**B**), positive likelihood ratio (PLR) (**C**), and negative likelihood ratio (NLR) (**D**) for T1a staging. The size of the each *circle* is proportional to the sample size for each study, and the *horizontal lines through the circles* indicate the 95% confidence interval (CI) for that study. For the pooled analysis, the *diamond* indicates the pooled value and the right and left ends of the *vertical dashed bar* indicate the 95% CI for the analysis.

early EC (T1a).[45] Esophageal surgery has a 30-day mortality rate between 3% and 13% and a morbidity rate as high as 40% to 50%.[44,46,47] The mortality rate after esophagectomy increases to more than 20% in low-volume centers (<5 esophagectomies per year) compared with national cancer institution hospitals and high-volume centers.[48-50] Secondary to high surgical morbidity and mortality, in past 2 decades, minimally invasive endoscopic therapy such as photodynamic therapy, argon plasma coagulation, EMR, and endoscopic submucosal dissection (ESD) has been investigated as treatment options for SEC limited to the mucosa. Compared with photodynamic therapy and argon plasma coagulation, EMR and ESD allow histological assessment of resected specimens to assess for depth of tumor invasion, tumor-free margins, lymphatic and venous invasion, and grade of differentiation. EMR has a significantly low morbidity rate (1%-3%) and mortality rate (0%) and preserves the organ and quality of life.[6,7] Preoperative differentiation of mucosal (T1a) and submucosal (T1b) invasion remains the most important question for the gastroenterologist, oncologist, and thoracic surgeon in deciding the best treatment option for patients with SEC.

Over the past several decades, staging accuracy of EUS has been studied in depth for EC. Compared with CT, which is the most commonly used imaging technique for staging most cancers, EUS has very high sensitivity and specificity for EC and its locoregional metastasis.[51,52] In past 2 decades, there have been many studies focusing on the accuracy of EUS in staging of SEC. The accuracy of EUS varied from as low as 33% to as high as 85% in staging submucosal EC.[24-42] Few studies reported that EUS underdiagnosed submucosal invasion in 12.5% to 67% of cases.[24-42] We conducted this systematic review and meta-

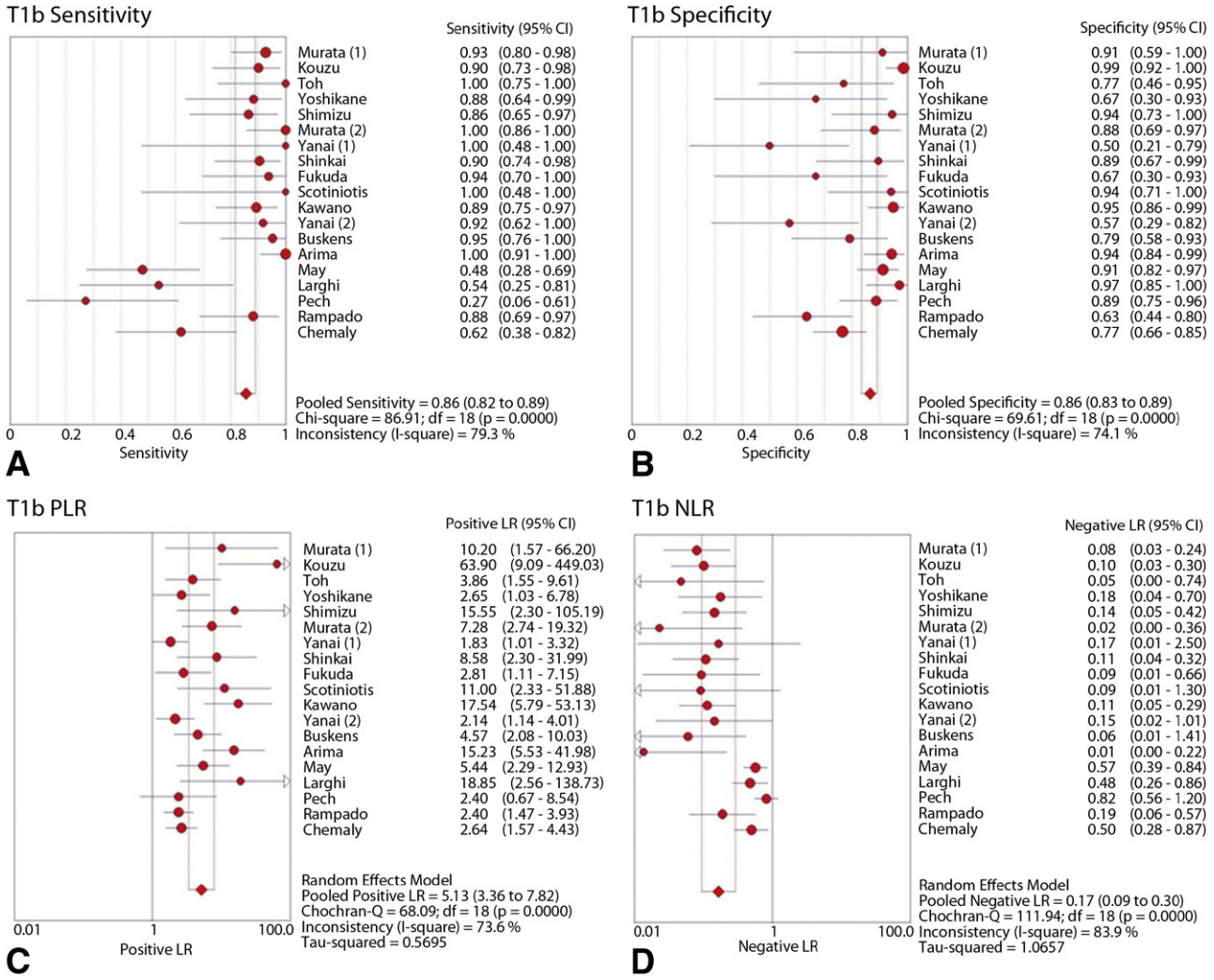

Figure 3. Pooled sensitivity (**A**), specificity (**B**), positive likelihood ratio (PLR) (**C**), and negative likelihood ratio (NLR) (**D**) for T1b staging. The size of the each *circle* is proportional to the sample size for each study, and the *horizontal lines through the circles* indicate the 95% confidence interval (CI) for that study. For the pooled analysis, the *diamond* indicates the pooled value, and the right and left ends of the *vertical dashed bar* indicate the 95% CI for the analysis.

analysis with the intent to evaluate the accuracy of EUS in staging SEC and to find factors that might explain heterogeneity among different studies.

Our meta-analysis shows that EUS has a pooled sensitivity of 0.85 (95% CI, 0.82-0.88) for T1a (mucosal) staging and 0.86 (95% CI, 0.82-0.89) for T1b (submucosal) staging. Similarly for both stages, EUS has a specificity of approximately 86% to 87%. The DOR is defined as the odds of having a positive test result in a patient with true disease compared with a patient who does not have the disease. EUS has a high DOR for mucosal (DOR 40) and submucosal (DOR 39) staging. The PLR is a measure of how well the test identified the disease, and the NLR assesses how well the same test performs in excluding the disease. Likelihood ratios greater than 10 and less than 0.1 provide strong evidence to rule a diagnosis in or rule out, respectively.[53] For EUS, the PLR and NLR were 6.62 (95% CI, 3.61-12.12) and 0.20 (95% CI, 0.14-0.30) for T1a staging and 5.13 (95% CI, 3.36-7.82), and 0.17 (95% CI, 0.09-0.30) for T1b staging, respectively.

SROC curves were constructed by using the Moses-Shapiro-Littenberg method.[15] The symmetrical curve shows a trade-off between sensitivity and specificity. An AUC of 1 for any test indicates that the test is excellent. EUS has AUC value of 0.93 for both T1a (mucosal) and T1b (submucosal) staging. We also performed subgroup analysis to further explore heterogeneity. We found that a subgroup of Japanese studies had higher a sensitivity, specificity, PLR, and DOR and lower NLR compared with a meta-analysis of all studies and a subgroup of studies performed outside Japan. The first reported study from Japan focusing on the staging of SEC was published in 1988, and the first study outside Japan was published in 2001. Both the subgroup analysis and the meta-regression analysis revealed a statistically significant difference in accuracy of EUS between studies done in Japan and those done out-

TABLE 2. Subgroup analysis to determine the source of heterogeneity

	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)	AUC
Mucosal invasion						
All studies (N = 19)	0.85 (0.82-0.88)	0.87 (0.84-0.90)	6.62 (3.61-12.12)	0.20 (0.14-0.30)	40.65 (18.55-89.05)	0.93
Japanese studies	0.87 (0.82-0.91)	0.95 (0.92-0.97)	10.48 (6.86-16.01)	0.17 (0.09-0.33)	86.97 (37.69-200.68)	0.97
Studies outside Japan	0.84 (0.79-0.88)	0.68 (0.59-0.76)	2.31 (1.41-3.78)	0.27 (0.18-0.42)	11.74 (5.17-26.69)	0.86
AC	0.90 (0.83-0.95)	0.70 (0.55-0.82)	3.12 (0.98-9.94)	0.18 (0.09-0.39)	25.39 (3.88-165.95)	0.94
SCC	0.69 (0.53-0.82)	0.92 (0.80-0.98)	8.27 (3.38-20.22)	0.36 (0.15-0.88)	31.78 (8.48-119.06)	0.96
Radial EUS only	0.84 (0.71-0.93)	0.93 (0.85-0.98)	9.95 (4.53-21.85)	0.23 (0.12-0.43)	50.18 (14.06-179.06)	0.96
High-frequency mini-probe	0.83 (0.79-0.87)	0.84 (0.78-0.89)	5.63 (2.27-13.96)	0.23 (0.12-0.42)	29.57 (8.83-99.05)	0.91
Japanese studies with high-frequency mini-probe	0.87 (0.81-0.91)	0.96 (0.90-0.99)	11.15 (4.93-25.19)	0.18 (0.06-0.55)	94.39 (19.67-452.92)	0.97
Submucosal invasion						
All studies (N = 19)	0.86 (0.82-0.89)	0.86 (0.83-0.89)	5.13 (3.36-7.82)	0.17 (0.09-0.30)	39.63 (18.38-85.42)	0.93
Japanese studies	0.93 (0.89-0.96)	0.89 (0.85-0.92)	6.17 (3.18-11.97)	0.10 (0.07-0.16)	84.76 (38.85-184.94)	0.97
Studies outside Japan	0.68 (0.59-0.76)	0.84 (0.79-0.88)	3.70 (2.41-5.68)	0.43 (0.26-0.71)	11.75 (5.17-26.69)	0.86
AC	0.70 (0.55-0.82)	0.90 (0.83-0.95)	5.43 (2.58-11.43)	0.32 (0.10-1.03)	25.39 (3.88-165.95)	0.94
SCC	0.92 (0.80-0.98)	0.69 (0.53-0.82)	2.78 (1.14-6.79)	0.12 (0.05-0.30)	31.79 (8.48-119.06)	0.96
Radial EUS only	0.93 (0.85-0.98)	0.84 (0.71-0.93)	4.42 (2.31-8.43)	0.10 (0.05-0.22)	50.18 (14.06-179.06)	0.96
High-frequency mini-probe	0.84 (0.78-0.89)	0.83 (0.79-0.87)	4.33 (2.36-7.95)	0.18 (0.07-0.44)	29.57 (8.83-99.04)	0.91
Japanese studies with high-frequency mini-probe	0.96 (0.90-0.97)	0.87 (0.81-0.92)	5.67 (1.83-17.62)	0.09 (0.04-0.20)	94.39 (19.67-452.92)	0.97

CI, Confidence interval; *PLR*, positive likelihood ratio; *NLR*, negative likelihood ratio; *DOR*, diagnostic odds ratio; *AUC*, area under the curve; *AC*, adenocarcinoma; *SCC*, squamous cell carcinoma.

TABLE 3. Metaregression analysis

Covariate	Coefficient	*P* value	RDOR	95% CI
EUS method (radial vs high-frequency mini-probe)	0.238	.6568	1.27	0.04-3.91
Disease type (SCC vs AC)	−0.211	.669	0.81	0.29-2.28
Country (Japan vs studies outside Japan)	−1.945	.0268	0.14	0.08-0.77

RDOR, Relative diagnostic ratio; *CI*, confidence interval; *SCC*, squamous cell carcinoma; *AC*, adenocarcinoma.

side Japan. These conflicting findings may be related to the overall volume of the patients, operator experience, and type of the lesions. However, none of the studies in this meta-analysis reported operator experience, so we were not able to perform any further statistical analysis. Because of a significant learning curve, accuracy of EUS directly correlates with the experience of the endosonographer. Rice et al[54] reported that their accuracy in T staging was only 59% in their first 28 patients, and it subsequently improved to 81% in their next 52 patients. A retrospective study by Van Vliet et al[55] showed that a low-volume center had a lower sensitivity (58% vs 75%-90%) and specificity (87% vs 94%-97%) for T1 and T2 stages than several high-volume centers.

We also considered that other factors such as the frequency of EUS probe use, EUS technology (radial vs linear or both), the location of a lesion within the esophagus, cancer type (adenocarcinoma vs SCC),and secondary information from other imaging modalities such as magnetic resonance imaging and CT before EUS might be contributing to heterogeneity among studies. We found that EUS had an overall low sensitivity of 0.69 (95% CI, 0.53-0.82)

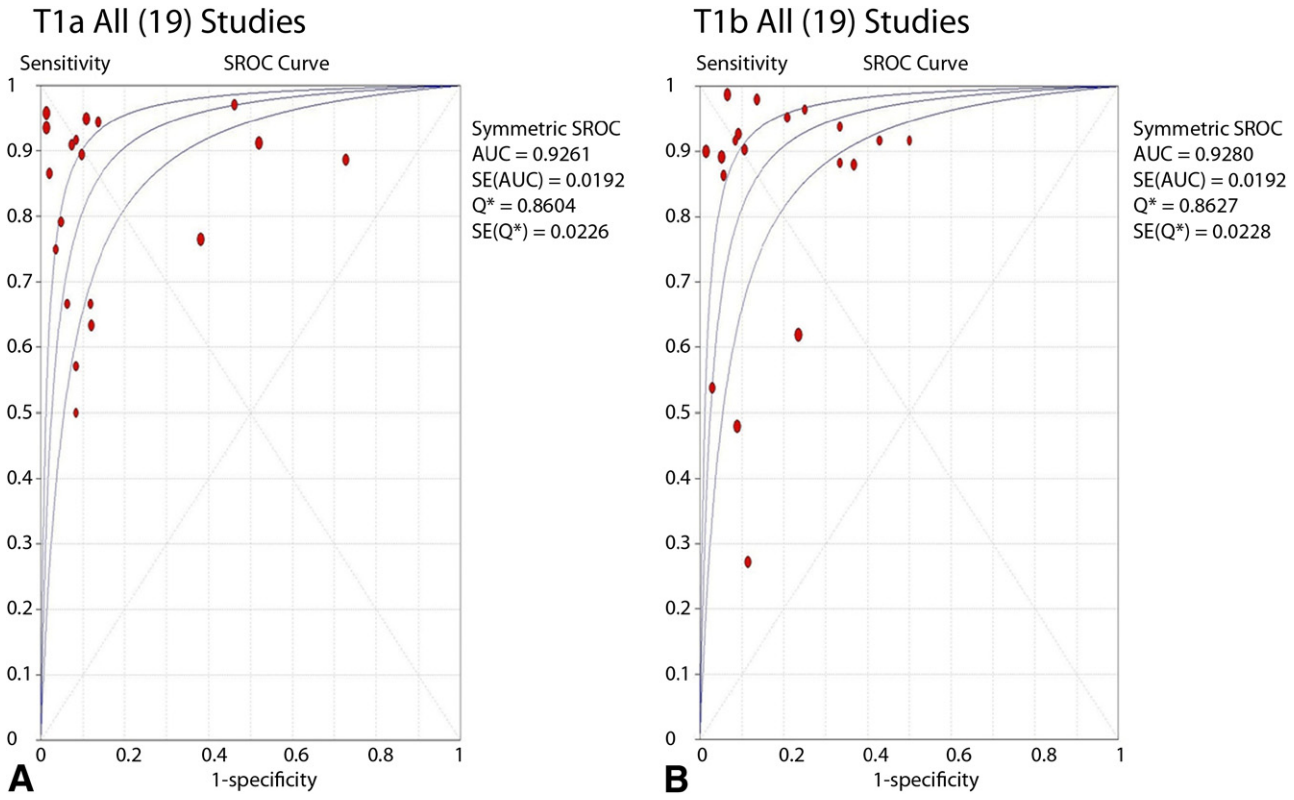

Figure 4. Summary receiver operating characteristics (SROC) curve for T1a and T1b staging. AUC, area under the curve; SE, standard error.

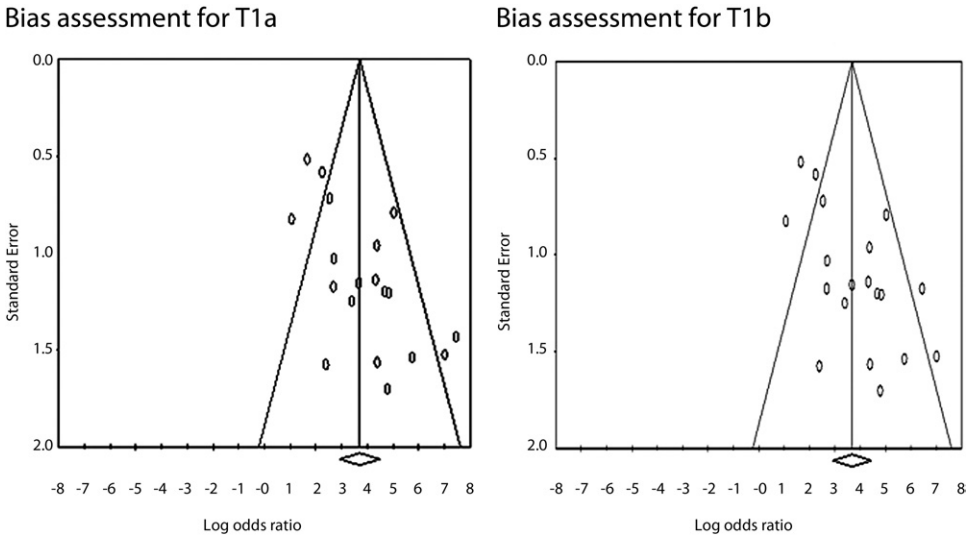

Figure 5. Funnel plots for bias assessment of T1a and T1b staging.

for T1a staging in cases of SCC; however, the AUC was higher (0.96) because of a relatively high specificity of 0.92 (95% CI, 0.80-0.98). However, only 3 studies focused solely on SCC, and the total number of patients was only 94 for this analysis. EUS with older echoendoscopes at frequencies of 7.5 or 12 MHz depicted the esophageal wall in 5 layers. EUS at higher frequencies (20 or 30 MHz) depicts the esophageal wall in 9 echo layers, allowing for more precise T staging. With high-frequency probes, the mucosa is visualized as 4 layers comprising the epithelium (m1 and m2), the lamina propria (m3), and the muscularis mucosa (m4). By using 15- to 20-MHz high-frequency probes, Murata et al[24] found an overall accuracy of 75% for T staging for early EC.[24] However, they reported superior accuracy in predicting cancers limited by the lamina propria (84%) and differentiating cancers limited to the mu-

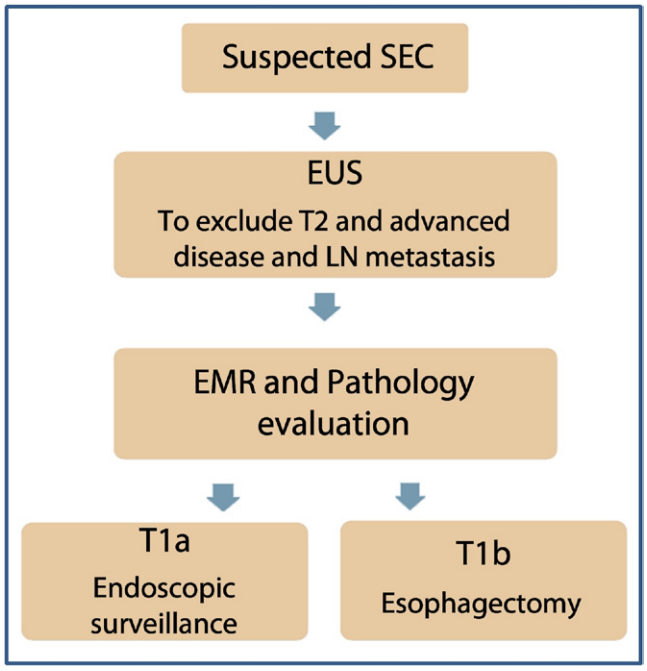

Figure 6. Proposed diagnostic algorithm for suspected superficial esophageal cancer (SEC). LN, lymph node.

cosa from those extending into the submucosa (94%).[24] Hasegawa et al[56] showed 86% and 94% accuracy rates for T1m and T1sm cancers, respectively, with an overall accuracy rate of 92% for all T lesions by using a 15-MHz high-frequency probe. In our analysis, we did not find a significant difference in overall accuracy between older echoendoscopes and high-frequency mini-probes. However, these results were affected based on the country. The subgroup of the studies using radial scanning EUS only had included 4 studies and all studies were from Japan only, whereas the subgroup of the studies using high-frequency EUS included studies from Japan and studies conducted outside Japan. When we performed a further subgroup analysis of studies from Japan only, using a high-frequency mini-probe, we found a higher sensitivity, specificity, PLR, and DOR and lower NLR compared with the subgroup of studies in which radial EUS only was used (Table 2).

A recent meta-analysis by Young et al[9] showed that EUS had a T staging accuracy of only 65% for early adenocarcinoma or high-grade dysplasia in the setting of Barrett's esophagus. When they analyzed individual patient-level data (132 patients from 8 studies), the accuracy of EUS was even lower (56%). They included only 12 studies in their analysis and excluded studies published in languages other than English. In 5 of the 12 included studies, the total number of patients was fewer than 20. Rather than focusing on mucosal versus submucosal invasion, most of the included studies in their meta-analysis looked at the differentiation between high-grade dysplasia, Tis (limited to superficial or deep mucosa but not invading lamina propria), T1a, and T1b. Compared with that, we included 1019 patients from 19 studies including studies published in languages other than English and also studies focusing on both cancer types: SCC and adenocarcinoma. In a subgroup analysis, we found that overall EUS accurately staged 143 of 170 lesions (84%) with adenocarcinoma and 75 of 93 lesions (81%) with SCC. We also found a statistically significant difference between understaging rates for adenocarcinoma (7%) and SCC (15%) while evaluating for submucosal invasion. Chemaly et al[42] found a statistically significant difference in the accuracy rate of EUS depending on the location of lesion in esophagus. When the lesion was located in the proximal and mid–esophagus, EUS accurately staged 61 of 70 (87.1%) of SECs compared with 10 of 21 of SECs (47.6%) when located in the distal esophagus. However, most of the studies did not report the location of lesions, and we were not able to evaluate the accuracy of EUS based on the location of the lesion: the proximal, mid, or distal esophagus or the gastroesophageal junction.

A recent study by Pouw et al[8] suggests that EUS has a limited role in staging SEC, and all patients with suspected SEC should undergo EMR as the first diagnostic, staging, and possible therapeutic option. Shami et al[57] found that almost 20% of patients referred for EMR for suspected SEC already had advanced lymph node metastasis and did not need EMR and diagnostic EUS before EMR dramatically changed the course of management. Most of the primary studies included in the meta-analysis did not report whether patients were excluded because of nodal positivity on EUS. Also, the role of neoadjuvant chemotherapy and chemoradiation is emerging in the treatment of localized resectable EC. A meta-analysis of 10 trials with 1209 patients undergoing neoadjuvant chemoradiation with surgery or surgery alone found that preoperative chemoradiation improved 2-year survival rates, and this improvement held for both SCC and adenocarcinoma.[58] A recent phase 3 trial[59] comparing the trimodality approach of neoadjuvant chemotherapy followed by surgery versus surgery alone found that the trimodality approach improved the median survival time (4.48 years vs 1.79 years) and 5-year overall survival rates (39% vs 16%). Because the therapeutic approach for localized EC is changing, the strategy of diagnostic EUS first will provide valuable information regarding T2 or advanced T stage and nodal metastasis as opposed to diagnostic EMR.

In conclusion, our meta-analysis shows that EUS has overall good sensitivity and specificity in detecting mucosal or submucosal invasion in SEC. Factors including operator experience and volume, EUS technology, cancer type, and location of lesion affect the overall accuracy of EUS. In the hands of experienced operators, EUS does change the diagnostic and therapeutic algorithm for suspected SEC. Further prospective studies are needed to explore the factors affecting the sensitivity and specificity of EUS in SEC staging and to standardize the EUS process

to improve overall accuracy. Our group believes that EUS plays a critical role in detecting T2 lesions that are not amenable to EMR or ESD. Also, at the same time, EUS can diagnose locoregional metastasis including lymph node involvement. In fact, at our institution, treatment of these lesions is individualized and includes either definitive chemoradiation or trimodality therapy including surgical resection. Thus, at present, for suspected early stage EC, we recommend examination with EUS first to rule out infiltration of the muscularis propria (T2) and regional lymph node metastasis (stage II and beyond), followed by EMR with histological examination of the resection specimen as the diagnostic algorithm for evaluation of SEC (Fig. 6).

REFERENCES

1. Holmes RS, Vaughan TL. Epidemiology and pathogenesis of esophageal cancer. Semin Radiat Oncol 2007;17:2-9.
2. Daly JM, Karnell LH, Menck HR. National Cancer Data Base report on esophageal carcinoma. Cancer 1996;78:1820-8.
3. Ide H, Nakamura T, Hayashi K, et al. Esophageal squamous cell carcinoma: pathology and prognosis. World J Surg 1994;18:321-30.
4. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. Gastrointest Endosc 2003;58(6 Suppl), S3-43 (2003).
5. Kodama M, Kakegawa T. Treatment of superficial cancer of the esophagus: a summary of responses to a questionnaire on superficial cancer of the esophagus in Japan. Surgery 1998;123:432-9.
6. Crumley AB, Going JJ, Mcewan K, et al. Endoscopic mucosal resection for gastroesophageal cancer in a U.K. population. Long-term follow-up of a consecutive series. Surg Endosc 2011;25:543-8.
7. Pech O, Behrens A, May A, et al. Long-term results and risk factor analysis for recurrence after curative endoscopic therapy in 349 patients with high-grade intraepithelial neoplasia and mucosal adenocarcinoma in Barrett's oesophagus. Gut 2008;57:1200-6.
8. Pouw RE, Heldoorn N, Herrero LA, et al. Do we still need EUS in the workup of patients with early esophageal neoplasia? A retrospective analysis of 131 cases. Gastrointest Endosc 2011;73:662-8.
9. Young PE, Gentry AB, Acosta RD, et al. Endoscopic ultrasound does not accurately stage early adenocarcinoma or high-grade dysplasia of the esophagus. Clin Gastroenterol Hepatol 2010;8:1037-41.
10. May A, Gossner L, Pech O, et al. Local endoscopic therapy for intraepithelial high-grade neoplasia and early adenocarcinoma in Barrett's oesophagus: acute-phase and intermediate results of a new treatment approach. Eur J Gastroenterol Hepatol 2002;14:1085-91.
11. Irwig L, Macaskill P, Glasziou P, et al. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol 1995;48:119-30; discussion 131-2.
12. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1559-73.
13. Zamora J, Abraira V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.
14. Der Simonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
15. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med 193;12:1293-316.
16. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
17. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
18. Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ 2001;323:157-62.
19. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
20. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
21. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455-63.
22. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 2001;54:1046-55.
23. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. BMJ 2001;323:101-5.
24. Murata Y, Suzuki S, Ohta M, et al. Small ultrasonic probes for determination of the depth of superficial esophageal cancer. Gastrointest Endosc 1996;44:23-8.
25. Murata Y, Suzuki S, Hashimoto H. Endoscopic ultrasonography of the upper gastrointestinal tract. Surg Endosc 1988;2:180-3.
26. Kouzu T, Arima M, Yamada H, et al. Endoscopic treatment of esophageal cancer [in Japanese]. Gan To Kagaku Ryoho 1992;19:1255-60.
27. Toh Y, Baba K, Ikebe M, et al. Endoscopic ultrasonography in the diagnosis of an early esophageal carcinoma. Hepatogastroenterology 1993; 40:212-6.
28. Yoshikane H, Tsukamoto Y, Niwa Y, et al. Superficial esophageal carcinoma: evaluation by endoscopic ultrasonography. Am J Gastroenterol 1994;89:702-7.
29. Simizu Y, Tsukagoshi H, Nakazato T, et al. Clinical evaluation of endoscopic ultrasonography (EUS) in the diagnosis of superficial esophageal carcinoma [in Japanese]. Rinsho Byori 1995;43:221-6.
30. Yanai H, Yoshida T, Harada T, et al. Endoscopic ultrasonography of superficial esophageal cancers using a thin ultrasound probe system equipped with switchable radial and linear scanning modes. Gastrointest Endosc 1996;44:578-82.
31. Shinkai M, Niwa Y, Arisawa T, et al. Evaluation of prognosis of squamous cell carcinoma of the oesophagus by endoscopic ultrasonography. Gut 2000;47: 120-5.
32. Fukuda M, Hirata K, Natori H. Endoscopic ultrasonography of the esophagus. World J Surg 2000;24:216-26.
33. Scotiniotis IA, Kochman ML, Lewis JD, et al. Accuracy of EUS in the evaluation of Barrett's esophagus and high-grade dysplasia or intramucosal carcinoma. Gastrointest Endosc 2001;54:689-96.
34. Kawano T, Ohshima M, Iwai T. Early esophageal carcinoma: endoscopic ultrasonography using the Sonoprobe. Abdom Imaging 2003; 28:477-85.
35. Yanai H, Harada T, Okamoto T, et al. Prognostic value and interobserver agreement of endoscopic ultrasonography for superficial squamous cell carcinoma of the esophagus: a prospective study. Int J Gastrointest Cancer 2003;34:1-8.
36. Buskens CJ, Westerterp M, Lagarde SM, et al. Prediction of appropriateness of local endoscopic treatment for high-grade dysplasia and early adenocarcinoma by EUS and histopathologic features. Gastrointest Endosc 2004;60:703-10.
37. Arima M, Tada M. Endosonographic assessment of the depth of tumor invasion by superficial esophageal cancer, using a high-frequency miniature US probe: difficulties in interpretation and misleading factors. Stomach Intest 2004;39:901-13.
38. May A, Gunter E, Roth F, et al. Accuracy of staging in early oesophageal cancer using high resolution endoscopy and high resolution endosonography: a comparative, prospective, and blinded trial. Gut 2004; 53:634-40.
39. Larghi A, Lightdale CJ, Memeo L, et al. EUS followed by EMR for staging of high-grade dysplasia and early cancer in Barrett;s esophagus. Gastrointest Endosc 2005;62:16-23.
40. Pech O, May A, Gunter E, et al. The impact of endoscopic ultrasound and computed tomography on the TNM staging of early cancer in Barrett's esophagus. Am J Gastroenterol 2006;101:2223-9.

41. Rampado S, Bocus P, Battaglia G, et al. Endoscopic ultrasound: accuracy in staging superficial carcinomas of the esophagus. Ann Thorac Surg 2008;85:251-6.
42. Chemaly M, Scalone O, Durivage G, et al. Miniprobe EUS in the pretherapeutic assessment of early esophageal neoplasia. Endoscopy 2008;40:2-6.
43. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med 2003;349: 2241-52.
44. Reed CE. Surgical management of esophageal carcinoma. Oncologist 1999;4:95-105.
45. Stahl M, Budach W, Meyer HJ, et al. Esophageal cancer: cinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2010; 21(Suppl 5):v46-9.
46. Millikan KW, Silverstein J, Hart V, et al. A 15-year review of esophagectomy for carcinoma of the esophagus and cardia. Arch Surg 1995;130: 617-24.
47. Svanes K, Stangeland L, Viste A, et al. Morbidity, ability to swallow, and survival, after oesophagectomy for cancer of the oesophagus and cardia. Eur J Surg 1995;161:669-75.
48. Swisher SG, Deford L, Merriman KW, et al. Effect of operative volume on morbidity, mortality, and hospital use after esophagectomy for cancer. J Thorac Cardiovasc Surg 2000;119:1126-32.
49. Miller JD, Jain MK, De Gara CJ, et al. Effect of surgical experience on results of esophagectomy for esophageal carcinoma. J Surg Oncol 1997; 65:20-1.
50. Wouters MW, Wijnhoven BP, Karim-Kos HE, et al. High-volume versus low-volume for esophageal resections for cancer: the essential role of case-mix adjustments based on clinical data. Ann Surg Oncol 2008;15: 80-7.
51. Ziegler K, Sanft C, Zeitz M, et al. Evaluation of endosonography in TN staging of oesophageal cancer. Gut 1991;32:16-20.
52. Catalano MF, Sivak MV Jr, Rice T, et al. Endosonographic features predictive of lymph node metastasis. Gastrointest Endosc 1994;40:442-6.
53. Gilbert R, Logan S, Moyer VA, et al. Assessing diagnostic and screening tests: Part 1. Concepts. West J Med 2001;174:405-9.
54. Rice TW, Boyce GA, Sivak MV. Esophageal ultrasound and the preoperative staging of carcinoma of the esophagus. J Thorac Cardiovasc Surg 1991;101:536-43; discussion 543-4.
55. Van Vliet EP, Eijkemans MJ, Poley JW, et al. Staging of esophageal carcinoma in a low-volume EUS center compared with reported results from high-volume centers. Gastrointest Endosc 2006;63:938-47.
56. Hasegawa N, Niwa Y, Arisawa T, et al. Preoperative staging of superficial esophageal carcinoma: comparison of an ultrasound probe and standard endoscopic ultrasonography. Gastrointest Endosc 1996;44: 388-93.
57. Shami VM, Villaverde A, Stearns L, et al. Clinical impact of conventional endosonography and endoscopic ultrasound-guided fine-needle aspiration in the assessment of patients with Barrett's esophagus and high-grade dysplasia or intramucosal carcinoma who have been referred for endoscopic ablation therapy. Endoscopy 2006;38:157-61.
58. Gebski V, Burmeister B, Smithers BM, et al. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. Lancet Oncol 2007;8:226-34.
59. Tepper J, Krasna MJ, Niedzwiecki D, et al. Phase III trial of trimodality therapy with cisplatin, fluorouracil, radiotherapy, and surgery compared with surgery alone for esophageal cancer: CALGB 9781. J Clin Oncol 2008;26:1086-92.