DR DIPTI M. KARAMCHANDANI (Orcid ID : 0000-0002-0351-2859)
DR RAUL S GONZALEZ (Orcid ID : 0000-0001-5526-4220)
DR NICOLE PANARELLI (Orcid ID : 0000-0002-1121-267X)

Article type : Original Article

Measuring Depth of Invasion of Submucosa- Invasive Adenocarcinoma in Esophageal Endoscopic Specimens: How Good are We?

Dipti M. Karamchandani[1] M.D., Raul S. Gonzalez[2] M.D., Maria Westerhoff[3] M.D., Lindsey M. Westbrook[4] M.D., Nicole C. Panarelli[5] M.D., Mayyadah Al-Nuaimi[1] M.D., Tonya King[6] PhD, Christina A. Arnold[4] M.D.

Department of Pathology, 1-Penn State Health Milton S. Hershey Medical Center/Penn State College of Medicine, Hershey, PA, USA, 2- Beth Israel Deaconess Medical Center, 3-University of Michigan, 4- University of Colorado, 5-Albert Einstein College of Medicine; 6-Department of Public Health Sciences, Penn State College of Medicine

Corresponding author
Dipti M. Karamchandani
Associate Professor
Department of Pathology, Division of Anatomic Pathology
Penn State Health Milton S. Hershey Medical Center
500 University Drive, H179, Hershey, PA 17033-0850.
Email: mahajan.dipti@gmail.com

Parts of this study were presented as a platform presentation at the Virtual 2021 annual United States and Canadian Academy of Pathologists (USCAP) meeting.

Declarations of interest: none

Short title: Depth of Invasion of pT1b Esophagus Adenocarcinoma

Word Count: 2670 words

Abstract

Aims: Emerging data support that submucosa-invasive (pT1b) esophageal adenocarcinomas are cured via endoscopic resection provided that invasion measures ≤500 microns (μm), they lack other histologic features predictive of nodal metastasis, and have negative margins. Hence, pathologists' measurement of depth of submucosal invasion in endoscopic resections may dictate further management (i.e. endoscopic follow-up vs. esophagectomy). In this study, we assessed the interobserver agreement in measuring the depth of submucosal invasion in esophageal endoscopic resections.

Methods and Results: Six subspecialized gastrointestinal (GI) pathologists from 5 academic centers independently measured the depth of submucosal invasion in μm from the deepest muscularis mucosae on 37 esophageal endoscopic resection slides (*Round 1 scoring*). A consensus meeting with a systematic approach for measuring and discussion of pitfalls was undertaken and re-measuring (*Round 2 scoring*) was done. Interobserver agreement was assessed by the `intraclass correlation coefficient (ICC) and Cohen's kappa statistics. A lack of agreement was seen amongst the six reviewers with a *poor* ICC for both rounds: 1 [0.40, 95% CI 0.26-0.56]; 2 [0.49 ,95%CI 0.34-0.63]. When measurements were categorized as < or >500 μm, the overall agreement amongst the 6 reviewers was only *fair* for both rounds: 1[Kappa 0.37, 95% CI 0.22-0.53]; 2 [Kappa 0.29, 95%CI 0.12- 0.46].

Conclusions: Our study shows a lack of agreement among GI pathologists in measuring depth of submucosal invasion in esophageal endoscopic resections despite formulating a consensus approach for scoring. If important management decisions continue to be based upon this parameter, more reproducible and concrete guidelines are needed.

Keywords: Esophagus, adenocarcinoma, depth, invasion

Introduction

Endoscopic management is now considered the standard of care for patients with Barrett's dysplasia and early esophageal adenocarcinoma. (1-6) Recent data support that submucosa-invasive (pT1b) esophageal adenocarcinomas are cured via endoscopic resection when the following criteria are met: submucosal invasion ≤500 microns (μm), lack of other histologic features predictive of nodal metastasis (i.e., poor tumor differentiation and lymphovascular invasion), and clear margins. (1, 4, 5, 7) In these cases of "low-risk adenocarcinoma", endoscopic management is a valid alternative to esophagectomy, given the significant morbidity associated with the latter. (1, 4, 5, 7-12) Hence, the pathologists' measurement of depth of submucosal invasion in esophageal endoscopic resections may be vital for guiding further management (i.e., endoscopic follow-up vs. esophagectomy). In this study, we assessed the interobserver agreement amongst gastrointestinal (GI) pathologists in measuring the depth of submucosal invasion in esophageal endoscopic resection specimens.

Methods

The study was approved by institutional review board of all the participating authors' institutions. Six subspecialized GI pathologists from 5 different academic centers cumulatively collected, and then independently reviewed hematoxylin and eosin (H&E) stained slides from esophageal endoscopic resections with submucosal invasive (pT1b) adenocarcinoma. Endoscopic resections included slides from endoscopic submucosal dissection specimens and endoscopic mucosal resection specimens. Five of the six pathologists had at least six years of clinical sign-out experience, while one pathologist had one year of clinical sign-out experience. For the purpose of this study, the slides were reviewed solely for the depth of submucosal invasion in μm (no other histopathologic parameters were assessed). All study slides had some discernible muscularis mucosae identified. Improperly oriented specimens including tangentially sectioned specimens were excluded.

Each pathologist individually selected an area on the slide for measuring the maximal depth of submucosal invasion using their best judgement, similar to what they would do in routine clinical practice. The depth of submucosal invasion was measured for each case via computer-captured photomicrograph of the H&E slide. The pathologists used their available measuring software

(Cell Sens for 5 reviewers and NIS elements for one reviewer) for their Olympus or Nikon cameras that were properly calibrated for each objective. All the pathologists were instructed to take one image using their preferred objective that allowed them to capture and best demonstrate the entire tumor depth in their field of view. The measurement was performed independently by each pathologist, per the recommended guidelines i.e., measurement in μm from the deepest aspect of the muscularis mucosae (MM) to the deepest extent of submucosal invasion. (4, 5) This was designated as **Round 1 Scoring.** Three months' after round 1 scoring, an online **consensus meeting** via Zoom (Zoom Video Communications, San Jose, CA) between all the pathologists was undertaken. Selected (n=7) cases were re-assessed together, and a consensus approach to measure the depth of invasion and potential pitfalls were discussed among the authors based upon the review of recent literature as well as the authors' own personal practices and experience. These selected cases included 4 cases with disagreement when measurements were categorized at 500 μm (i.e. same case scored as > 500 μm and < 500 μm by different pathologists), 2 cases which were diagnosed as intramucosal carcinoma (IMC) by some pathologists and as submucosal invasion by the others, and 1 case which was measured as > 500 μm by all pathologists. The consensus approach and pitfalls are tabulated in **Table 1** and shown in **Figure 1**. Following the consensus meeting, re-measurements was performed by all the pathologists independently on all the cases on the same H&E slides, using the consensus approach developed in the meeting. This was designated as **Round 2 Scoring**. A photograph displaying the measurement was taken by each reviewer for every case scored in Round 2, again using their preferred objective. The pathologists were blinded from each other's results for both rounds of scoring.

Statistical analysis

Interobserver agreement was assessed by the intraclass correlation coefficient (ICC) for continuous variables and by Cohen's kappa statistic for categorical variables. When conducting a reliability study, it is recommended to obtain at least 30 heterogeneous samples and involve at least 3 raters whenever possible. Under such conditions, ICC values less than 0.5 are indicative of poor reliability, values of 0.5-0.75 indicate moderate reliability, values of 0.75-0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability. (13) Cohen suggested the Kappa results be interpreted as follows: values ≤ 0 as indicating no agreement and 0.01–0.20

as none to slight, 0.21–0.40 as fair, 0.41– 0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement. (14)

Results

The six participating pathologists cumulatively collected 37 H&E stained slides from 34 esophageal endoscopic specimens originally diagnosed as submucosa-invasive (pT1b) adenocarcinoma by the contributing pathologist. For the sake of numbers, different slide(s) of the same lesion from the same patient were accepted as being separate measurable test cases (6 slides from 3 endoscopic specimens; 1 slide each from 31 endoscopic specimens).

Round 1 Scoring

The scoring results for all six reviewers for Round 1 are tabulated in **Table 2**. Of the 37 cases, 28 were scored as demonstrating submucosal invasion by all six reviewers, and 9 were scored as submucosal invasion by 4 or 5 reviewers and as IMC by the remaining one or two reviewers **(Supplementary Figure 1)**. The lack of agreement among the 6 reviewers can be seen by the variation around the line of agreement in each plot (**Figure 2)**. The overall estimate of agreement (ICC) among the 6 reviewers in measuring the depth of invasion was moderate [0.64, 95% CI 0.51-0.76]; however, this agreement was skewed higher due to a single outlying scoring observation (with great agreement; all measured >2000 μm) that was identified along the line of agreement, seen in the plots (**Figure 2**). When this outlying observation was excluded, the ICC was poor [0.48, 95% CI 0.33-0 .63]. When the measurements were categorized as either > or ≤ 500 μm, 1 case (3%) was measured as ≤500 μm (including IMC by 2 reviewers) and 14 cases (38%) were measured >500 μm by all reviewers **(Figure 3)**. There was disagreement in 22 cases (59%), i.e., a case measured variably as >500 μm and ≤ 500 μm by six reviewers **(Figure 4)**. The percentage of measurements scored as >500 μm varied among the reviewers from 49% to 90% (**Table 2).** The overall agreement was only fair [Cohen Kappa 0.37, 95% CI 0.22-0.53] among the reviewers when measurements were categorized at 500 μm.

Round 2 Scoring

The re-scoring results for all six reviewer's post-consensus meeting are shown in **Table 2**. Of the 37 cases, 34 were scored as submucosal invasion by all six reviewers, and 3 were

scored as submucosal invasion by 3 or more reviewers and as IMC by the remaining reviewers **(Supplementary Figure 2)**. The lack of agreement among the measurements between reviewers can be again seen by the variation around the line of agreement in each plot (**Figure 2)**. The overall ICC among the 6 reviewers remained moderate [0.60, 95%CI 0.46-0.72]. There still was an outlying observation (with great agreement; all measured >2000 μm) (**Figure 2**) and when the furthest outlying observation along the line of agreement was removed, the ICC indicated poor agreement [0.49, 95%CI 0.34- 0.63]. The overall estimate of agreement among the 6 reviewers when measurements are categorized at 500 μm indicated only fair agreement [0.29, 95% CI 0.12-0.46]. The percentage of measurements scored as > 500 μm varied among the reviewers from 65% to 95%. When the measurements were categorized at 500 μm, there was disagreement in 17 cases (46%) **(Figure 5)**, while in the remaining 20 cases, all reviewers scored as >500 μm (54%) **(Figure 6)**.

Agreement between each pathologist's scoring for rounds 1 and 2

The agreement for each pathologist between their round 1 and round 2 measurements ranged from 0.45 (poor; reviewer 2) to 0.87 (good, reviewer 1); both these reviewers had at least 6 years of clinical experience. The agreement for other reviewers were 0.67 (reviewers 3 and 4), 0.84 (reviewer 5), and 0.48 (reviewer 6). Of note, since round 2 measurements were undertaken after discussion and establishment of a "consensus" approach, this was not considered as a true "intra-observer reliability".

Discussion

Many recent papers have emphasized that the curative status of an endoscopic resection depends upon adequacy of resection and risk of nodal metastases. The latter cannot be evaluated upon histologic examination of endoscopic resection specimens; however, certain histopathologic features are predictive of lymph node metastases, i.e., depth of invasion, poor differentiation, and lymphovascular invasion. (4, 5, 15, 16) Although recently published studies suggest that tumor budding may be an independent prognostic factor in resected esophageal adenocarcinoma, unlike similar colonic specimens, (17) currently there are no strong guidelines to report tumor budding in esophageal endoscopic resections. (18, 19)

The current guidelines support that esophageal adenocarcinoma invading into the upper one-third of the submucosa with the absence of other negative histological risk factors has a low-risk for lymph node metastasis and is regarded as low-risk adenocarcinoma.(7-11) The pragmatic classification of dividing submucosa into three equal thirds (sm1, superficial one-third submucosa; sm2, intermediate one-third submucosa; sm3, outer one-third submucosa) is considered reliable in esophagectomy specimens wherein the whole submucosa can be visualized and muscularis propria is present as a landmark for subdivision of submucosa; however this classification scheme can be challenging and inaccurate in endoscopic resections. Therefore, the Paris endoscopic classification of superficial neoplasms recommends measurement of submucosal invasion in microns as an alternative, wherein 500 μm roughly corresponds to sm1. (20-22). Recent endoscopic data report a lower risk of nodal metastases (0–2%) when submucosal invasion is ≤500 μm and there are no other associated negative histological risk factors. (4, 5, 7, 10, 22-25) Given this recent data, measuring depth of invasion within the submucosa has become an increasingly reported histopathologic parameter in esophageal endoscopic resections. However, if significant treatment decisions are to be based upon this parameter, it needs to be accurate, and reproducible amongst pathologists. Our study shows a lack of agreement among GI pathologists in assessing the depth of submucosal invasion in esophagus endoscopic resections, even after developing a consensus approach for measurement, and discussing potential pitfalls. Our results are similar to a recent study wherein the authors showed a substantial discordance among three GI pathologists for the assessment of depth of submucosal invasion depth in pT1b esophageal adenocarcinoma. The authors concluded that the discordance may potentially lead to a false estimation of risk of nodal metastases with serious implications for further therapy. (26) The same study found that there was good to excellent agreement between pathologists for the histological assessment of differentiation grade and lymphovascular invasion in pT1b adenocarcinoma. Of note, their study included both esophageal endoscopic resection specimens (n=35) and esophagectomy specimens (n=50). Additionally, only one round of scoring was done and no consensus guidelines were discussed.

Given the crucial cut-off of 500 μm, wherein close surveillance is recommended for ≤500 μm invasion and additional esophagectomy with lymphadenectomy is recommended for invasion depth of >500 μm (7, 27-33), the above results raise concern and question the validity of using this guideline for clinical decision making purposes. Although the guideline for measurement

(i.e., measure from deepest aspect of MM to deepest extent of submucosal invasion) may appear straightforward, it can be difficult in routine clinical sign-out. The substantial discordance may be explained by subjective interpretation of the deepest layer of MM and deepest invasion, along with several other practical difficulties associated with the measurement of submucosal invasion. Even in properly fixed and well-oriented specimens, the deepest layers of MM may be destroyed by the invading tumor, leaving no discernible MM present above the deepest extent of invasion. In these cases, one approach used is to approximate the deepest extent of MM by drawing an imaginary line from the base of the existing deepest MM (which is usually comprised of a thicker muscle bundle, rather than wisps of smooth muscle) at the shoulder(s) of the invasive carcinoma and using this line as the deepest point of MM. Even this approach can be very challenging given the well-known MM abnormalities reported to be present in more than 90% of the Barrett's resections (34-36). These include distortion and duplication of MM, and the presence of a discontinuous or hypertrophic MM, which further adds to the wide subjective assessment of the deepest layer of MM, as seen in our study. These abnormalities of the MM, along with variability in method of measurement (whether from the deepest visible MM or from an imaginary line of MM) have been postulated to be a cause of disagreement in other sites within the GI tract as well. (37, 38). We also found that the pathologists were not uniform in selecting the area with deepest invasion on the same slide, adding to the discordance in measurements amongst pathologists. Lastly, tissue artifacts during fixation and processing along with specimen curling may make it difficult to assess the focus of deepest invasion as well as the orientation plane, further adding to subjectivity in scoring.

Despite the numerous challenges that lead to poor reproducibility even after a consensus meeting, we did identify several important pitfalls which pathologists should be aware of. First, the space between the duplicated muscle layers may resemble submucosa. We found it helpful to compare the wall thickness of blood vessels in this space to those in areas that were clearly submucosal in order to more readily distinguish lamina propria from submucosa, given that submucosal vessels are thick-walled. Second, we found that fibromuscularization of the submucosa produced by tumoral infiltration can mimic MM, particularly in a background of diathermy artifact. Third, that the muscular layer of thick wall submucosal vessels can simulate wisps of MM when sectioned tangentially. Of note, prior studies have shown that distinction of IMC from pT1b tumors itself is less than perfect with an overall kappa value of 0.71 and a kappa

of 0.76 between the GI pathologists even in esophagectomy specimens. (39, 40) In our study, there were a few endoscopic resection cases diagnosed as submucosal invasion by most reviewers, but as IMC by some.

A strength of our study is that we performed two rounds of measurements: the first round simulated authors' routine clinical practice, and the second round was after discussion of a consensus approach and pitfalls. Despite that, there remained a wide variation in scoring amongst pathologists, emphasizing that measuring depth of submucosal invasion remains challenging in clinical practice. Also, esophagectomy specimens were not included in our study in an attempt to mimic the real-life decision process on endoscopic specimens. A limitation of the study was that these endoscopic resection specimens were collected, fixed and processed at 5 different institutions with possibly different protocols. However, given that all pathologists reviewed all study slides, we believe that this would not affect the interobserver agreement for scoring. Also, this simulates real-world clinical practice, as even many tertiary academic centers in United States do not pin endoscopic specimens before fixation and frequently evaluate slides prepared at other institutions. A recent study compared the pinning and non-pinning methods and failed to demonstrate a statistically significant difference for clinically relevant histopathologic parameters, including depth of tumor infiltration. (41) Also, since the authors of this study typically measure the depth of invasion on H&E stained slides in routine clinical sign-out practice; ancillary stains were not performed in this study to simulate routine sign-out decision making process. Of note, the agreement for scoring depth of invasion did not improve by use of immunostains in a recent study. (26) We also didn't attempt to look at interobserver agreement of other histologic features typically reported in endoscopic resections, given that studies have already documented good to excellent agreement between pathologists for the histological assessment of differentiation grade and lymphovascular invasion in pT1b adenocarcinoma. (26)

To conclude, our study shows a lack of agreement among GI pathologists in measuring the depth of submucosal invasion in esophageal endoscopic resections performed for pT1b adenocarcinoma. If important clinical management decisions continue to be based upon this parameter, the recommendations need to be revisited with more reproducible and concrete guidelines.

Acknowledgements

None

Funding

Nothing to disclose.

References

1. Bennett C, Vakil N, Bergman J, et al. Consensus statements for management of Barrett's dysplasia and early-stage esophageal adenocarcinoma, based on a Delphi process. Gastroenterology. 2012;143:336-346.

2. DeMeester SR, DeMeester TR. The diagnosis and management of Barrett's esophagus. Adv Surg. 1999;33:29-68.
3. Gockel I, Exner C, Junginger T. Morbidity and mortality after esophagectomy for esophageal carcinoma: a risk analysis. World J Surg Oncol. 2005;3:37.
4. Kumarasinghe MP, Armstrong M, Foo J, et al. The modern management of Barrett's oesophagus and related neoplasia: role of pathology. Histopathology. 2021;78:18-38.
5. Kumarasinghe MP, Bourke MJ, Brown I, et al. Pathological assessment of endoscopic resections of the gastrointestinal tract: a comprehensive clinicopathologic review. Mod Path. 2020;33:986-1006.
6. Seewald S, Ang TL, Pouw RE, et al. Management of Early-Stage Adenocarcinoma of the Esophagus: Endoscopic Mucosal Resection and Endoscopic Submucosal Dissection. Dig Dis Sci. 2018;63:2146-2154.
7. 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.
8. Ancona E, Rampado S, Cassaro M, et al. Prediction of lymph node status in superficial esophageal carcinoma. Ann Surg Oncol. 2008;15:3278-3288.
9. Gockel I, Sgourakis G, Lyros O, et al. Risk of lymph node metastasis in submucosal esophageal cancer: a review of surgically resected patients. Expert Rev Gastroenterol Hepatol. 2011; 5:371-384.
10. Manner H, Pech O, Heldmann Y, et al. Efficacy, safety, and long-term results of endoscopic treatment for early stage adenocarcinoma of the esophagus with low-risk sm1 invasion. Clinical Gastroenterol Hepatol. 2013;11:630-635.
11. Peters FP, Brakenhoff KP, Curvers WL, et al. Histologic evaluation of resection specimens obtained at 293 endoscopic resections in Barrett's esophagus. Gastrointest Endosc. 2008;67:604-609.
12. Metzger R, Bollschweiler E, Vallböhmer D, et al. High volume centers for esophagectomy: What is the number needed to achieve low postoperative mortality? Dis Esophagus. 2004;17:310-314.
13. Portney LG, Watkins MP. Prentice Hall; New Jersey: 2000. Foundations of clinical research: applications to practice.585-618.

14. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37-46.

15. Sgourakis G, Gockel I, Lang H. Endoscopic and surgical resection of T1a/T1b esophageal neoplasms: a systematic review. World J Gastroenterol. 2013;19:1424-1437.

16. Vieth M, Stolte M. Pathology of early upper GI cancers. Best Pract Res Clin Gastroenterol. 2005;19:857-869.

17. Lohneis P, Hieggelke L, Gebauer F, et al. Tumor budding assessed according to the criteria of the International Tumor Budding Consensus Conference determines prognosis in resected esophageal adenocarcinoma. Virchows Archiv. 2021;478:393-400.

18. Berg KB, Schaeffer DF. Tumor budding as a standardized parameter in gastrointestinal carcinomas: more than just the colon. Mod Path. 2018;31:862-872.

19. Lugli A, Kirsch R, Ajioka Y, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Path. 2017;30:1299-1311.

20. 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.

21. 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.

22. Update on the paris classification of superficial neoplastic lesions in the digestive tract. Endoscopy. 2005;37:570-578.

23. Liu L, Hofstetter W, Rashid A, et al. Significance of the depth of tumor invasion and lymph node metastasis in superficially invasive (T1) esophageal adenocarcinoma. Am J Surg Pathol. 2005;29:1079-1085.

24. Hölscher AH, Vallböhmer D, Bollschweiler E. Early Barrett's carcinoma of the esophagus. Ann Thorac Cardiovasc Surg. 2008;14:347-354.

25. Bourke MJ Neuhaus H, Bergman JJ. Endoscopic submucosal dissection: indications and application in western endoscopy practice. Gastroenterology. 2018;154:1887–1900.

26. Gotink AW, Ten Kate FJ, Doukas M, et al. Do pathologists agree with each other on the histological assessment of pT1b oesophageal adenocarcinoma? United European Gastroenterol J. 2019;7:261-269.

27.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.
28. 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.
29. 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.
30. 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.
31. 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.
32. Shaheen NJ, Falk GW, Iyer PG, et al. American College of Gastroenterology. ACG Clinical Guideline: Diagnosis and management of Barrett's esophagus. American J Gastroenterol. 2016;111:30-50.
33. Künzli 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.
34. Abraham SC, Krasinskas AM, Correa AM, et al. Duplication of the muscularis mucosae in Barrett esophagus: an underrecognized feature and its implication for staging of adenocarcinoma. Am J Surg Pathol. 2007;31:1719-1725.
35. Kaneshiro DK, Post JC, Rybicki L, et al. Clinical significance of the duplicated muscularis mucosae in Barrett esophagus-related superficial adenocarcinoma. Am J Surg Pathol. 2011;35:697-700.
36. Lewis JT, Wang KK, Abraham SC. Muscularis mucosae duplication and the musculo-fibrous anomaly in endoscopic mucosal resections for barrett esophagus: implications for staging of adenocarcinoma. Am J Surg Pathol. 2008;32:566-571.

37. 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.
38. Kouyama Y, Kudo SE, Miyachi H, et al. Practical problems of measuring depth of submucosal invasion in T1 colorectal carcinomas. Int J Colorectal Dis 2016;31:137-146.
39.Ormsby AH, Petras RE, Henricks WH, et al. Observer variation in the diagnosis of superficial oesophageal adenocarcinoma. Gut. 2002;51:671-676.
40. Worrell SG, Boys JA, Chandrasoma P, et al. Inter-Observer Variability in the Interpretation of Endoscopic Mucosal Resection Specimens of Esophageal Adenocarcinoma: Interpretation of ER specimens. J Gastrointest Surg. 2016;20:140-144; discussion 144-145.
41. Overwater A, van der Meulen KE, Künzli HT, et al. Optimizing histopathologic evaluation of EMR specimens of Barrett's esophagus-related neoplasia: a randomized study of 3 specimen handling methods. Gastrointest Endosc. 2019;90:384-392.e5.

Figure Legends:

Figure 1. Approach and Pitfalls discussed during consensus meeting. A. Submucosal thick walled vessels and stromal desmoplasia confirms submucosal invasion (lower thin bracket). Duplication of muscularis mucosae (MM) is seen (arrows), and the carcinoma within the duplicated muscle represents intramucosal carcinoma (upper thick bracket). The black thin curved line represents deepest aspect of visible MM and the blue thick line represents the imaginary (approximated) site of deepest MM destroyed by the tumor; B. Smooth muscle wall of the thick walled submucosal blood vessels (in tangential sections) can mimic the wisps of MM (black arrows); C. Wisps of muscle identified in the submucosal tissue (highlighted by thick black arrows) can cause problems with staging and scoring depth of submucosal invasion. This carcinoma is clearly submucosal given the thick-walled submucosal blood vessel (single long thin arrow) and the presence of submucosal esophageal glands (double long thin arrows); D.

Diathermy artifact of the submucosal tissue can mimic MM (double arrows). Dilated, large caliber vessels can be seen in lamina propria and cause confusion with submucosa (long thin arrow); however, they are not thick walled as submucosal vessels (short thick black arrow). (A. Hematoxylin &eosin x 40; B. D. Hematoxylin &eosin X100; C.- Hematoxylin &eosin x200)

Figure 2: Matrix of scatterplots showing the lack of agreement (i.e. variation around the line of agreement in each plot) among the six reviewers for the measurements of depth of submucosal invasion in esophageal endoscopic resections for both Round 1 (left panel) and Round 2 (right panel). The circle highlights the single outlying observation which was close to the line of agreement in each plot.

Figure 3: Example of a case in Round 1 where all six reviewers measured > 500 microns; however, measurements varied from 707 microns up to 993 microns. (R- reviewer; R1, R3-R5- Hematoxylin &eosin x 100; R2, R6- Hematoxylin &eosin x 40)

Figure 4: Example of a case with disagreement in Round 1 among the six reviewers when the measurements were categorized at 500 microns. Three reviewers (R1, R2, and R5) measured > 500 microns and three reviewers (R3, R4 and R6) measured < 500 microns. (R- reviewer; R1, R3- R6- Hematoxylin &eosin x 100; R2- Hematoxylin &eosin x 200)

Figure 5: Example of a case with disagreement in Round 2 among the six reviewers when the measurements were categorized at 500 microns. Four reviewers (R1, R2, R3 and R5) measured < 500 microns and two reviewers (R4 and R6) measured >500 microns. The blue curved line in R4 represents the reviewer's assessment of the approximated (imaginary) muscularis mucosae (R- reviewer; R1, R3, R5- Hematoxylin &eosin x 100; R2- Hematoxylin &eosin x 40, R4, R6- Hematoxylin &eosin x 20)

Figure 6: Round 2 measurements (after the consensus meeting) for the same case as Figure 3. All six reviewers still measured > 500 microns; however, measurements varied from 708 microns up to 1675 microns. (R- reviewer; R1, R2, R4, R5, R6- Hematoxylin &eosin x 40; R3- Hematoxylin &eosin x 100)

Supplementary Figure 1: Example of a case in Round 1 wherein measurements of submucosal invasion varied from 227 microns to 1629 microns among five reviewers and one diagnosed intramucosal adenocarcinoma. (R- reviewer; R1, R4- Hematoxylin &eosin x 100; R2, R5- Hematoxylin &eosin x 40x, R6- Hematoxylin &eosin x 200)

Supplementary Figure 2: Example of a case in Round 2 wherein measurements of submucosal invasion varied from 269 microns to 1190 microns among five reviewers and one diagnosed intramucosal adenocarcinoma. (R- reviewer; R1- Hematoxylin &eosin X100; R2- Hematoxylin &eosin X20; R3, R5- Hematoxylin &eosin x 200x; R4- Hematoxylin &eosin x 40)

Table 1: Consensus approach formulated for the re-measuring (Round 2 scoring) of esophageal endoscopic resection specimens

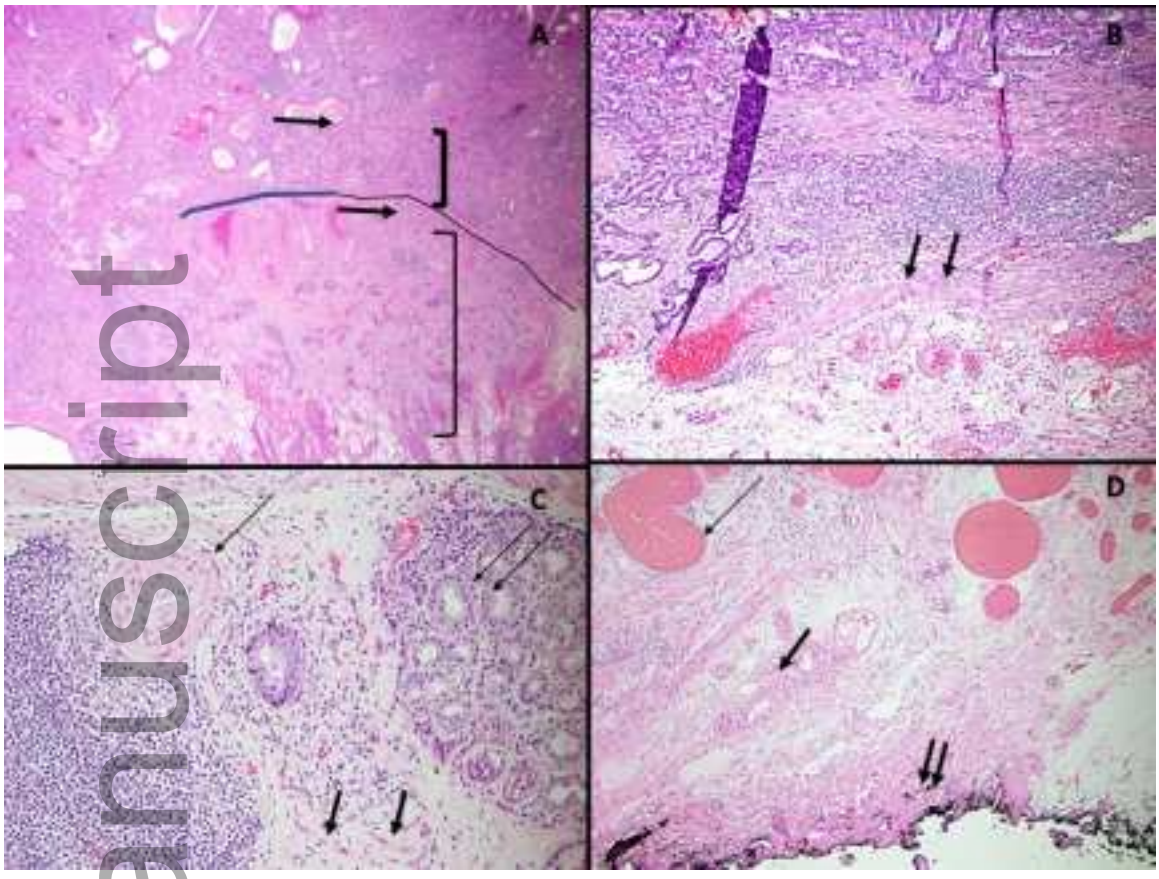
1.	Confirm the diagnosis of submucosal invasive (pT1b) esophageal adenocarcinoma on the endoscopic resection specimen. Stromal desmoplasia, tumor adjacent to thick-walled submucosal vessels and/or esophageal ducts and/or submucosal glands are the best indicators of submucosal invasion (**Figure 1A**).
2.	Look at all the fragments on the slides. Choose the fragment with the deepest extent of submucosal invasion by eyeballing and/or by measuring if there is confusion as to which fragment shows the greatest depth of invasion.
3.	Be wary of muscularis mucosae (MM) abnormalities that are commonly seen in endoscopic resections (**Figure 1A).** Depth of invasion in the submucosa should be measured in microns from the outermost extent of the outer (deeper) MM to the deepest point of invasion
4.	Deepest layers of MM are often destroyed by the invading tumor. In these cases, try to approximate the deepest extent of MM by drawing an imaginary line from the base of the existing deepest MM (which is usually comprised of a thicker muscle bundle, rather than wisps of smooth muscle) at the shoulder(s) of the invasive carcinoma. Use this approximated (imaginary) line as the deepest point of MM (**Figure 1A**).
5.	Draw a vertical line from visible or approximated (imaginary) deepest MM to the deepest extent of tumor in the submucosa and measure.
6.	If neoplastic glands do not extend beyond the bottom aspect of the imaginary MM compared to that of the adjacent intact MM, they are best classified as intramucosal carcinoma (pT1a).

PITFALLS

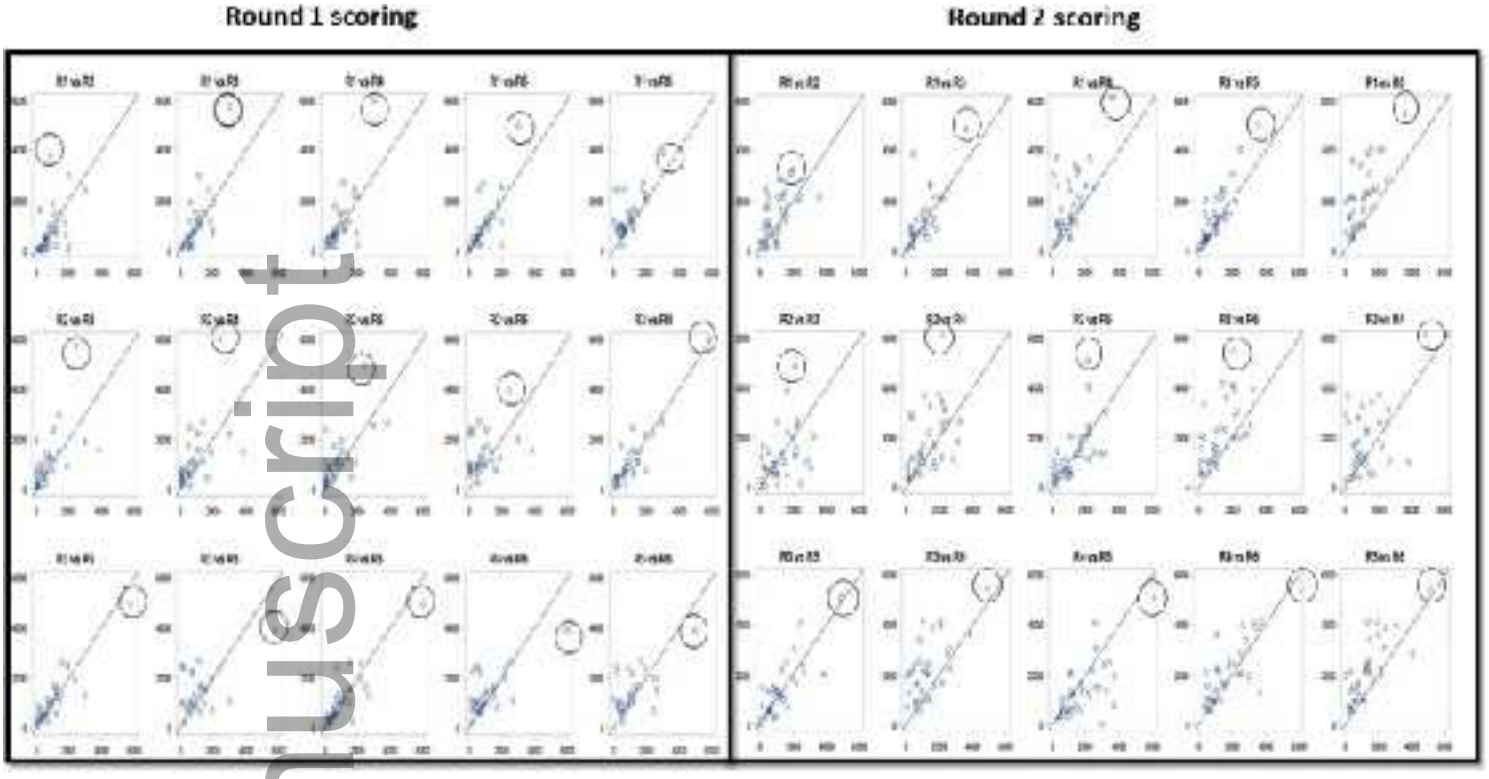
1. Smooth muscle walls of thick-walled submucosal blood vessels (in tangential sections) can mimic the wisps of MM. **(Figure 1B)**
2. Wisps of muscle when identified in the submucosal tissue can cause problems with depth of invasion. **(Figure 1C)** A better indicator is to always look at the shoulder of intact deepest MM adjacent to the invasive tumor to get the best estimate of the location of deepest MM.
3. Diathermy artifact of the submucosal tissue can mimic MM. (**Figure 1D)**
4. Dilated, large-caliber vessels are seen in lamina propria within duplicated MM. Do not confuse them with submucosal vessels, which are thick-walled. (**Figure 1D**)

Table 2: Results of the depth of submucosal invasion measurements in esophageal endoscopic resection specimens among the six reviewers for Round 1 Scoring and Round 2 scoring

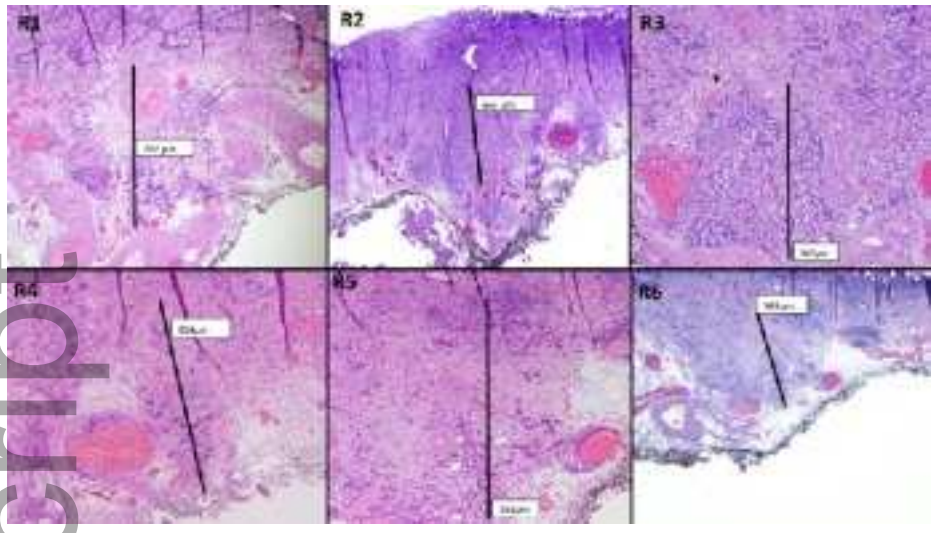
Reviewer No.(R)	ROUND 1			ROUND 2		
	Number of cases diagnosed as intramucosal adenocarcinoma	Number of cases scored as submucosal invasion ≤ 500 microns	Number of cases scored as submucosal invasion > 500 microns	Number of cases diagnosed as intramucosal adenocarcinoma	Number of cases scored as submucosal invasion ≤ 500 microns	Number of cases scored as submucosal invasion > 500 microns
R1	0 (0%)	9 (24%)	28 (76%)	0 (0%)	13 (35%)	24 (65%)
R2	2 (6%)	17 (46%)	18 (49%)	0 (0%)	9 (24%)	28 (76%)
R3	5 (14%)	10 (27%)	22 (59%)	2 (5%)	6 (16%)	29 (78%)
R4	3 (8%)	9 (24%)	25 (68%)	0 (0%)	3 (8%)	34 (92%)
R5	3 (8%)	9 (24%)	25 (68%)	1 (3%)	7 (19%)	29 (78%)
R6	3 (8%)	1 (3%)	33 (90%)	2 (5%)	0 (0%)	35 (95%)
Total reads	16	55	151	5	38	179



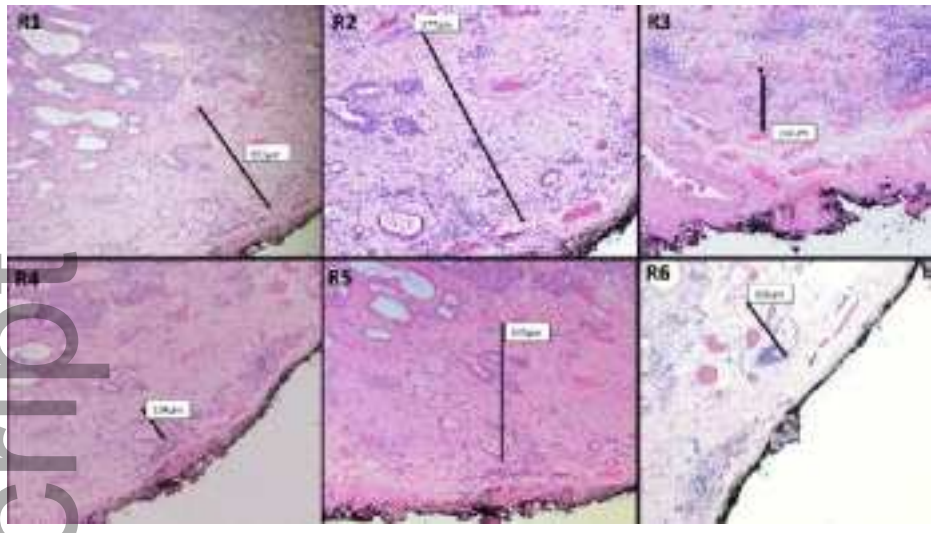
his_14566_f1.tiff



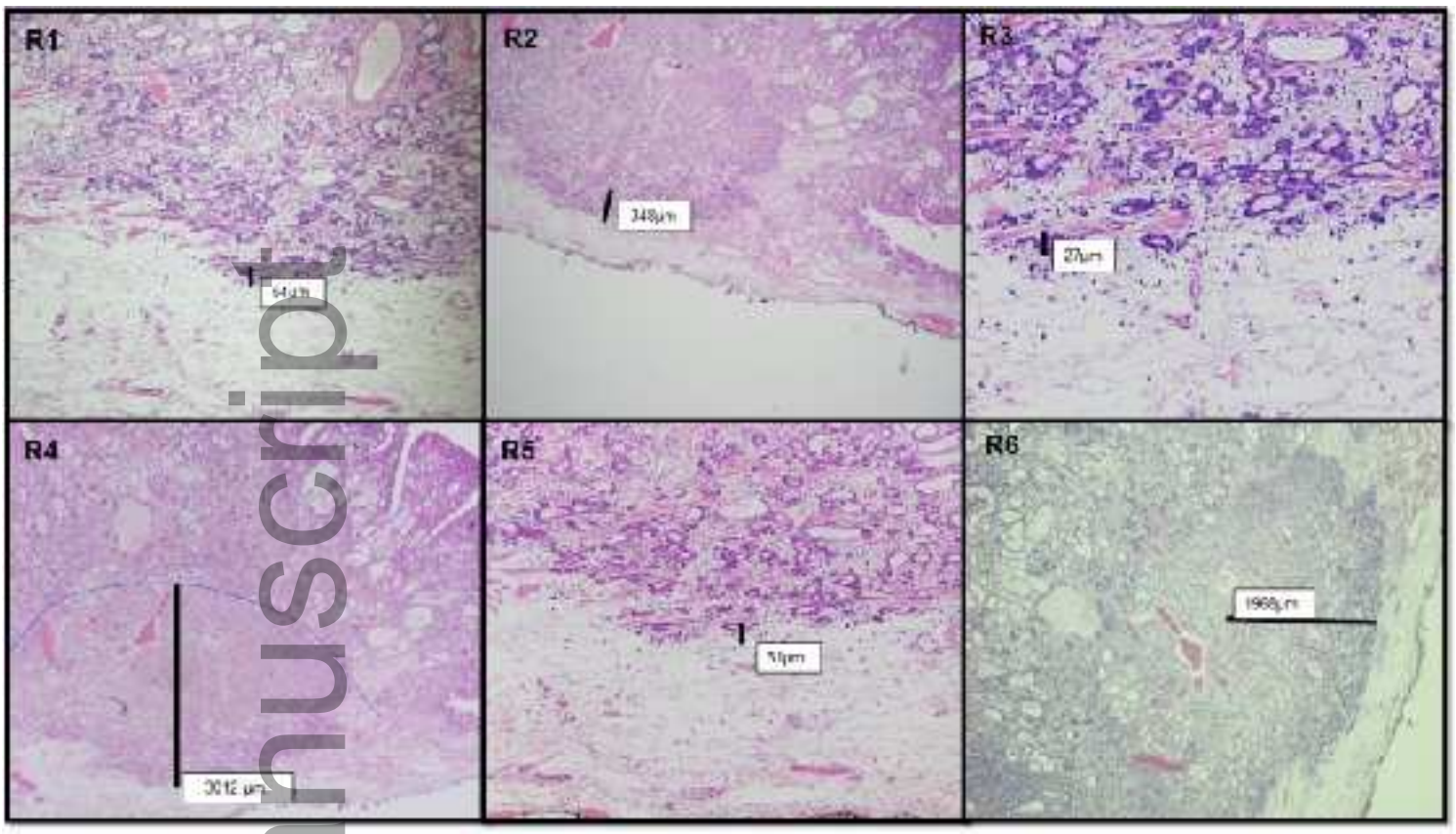
his_14566_f2.tif



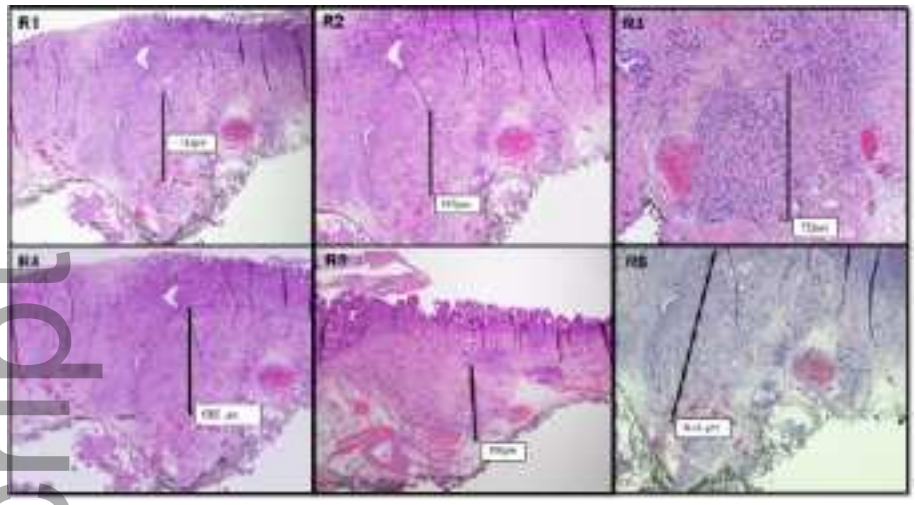
his_14566_f3.tif



his_14566_f4.tif



his_14566_f5.tif



his_14566_f6.tif