uegjournal

Original Article

Do pathologists agree with each other on the histological assessment of pT1b oesophageal adenocarcinoma?

Annieke W Gotink[1] ⓘ, Fiebo JC ten Kate[2], Michael Doukas[2], Bas PL Wijnhoven[3], Marco J Bruno[1], Leendert HJ Looijenga[2], Arjun D Koch[1] and Katharina Biermann[2]

Abstract

Background: In early (T1) oesophageal adenocarcinoma (OAC), the histological profile of an endoscopic resection specimen plays a pivotal role in the prediction of lymph node metastasis and the potential need for oesophagectomy with lymphadenectomy.

Objective: To evaluate the inter-observer agreement of the histological assessment of submucosal (pT1b) OAC.

Methods: Surgical and endoscopic resection specimens with pT1b OAC were independently reviewed by three gastrointestinal pathologists. Agreement was determined by intraclass correlation coefficient for continuous variables, and Fleiss' kappa (κ) for categorical variables. Bland–Altman plots of the submucosal invasion depth were made.

Results: Eighty-five resection specimens with pT1b OAC were evaluated. The agreement was good for differentiation grade (κ=0.77, 95% confidence interval (CI) 0.68–0.87), excellent for lymphovascular invasion (κ=0.88, 95% CI 0.76–1.00) and moderate for submucosal invasion depth using the Paris and Pragmatic classifications (κ=0.60, 95% CI 0.49–0.72 and κ=0.42, 95% CI 0.33–0.51, respectively). Systematic mean differences between pathologists were detected for the measurement of submucosal invasion depth, ranging from 297 μm to 602 μm.

Conclusions: A substantial discordance was found between pathologists for the measurement of submucosal invasion depth in pT1b OAC. Differences may lead to an over- or underestimation of the lymph node metastasis risk, with grave implications for the treatment strategy. Review by a second gastrointestinal pathologist is recommended to improve differentiating between a favourable and an unfavourable histological profile.

Key summary

Established knowledge on this subject:
- The histological profile of an endoscopic resection specimen plays a pivotal role in the prediction of lymph node metastasis (LNM) in early (T1) oesophageal adenocarcinoma (OAC).
- Patients with a mucosal OAC can be treated curatively by endoscopic resection, the risk for LNM being considered negligible.

[1]Department of Gastroenterology and Hepatology, Erasmus MC, University Medical Centre Rotterdam, The Netherlands
[2]Department of Pathology, Erasmus MC, University Medical Centre Rotterdam, The Netherlands
[3]Department of Surgery, Erasmus MC, University Medical Centre Rotterdam, The Netherlands

Corresponding author:
Annieke W Gotink, Department of Gastroenterology and Hepatology, Erasmus MC, University Medical Centre Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands.
Email: a.gotink@erasmusmc.nl

- Patients with a submucosal OAC are generally advised to undergo oesophagectomy with lymphadenectomy for the presumed risk of LNM.
- New studies now cautiously propose a conservative policy with surveillance after endoscopic resection for low-risk submucosal OAC, given that the risk of LNM remains low.

New findings:

- There is good to excellent agreement between pathologists for the histological assessment of differentiation grade and lymphovascular invasion.
- There is a substantial discordance between pathologists for the assessment of submucosal invasion depth in OAC, highlighting the practical difficulties associated with the measurement of this parameter.
- The found discordance may potentially lead to an over- or underestimation of the LNM risk, with grave implications for the treatment strategy of choice.

Introduction

In the past three decades, the incidence of oesophageal adenocarcinoma (OAC) has risen more rapidly than any other malignant neoplasm in the western world.[1,2] Radical oesophagectomy has long been the cornerstone treatment for OAC or high-grade dysplasia in Barrett's oesophagus, with mortality rates ranging from 0% to 22%, and morbidity rates of up to 80%.[3] In current clinical practice, however, minimally invasive endoscopic treatment modalities have been widely accepted and integrated into the management of Barrett's oesophagus and early (T1) OAC. A major drawback of these endoscopic procedures, however, is the inability to perform lymphadenectomy. Surgery is advocated when the risk of lymph node metastasis (LNM) outweighs the risk for mortality associated with oesophagectomy.[4–6]

The histological profile of an endoscopic resection specimen plays a pivotal role in assessing the risk of LNM in early (T1) OAC. Endoscopic resection was initially only accepted for tumours that were limited to the mucosa, based on the premise that the risk of LNM in these superficial lesions is negligible. However, more recent publications have shown that when the tumour extends into the submucosa with a maximum invasion depth of 500 μm (sm1), in combination with a well to moderate grade of differentiation (G1/G2) without lymphovascular invasion (LVI), the risk of LNM remains low.[7–13] Hence, for these low-risk sm1 OACs, multiple studies now cautiously propose a conservative approach with surveillance after endoscopic resection.[9,11–14] Guidelines in fact consider it a valid alternative for patients with a high surgical risk.[4,15,16] An accurate and reproducible histological assessment of submucosal OAC is therefore essential, when this is the fundament on which the appropriate treatment strategy is determined.

The depth of invasion in the submucosa was originally classified by subdividing the submucosa into three equal layers (sm1/sm2/sm3). This so-called Pragmatic classification was developed based on oesophagectomy specimens.[17] In endoscopic resection specimens, however, this classification may not be reliable, as subtotal resection of the submucosa may lead to over-staging. This led to the development of a quantitative classification where submucosal invasion was measured in micrometres, with sm1 defined as invasion depth of $\leq 500\,\mu m$ and sm2/3 as invasion depth of $> 500\,\mu m$ (Paris classification).[7,18] Although this new classification has already been integrated into current clinical practice,[19] the agreement of this classification has not been determined.

Differentiation grade and LVI are two other histological characteristics associated with the risk of LNM in submucosal OAC. A high inter-observer variation was detected for these parameters in earlier studies on breast and colorectal cancer.[20–23] To date, the agreement for the assessment of these parameters in OAC has not been investigated.

The aim of this study was to determine the agreement between pathologists for the histological assessment of submucosal OAC, which includes the depth of submucosal invasion, differentiation grade and the presence of LVI.

Materials and methods

Study design and patients

We conducted a retrospective cohort study, including patients from a tertiary referral centre in The Netherlands. All consecutive patients who underwent an oesophagectomy or endoscopic resection for a pathologically staged submucosal (pT1b) OAC between 1989 and 2014 in the Erasmus University Medical Centre were identified using the Registry of the Netherlands Comprehensive Cancer Organization. Patients were included when pT1b OAC was reported in the original pathology report. Patients were excluded when treated by neoadjuvant chemo- and/or radiation therapy, or when histopathology was unavailable for revision.

Histopathological evaluation

After eligible patients were identified, hematoxylin and eosin stained (H&E) slides and formalin-fixed paraffin-embedded (FFPE) tissue blocks of the resection specimen were retrieved from the archives of the department of pathology. All samples were independently reviewed by three gastrointestinal (GI) pathologists from the Erasmus University Medical Centre, who were blinded from the original diagnosis and clinical outcome (development of LNM), as well as from each other's results.

Submucosal invasion depth. All H&E-stained slides were initially reexamined by one pathologist to verify that there was submucosal invasion, and to identify the slide and corresponding FFPE block with the deepest invasion of tumour in the submucosa. Subsequently, immunohistochemistry (IHC) using desmin and pankeratin (D&P) double staining was done to highlight the lamina muscularis mucosae as well as the adenocarcinoma. The H&E and D&P-IHC slides were digitally scanned and made available for review to three GI pathologists. Measurements on submucosal invasion depth were done in both the H&E and D&P-IHC slides respectively.

The deepest level of submucosal invasion was measured in micrometres (μm), measured from below the lamina muscularis mucosa (Figure 1). Where the lamina muscularis mucosae was duplicated, submucosal invasion was measured from the deepest layer.[24] Based on the invasion depth in μm, the Paris classification was derived (sm1 invasion ≤500 μm; sm2/3 invasion > 500 μm).

In order to determine the Pragmatic classification, the total thickness of the submucosa was additionally measured in μm. In surgical resection specimens this was measured at the point of deepest submucosal tumour invasion, and in endoscopic resection specimens this was measured adjacent to the tumour. The Pragmatic classification was derived by correlating the depth of submucosal tumour invasion with the total
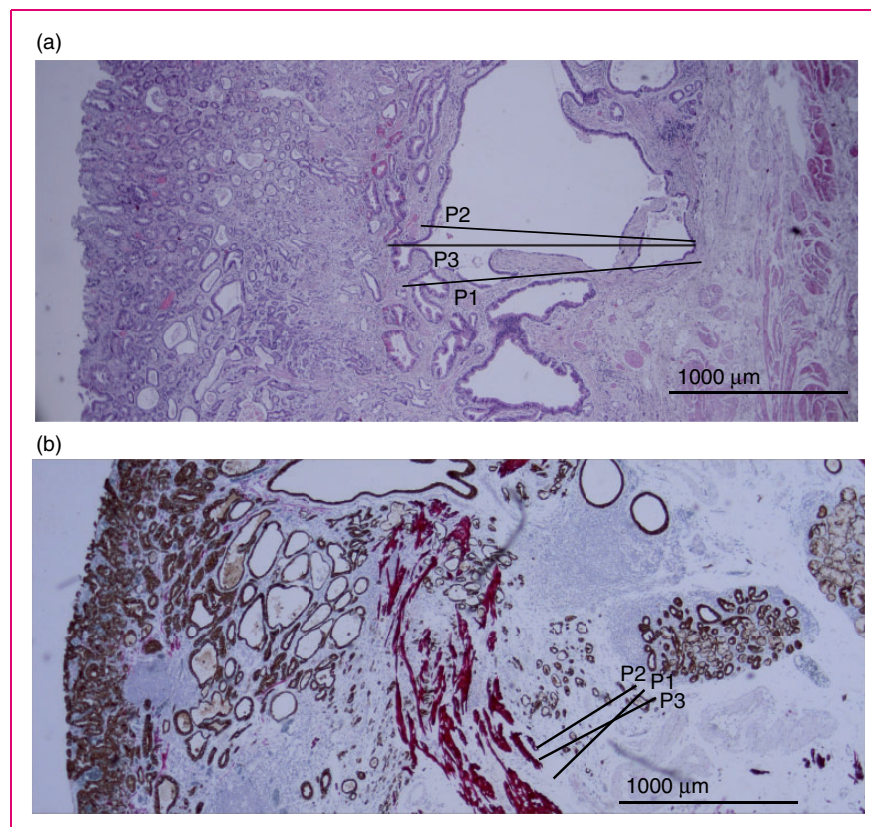
Figure 1. Histological assessment of pT1b OAC, scored independently by three gastrointestinal pathologists. (a) Surgical resection specimen, stained by hematoxylin and eosin: all three pathologists scored a moderate grade of differentiation (G2) and no lymphovascular invasion (LVI). Submucosal invasion depth for pathologist 1 (P1) was 1720 μm, pathologist 2 (P2) was 1525 μm and pathologist 3 (P3) was 1890 μm. (b) Endoscopic resection specimen, with desmin and pankeratin immunohistochemistry: all three pathologists scored no LVI. The grade of differentiation was moderate for P1, good for P2 and moderate for P3. Submucosal invasion depth for P1 was 581 μm, P2 was 340 μm and P3 was 620 μm.

submucosal thickness, where the submucosa was sub-divided into three equal layers (sm1/sm2/sm3).

Grade of differentiation and LVI. Grade of differentiation and presence of LVI were assessed independently by the same three pathologists. For these parameters, whole case H&E slide sets were re-examined by microscopy. Grading was classified according to the World Health Organization classification for colorectal adenocarcinoma[21] and included G1: well differentiated, G2: moderately differentiated, G3: poorly differentiated, G4: undifferentiated. LVI was scored as being either present or absent.

Combination of histological characteristics. Histological characteristics of the tumours were combined and classified as 'favourable' or 'unfavourable'. Favourable was defined as submucosal invasion depth of $\leq 500\,\mu m$ (sm1 according to the Paris classification), G1 or G2 and no LVI. Unfavourable was defined as any poor histological characteristic, such as submucosal invasion depth of $>500\,\mu m$ (sm2/3 according to the Paris classification), G3 or G4 or the presence of LVI.

Clinical follow-up

The primary clinical endpoint for patients in this study was the development of LNM, as this is the main cause for concern in patients with submucosal OAC. After oesophagectomy with lymph node dissection, a threshold of at least 12 lymph nodes was used to assess the lymph node status.[25] However, due to different surgical resection techniques over time with less extensive lymphadenectomy, the development of suspected LNM during follow-up, either pathologically confirmed or not, was used as a surrogate endpoint for LNM. This surrogate endpoint was also used for patients that were treated by primary endoscopic resection only. Patient charts were assessed to review the clinical follow-up. The date of established LNM and the date of the last known clinical follow-up were recorded.

Statistical analysis

Statistical analysis was performed using SPSS software (version 21, SPSS Inc., Chicago, Illinois, USA), and RStudio Team (2016; RStudio: Integrated Development for R. RStudio, Inc., Boston, Massachusetts, USA, version 1.0.44). For categorical variables, the observed agreement was summarized as a percentage agreement between the three pathologists. The degree of agreement was determined using the Fleiss' kappa (κ) for categorical variables and intraclass correlation coefficient (ICC) for continuous variables. Coefficients < 0.21, 0.21–0.40, 0.41–0.60, 0.61–0.80 and 0.81–1.00 represent a poor, fair, moderate, good and excellent agreement, respectively.[26] Corresponding 95% confidence intervals (95% CI) were calculated. Continuous variables were plotted in a Bland–Altman plot in order to assess for systematic differences between pathologists.

Results

Patients

A total of 89 patients with a pT1b OAC were identified. Four patients were excluded because the H&E slides and FFPE material were not available (three patients) or were of poor quality (one patient). Hence, 85 patients with a pT1b OAC were included in this study. An endoscopic resection had been performed in 35/85 patients (41%) and 50/85 patients (59%) were treated by primary oesophagectomy. Patient characteristics are shown in Table 1.

In 74/85 patients (87%) the clinical follow-up concerning LNM was known. LNM had occurred in 24/74 patients (33%) with pT1b OAC, found either in surgically resected specimens (21 patients) or during follow-up (three patients). The median follow-up time was 4.98 years (interquartile range 2.8–5.9 years).

In 11/85 patients (13%) the LNM status during follow-up was unknown. In nine patients this was because of loss to follow-up after resection, one patient was treated by chemotherapy for a second primary tumour, and one patient had a second primary tumour with LNM of unknown origin.

Tumour characteristics of the 85 patients included in the study are shown in Supplementary Material Table 1 online.

Table 1. Patient characteristics of the 85 patients with a submucosal oesophageal adenocarcinoma.

Gender, male, *n* (%)	69 (81)
Age at resection, median (range), years	66 (41–85)
Treatment regime	
Endoscopic resection, *n* (%)	35 (41)
Primary surgery, *n* (%)	50 (59)
Surgical approach	
Transhiatal oesophagectomy, *n* (%)	45 (90)
Transthoracic oesophagectomy, *n* (%)	3 (6)
Gastrectomy, *n* (%)	1 (2)
Unknown, *n* (%)	1 (2)
Lymph node metastasis during follow-up[a]	24 (33)

Continuous data is presented as median (range). Categorical data as *n* (%).
[a]Known in only 74 (87%) patients.

Histopathological assessment

Submucosal invasion depth. There was a good degree of agreement between pathologists for the assessment of the depth of tumour invasion in micrometres (ICC = 0.649, 95% CI 0.536–0.746) and an overall good degree of agreement for the assessment of the submucosal thickness (ICC = 0.659, 95% CI 0.543–0.755) (Table 2).

When the Paris and Pragmatic classification for submucosal invasion depth were derived from these measurements, there was agreement between pathologists in 69/85 patients (82%) and 40/84 patients (47%) respectively, with a moderate degree of agreement (κ = 0.603, 95% CI 0.489–0.717; κ = 0.419, 95% CI 0.333–0.505) (Table 2).

Importantly, when comparing measurements of submucosal invasion depth, significant differences were found between pathologists. See Figure 1 for a representative example of how pathologists measured submucosal invasion depth differently. Between pathologists 1 and 2 there was a systematic mean difference of 297 μm (*p* = 0.073), between pathologists 1 and 3 this was 304 μm (*p* = 0.065) and between pathologists 2 and 3 this was 602 μm (*p* < 0.0001) (Figure 2, Table 3). The discordance was further supported by Bland–Altman plots (Figure 2), where a funnel-shaped trend between the measurements of the different pathologists was seen. As the mean invasion depth increased, the difference in measurement of submucosal invasion depth between pathologists became larger. In other words, smaller differences were observed in superficial tumours, and larger differences were observed for deeper tumours.

Additional D&P-IHC could be performed in 52/85 patients (61%). In the remaining 33 patients, FFPE blocks were not available, were of poor quality or the invasive component was not present after sectioning. The agreement between pathologists for the assessment of submucosal invasion depth in D&P-IHC, was comparable to the results derived from H&E slides, as seen in Supplementary Material Table 2 Online. The agreement for both the Paris and Pragmatic classifications did not improve by usage of D&P-IHC.

Grade of differentiation and LVI. For the assessment of grade of differentiation, agreement between pathologists was seen in 68/85 patients (80%), with a good degree of agreement (κ = 0.774, 95% CI 0.680–0.868). For the assessment of LVI, agreement between pathologists was seen in 80/85 patients (94%), with an excellent degree of agreement (κ = 0.879, 95% CI 0.756–1.002) (Table 2).

*Combination of histological characteristics.* Agreement between pathologists for favourable or unfavourable histological characteristics was seen in 76/85 patients (89%), with a good degree of agreement (κ = 0.728, 95% CI 0.605–0.851) (Table 2). The tumour was classified as unfavourable in 67/76 patients (88%) and favourable in 9/76 patients (12%).

In 66/76 patients (87%) the clinical follow-up concerning LNM was known. The tumour was classified as unfavourable in 58/66 patients (88%), and 24/58

Table 2. Inter-observer agreement between three gastrointestinal pathologists, for the histological assessment of pT1b oesophageal adenocarcinoma in 85 patients. The histological parameters were established using hematoxylin and eosin stained slides.

Histological parameter (continuous variable)	*n*	Observed agreement	ICC	95% CI	Interpretation
Submucosal invasion in μm	85	N/A	0.657	(0.547–0.750)	Good
Total thickness of submucosa in μm	84	N/A	0.659	(0.543–0.755)	Good
	35 (ER)	N/A	0.704	(0.515–0.833)	Good
	49 (surgical)	N/A	0.553	(0.389–0.698)	Moderate
Histological parameter (categorical variable)	*n*	Observed agreement	Fleiss' kappa	95% CI	Interpretation
Paris classification	85	82%	0.603	(0.489–0.717)	Moderate
Pragmatic classification	84	47%	0.419	(0.333–0.505)	Moderate
Tumour differentiation	85	80%	0.774	(0.680–0.868)	Good
LVI	85	94%	0.879	(0.756–1.002)	Excellent
Combined histological characteristics[a]	85	89%	0.728	(0.605–0.851)	Good

[a]Favourable: pT1a/sm1, G1/2, and no lymphovascular invasion (LVI); unfavourable: pTsm2/3/pT2, or G3/4, or LVI.
N/A: not applicable; ICC: intraclass correlation coefficient; CI: confidence interval; ER: endoscopic resection.
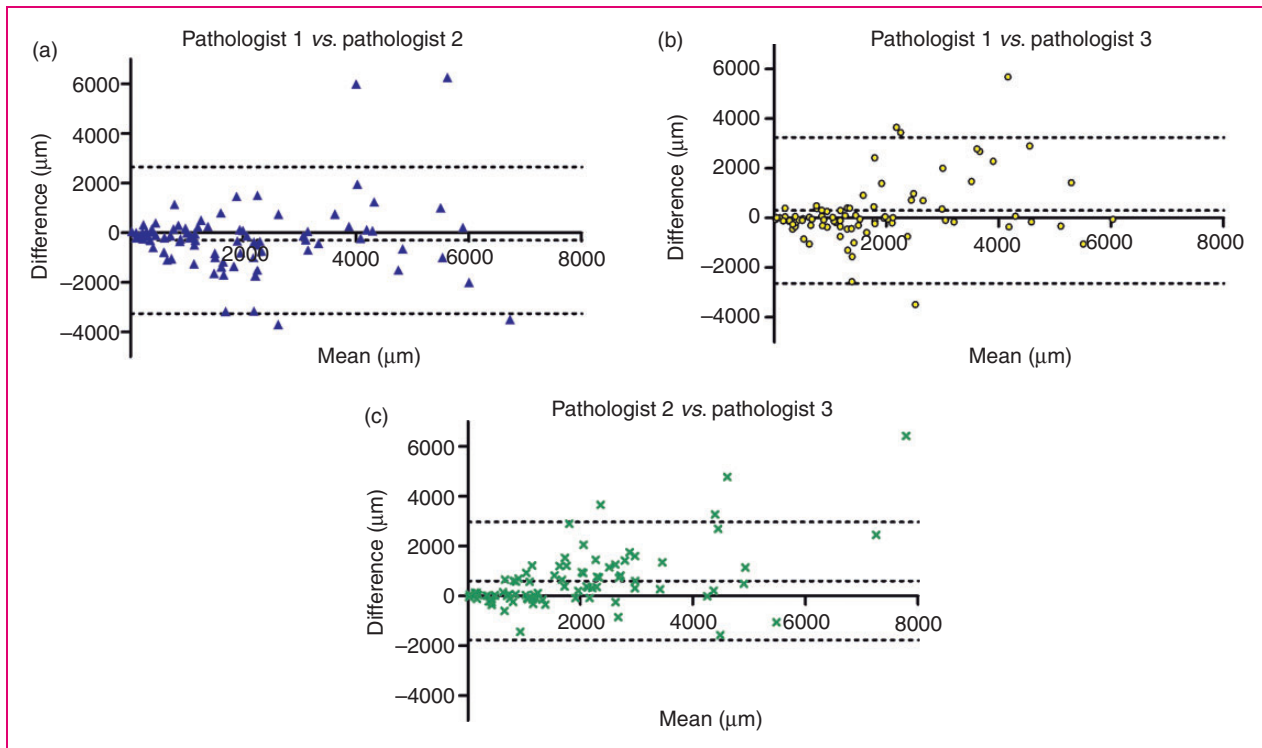
Figure 2. Bland-Altman plots, illustrating systematic mean differences between pathologists for the measurement of submucosal invasion depth. Differences between measurements were plotted against the mean of two measurements. The mean difference between pathologists is represented by the parallel line adjacent to the 0-line in the x-axis. A: mean difference of submucosal invasion depth is $-297\,\mu m$ ($p = 0.073$), B: mean difference of submucosal invasion depth is $304\,\mu m$ ($p = 0.065$), C: mean difference of submucosal invasion depth is $602\,\mu m$ ($p < 0.05$). Please note the funnel-shaped trend between the pathologists; as the mean invasion depth increased, differences in measurement of invasion depth became more prominent.

Table 3. Systematic mean differences between pathologists for the measurement of submucosal invasion depth.

	Mean difference in measurement of submucosal invasion depth (μm)	*p*-value	95% CI
Pathologist 1 *vs.* pathologist 2	−297	0.073	(−3254; 2659)
Pathologist 1 *vs.* pathologist 3	304	0.065	(−2637; 3246)
Pathologist 2 *vs.* pathologist 3	602	0.00002	(−1768; 2972)

CI: confidence interval.

patients (41%) developed LNM. In contrast, no LNM developed in the 8/66 patients (12%) that were classified as favourable.

Disagreement between pathologists for favourable or unfavourable histological characteristics was seen in 9/85 patients (11%). In all these patients, the disagreement concerned the Paris classification of tumour depth (sm1 invasion $\leq 500\,\mu m$ or sm2/3

invasion $> 500\,\mu m$). It was agreed that all other characteristics were favourable, such as a G1 or G2 and no LVI. In 8/9 (89%) patients the clinical follow-up concerning LNM was known and none developed LNM.

Discussion

In early (T1) OAC, the risk for LNM is predicted based on the histological profile of an endoscopic resection specimen. An accurate and reliable histological assessment is crucial when this is the fundament on which the appropriate treatment strategy is determined. In this study on patients with submucosal OAC, we found that there was a good to excellent degree of agreement between GI pathologists for the assessment of the differentiation grade and the presence of LVI. However, we did detect a significant discordance between pathologists for the measurement of submucosal invasion depth, which highlights the practical difficulties associated with the assessment of this parameter. This is a highly relevant observation, given the decisive nature of the $500\,\mu m$ cut-off level, where surveillance is recommended for $\leq 500\,\mu m$ invasion and additional

oesophagectomy with lymphadenectomy is recommended for invasion depth of $>500\,\mu m$.[4,9,11–16]

The agreement for submucosal invasion depth alone was good (ICC = 0.649). However, when the Paris classification for tumour depth invasion was derived, the agreement between pathologists became moderate (κ = 0.603). This aggravation can be explained by the systematic discordance in measurement between pathologists, ranging between $297\,\mu m$ and $602\,\mu m$, which is of great impact for values in close proximity to the $500\,\mu m$ cut-off level. The inter-observer variation was even higher for the Pragmatic classification (κ = 0.419), which may be explained by the fact that a kappa value tends to be lower with increasing number of categories.[27] Another explanation may be the fact that we calculated the invasion category by correlating two continuous variables, each with its own variation, whilst in clinical practice this classification is often made by eyeballing. In a previous inter-observer agreement study on the combined assessment of mucosal and submucosal OAC, the Pragmatic classification was also found to be moderate.[28] Although both classifications should be used for clinical decision making,[7,29] we cannot reliably demonstrate superiority of either system in terms of agreement. Moreover, D&P-IHC did not substantially improve agreement in Paris and Pragmatic classifications. A consensus assessment by a second GI pathologist may be of added value for pT1b lesions with submucosal invasion in close proximity to $500\,\mu m$ cut-off.

The substantial discordance found in this study may be explained by subjective anticipation of the deepest invasion, but also by several practical difficulties associated with the measurement of submucosal invasion. These include specimen curling and associated artefacts, which may make it difficult to assess the orientation plane and area of deepest invasion. Additionally, morphologic characteristics of the muscularis mucosa, such as the presence of a discontinuous or hypertrophic muscularis mucosa, the presence of a duplicated muscularis mucosa or destruction by the carcinoma, could further impair a valid assessment (Figure 1). These properties of the muscularis mucosa, but also the method of measurement (from the bottom of the muscularis mucosa or from an imaginary line of muscularis mucosa in the case of a discontinuous muscularis mucosa) have been named as a cause of discrepancy.[30,31] Tissue preservation in this retrospective study might also have had an impact on the slide quality and accuracy of assessment.

The validity of the $500\,\mu m$ cut-off for the prediction of LNM has been challenged in various gastric cancer studies, but also in colorectal cancer studies where a cut-off of $1000\,\mu m$ is used.[30–33] Historically, the quantitative cut-off was determined based on surgical resection specimens; however, specimen handling could artificially affect the depth of invasion. Stretching of the specimen after resection could decrease the thickness of the submucosa[32] and inadequate pinning of the specimen before rapid fixation may lead to submucosal shrinkage.[33] Our results on the discordant measurement between pathologists form an additional important argument to question the validity of the $500\,\mu m$ cut-off for the prediction of LNM.

Our results on the agreement for differentiation grade (κ = 0.774) and LVI (κ = 0.879) are in contrast with previous studies in breast and colorectal cancer. These studies reported a moderate inter-observer agreement for differentiation grade[20–22] and a substantial variability for LVI.[23] A possible explanation is that the pathologists involved in this study are highly experienced, and all from the same centre, with a close collaboration over the years.

Clinical management decisions are made based on the combined histological profile of a pT1b OAC. Our results show a good degree of agreement (κ = 0.728) between pathologists when tumours were classified as favourable or unfavourable. All patients that developed LNM had tumours that the pathologists agreed were unfavourable, and no LNM developed in patients with tumours that the pathologists agreed were favourable. This supports the proposed shift to a conservative policy after endoscopic resection when favourable characteristics are found. Where there was disagreement between pathologists (11%), this always concerned the Paris classification of tumour depth. Fortunately, no LNM was found in these patients. We should caution the validity of the $500\,\mu m$ cut-off; however, considering no LNM was found, we may also be less apprehensive of the clinical consequence.

Ultimately, reliable conclusions on the correlation with LNM cannot be made based on our small sample size. It is also questionable whether histological data from surgical and endoscopic resection specimens combined are comparable for the prediction of LNM. Specimen handling and different sectioning intervals between endoscopic resection and surgical specimens (2 mm vs. 5 mm) may result in an underestimation of the deepest invasion in surgical resection specimens.

The strength of this study is that, for the first time, the quantitative assessment for submucosal invasion depth in early OAC is analysed systematically, highlighting the practical difficulties and inaccuracy associated with the measurement of this parameter. Other strengths include the assessment of a large cohort of patients with submucosal OAC and the blinding of GI pathologists involved in this study. This study also has limitations. The involved pathologists were all from the same centre, having learned from each other over

the years. Additionally, the intra-observer agreement was not evaluated. Due to the retrospective study design, the clinical follow-up on LNM was less standardized and patients were also lost to follow-up.

In conclusion, there is a good to excellent degree of agreement for the histological assessment of differentiation grade and LVI in submucosal OAC. However, for the measurement of submucosal invasion depth, substantial differences exist between pathologists. This potentially leads to an over- or underestimation of the LNM risk, with grave implications on the treatment strategy of choice. In the current innovative climate, where we explore the limits to endoscopic therapy in patients with early OAC, this is a highly relevant issue to be aware of.

Declaration of conflicting interests

The authors declare that there is no conflict of interest.

Ethics approval

This retrospective study was exempted from institutional review board approval after assessment of the study design and protocol by the Medical Ethics Committee of the Erasmus University Medical Centre Rotterdam, as is in accordance with our national guidelines (www.ccmo.nl).

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Informed consent

Informed consent was not required for this study.

ORCID iD

Annieke W Gotink http://orcid.org/0000-0003-0509-3472

References

1. Pohl H and Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J Natl Cancer Inst* 2005; 97: 142–146.
2. Pennathur A, Gibson MK, Jobe BA, et al. Oesophageal carcinoma. *Lancet* 2013; 381: 400–412.
3. Metzger R, Bollschweiler E, Vallbohmer D, et al. High volume centers for esophagectomy: What is the number needed to achieve low postoperative mortality? *Dis Esophagus* 2004; 17: 310–314.
4. Fitzgerald RC, di Pietro M, Ragunath K, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* 2014; 63: 7–42.
5. Spechler SJ, Sharma P, Souza RF, et al. American Gastroenterological Association technical review on the management of Barrett's esophagus. *Gastroenterology* 2011; 140: e18–e52.
6. Evans JA, Early DS, Chandraskhara V, et al. The role of endoscopy in the assessment and treatment of esophageal cancer. *Gastrointest Endosc* 2013; 77: 328–334.
7. Fotis D, Doukas M, Wijnhoven BP, et al. Submucosal invasion and risk of lymph node invasion in early Barrett's cancer: Potential impact of different classification systems on patient management. *United European Gastroenterol J* 2015; 3: 505–513.
8. Boys JA, Worrell SG, Chandrasoma P, et al. Can the risk of lymph node metastases be gauged in endoscopically resected submucosal esophageal adenocarcinomas? A multi-center study. *J Gastrointest Surg* 2016; 20: 6–12.
9. Manner H, Pech O, Heldmann Y, et al. The frequency of lymph node metastasis in early-stage adenocarcinoma of the esophagus with incipient submucosal invasion (pT1b sm1) depending on histological risk patterns. *Surg Endosc* 2015; 29: 1888–1896.
10. Sepesi B, Watson TJ, Zhou D, et al. Are endoscopic therapies appropriate for superficial submucosal esophageal adenocarcinoma? An analysis of esophagectomy specimens. *J Am Coll Surg* 2010; 210: 418–427.
11. Alvarez Herrero L, Pouw RE, van Vilsteren FG, et al. Risk of lymph node metastasis associated with deeper invasion by early adenocarcinoma of the esophagus and cardia: Study based on endoscopic resection specimens. *Endoscopy* 2010; 42: 1030–1036.
12. Schölvinck D, Künzli H, Meijer S, et al. Management of patients with T1b esophageal adenocarcinoma: A retrospective cohort study on patient management and risk of metastatic disease. *Surg Endosc* 2016; 30: 4102–4113.
13. Kunzli HT, Belghazi K, Pouw RE, et al. Endoscopic management and follow-up of patients with a submucosal esophageal adenocarcinoma. *United European Gastroenterol J* 2018; 6: 669–677.
14. Manner H, May A, Pech O, et al. Early Barrett's carcinoma with "low-risk" submucosal invasion: Long-term results of endoscopic resection with a curative intent. *Am J Gastroenterol* 2008; 103: 2589–2597.
15. Weusten B, Bisschops R, Coron E, et al. Endoscopic management of Barrett's esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy* 2017; 49: 191–198.
16. Shaheen NJ, Falk GW, Iyer PG, et al. ACG Clinical Guideline: Diagnosis and management of Barrett's esophagus. *Am J Gastroenterol* 2016; 111: 30–50.
17. Japan Esophageal Society. Japanese classification of esophageal cancer, 11th edition: part I. *Esophagus* 2017; 14: 1–36.
18. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. *Gastrointest Endosc* 2003; 58: S3–43.
19. Endoscopic Classification Review Group. Update on the Paris classification of superficial neoplastic lesions in the digestive tract. *Endoscopy* 2005; 37: 570–578.
20. Quirke P, Dyson JE, Dixon MF, et al. Heterogeneity of colorectal adenocarcinomas evaluated by flow cytometry and histopathology. *Br J Cancer* 1985; 51: 99–106.
21. Bosman FT, Carniero F, Hruban RH, et al. *WHO classification of tumors of the digestive system*, 4th ed. Lyon, France: IARC, 2010.

22. Thomas GD, Dixon MF, Smeeton NC, et al. Observer variation in the histological grading of rectal carcinoma. *J Clin Pathol* 1983; 36: 385–391.

23. Harris EI, Lewin DN, Wang HL, et al. Lymphovascular invasion in colorectal cancer: An interobserver variability study. *Am J Surg Pathol* 2008; 32: 1816–1821.

24. Takubo K, Sasajima K, Yamashita K, et al. Double muscularis mucosae in Barrett's esophagus. *Hum Pathol* 1991; 22: 1158–1161.

25. Dutkowski P, Hommel G, Bottger T, et al. How many lymph nodes are needed for an accurate pN classification in esophageal cancer? Evidence for a new threshold value. *Hepatogastroenterology* 2002; 49: 176–180.

26. Landis JR and Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.

27. Sim J and Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther* 2005; 85: 257–268.

28. Grotenhuis BA, van Heijl M, ten Kate FJ, et al. Inter- and intraobserver variation in the histopathological evaluation of early oesophageal adenocarcinoma. *J Clin Pathol* 2010; 63: 978–981.

29. Manner H and Pech O. Measurement of the tumor invasion depth into the submucosa in early adenocarcinoma of the esophagus (pT1b): Can microns be the new standard for the endoscopist? *United European Gastroenterol J* 2015; 3: 501–504.

30. Kim JY, Kim WG, Jeon TY, et al. Lymph node metastasis in early gastric cancer: Evaluation of a novel method for measuring submucosal invasion and development of a nodal predicting index. *Hum Pathol* 2013; 44: 2829–2836.

31. Kouyama Y, Kudo S-e, Miyachi H, et al. Practical problems of measuring depth of submucosal invasion in T1 colorectal carcinomas. *Int J Colorectal Dis* 2016; 31: 137–146.

32. Cho JY, Kim YS, Jung IS, et al. Controversy concerning the cutoff value for depth of submucosal invasion after endoscopic mucosal resection of early gastric cancer. *Endoscopy* 2006; 38: 429–430; author reply 430.

33. Haboubi N and Salmo E. Are we accurately measuring the depth of the submucosal invasion in early colorectal cancer by equating the Kikuchi submucosa levels with distances measured in fractions of a millimetre? *Colorectal Dis* 2013; 15: 775–777.