Summary

Initial Staging of Esophageal Cancer: Systematic Review of the Performance of Diagnostic Methods

Introduction

In Québec, it is estimated that 339 people were diagnosed with esophageal cancer and that 319 died from the disease in 2007. In roughly 50% of patients, the disease is at an advanced stage at the time of diagnosis owing to the lack of early symptoms and the rapid locoregional invasion of the tumour. Staging esophageal cancer is a key step to determine the feasibility of curative treatment and, if so, to establish an optimal therapy plan. In the absence of distant metastases, it is important to determine the extent of tumour infiltration into the wall of the esophagus and adjacent structures, along with lymph node involvement, an important prognostic factor. However, certain elements complicate initial staging, especially the fact that neoadjuvant therapy or palliative therapy is administered to many patients because of the advanced stage of the disease at diagnosis, which means that surgically resected samples are not available or are not valid to confirm the histopathological stage. Furthermore, the epidemiology of esophageal cancer has changed in recent years: there has been a noticeable increase in the incidence of adenocarcinoma of the distal esophagus or gastroesophageal junction compared with that of squamous cell carcinoma of the middle third and upper third of the esophagus. This alters esophageal cancer management because the position and mode of spread of these two histological types of esophageal tumours differ considerably. The presence of a stenosing tumour also complicates clinical staging.

Currently in Québec, there is no consensus on the methods to use or strategy to follow for clinically staging esophageal tumours. As a result, the *Comité de l'évolution des pratiques en oncologie* (CEPO), which reports to the *Direction de la lutte contre le cancer*, asked the *Agence d'évaluation des technologies et des modes d'intervention en santé* (AETMIS) to conduct a systematic review of the literature on the performance of the diagnostic methods used for clinically staging esophageal cancer. This systematic review will serve as the basis for the CEPO to develop clinical practice guidelines for Québec. With the help of experts mandated by the CEPO, we defined the following assessment questions:

- What is the diagnostic performance of various techniques – computed tomography (CT), positron emission tomography with CT (PET-CT), endoscopic ultrasound (EUS) with and without ultrasound-guided fine-needle aspiration, magnetic resonance imaging (MRI) and minimally invasive surgical procedures (thoracoscopy and laparoscopy) – according to tumour TNM stage, histological type and location?
- What is the comparative diagnostic performance of these methods?
- What is the optimal staging sequence?
- In the case of endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA), what is the procedure to follow to biopsy lymph nodes adjacent to the tumour?
- What is the procedure to follow in the case of stenosing esophageal carcinomas?
- Which clinically significant diagnostic biomarkers, either immunohistochemical or other (lymph node or esophageal biopsy), could improve the ability to establish a more reliable prognosis?

This report is limited to evaluating diagnostic test performance. It does not address the impact of the different staging methods on patients' clinical outcomes or survival, for example. Nor does it cover organizational and economic issues.

Technologies used for the clinical staging of esophageal cancer

Conventional technologies used for clinically staging esophageal cancer include computed tomography (CT) and endoscopic ultrasound (EUS). Computed tomography generates anatomical images of the body's internal structures through computerized X-rays and is mainly used for evaluating distant metastases (Stage M). There are three types of CT scanners: incremental, spiral, and multi-detector (multi-slice) spiral. The last type offers the best spatial resolution and allows for multi-planar image reconstruction during a single breath-hold.

Endoscopic ultrasound couples ultrasound technology with endoscopy and is mostly used to determine primary tumour extension (Stage T) and to evaluate metastases in nearby lymph nodes (Stage N). There are two types of endoscopic ultrasound probes: radial and linear. Radial probes are the more widely used because they provide a cross-section of the esophagus and surrounding structures. Linear probes provide parallel images of the intestinal tract and are mostly used for endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA). Mini-probes use higher frequencies; however, since they are used only infrequently or not at all in Québec because of their short lifespan, they were not evaluated in this report. The performance of endoscopic ultrasound is operator-dependent, among other factors. Some authors have indicated that it is better after 100 examinations, while others have shown that it is better in centres with more than 50 examinations per year per examiner than in centres with fewer than 50 examinations per year per examiner.

The EUS-FNA method allows for the insertion of a needle to sample cytopathological tissue in the lymph nodes and tumour masses for cytopathological or histopathological analysis. This technique is not recommended in the case of suspicious lymph nodes located near the tumour because the needle passing through the tumour may contaminate the sample and produce a false-positive result.

The fusion of positron emission tomography with computed tomography (PET-CT) is of great benefit for esophageal cancer staging. It uses a glucose analogue, the ^{18}F-fluorodeoxyglucose tracer, or ^{18}F-FDG, to measure tissue and cell activity. The PET-CT combination improves the spatial resolution of the signal from the positron emitter tracer and contributes to optimizing the clinical performance of PET in oncology.

Magnetic resonance imaging (MRI) yields anatomical images of the body's internal structures using a magnetic field and non-ionizing radiation like radiofrequency waves. Its contrast resolution is considered better than that of CT in some locations, such as soft tissue. Some factors such as magnetic field strength (measured in teslas), the use of contrast agents with magnetic properties, and an antenna coupled with a gastroscope influence its performance.

Minimally invasive surgical procedures (MISP), such as laparoscopy and thoracoscopy, involve inserting a probe with a camera into the abdominal or thoracic cavity to view tissues more directly and to perform biopsies of suspicious lesions and lymph node dissections for histopathological analysis. They require general anesthesia.

Methodology

This report is a systematic review of the scientific literature published on the performance of the diagnostic tests used for the clinical staging of cancer of the esophagus and gastroesophageal junction not associated with stomach cancer. The review of primary studies covered the period from January 1999 to December 2007 in the Medline (PubMed) and Cochrane Library databases. A literature watch for systematic reviews and guidelines continued until July 2008. Study inclusion criteria were based on subject relevance, reference standards used according to the different categories in the TNM classification system, and availability of data allowing for performance assessment. Insofar as possible, we performed stratified analyses according to tumour TNM stage, histological type and position. Studies clearly indicating that they had included patients who had received neoadjuvant therapy were excluded, except for those addressing technologies for which there were very few primary data.[6] Also excluded were studies with 10 or fewer patients, those published in a language other than French

6. PET-CT, EUS-FNA and minimally invasive surgical procedures (MISP).

or English, and those dealing with Barrett's esophagus.

Diagnostic performance was expressed as a weighted mean or as a median. For Stage T, when data were available, we also calculated the rate of correctly staged, understaged and overstaged tumours. The comparative performance of the technologies was assessed directly or indirectly, depending on the results available. The quality of the studies was evaluated by means of scales adapted from recognized assessment tools.

Results

Most of the studies we reviewed were of poor methodological quality and had small sample sizes. Several were retrospective studies and did not mention if the results of the histopathological analysis or other diagnostic tests where blinded to those interpreting the imaging results. There was considerable heterogeneity in the results, likely due to variability in subject inclusion criteria, lack of consensus on the criteria for assessing lymph node stages, and differing definitions of Stages N and M. Examiner experience also probably contributed to the variability in the results.

This literature review has several limitations. The fact that it generally excluded the studies that analyzed the outcomes of patients who had received neoadjuvant therapy may have favoured the inclusion of cases of less advanced stages, but this was probably less significant than the bias that would have been introduced by the inclusion of patients who had received that type of treatment. The performance assessment of the technologies for detecting distant metastases was not complete since it was limited exclusively to evaluating their performance in staging esophageal cancer and did not take into account their performance in evaluating any other primary cancer. The comparison of the different technologies was primarily done indirectly. Lastly, the impact on clinical outcomes was not analyzed.

Performance of CT

Most of the studies on CT dealt with single-slice spiral CT or multi-detector spiral CT technology, but the latest generation of scanners (64- and 256-slice) was not evaluated. Most of the studies used abdominal and thoracic CT. In evaluating Stage N, CT had low sensitivity (median, 41.7%) but fairly high specificity (median, 82.4%). Its specificity for diagnosing distant lymph node metastases and distant organ metastases (Stage M) was relatively high (median, 87.2%), but its sensitivity was low (median, 49.2%). Its sensitivity was likely underestimated because the selected studies did not separately evaluate metastases in distant organs and distant lymph nodes, and several of the studies likely had a selection bias in favour of cases at earlier stages.

Computed tomography (CT) with 3D image reconstruction could be useful for locoregional assessment in the presence of stenosing tumours.

Performance of PET with and without CT fusion

Positron emission tomography (PET) was analyzed separately from PET-CT. Data on PET were derived from systematic reviews and meta-analyses. The PET primary tumour detection rate was high (95–100%), except for early-stage tumours. In evaluating Stage N, the sensitivity of PET was 57% (95% CI: 43–70%) and its specificity was 85% (95% CI: 76–95%). However, PET had a better performance in evaluating metastases in distant lymph nodes and organs (Stage M): its sensitivity was 71% (95% CI: 62–79%) and its specificity was 93% (95% CI: 89–97%).

Data on PET-CT for initial staging of esophageal cancer are scarce. They were based on two primary studies: one analyzed groups of lymph nodes and the other included patients who had received neoadjuvant therapy. For assessing locoregional lymph nodes and Stage M1a, PET-CT had a sensitivity of 83.3% to 93.9% and a specificity of 92.1%. It was very sensitive (100%) in evaluating Stage M, but there were only four patients presenting with Stage M1 in the analyzed studies.

Performance of endoscopic ultrasound with and without EUS-guided fine-needle aspiration

Endoscopic ultrasound performed well in determining patients who needed more aggressive treatment (≥ T2), with a median sensitivity of 97.1%, but it performed less well in determining potential candidates for mucosal resection (≤ T1), with a median specificity of 75%. It was not possible to appraise the sensitivity of EUS in detecting Stage T4 tumours because of the limited number of patients in the studies, but false diagnoses of spread to nearby organs were infrequent (median specificity, 98.8%).

The performance of EUS in evaluating Stage N was limited (median sensitivity, 76.2%; median specificity, 66.7%). However, it was higher if to the four pre-set practice standard criteria (size greater than 5–10 mm; round shape; hypoechogenic lymph node; clear border), we added three criteria (visualization of celiac lymph nodes; visualization of more than five lymph nodes; Stage T3 or T4 tumour).

Its median sensitivity was 75% and its median specificity was 93.7% in diagnosing celiac lymph node metastasis (Stage M1a). There was a low risk of confusing celiac lymph nodes and left gastric artery lymph nodes (Stage N). Very few studies specifically evaluated metastases to the liver.

Endoscopic ultrasound is more effective for evaluating squamous cell carcinomas than adenocarcinomas, but there were very few studies on that topic and tumour position may be a confounding factor. According to a meta-analysis, the effectiveness of EUS for distinguishing between Stage T1/T2 and Stage T3/T4 tumours located in the cardia (Q*[7] = 0.85) was less than for such tumours located in the esophagus (Q* = 0.90). The performance of EUS in detecting cardia tumours gradually declined from Type I to Type III tumours (Siewert classification), but that outcome was based on a single study.

In the case of stenosing tumours, locoregional staging was more difficult to perform with EUS. Dilation seemed associated with a low risk of esophageal perforation (between 0% and 2.9%), but these data came from centres that had a high EUS volume, performed gradual dilation and generally avoided dilation in cases of severe stenosis. The clinical effect of potential cancer cell seeding from handling the tumour during dilation is not known. The Esophagoprobe had a fairly good passage rate, and its performance in cases of stenosis (accuracy in evaluating Stage T, 62.3–90%) seemed to be equivalent to that of conventional radial probes in the absence of stenosis (accuracy in evaluating Stage T, 67.1–90%). Nevertheless, this probe cannot be used for needle aspiration and is not used in Québec.

The possibility of performing biopsies with EUS is very attractive for optimizing the evaluation of lymph node metastases. The sensitivity of EUS-FNA in evaluating Stage N ranged from 83.3% to 93.3%, and its specificity was 92.9%. In evaluating celiac lymph nodes (Stage M), its sensitivity ranged from 92.9% to 97.8%, and its specificity was 100%. However, these data were based on very few studies, and the results must be interpreted with caution because of the small number of patients who presented with Stage N0 or M0 and the reference standard used: in one study, the standard consisted of the histopathological results of patients who had received neoadjuvant therapy, while in three other studies, the positive results from the cytopathological examination used as the reference standard were all presumed to be true positives. In the majority of the studies, a cytopathologist was present to ensure that the biopsy was done correctly.

According to expert opinion, EUS-FNA is not recommended for evaluating lymph nodes adjacent to the tumour because of the risk that the primary tumour may contaminate the sample. The performance of EUS with modified criteria[8] in distinguishing between affected and non-affected lymph nodes was sufficient to avoid needle biopsy, but the definition of malignancy threshold needs to be validated before this technique can be used clinically.

7. Q* is the ROC (Receiver Operating Characteristic) value where sensitivity is equal to specificity. The ROC curve relates the sensitivity and the specificity of a diagnostic test.

8. Size > 5–10 mm; round shape; hypoechogenic lymph node; clear border; visualization of celiac lymph nodes; visualization of > 5 lymph nodes; Stage T3 or T4 tumour.

Performance of minimally invasive surgical procedures and MRI

Very few studies on minimally invasive surgical procedures (MISP) and MRI were identified. For Stage N, the specificity of thoracoscopy and laparoscopy was very high (100%), but the sensitivity of thoracoscopy (45.5%) was lower than that of laparoscopy (90.9%). The reference standard used in the studies included the histopathological results of patients who had received neoadjuvant therapy.

In evaluating Stage T, MRI with an endoscopic coil performed better than MRI with an external coil, but the strength of the magnetic field (0.15 tesla) used with the external coil is no longer part of current practice. We did not identify any studies on MRI diagnosis of distant metastases for the initial staging of esophageal cancer.

Comparative performance of the different technologies

We compared the performance of the different technologies for each tumour stage (T, N and M), except for minimally invasive and non-invasive surgical techniques, since most of the studies on that topic did not use the TNM classification system.

Stage T

The rate of correctly staged tumours was higher with EUS (median, 79.5%) than with spiral CT (median, 61.8%), but the latest generation of scanners (64- and 256-slice) was not evaluated.

Stage N

The sensitivity of EUS was significantly higher (80% [95% CI: 75–84%]) than that of CT (50% [95% CI: 41–60%]) and that of PET (57% [95% CI: 43–70%]). However, its specificity was lower (70% [95% CI: 65–75%]) than that of CT (83% [95% CI: 77–89%]) and of PET (85% [95% CI: 76–95%]), and the difference was also significant. Spiral CT did not perform very well in detecting Stage N because lymph node assessment relied only on lymph node size, and the malignancy threshold was generally a short-axis diameter of 10 mm. However, for EUS diagnosis, the malignancy threshold was sometimes a short-axis diameter of 5 mm, and other morphological criteria were used, such as lymph node shape or clear border. The lower sensitivity of PET in evaluating locoregional lymph nodes may have been due to the fact that the positive signal of the primary tumour masked that of nearby affected lymph nodes.

Data indicate that EUS-FNA performed better than EUS alone in evaluating locoregional lymph nodes, independently of the criteria used to differentiate affected and non-affected lymph nodes by means of EUS.

The PET-CT fusion staged lesions that PET had rated as uncertain. However, none of the selected studies compared PET-CT with EUS-FNA in evaluating locoregional lymph nodes.

Stage M

Very few studies compared EUS and CT in evaluating celiac lymph nodes (Stage M). Combining EUS and CT would yield a higher metastasis detection rate than CT alone. Only one of the selected studies compared EUS-FNA and EUS in diagnosing celiac lymph nodes. Outcomes indicated that the sensitivity of EUS-FNA (92.9%) was superior to that of EUS alone (75%). However, their specificity was not compared because there was a lack of patients presenting with Stage M0; the cytopathological examination results were considered the reference standard since they could not be compared with the results from examination of the resected samples, and the authors did not compare the same groups of patients.

In evaluating distant lymph node and distant organ metastases, the diagnostic performance of PET was significantly better than that of CT (value of the diagnostic odds ratios[9] = 2.26 [95% CI: 1.09–4.71%]). The only selected study on this topic indicated that PET-CT had a higher sensitivity than CT, but the study had only four patients who presented with Stage M1. We did not identify any

9. The diagnostic odds ratio is a measure describing the odds of positive test results in people with a disease compared with the odds of positive test results in people without the disease. It allows indirect comparisons between two diagnostic tests.

studies comparing MRI with other methods for diagnosing distant metastases in esophageal cancer.

Minimally invasive surgical procedures

Minimally invasive surgical procedures (MISP) help localize lymph node metastases and distant metastases undetected by a combination of non-invasive techniques (EUS, CT and MRI). In the selected studies, several patients had adenocarcinomas in the lower part of the thoracic esophagus. However, several of those studies used the biopsies performed with MISP as the reference standard because the results from the different clinical staging methods could not be compared with those from the histopathological examination of a resected specimen.

Conclusion

This report indicates that it is necessary to use a combination of techniques for optimal clinical staging of esophageal cancer. Our conclusions are consistent with several existing practice guidelines. However, the many methodological limitations of the studies we examined did not allow us to draw any firm conclusions.

Taking these limitations into account, AETMIS has drawn several conclusions and proposes the following diagnostic test sequence for the clinical staging of esophageal cancer:

- Start with a CT scan of the neck, thorax and abdomen to determine if there are distant metastases.
- If no distant metastases are present, use EUS to evaluate locoregional invasion (Stages T and N) and celiac lymph nodes and EUS-FNA if the tumour does not obstruct the needle; if a stenosing tumour is present, the optimal approach is not known, but dilation is indicated, except in the case of severe stenosis, and it should be done in centres with considerable expertise.
- Add PET-CT to the cancer staging if the patient is judged eligible for curative treatment after a CT scan and EUS (however, further research is needed to confirm the utility of this approach).
- Use MRI if CT cannot be performed (even though this option is mentioned in only some practice guidelines).
- Perform MISPs (laparoscopy, thoracoscopy) in certain situations, such as laparoscopy to evaluate abdominal metastases (e.g., in the peritoneum) when the cancer is located in the gastroesophageal junction.

The biomarkers used to establish prognosis at the time of diagnosis are still at the experimental stage and not currently in clinical use.

Several authors have mentioned that EUS examiner experience is a major cause for the variability in the performance outcomes achieved with this technique. It would therefore be important to centralize the different exploratory tests for clinically staging esophageal cancer so that examiners can maintain their expertise and level of proficiency in using rapidly evolving diagnostic technologies.