Gastroenterology 2017;152:564–570

Discordance Among Pathologists in the United States and Europe in Diagnosis of Low-Grade Dysplasia for Patients With Barrett's Esophagus

Prashanth Vennalaganti,[1,2] Vijay Kanakadandi,[1,2] John R. Goldblum,[3] Sharad C. Mathur,[4] Deepa T. Patil,[3] G. Johan Offerhaus,[5] Sybren L. Meijer,[6] Michael Vieth,[7] Robert D. Odze,[8] Saligram Shreyas,[1,2] Sravanthi Parasa,[1,2] Neil Gupta,[9] Alessandro Repici,[10] Ajay Bansal,[1,2] Titi Mohammad,[1,2] and Prateek Sharma[1,2]

[1]Gastroenterology and Hepatology, Veterans Affairs Medical Center, Kansas City, Missouri; [2]Gastroenterology Department, Kansas University Medical Center, Kansas City, Kansas; [3]Pathology Department, Cleveland Clinic, Cleveland, Ohio; [4]Department of Pathology, Veterans Affairs Medical Center, Kansas City, Missouri; [5]Department of Pathology, University Medical Center, Utrecht, the Netherlands; [6]Department of Pathology, Academic Medical Center, Amsterdam, the Netherlands; [7]Institute of Pathology, Klinikum Bayreuth, Bayreuth, Germany; [8]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts; [9]Gastroenterology, Loyola University Medical Center, Maywood, Illinois; and [10]Department of Gastroenterology, Instituto Clinico Humanitas, Milano, Italy

BACKGROUND & AIMS: There is suboptimal inter-observer agreement, even among expert gastrointestinal pathologists, in the diagnosis of low-grade dysplasia (LGD) in patients with Barrett's esophagus (BE). We analyzed histopathologic criteria required for a diagnosis of LGD using the new subcategories of LGD with inflammatory and dysplastic features. We categorized each diagnosis based on the level of confidence and assessed inter-observer agreement among gastrointestinal pathologists from 5 tertiary centers in the United States and Europe. **METHODS:** In the first phase of the study, 3 pathologists held a consensus conference at which they discussed the diagnostic criteria for LGD. In the second phase, 79 slides from patients with BE (23 samples of non-dysplastic BE, 22 samples of LGD, and 34 samples of high-grade dysplasia) were identified, randomly assigned to 7 pathologists (4 from the United States and 3 from Europe), and interpreted in a blinded fashion. κ Values were calculated for inter-observer agreement. We performed multinomial logistic regression analysis to assess the weighting of histologic features with the diagnosis. **RESULTS:** The overall κ value for diagnosis was 0.43 (95% confidence interval [CI], 0.42−0.48). When categorized based on degree of dysplasia, the κ value was 0.22 (95% CI, 0.11−0.29) for non-dysplastic BE, 0.11 (95% CI, 0.004−0.15) for LGD, and 0.43 (95% CI, 0.36−0.46) for high-grade dysplasia. When all pathologists made a diagnosis with high confidence, the inter-observer agreement was substantial among the US pathologists (κ, 0.63; 95% CI, 0.61−0.66) and European pathologists (κ, 0.80; 95% CI, 0.74−0.97). The κ values for all diagnoses made by European pathologists were higher than those made by US pathologists. **CONCLUSIONS:** In an analysis of criteria used in histopathologic diagnosis of LGD, we did not observe improvement in level of agreement among experienced pathologists, even after accounting for inflammation. The level of inter-observer agreement increased with level of pathologist confidence. There was also a difference in reading of histopathology samples of BE tissues between US and European pathologists.

Keywords: Barrett's Esophagus; Low-Grade Dysplasia; Interobserver Agreement; κ Values.

The incidence of esophageal adenocarcinoma (EAC) continues to rise every year.[1] The prognosis of advanced EAC is dismal, with 5-year survival close to 10%.[1] Barrett's esophagus (BE), a condition in which metaplastic intestinal epithelium replaces normal squamous epithelium in the distal esophagus is the only known precursor lesion for development of EAC.[2] BE is believed to progress through stages of non-dysplastic Barrett's esophagus (NDBE), low-grade dysplasia (LGD), high-grade dysplasia (HGD), and EAC.

Diagnosis of BE requires the presence of endoscopically visible esophageal metaplasia combined with histologic evidence of metaplasia. The degree of dysplasia is an important determinant of disease progression and current guidelines recommend management decisions based on presence and histologic grade of dysplasia in BE.[2–4] The diagnosis of dysplasia in BE is graded based on architectural, cytologic, and nuclear abnormalities based on consensus criteria of 1988 (Supplementary Table 1) and 2001. However, inflammation-mediated epithelial injury can induce regenerative cytologic changes that can be difficult to differentiate from dysplasia.[5]

Abbreviations used in this paper: **BE, Barrett's esophagus; CRF, case report form; EAC, esophageal adenocarcinoma; HGD, high-grade dysplasia; IND, indefinite for dysplasia; LGD, low-grade dysplasia; LGD-D, low-grade dysplasia with prominent dysplastic features; LGD-I, low-grade dysplasia with prominent inflammatory features; NDBE, non-dysplastic Barrett's esophagus.**

Most current article

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2016.10.041

CLINICAL AT

The diagnosis of LGD is limited by poor inter-observer agreement. Estimates of inter-observer agreement measured using κ statistics have varied from −0.17 to 0.32 on standard biopsies.[6–9] Due to the poor inter-observer agreement, rates of progression from LGD to EAC have varied between 0.02%[10,11] and 11.4%.[12] In addition, while some studies have shown increased rates of progression when 2 or more pathologists confirm the diagnosis, others have failed to show the same association, adding to the overall uncertainty.[8,9,12–14]

The aim of the study was to re-evaluate the reproducibility of dysplasia diagnosis among pathologists, particularly the histopathologic criteria for LGD with prominent inflammatory features (LGD-I) and LGD with prominent dysplastic features (LGD-D), in an attempt to improve inter-observer agreement. We also measured inter-observer agreement between US and European pathologists based on these classification and the level of confidence.[15]

Materials and Methods

Study Design

The study was conducted in 2 phases. During the first phase, a consensus meeting was held among the pathologists for refining the criteria for LGD diagnosis. Thereafter, these criteria were prospectively applied to a separate set of slides in a blinded fashion. The study was approved by the Institutional Review Board.

Consensus Meeting

During the first phase, a consensus meeting was held with 3 expert GI pathologists (1 from Kansas City, 2 from Cleveland Clinic Foundation) before initial slide evaluation. The 1988 consensus criteria to grade BE along with modified 2001 consensus criteria were discussed. Because the aim was to address challenges in diagnosing LGD and differentiating this entity from distal esophageal mucosal inflammatory changes, a consensus was reached to subdivide LGD into LGD with predominant inflammatory (LGD-I, Figure 1, Supplementary Figure 1) and dysplastic features (LGD-D, Figure 2, Supplementary Figure 2).

The pathologists also agreed on the following criteria to be considered during the assessment of dysplastic changes: glandular crowding; cribriform glands; cytologic atypia extending to the surface; nuclear enlargement; nuclear hyperchromasia; nuclear crowding or pseudostratification; irregular nuclear contours; and mucin depletion (Supplementary Figure 3). The pathologists separated cases with atypia into primarily dysplasia or primarily inflammation, and degree of confidence was indicated for each slide. In addition, pathologists also commented whether any of these criteria were highly weighted in arriving at the diagnosis.

Slide Selection and Interpretation

For the second phase of the study, histopathology slides with varying degrees of dysplasia (NDBE, LGD, and HGD) were identified from 2 tertiary referral centers (Veterans Affairs Medical Center, Kansas City and Cleveland Clinic Foundation, Cleveland Clinic Foundation, Cleveland). The slides were randomly selected to meet the sample size requirements of the a priori power analysis. No effort was made to select particularly challenging slides and the cases reflect those normally seen in a tertiary care center. The slides were randomized using computer-assisted randomization and interpreted in a blinded fashion. There was no area of interest marked on the slide and the pathologists were provided with no clinical information. The slide-processing methodology at both sites was similar and both used formalin-based processing. Both sites used 10% neutral buffered formalin as tissue fixative. Tissues were processed using standard protocols, embedded in paraffin, sectioned at 3−4 μm, and stained with H&E. No special fixatives (eg, Bouin's solution) or staining techniques were used to enhance nuclear morphology, thus ensuring that slides from both institutions were fully comparable. For each slide, a separate case report form (CRF) was completed by the participating pathologist. The following information was recorded on each CRF: the distinct histopathologic criteria used for arriving at the particular diagnosis; the degree of weighting placed (low vs high) on each of these criteria; the highest grade of dysplasia for that particular slide (non-dysplastic vs LGD-D vs LGD-I vs HGD); and the level of confidence (low or high) for that diagnosis. The pathologists were asked to mark high confidence if they were >90% confident in their diagnosis.

Data Collection, Sample Size Estimates, and Statistical Analysis

All CRFs were completed for each slide by the 7 assessors (pathologists). Data from the individual CRFs were then transferred into Microsoft Excel (Microsoft Corp, Redmond, WA). Inter-observer agreement was calculated using κ statistics and graded based on Landis and Koch scale (κ values: <0 indicates poor agreement, 0.01−0.20 indicates slight agreement, 0.21−0.40 indicates fair agreement, 0.41−0.60 indicates moderate agreement, 0.61−0.80 indicates substantial agreement, and >0.80 indicate nearly perfect agreement).[15] With the histologic features mentioned serving as covariates, and the grade of dysplasia as the outcome variable, multinomial logistic regression analysis was performed to evaluate influence of presence of the these characteristics on the final diagnosis. Analysis was performed using STATA, version 10 (StataCorp, College Station, TX). In order for the revisited criteria to improve the κ values from an estimated baseline of 0.19−0.31 with a power of 80%, a sample size of 72 slides, each read by 7 pathologists, would be required.

Results

Pathologists and Slide Distribution

Seven experienced gastrointestinal pathologists (4 from the United States and 3 from Europe) participated in the study. The breakdown of the submitted diagnosis were as follows: NDBE (n = 23), 29%; LGD (n = 22), 28%; and HGD (n = 34), 43%.

Inter-observer Agreement

The overall mean inter-observer agreement for all diagnoses was moderate at 0.43 (95% CI, 0.42−0.48). When categorized by the degree of dysplasia, the κ values were as

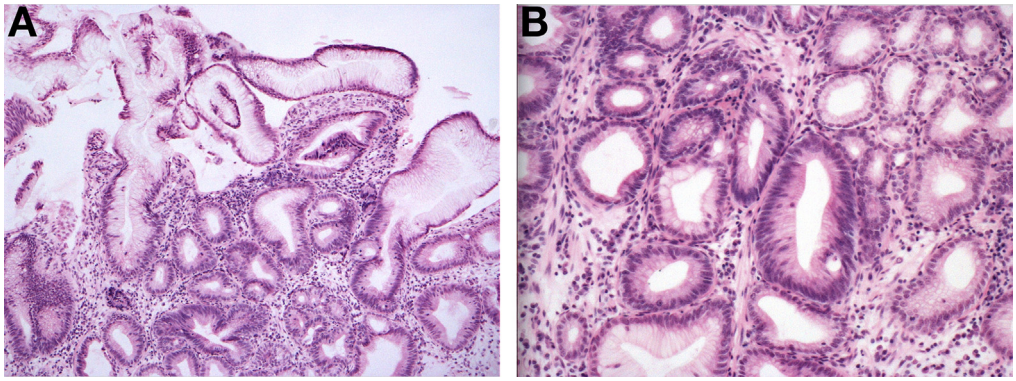


Figure 1. LGD with predominantly inflammatory features. (*A*) Low magnification. Inflammatory infiltrate can be seen in the low-power field. (*B*) High magnification. The changes of nuclear hyperchromasia, pseudostratification and nuclear crowding is noted only in some cells while others remain normal-appearing.

follows: non-dysplastic BE: 0.22 (95% CI, 0.11−0.29); LGD: 0.11 (95% CI, 0.004−0.15), and HGD: 0.43 (95% CI, 0.36−0.46) (Table 1).

Level of Confidence and Inter-observer Agreement

When the number of pathologists who diagnosed with higher confidence increased, the inter-observer agreement increased as well. When only the diagnoses marked with high confidence from all 7 pathologists were included, the κ value for overall diagnoses was 0.57 (95% CI, 0.45−0.62) and when only 1 or more pathologists reported high confidence, the κ value was 0.43 (95% CI, 0.42−0.48) (Table 2). Similarly, among US-based pathologists, when the diagnoses were reported with high confidence by all 4 pathologists, the κ value was 0.63 (95% CI, 0.61−0.66) and when the diagnoses were reported by 1 or more pathologists with high confidence, the κ value was 0.44 (95% CI, 0.39−0.49) and among European pathologists, the κ values were 0.80 (95% CI, 0.74−0.97) and 0.66 (95% CI, 0.60−0.71) when the diagnoses were reported with high confidence by all 3 pathologists and by 1 or more pathologists, respectively.

United States and European Inter-observer Agreement

We measured the inter-observer agreement for the 4 US-based and 3 European pathologists. European pathologists had higher κ values for all levels of dysplasia (Table 3). The overall inter-observer agreement among European pathologists was 0.65 (95% CI, 0.64−0.71) and for US-based pathologists was 0.44 (95% CI, 0.39−0.48). The corresponding values for NDBE, LGD, and HGD were 0.37 (95% CI, 0.26−0.51), 0.32 (95% CI, 0.08−0.07), and 0.63 (95% CI, 0.51−0.69) among European pathologists and 0.21 (95% CI, 0.05−0.35), 0.14 (95% CI, 0.09−0.22), and 0.45 (95% CI, 0.42−0.49) among US-based pathologists.

Assessment of Individual Histopathologic Criteria

A multinomial logistic regression was performed to evaluate whether highly weighting the recorded individual histologic features was independently associated with the final diagnosis (Table 4). The analysis used NDBE as a reference category. The diagnosis of LGD was associated with weighting the presence of cytologic atypia, nuclear hyperchromasia, and nuclear crowding highly. The diagnosis of HGD was associated with highly weighting glandular crowding, cytological atypia, nuclear enlargement, and irregular nuclear contours. We also found that an increase in the number of criteria was associated with an increase in the grade of dysplasia (Table 5). The median number of positive criteria for NDBE among pathologists of both regions was zero. US-based pathologists diagnosed LGD-I, LGD-D, and HGD based on the presence of a median of 5, 6, and 7 criteria, respectively. Their European counterparts diagnosed LGD-I, LGD-D, and HGD based on the presence of a median of 3, 4, and 5 criteria.

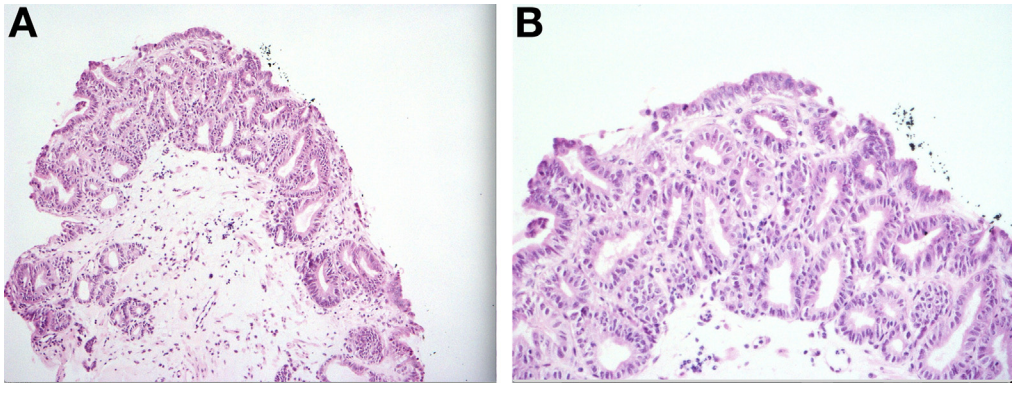


Figure 2. LGD with predominantly dysplastic features. (*A*) Low magnification. There is no significant inflammatory infiltrate. (*B*) High magnification. The changes of nuclear hyperchromasia, pseudostratification and nuclear crowding is seen in all the cells in the high-power field.

Table 1. κ Values for Inter-observer Agreement Among All 7 Pathologists From the United States and Europe

Histologic diagnosis (no. of slides)	Overall κ (95% CI)
Overall (79)	0.43 (0.42–0.48)
NDBE (23)	0.22 (0.11–0.29)
LGD (22)	0.11 (0.004–0.15)
HGD (34)	0.43 (0.36–0.46)

Table 3. Inter-observer Agreement for the US-Based and European Pathologists

Diagnosis	US pathologists, κ (95% CI)	European pathologists, κ (95% CI)
Overall	0.44 (0.39–0.48)	0.65 (0.64–0.71)
NDBE	0.21 (0.05–0.35)	0.37 (0.26–0.51)
LGD	0.14 (0.09–0.22)	0.32 (0.08–0.73)
HGD	0.45 (0.42–0.49)	0.63 (0.51–0.69)

Finally, we also measured the number of slides classified by each pathologist in individual dysplasia categories. Pathologists from Europe diagnosed fewer patients with dysplasia, particularly LGD-I, and a greater number of patients were diagnosed with NDBE (Supplementary Table 2). There were no significant differences in the criteria highly weighted by the groups (Supplementary Table 3).

Discussion

At the current time, dysplasia remains the most widely used tool for risk stratification of patients with BE and drives management decisions.[2–4] Therefore, it is of critical importance to have a diagnosis that is reproducible and impacts disease progression. It has been known that the diagnosis of LGD is fraught with poor inter-observer agreement.[6,7,16] This is probably reflected in the wide ranges of progression to esophageal adenocarcinoma in patients with LGD. Although recent meta-analysis reported that the overall risk of progression to cancer in BE patients with LGD was 0.12%–0.5% per year,[11,17] other investigators have reported an 8%–42% risk of progression to cancer in LGD patients.[13,18] We, therefore, undertook this study with 7 experienced pathologists (4 from the United States and 3 from Europe) with the final goal of improving inter-observer agreement for LGD diagnosis. To achieve this goal, we set to revisit the histopathologic criteria for a diagnosis of LGD, taking into account the inflammatory component, stratifying the level of confidence in making this diagnosis (this is something that clinicians do in daily clinical practice), and to assign weighting to each criterion to evaluate which criteria drives the diagnosis of dysplasia.

However, despite refining the criteria, we were unable to detect the expected improvement in the overall inter-observer agreement for this difficult diagnosis; overall κ value was 0.43 (95% CI, 0.42–0.48) and LGD κ value was 0.11 (95% CI, 0.004–0.15) indicating slight agreement. As more pathologists made the diagnosis with high confidence, we found an increasing inter-observer agreement. When all

Table 2. κ Values and the Level of Confidence

Variable	κ (95% CI)
All pathologists	
7	0.57 (0.45–0.62)
≥6	0.62 (0.58–0.64)
≥5	0.59 (0.53–0.64)
≥4	0.52 (0.47–0.55)
≥3	0.47 (0.42–0.50)
≥2	0.44 (0.38–0.49)
≥1	0.43 (0.42–0.48)
US-based pathologists	
4	0.63 (0.61–0.66)
≥3	0.53 (0.4–0.54)
≥2	0.46 (0.43–0.52)
≥1	0.44 (0.39–0.49)
Europe-based pathologists	
3	0.80 (0.74–0.97)
≥2	0.74 (0.71–0.80)
≥1	0.66 (0.60–0.71)

Table 4. Multinomial Logistic Regression Measuring the Association of Histologic Features With Diagnosis

Weighting pathologic criteria	*P* value	Odds ratio	95% CI for odds ratio	
			Lower bound	Upper bound
LGD				
Glandular crowding	.13	5.7	0.6	54
Cribriform glands	.26	0.2	0	3.3
Cytological atypia	.00	4.8	2.4	9.9
Nuclear enlargement	.24	2.1	0.6	7.7
Nuclear hyperchromasia	.01	4.9	1.5	16
Nuclear crowding	.04	3.3	1	10.4
Irregular nuclear contours	.30	2	0.5	7.7
Mucin depletion	.17	2.7	0.6	11.6
HGD				
Glandular crowding	.00	54.1	5.9	497.4
Cribriform glands	.22	4.8	0.4	58.2
Cytological atypia	.00	11	4.4	27.2
Nuclear enlargement	.03	4.9	1.2	20.7
Nuclear hyperchromasia	.17	2.6	0.7	9.8
Nuclear crowding	.80	0.8	.2	3.3
Irregular nuclear contours	.02	5.4	1.3	23.2
Mucin depletion	.35	2.1	0.4	9.7

The reference category is NDBE.

Table 5. Number of Positive Criteria Associated With Each Diagnosis

Diagnosis	United States, median (IQR)	Europe, median (IQR)	*P* value
NDBE	0 (1)	0 (1)	.974
LGD-I	5 (2)	3 (1.5)	<.001
LGD-D	6 (1)	4 (2)	<.001
HGD	7 (1)	5 (3)	<.001

IQR, interquartile range.

pathologists made the diagnosis with high confidence, the inter-observer agreement was substantial among both US (κ, 0.63; 95% CI, 0.61−0.66) and European (κ, 0.80; 95% CI, 0.74−0.97) pathologists. This has been shown in imaging studies as well, wherein gastroenterologists are more accurate in differentiating tubular adenomas from hyperplastic polyps on endoscopy of the colon when they are "confident: of their diagnosis based on specific endoscopic features noted."[19,20] Therefore, confidence level in diagnosis should be considered when expert panels provide consensus diagnoses for Barrett's-associated neoplasia.

The finding of a lower than expected rate of concordance in cases of NDBE (European pathologists had a higher concordance rate than US pathologists) may be due to a higher background prevalence of dysplasia in the study samples (perhaps expected by pathologists). In clinical practice, the pathologists might exhibit higher concordance, as the majority of the patients have non-dysplastic BE.

We found an important difference in the agreement between US and European pathologists. The European pathologists appear to have higher inter-observer agreement when compared with US-based pathologists for all diagnoses. Although we do not know the reason for this, the European pathologists tended to diagnose fewer cases of LGD-I, more cases of NDBE, and used fewer criteria for making a diagnosis of dysplasia when compared with US pathologists. The increased agreement seen among European pathologists can contribute to the increased rates of progression that have been reported in European studies.[18] This is particularly important because the rates of progression to HGD or EAC were higher when 2 or more pathologists agreed on the diagnosis of LGD.[8,13] Consistency and reproducibility of diagnoses are the first steps in harmonizing diagnostic criteria. At this very moment, more consistency does not mean right or wrong without follow-up toward a higher probability for progression. Additional studies are needed to understand the reason for the differences in agreement between US-based and European pathologists.

The presence and severity of architectural and cytologic features were used to diagnose and grade dysplasia as outlined by Reid et al[21] in 1988. In 2001, Montgomery et al[7,12] performed a large multi-institutional study and developed a new algorithm that included the presence of surface maturation, inflammation, and erosions/ulcers along with cytologic and architectural features to grade. Even with additional histologic features, inter-observer agreement for LGD was fair (κ, 0.32), due in large part to subjective nuclear and architectural changes that distinguished NDBE from LGD. Similarly, over the years, the inter-observer agreement seen in multiple studies has remained variable and inconsistent.[6,8,16,22,23] From a clinical standpoint, it is important to accurately diagnose LGD, as current BE guidelines risk-stratify patients based on grade of dysplasia with attendant recommend management guidelines.[2–4] Some authors who have shown high rates of progression of LGD have successfully shown that LGD ablation leads to decreased rates of progression to HGD/cancer.[18] However, in the absence of diagnostic reproducibility, the most appropriate management of LGD remains uncertain. How can we decide on invasive therapies for management of a disease that cannot be appropriately defined and diagnosed?

The results of our study coupled with previous data call for newer techniques that can improve the diagnostic reproducibility of LGD and predict progression to HGD and EAC. A recent study demonstrated that a combination panel of LGD, abnormal DNA ploidy, and *Aspergillus oryzae* lectin most accurately predicted progression from BE to HGD and EAC.[24] Aberrant p53 overexpression is another such marker that, when combined with LGD, demonstrated higher rates of progression.[25] This has led the British Society of Gastroenterology to recommend routine use of p53 staining to histopathologic assessment to improve diagnostic reproducibility of dysplasia.[26] But at the same time, that group also proposed expert pathologists consensus as the strongest criterion for a correct diagnosis. The ideal standard is seen in a combination of expert consensus and these new techniques.

The present study has limitations that merit discussion. Centers that have chosen cases for evaluation are highly specialized tertiary referral centers and the population of BE and dysplasia were artificially enriched. Although we did our best to replicate a "real-life" scenario for the pathologists in the study, the pathologists were aware that they were evaluating slides as part of dysplasia study. Also, given that there was a higher prevalence of dysplasia in the study slides, the level of agreement among the slides without dysplasia was noted to be lower than expected, possibly due to "overdiagnosing" of dysplasia. The study was performed among experienced pathologists in the United States and Europe, therefore, results might be not be generalizable to community pathologists. This, however, suggests that the agreement in the hands of nonexperienced pathologists might be even lower.[16] Also, we did not have a comparison group with slides categorized by the standard criteria without the subdivision into LGD-I and LGD-D. Subdivision of LGD and increasing the categories of diagnosis might have contributed to increasing the complexity and reduction in κ values. Finally, we did not evaluate the role of agreement in disease progression, that is, whether agreement among the pathologists in a diagnosis of LGD was associated with increased progression to cancer.[12–14]

In conclusion, the use of consensus histopathologic criteria did not improve overall agreement, despite

accounting for inflammation. There was low inter-observer agreement for dysplasia diagnoses even among expert pathologists, particularly for LGD, with significant differences between US and European pathologists. Higher levels of diagnostic confidence were associated with greater agreement among pathologists. Therefore, consistently and reliably diagnosing LGD remains challenging, even to expert gastrointestinal pathologists questioning the clinical utility of current BE management guidelines pertaining to LGD. The sole use of LGD as a risk-stratification strategy for the management of BE should be re-evaluated and new markers of progression are sorely needed. Therefore, at the current time, expert pathologist level of confidence and consensus is highly recommended for cases with LGD.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2016.10.041.

References

1. Hur C, Miller M, Kong CY, et al. Trends in esophageal adenocarcinoma incidence and mortality. Cancer 2013;119:1149–1158.
2. American Gastroenterological A, Spechler SJ, Sharma P, et al. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. Gastroenterology 2011;140:1084–1091.
3. Wang KK, Sampliner RE; Practice Parameters Committee of the American College of Gastroenterology. Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. Am J Gastroenterol 2008;103:788–797.
4. ASGE Standards of Practice Committee, Evans JA, Early DS, et al. The role of endoscopy in Barrett's esophagus and other premalignant conditions of the esophagus. Gastrointest Endosc 2012;76:1087–1094.
5. Goldblum JR. Controversies in the diagnosis of Barrett esophagus and Barrett-related dysplasia: one pathologist's perspective. Arch Pathol Lab Med 2010;134:1479–1484.
6. Pech O, Vieth M, Schmitz D, et al. Conclusions from the histological diagnosis of low-grade intraepithelial neoplasia in Barrett's oesophagus. Scand J Gastroenterol 2007;42:682–688.
7. Montgomery E, Bronner MP, Goldblum JR, et al. Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. Hum Pathol 2001;32:368–378.
8. Skacel M, Petras RE, Gramlich TL, et al. The diagnosis of low-grade dysplasia in Barrett's esophagus and its implications for disease progression. Am J Gastroenterol 2000;95:3383–3387.
9. Wani S, Falk GW, Post J, et al. Risk factors for progression of low-grade dysplasia in patients with Barrett's esophagus. Gastroenterology 2011;141:1179–1186; 1186 e1.
10. Schnell TG, Sontag SJ, Chejfec G, et al. Long-term nonsurgical management of Barrett's esophagus with high-grade dysplasia. Gastroenterology 2001;120:1607–1619.
11. Singh S, Manickam P, Amin AV, et al. Incidence of esophageal adenocarcinoma in Barrett's esophagus with low-grade dysplasia: a systematic review and meta-analysis. Gastrointest Endosc 2014;79:897–909.e4; quiz 983.e1, 983.e3.
12. Montgomery E, Goldblum JR, Greenson JK, et al. Dysplasia as a predictive marker for invasive carcinoma in Barrett esophagus: a follow-up study based on 138 cases from a diagnostic variability study. Hum Pathol 2001;32:379–388.
13. Curvers WL, ten Kate FJ, Krishnadath KK, et al. Low-grade dysplasia in Barrett's esophagus: overdiagnosed and underestimated. Am J Gastroenterol 2010;105:1523–1530.
14. Duits LC, Phoa KN, Curvers WL, et al. Barrett's oesophagus patients with low-grade dysplasia can be accurately risk-stratified after histological review by an expert pathology panel. Gut 2015;64:700–706.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
16. Kerkhof M, van Dekken H, Steyerberg EW, et al. Grading of dysplasia in Barrett's oesophagus: substantial interobserver variation between general and gastrointestinal pathologists. Histopathology 2007;50:920–927.
17. Hvid-Jensen F, Pedersen L, Drewes AM, et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. N Engl J Med 2011;365:1375–1383.
18. Phoa KN, van Vilsteren FG, Weusten BL, et al. Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia: a randomized clinical trial. JAMA 2014;311:1209–1217.
19. Repici A, Hassan C, Radaelli F, et al. Accuracy of narrow-band imaging in predicting colonoscopy surveillance intervals and histology of distal diminutive polyps: results from a multicenter, prospective trial. Gastrointest Endosc 2013;78:106–114.
20. Hewett DG, Huffman ME, Rex DK. Leaving distal colorectal hyperplastic polyps in place can be achieved with high accuracy by using narrow-band imaging: an observational study. Gastrointest Endosc 2012;76:374–380.
21. Reid BJ, Haggitt RC, Rubin CE, et al. Observer variation in the diagnosis of dysplasia in Barrett's esophagus. Hum Pathol 1988;19:166–178.
22. Wani SB, Goldblum JR, Post J, et al. Agreement among expert gastrointestinal pathologists for low-grade dysplasia (LGD) in Barrett's esophagus (BE) and implications for progression: results from a large, multicenter cohort study. Gastroenterology 2011;140:S80.
23. Wani S, Mathur SC, Curvers WL, et al. Greater interobserver agreement by endoscopic mucosal resection than biopsy samples in Barrett's dysplasia. Clin Gastroenterol Hepatol 2010;8:783–788.

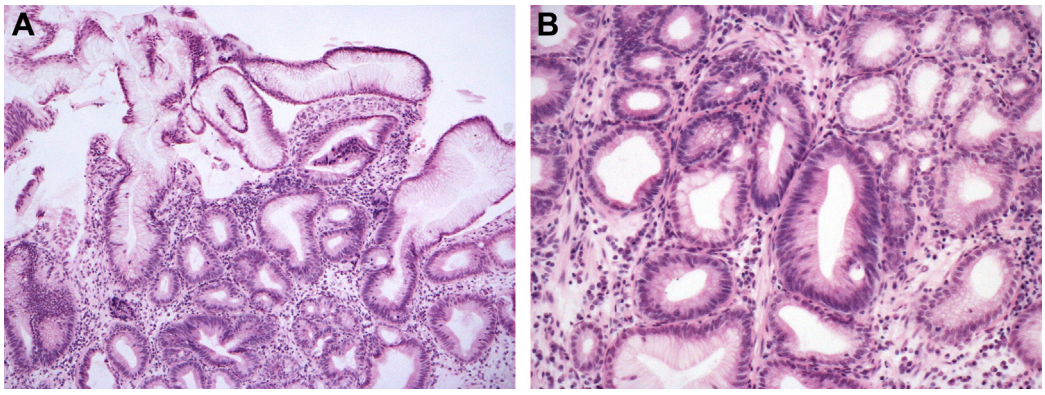
24. Bird-Lieberman EL, Dunn JM, Coleman HG, et al. Population-based study reveals new risk-stratification biomarker panel for Barrett's esophagus. Gastroenterology 2012;143:927–935.e3.
25. Kastelein F, Biermann K, Steyerberg EW, et al. Aberrant p53 protein expression is associated with an increased risk of neoplastic progression in patients with Barrett's oesophagus. Gut 2013;62:1676–1683.
26. 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.

Received September 18, 2014. Accepted October 20, 2016.

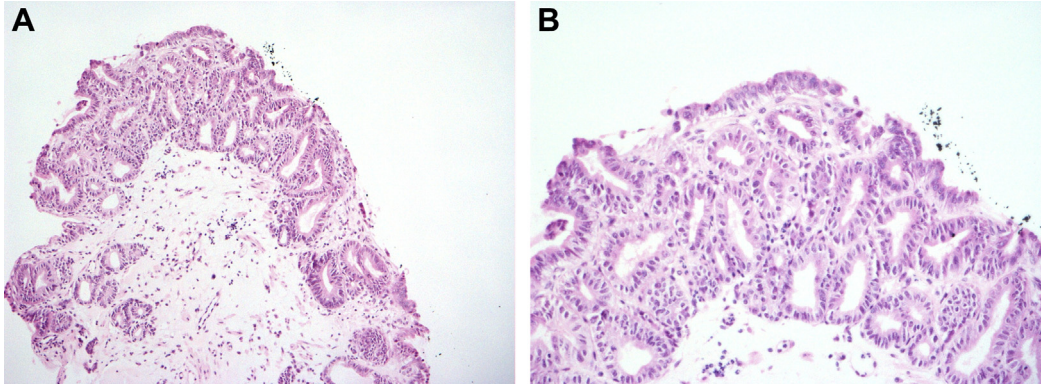
Reprint requests
Address requests for reprints to: Prateek Sharma, MD, Gastroenterology and Hepatology (111), Department of Veterans Affairs Medical Center, 4801 E Linwood Boulevard, Kansas City, Missouri 64128-2295. e-mail: psharma@kumc.edu; fax: (816) 922-3362.

Acknowledgments
Presented in part at the Barrett's esophagus research forum at Digestive Diseases Week 2014, Chicago, IL.

Conflicts of interest
This authors discloses the following: Prateek Sharma received grant support from Olympus, Effexus Pharmaceuticals, Cosmo Pharmaceuticals, and CDx Diagnostics. The remaining authors disclose no conflicts.



Supplementary Figure 1. Low-grade dysplasia with predominantly inflammatory features. (*A*) Low magnification. (*B*) High magnification.



Supplementary Figure 2. Low-grade dysplasia with predominantly dysplastic features. (*A*) Low magnification. (*B*) High magnification.

Revision of Pathologic Criteria For LGD in Barrett's Disease

Slide No. ☐ Reviewer's Initials: ☐ Date of Review: / /20

1. Kindly check for all criteria. Please mark if it is highly weighed.

Pathology Criteria	Yes	No	Highly weighed
Glandular crowding	☐	☐	☐
Cribriform glands	☐	☐	☐
Cytologic atypia (criteria below) extending to the surface	☐	☐	☐
Nuclear enlargement	☐	☐	☐
Nuclear hyperchromasia	☐	☐	☐
Nuclear crowding or pseudostratification	☐	☐	☐
Irregular nuclear contours	☐	☐	☐
Mucin depletion	☐	☐	☐

2. What is your overall impression of the Slide?

☐ Non Dysplastic ☐ Low Grade Dysplasia(DDx-Reactive disease)

☐ Low Grade Dysplasia(DDx-High Grade Dysplasia) ☐ High Grade Dysplasia

3. How confident are you about your diagnosis?

☐ Low Confidence ☐ High Confidence

Supplementary Figure 3. Case report form.

Supplementary Table 1. 1988 Consensus Criteria for Grading Dysplasia in Barrett's Criteria

Negative for dysplasia	The architecture is within normal limits. The nuclei do not vary greatly in size or shape and are located basally. The nuclear-to-cytoplasmic ratio is not increased. The nuclear envelope is generally smooth. Nucleoli are not markedly enlarged. Focal nuclear stratification is acceptable, as are small numbers of "dystrophic" goblet cells, the apical aspect of which does not communicate with the luminal surface. Greater nuclear alterations are acceptable when associated with evidence of inflammation, erosion, or ulceration. Numbers of abnormal-appearing mitoses are variable. Apical cytoplasmic mucus is usually present, but can be reduced or absent in inflammation. Normal nuclei appear more vesicular with more prominent nucleoli in Bouin and Hollande fixatives than in formalin. Fixatives, therefore, must be considered in interpretation.
Indefinite for dysplasia (IND)	The architecture may be moderately distorted. Nuclear abnormalities are less marked than those seen in dysplasia. Other features that can lead to a diagnosis of IND include more numerous dystrophic goblet cells, more extensive nuclear stratification, diminished or absent mucus production, increased cytoplasmic basophilia, and increased mitoses. The diagnosis of IND should be limited to cases in which the changes are too marked for negative but not sufficient for the diagnosis of dysplasia.
Positive for dysplasia (LGD and HGD)	The diagnosis of LGD or HGD is based on the severity of both architectural and cytologic criteria that suggest neoplastic transformation of the columnar epithelium. Although either architectural or cytologic abnormalities may predominate, HGD is diagnosed if either one is sufficiently prominent. Architectural abnormalities may include budded, branched, crowded, or irregularly shaped glands; papillary extensions into gland lumina; and villiform configuration of the mucosal surface. Nuclear features may include marked variation in size and shape, nuclear and/or nucleolar enlargement, increased nuclear-to-cytoplasmic ratio, hyperchromatism, and increased numbers of abnormal mitoses. Nuclear alterations are especially noteworthy if they involve the mucosal surface. Diagnostic features easily recognizable at lower power are cytoplasmic basophilia with loss of mucus and excessive nuclear stratification, often extending from the epithelial basement membrane to the luminal surface.
Intramucosal carcinoma	Intramucosal carcinoma is defined as carcinoma that has penetrated through the basement membrane of the glands into the lamina propria, but has not yet invaded through the muscularis mucosae into the submucosa. Most biopsy specimens will not be deep enough to rule out submucosal invasion.

Grading dysplasia in BE: 2001 consensus criteria

Surface maturation	+	+	−	−
Architecture	Normal	Normal or mild alteration	Mild alteration	Marked alteration
Cytology	Normal or reactive	Mild alterations or focal marked atypia with inflammation	Mild alterations, diffuse or marked alterations, focal; maintained polarity HGD	Marked alterations; loss of polarity

Supplementary Table 2. Reviewer Classification of Dysplasia

Reviewer no.	Dysplasia classification, n
1	NDBE 21
	LGD-1 23
	LGD-D 24
	HGD 10
2	NDBE 23
	LGD-I 18
	LGD-D 17
	HGD 19
3	NDBE 30
	LGD-I 16
	LGD-D 14
	HGD 19
4	NDBE 35
	LGD-I 8
	LGD-D 12
	HGD 22
5	NDBE 41
	LGD-I 8
	LGD-D 16
	HGD 12
6	NDBE 46
	LGD-I 6
	LGD-D 15
	HGD 12
7	NDBE 53
	LGD-I 1
	LGD-D 12
	HGD 13

Supplementary Table 3. Factors That Were Highly Weighted in Each Dysplasia Category

Grade of dysplasia	US criteria (percentage of cases highly weighted)	Europe (percentage of cases highly weighted)
LGD	Cytologic atypia 53.7	Cytologic atypia 43.1
	Nuclear crowding 44	Nuclear hyperchromasia 32.8
HGD	Cytological atypia 12.8	No criterion highly weighted