Which factors make Barrett’s esophagus lesions difficult to diagnose?

Authors
Ippei Tanaka1, Dai Hirasawa1, Kenjiro Suzuki1, Syuhei Unno1, Shin Inoue1, Satoshi Ito1, Jyunichi Togashi1, Junichi Akahira2, Fumiyoshi Fujishima3, Tomoki Matsuda1

Institutions
1 Department of Gastroenterology, Sendai Kousei Hospital, Miyagi, Japan
2 Department of Anatomic Pathology, Sendai Kousei Hospital, Miyagi, Japan
3 Department of Anatomic Pathology, Tohoku University School of Medicine, Miyagi, Japan

Background and study aims
Although the Japan Esophageal Society’s magnifying endoscopic classification for Barrett’s epithelium (JES-BE) offers high diagnostic accuracy, some cases are challenging to diagnose as dysplastic or non-dysplastic in daily clinical practice. Therefore, we investigated the diagnostic accuracy of this classification and the clinicopathological features of Barrett’s esophagus cases that are difficult to diagnose correctly.

Patients and methods
Five endoscopists with experience with fewer than 10 cases of magnifying observation for superficial Barrett’s esophageal carcinoma reviewed 132 images of Barrett’s mucosa or carcinoma (75 dysplastic and 57 non-dysplastic cases) obtained using high-definition magnification endoscopy with narrow-band imaging (ME-NBI). They diagnosed each image as dysplastic or non-dysplastic according to the JES-BE classification, and the diagnostic accuracy was calculated. To identify risk factors for misdiagnosed images, images with a correct rate of less than 40% were defined as difficult-to-diagnose, and those with 60% or more were defined as easy-to-diagnose. Logistic regression analysis was performed to identify risk factors for difficult-to-diagnose images.

Results
The sensitivity, specificity and overall accuracy were 67%, 80% and 73%, respectively. Of the 132 ME-NBI images, 34 (26%) were difficult-to-diagnose and 99 (74%) were easy-to-diagnose. Logistic regression analysis showed low-grade dysplasia (LGD) and high-power magnification images were each significant risk factors for difficult-to-diagnose images (OR: 6.80, \( P = 0.0017 \) and OR: 3.31, \( P = 0.0125 \), respectively).

Conclusions
This image assessment study suggested feasibility of the JES-BE classification for diagnosis of Barrett’s esophagus by non-expert endoscopists and risk factors for difficult diagnosis as high-power magnification and LGD histology. For non-experts, high-power magnification images are better evaluated in combination with low-power magnification images.

Introduction
The incidence of Barrett’s esophageal adenocarcinoma (BAC) has increased rapidly in the Western countries over the past few decades and it now accounts for more than half of all esophageal cancers [1,2]. In contrast, BAC remains rare in Japan. However, the number of Japanese patients with BAC is expected to increase in the future due to the decreasing rate of Helicobacter pylori infections and the increasing incidence of
gastroesophageal reflux disease [3,4]. The 5-year survival rate for patients with BAC, including those with locally advanced disease, is reportedly less than 20 % [5]. The poor survival rate for patients with BAC necessitates its early detection [6,7]. To detect a BAC lesion in the early stage, the Seattle protocol is used for surveillance of Barrett’s lesions in Western countries [8]. However, it is expensive and time-consuming, and involves a high risk of sampling error [9]. Targeted biopsy using narrow-band imaging (NBI) may overcome these weaknesses, and several NBI classification systems have been proposed [10–14]. However, these classifications have not gained worldwide acceptance, as they involve complicated and diverse criteria.

The Japan Esophageal Society (JES) proposed a new magnifying endoscopic classification for predicting the histology of Barrett’s epithelium (JES-BE classification) in 2018 [15]. The diagnostic flow of this classification is simply described as follows: magnifying endoscopic findings are composed of mucosal and vascular patterns, and each are classified as “regular” or “irregular.” Then, the histology, non-dysplastic or dysplastic, was comprehensively predicted. Compared to the other NBI classification, the JES-BE classification has more simplified criteria including a flat pattern. These characteristics lead to the high diagnostic accuracy of this classification. The sensitivity and specificity of this classification were reported to be 87% and 97%, respectively [16]. Consequently, this classification is now widely accepted by Japanese endoscopists. However, in daily clinical practice, we often observe Barrett’s lesions that are difficult to diagnose as non-dysplastic or dysplastic by magnifying endoscopy with NBI (ME-NBI), even using the JES-BE classification. In addition, to date, it is unclear for which cases additional care should be taken in the diagnosis of BAC.

**Patients and methods**

**Study design**

This image assessment study comprised the following two stages: 1) diagnosis by endoscopists of images as dysplasia or non-dysplasia based on the JES-BE classification; and 2) assessment of the diagnostic accuracy of the JES-BE classification and investigation of the features of the misdiagnosed images, such as histology, magnification power, and location of captured images.

**Image selection**

First, 1372 ME-NBI images from 98 successive patients with Barrett’s epithelium (BE) or Barrett’s dysplasia/early cancer obtained between January 2012 and July 2019 were retrieved from the endoscopic database of the Sendai Kousei Hospital. All images were captured using a high-definition magnification endoscope (GIF-H260Z, Olympus Corp., Tokyo, Japan) and a videendoscopy system (EVIS LUCERA SPECTRUM, CV-260SL and CLV-260SL, Olympus Corp.). This scope has optical zoom function that enables adjustment of magnification level. Second, among these images, only the images captured just before taking biopsy samples were selected. Therefore, the site of the selected images and the site of biopsies were in one-to-one correspondence. The images of dysplastic lesion were captured before taking biopsies for tumor diagnosis. In contrast, the images of non-dysplastic epithelium were captured before taking negative biopsies necessary for diagnosis of tumor demarcation. Images of the biopsy site that were taken because of suspicion for dysplastic lesions but that were pathologically non-dysplastic were not included in this study. Images were excluded from the study if the view was indistinct (i.e., with blood or mucus or out of focus). If several images were taken from a single biopsy site, the best quality image was selected. Finally, 142 ME-NBI images from 50 patients diagnosed histopathologically were chosen for the present study. Image collection and selection were performed by the principal investigator (IT).

**JES-BE classification**

*Fig. 1* shows the diagnostic flowchart for the JES-BE classification [15]. First, a mucosal pattern is assessed under low-power magnifying observation (regular/irregular). Then, a vascular pattern is evaluated under fully magnifying observation (regular/irregular). Finally, a comprehensive diagnosis to predict histology is made based on mucosal and vascular patterns. If both mucosal and vascular patterns are regular, the lesion is considered non-dysplastic; if irregular, it is considered dysplastic. If judgment about regularity of the mucosal and vascular patterns differs, the overall judgment should be irregular and the lesion considered to dysplastic. In addition, this classification includes the modified flat pattern, which is defined as flat mucosa (none of pits and villi: i.e., absent pattern) without clear demarcation with green thick vessels. It is classified into a “regular” mucosal/vascular pattern and suggested to be non-dysplastic BE.

**Reviewer recruitment**

Five endoscopists at Sendai Kousei Hospital were enrolled as image reviewers. They had experience with more than 100 cases of magnification endoscopy for superficial gastric cancer but had fewer than 10 cases of NBI magnifying observation for superficial BAC, and therefore, were considered general endoscopists in Japan. The principal investigator (IT) who prepared the ME-NBI images was not included as a reviewer. In addition, to reflect real-world clinical practice, experts in the field of esophageal cancer were not included as reviewers.

**Image preparation**

First, the 142 ME-NBI images were divided into two groups: training images (n = 10) and test images (n = 132). The 132 images were pooled into a single dataset and inserted into Microsoft PowerPoint (Windows 2010; Microsoft, Santa Clara, California, United States) presentation slides against a black background. Each of the 132 images was numbered and randomly arranged according to random number tables created by Microsoft Excel (Windows 2010; Microsoft, Santa Clara, California, United States).
Image assessment

A short lecture about JES-BE classification using 10 sheets of training images was conducted by an expert in magnifying endoscopic diagnosis of Barrett’s esophageal cancer (DH). After the lecture, each of the five endoscopists diagnosed the mucosal and vascular pattern of 132 images according to the JES-BE classification, and predicted the histology of these images (i.e., non-dysplastic or dysplastic). The reviewers had no access to the patient’s clinical information, histological data, or other imaging material.

Histology

Biopsy samples were embedded in paraffin and stained with hematoxylin and eosin, and immunohistologically, such as p53 and Ki-67. All samples were evaluated by two expert pathologists specializing in BAC diagnosis who were blinded to any endoscopic findings. When the diagnoses of the lesions by the two pathologists differed, the final diagnosis was reached after they had a discussion and reach consensus. The histological diagnosis was established based on the Vienna classification [17]. The pathologist graded the samples as non-dysplastic, low-grade dysplasia (LGD), high-grade dysplasia (HGD), or adenocarcinoma. Because pathologists in Japan are not familiar with the diagnosis of LGD, the description and reference images of LGD in the Vienna classification were always reviewed before diagnosing LGD (Fig. 2a,b) [17]. In the current study, LGD, HGD, and adenocarcinoma were classified as dysplastic.

Outcomes

We analyzed the diagnostic accuracy of the JES-BE classification based on the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy of the predicted histology of BE. The diagnostic accuracy of the mucosal pattern alone was also evaluated. We analyzed interobserver agreement of the five reviewers in prediction of histology and interpretation of NBI surface patterns.

To identify risk factors for misdiagnosed images, images with a correct rate of less than 40% were defined as difficult-to-diagnose, and those with 60% or more were defined as easy-to-diagnose. Logistic regression analysis was performed to identify risk factors for difficult-to-diagnose images. We hypothesized that histology, magnification power, and location of captured images are factors involved in misdiagnosis. The magnification power of each image was evaluated based on the size of the mucosa and vessels (Fig. 3a,b). Based on the length of the background BE, each case was classified into the following two groups: long-segment Barrett’s esophagus (LSBE) or short-segment Barrett’s esophagus (SSBE). For instance, when an image of a site within 1 cm from the esophago-gastro-jejunum was evaluated in a patient with LSBE, it was categorized as LSBE.

This study was conducted in accordance with the standards of the Declaration of Helsinki and the current ethical guidelines, and was approved by the institutional Ethics Board at Sendai Kousei Hospital.
Statistical analysis

Diagnostic accuracy was calculated based on per image analysis. Observer agreement between reviewers was evaluated using Fleiss’ kappa statistics for the predicted histology and surface pattern of the JES-BE classification [18, 19]. Kappa values were interpreted as previously described [20], as follows: kappa = 0, absence of agreement; kappa < 0.20, slight agreement; kappa = 0.21–0.40, fair agreement; kappa = 0.41–0.60, moderate agreement; kappa = 0.61–0.80, substantial agreement; and kappa > 0.81, almost perfect agreement. Continuous variables were compared using the Mann-Whitney rank-sum test. The ratios of nominal variables were compared using Fisher’s exact test. Risk factors for misdiagnosis were evaluated using logistic regression analysis. Candidate variables were histology, magnification power, and location of the captured image.

Statistical analyses were performed using JMP 13 (SAS Institute Inc., Cary, North Carolina, United States) and P < 0.05 was considered to indicate statistical significance.

Results

Baseline characteristics

Characteristics of patients and ME-NBI images are shown in Table 1. One hundred thirty-two ME-NBI images were retrieved from 44 patients with BE. Twenty-nine patients had Barrett’s esophageal dysplastic lesion, whereas 15 patients had
Management of patients was as follows: endoscopic submucosal dissection (ESD) (n = 28), surgery (n = 1), and no treatment (n = 15). Of the 44 patients, 35 (80%) were male with a mean age of 71.0 years (interquartile range [IQR] 66–78 years). The median maximum length of BE was 2.0 cm (IQR 1.0 cm–3.0 cm). Thirty-six patients (81%) had hiatal hernia, and 31 (70%) had reflux esophagitis.

Of the 132 ME-NBI images, 75 (57%) were dysplastic (LGD, n = 26; HGD or adenocarcinoma, n = 49) and 57 (43%) were non-dysplastic. The invasion depth of all adenocarcinomas was limited to the mucosal or submucosal layers. Forty-six images (35%) were high magnification, and 86 were low magnification. Eighty-one images (61%) were captured from LSBE, and 51 were from SSBE.

Diagnostic accuracy and interobserver agreement of the JES-BE classification

The diagnostic accuracy of the predicted histology with both mucosal and vascular pattern and with mucosal pattern alone is shown in Table 2. Mean values for sensitivity, specificity, PPV, and NPV were 67%, 80%, 82%, and 65%, respectively. The overall accuracy of the JES-BE classification was 73%. The diagnostic accuracy of the predicted histology with mucosal pattern alone was as follows. Mean values of sensitivity, specificity, PPV, NPV, and overall accuracy were 79%, 72%, 79%, 72%, and 76%, respectively.

Interobserver agreement for predicted histology and NBI surface patterns was moderate and fair, respectively (κ = 0.720 and 0.573, respectively).

Risk factor analysis for misdiagnosed images

Of the 132 ME-NBI images, 34 (26%) were difficult-to-diagnose and 99 (74%) were easy-to-diagnose. The characteristics of these two groups are shown in Table 3. Table 4 shows the results of logistic regression analysis. LGD and high-power magnification images were each significant risk factors for difficult-to-diagnose images (OR 6.80, P = 0.0017 and OR 3.31, P = 0.0125, respectively). Images captured from LSBE were more likely to be difficult-to-diagnose than those from SSBE, but the difference was insignificant (OR 2.13, P = 0.157).

Discussion

In the current study, five endoscopists reviewed 132 ME-NBI images of BE using the JES-BE classification, and the diagnostic accuracy of this classification was relatively high (73%). Moreover, this study revealed that LGD and high-power magnification images were significant risk factors for misdiagnosis. The JES-BE classification was proposed in 2018 [15]. Although the diagnostic accuracy of this classification has not been evaluated on a large scale, a review article described a validation study in which 156 images were reviewed by 10 endoscopists [16]. This study reported sensitivity and specificity for this classification of 87% and 97%, respectively. The high diagnostic accuracy evaluated in the review article indicates that
Table 3: Characteristics of difficult-to-diagnose and easy-to-diagnose groups.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Difficult-to-diagnose</th>
<th>Easy-to-diagnose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histology (non-dysplastic/LGD/HGD and adenocarcinoma)</td>
<td>10/17/7</td>
<td>47/11/40</td>
</tr>
<tr>
<td>Magnification power (high/low)</td>
<td>19/15</td>
<td>27/71</td>
</tr>
<tr>
<td>Location of the captured image (LSBE/SSBE)</td>
<td>26/8</td>
<td>55/43</td>
</tr>
</tbody>
</table>

LGD, low-grade dysplasia; HGD, high-grade dysplasia; LSBE, long-segment Barrett’s esophagus; SSBE, short-segment Barrett’s esophagus.

Table 4: Risk factors of misdiagnosis based on logistic regression analysis.

<table>
<thead>
<tr>
<th>Risk Factors of Misdiagnosis</th>
<th>Difficult-to-diagnose images</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR 95% CI</td>
<td>P value</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
</tr>
<tr>
<td>LGD</td>
<td>6.81</td>
</tr>
<tr>
<td>Non-dysplastic</td>
<td>1.36</td>
</tr>
<tr>
<td>HGD and adenocarcinoma</td>
<td>Reference</td>
</tr>
<tr>
<td>Magnification power</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>3.31</td>
</tr>
<tr>
<td>Low</td>
<td>Reference</td>
</tr>
<tr>
<td>Location of captured image</td>
<td></td>
</tr>
<tr>
<td>LSBE</td>
<td>2.13</td>
</tr>
<tr>
<td>SSBE</td>
<td>Reference</td>
</tr>
</tbody>
</table>

LGD, low-grade dysplasia; HGD, high-grade dysplasia; LSBE, long-segment Barrett’s esophagus; SSBE, short-segment Barrett’s esophagus.

The JES-BE classification is practical in a clinical setting. However, in our study, the diagnostic accuracy of this classification was lower. The main reason may be the differences in the reviewers. The previous article included experts in the field of esophageal cancer, which may have contributed to the high diagnostic accuracy. On the other hand, our study did not include experts, and all reviewers were general and non-expert endoscopists working in a city hospital. We believe that our study reflects a real-world situation more accurately. Another reason for the low diagnostic accuracy is that some patients in whom the diagnosis was extremely difficult to make were included in this study. Moreover, because patients with difficult-to-diagnose cases undergo biopsy preoperatively multiple times to check lesion demarcation, the number of difficult-to-diagnose images would have increased. These are the main reasons why the diagnostic accuracy in this study was lower than that in the previous study. LGD was a significant risk factor for misdiagnosis compared to HGD and adenocarcinoma in the current study (OR: 6.80, P = 0.0017). Dysplasia is defined as a neoplastic epithelium that is confined to the basement membrane of the gland [21, 22]. In the Vienna classification, LGD in BE is characterized by relative preservation of glandular architecture but with cellular atypia, including nuclear hyperchromatism, pleomorphism, mucin depletion and absence of goblet cells [17]. It is familiar to endoscopists and pathologists in Western countries, but not in Eastern countries including Japan. Indeed, Japanese pathology guidelines for esophageal cancer do not include a description of dysplasia in BE. Therefore, in the current study conducted in Japan, pathological diagnosis of LGD was made carefully. Two expert pathologists in the field of gastrointestinal cancer made the diagnosis of LGD while reviewing the description and reference images for LGD in the Vienna classification.

Previous reports showed that histopathological and endoscopic diagnosis of LGD was difficult [23]. The main reason for this is the presence of inflammation, as the non-dysplastic epithelium in BE may mimic that of LGD [24]. Moreover, because LGD is a low-grade atypical tumor, it often shows no or weak irregularity endoscopically [25]. Therefore, it may be difficult to differentiate LGD from non-dysplastic epithelium even with ME-NBI. Endoscopists should recognize that LGD is a dysplastic lesion that is difficult to differentiate from non-dysplastic epithelium in Barrett’s mucosa. Consequently, when encountering a difficult lesion to diagnose, we should actively check pathological diagnosis based on biopsy specimen rather than trying to diagnose with endoscopy alone. Adopting this practice may help reduce the number of misdiagnoses of Barrett’s lesions.

In this study, obtaining images with high-power magnification was a significant risk factor for misdiagnosis. According to the JES-BE classification, high-power magnifying observation should be performed to evaluate a vascular pattern, and a mucosal pattern should be assessed under low-power magnifying observation. That is because when Barrett’s mucosa is observed at high-power magnification, the target mucosa is stretched by the scope, and the mucosal pattern becomes obscured. Unclear mucosal pattern may be the reason why high-power magnification images were risk factor for misdiagnosis. In this study, there was no significant difference in accuracy between diagnosis using mucosal patterns alone and diagnosis including both mucosal and vascular patterns (Table 2). In other words, the additional effect of diagnosis of a vascular pattern was insignificant. Based on these results, high-power magnification images may be better evaluated in combination with low-power magnification images, or it might be better to consider simplifying the diagnostic strategy of the JES-BE classification, such as using just mucosal pattern.

Images captured in LSBE were more likely to be misdiagnosed than those in SSBE, but there was no significant difference. Probst et al. reported that the R0 resection rate for BAC in LSBE was significantly lower than that in SSBE (70.4% vs. 90.0%, P = 0.029) [26]. The main reason for this difference might be difficulty in recognizing lateral extension of the lesion. LSBE is often associated with background inflammation.
due to reflux esophagitis. Inflamed mucosa sometimes shows mild atypia, which makes it extremely difficult to distinguish dysplastic mucosa from non-dysplastic mucosa [24]. However, some previous articles reported that R0 resection rates of ESD cases were comparable between SSBE and LSBE [27, 28]. Therefore, larger-scale prospective studies should be conducted to evaluate the difficulty of endoscopic diagnosis of lesions in SSBE and LSBE.

The present study had some limitations. First, it was a single-center study conducted in Asia. Thus, fewer patients with LSBE were enrolled, which led to the small number of LSBE images. As a result, the odds ratio of 2.13 for LSBE was relatively high but it caused a wide 95% confidence interval and was not statistically significant. Further studies at multiple institutions with large sample size are recommended. Second, because the decision about whether an image is high or low magnification was made by a single endoscopist who prepared the images, which may have added bias to the study. No fair comparison may be possible, because both images were not selected from the same lesion as paired views. In addition, some images were taken from the same patient, which would lead to selection bias and affect the results of statistical analysis. Third, easy- and difficult-to-diagnose images were defined based on the proportion of accurate diagnoses (less than 40% or more than 60%). This means the definition of these two groups were relatively determined, not by an absolute criterion. In addition, the correct rate of 60%, which was the definition of easy-to-diagnose images, is not very high compared to that for diagnosis of early gastric cancer [29]. However, considering the difficulty of diagnosing BAC due to background inflammation and existence of LGD, this criterion may be acceptable. Fourth, LGD is not well known in Eastern countries, including Japan, either pathologically or endoscopically. In addition, the diagnosis of LGD can vary among pathologists [30]. Therefore, in the current study, we asked the two pathologists to check the definition of LGD in the Vienna classification every time before diagnosing LGD. As a result, the diagnosis of LGD may be nearly identical.

Conclusions

In conclusion, this study elucidated risk factors for misdiagnosis of BE based on the JES-BE classification. We found that LGD and high-power magnification images were risk factors for misdiagnosis. Although a prospective validation study is necessary to confirm these results, we believe that endoscopists keep these risk factors in mind during endoscopic examination of patients suspected of having BE.

Competing interests

The authors declare that they have no conflict of interest.

References

[18] Cohen J. Weighted kappa: nominal scale agreement with provision for squared disagreement or partial credit. Psychol Bull 1968; 70: 213–220


