

Inter-observer Reproducibility in the Pathologic Diagnosis of Gastric Intraepithelial Neoplasia and Early Carcinoma in Endoscopic Submucosal Dissection Specimens: A Multi-center Study

Joon Mee Kim, MD¹
Jin Hee Sohn, MD²
Mee-Yon Cho, MD³
Woo Ho Kim, MD⁴
Hee Kyung Chang, MD⁵
Eun Sun Jung, MD⁶
Myeong-Cherl Kook, MD⁷
So-Young Jin, MD⁸
Yang Seok Chae, MD⁹
Young Soo Park, MD¹⁰
Mi Seon Kang, MD¹¹
Hyunki Kim, MD¹²
Jae Hyuk Lee, MD¹³
Do Youn Park, MD¹⁴
Kyoung Mee Kim, MD¹⁵
Hoguen Kim, MD¹²
Young Ju Suh, PhD¹⁶
Sang Yong Seol, MD¹⁷
Hwoon-Yong Jung, MD¹⁸
Deuck-Hwa Kim, BS¹⁹
Na Rae Lee, MPH²⁰
Seung-Hee Park, BS²⁰
Ji Hye You, MSN²⁰

*A list of author's affiliations appears at the end of the paper.

Correspondence: Jin Hee Sohn, MD
Department of Pathology, Kangbuk Samsung
Hospital, Sungkyunkwan University School
of Medicine, 29 Saemooan-ro, Jongro-gu,
Seoul 03181, Korea
Tel: 82-2-2001-2391
Fax: 82-2-2001-2398
E-mail: jhpath.sohn@samsung.com

Received January 9, 2019
Accepted March 29, 2019
Published Online April 1, 2019

Purpose

The diagnostic criteria of gastric intraepithelial neoplasia (IEN) are controversial across the world. We investigated how many discrepancies occur in the pathologic diagnosis of IEN and early gastric carcinoma in endoscopic submucosal dissection (ESD) specimens, and evaluated the reasons of the discordance.

Materials and Methods

We retrospectively reviewed 1,202 ESD specimens that were originally diagnosed as gastric IEN and early carcinoma at 12 institutions.

Results

The final consensus diagnosis of carcinoma were 756 cases, which were originally 692 carcinomas (91.5%), 43 high-grade dysplasias (5.7%), 20 low-grade dysplasias (2.6%), and 1 others (0.1%), respectively. High- and low-grade dysplasia were finally made in 63 and 342 cases, respectively. The diagnostic concordance with the consensus diagnosis was the highest for carcinoma (91.5%), followed by low-grade dysplasia (86.3%), others (63.4%) and high-grade dysplasia (50.8%). The general kappa value was 0.83, indicating excellent concordance. The kappa values of individual institutions ranged from 0.74 to 1 and correlated with the proportion of carcinoma cases. The cases revised to a final diagnosis of carcinoma exhibited both architectural abnormalities and cytologic atypia. The main differential points between low- and high-grade dysplasias were the glandular distribution and glandular shape. Additional features such as the glandular axis, surface maturation, nuclear stratification, and nuclear polarity were also important.

Conclusion

The overall concordance of the diagnosis of gastric IEN and early carcinoma in ESD specimens was excellent. It correlated with the proportion of carcinoma cases, demonstrating that the diagnostic criteria for carcinoma are more reproducible than those for dysplasia.

Key words

Intraepithelial neoplasia, Stomach neoplasms,
Endoscopic mucosal resection, Discrepancy, Reproducibility

Introduction

Endoscopic submucosal dissection (ESD) is an important treatment modality for intraepithelial neoplasias (IENs) including adenoma, dysplasia, and early gastric carcinoma (EGC) with minimal risk of lymph node metastasis. The histologic diagnosis of ESD specimens is critical for decision-making for further treatments. However, there is considerable discordance among pathologists, especially between the East and West, in the diagnosis of IEN and EGC. Although the Vienna and modified Vienna classifications [1,2] were suggested to reduce discrepancies in the pathologic interpretation of IEN of the gastrointestinal tract, there is still intra- and inter-observer variability. Standardization of the morphologic criteria for the pathologic diagnosis of gastric IEN is necessary to minimize confusion in clinical practice. The aim of this study was to assess the inter-observer reproducibility of IEN and early cancer diagnoses in gastric ESD specimens and to evaluate the pathologic findings resulting in diagnostic discrepancies.

Materials and Methods

1. Materials

Twelve institutions in Korea were involved in this study that actively perform ESD, and a study group (NECA-Korea ESD for early gastric cancer prospective study: N-Keep study) was organized. From June 2010 to May 2011, experienced endoscopists performed ESD for gastric adenomas (dysplasias) and EGCs that satisfied all the following inclusion criteria: (1) patients 20 years of age or older, (2) lesions less than 3 cm in length based on endoscopic findings, (3) adenoma or differentiated type (well or moderately differentiated) adenocarcinoma based on histologic examination of endoscopic biopsy tissue, (4) absence of ulcers in the lesion, and (5) no metastasis based on abdominal computed tomography findings prior to the procedure. ESD specimens were sent to the pathology lab at each institution, where they were fixed immediately in 10% buffered formalin for more than 4 hours. Then, paraffin blocks and hematoxylin and eosin-stained slides were made according to standard protocols.

2. Methods

Consensus diagnoses were made in a consensus board meeting composed of 16 pathologists with expertise in gas-

trointestinal pathology, the same method described in a previous manuscript [3]. In total, 1,202 ESD specimens collected from 1,138 patients were used. The final pathologic diagnoses were grouped as carcinoma (CA), high-grade dysplasia (adenoma with high-grade dysplasia [HG]), low-grade dysplasia (adenoma with low-grade dysplasia [LG]), and others (regeneration, gastritis, no residual tumor, etc.). The diagnostic criteria for CA were based on invasion, per Western standards.

After all the consensus meetings had been held, four expert pathologists (J.H.S., S.Y.J., M.Y.C., and J.M.K.) reviewed the 157 cases with discrepancies between the original and final diagnoses and analyzed the cause of discrepancy.

3. Statistical analysis

The agreement between the original and final diagnosis at each hospital was assessed with Cohen's kappa or weighted kappa statistic (for ordinal data with ≥ 3 categories). The estimated kappa statistic was interpreted as follows: < 0.20 , poor; 0.20 to < 0.40 , fair; 0.40 to < 0.60 , moderate; 0.60 to < 0.80 , good; and 0.80 to 1.00 , excellent agreement. The relationship between the kappa statistic and the cancer rate was examined by Spearman's correlation analysis. The differences in kappa values between hospitals were tested with the Wald statistic, $z = \frac{\widehat{K}_i - \widehat{K}_j}{\sqrt{\text{var}(\widehat{K}_i) + \text{var}(\widehat{K}_j)}}$ for an estimated kappa \widehat{K}_i at hospital i , which asymptotically follows a standard normal distribution. For adjustments for multiple comparisons, we used the Bonferroni method [4]. Statistical significance was accepted for p -values < 0.05 (two-tailed). Statistical analysis was performed with R software ver. 3.5.0 (<https://www.r-project.org/>, 'rel' package).

4. Ethical statement

This study was performed after approval from the Institutional Review Board of The National Evidence-based Healthcare Collaborating Agency (NECA) (approval number: NECA IRB09-013) in accordance with the principles of the Declaration of Helsinki. Written informed consents were obtained.

Results

1. Comparison between original diagnosis and final consensus diagnosis

There were discrepancies between the original and consensus diagnoses in 157 of the 1,202 cases (13.1%). The final

Table 1. Comparison between the original and final pathologic diagnoses

Original Dx	Final Dx				Total
	CA	HG	LG	Others	
CA	692 (91.5)	12 (19.0)	13 (3.8)	2 (4.9)	719
HG	43 (5.7)	32 (50.8)	31 (9.1)	3 (7.3)	109
LG	20 (2.6)	19 (30.2)	295 (86.3)	10 (24.4)	344
Others	1 (0.1)	0	3 (0.9)	26 (63.4)	30
Total	756 (100)	63 (100)	342 (100)	41 (100)	1,202

Values are presented as number (%). Dx, diagnosis; CA, carcinoma; HG, high-grade dysplasia; LG, low-grade dysplasia.

diagnoses of the patients included 756 CA (62.9%), 63 HG (5.2%), 342 LG (28.5%) and 41 other (3.4%) cases. Among the 756 CA cases, the original diagnoses included 692 CA (91.5%), 43 HG (5.7%), 20 LG (2.6%), and one regeneration (0.1%). Among the 63 cases of HG, the original diagnoses were CA, HG, LG and regeneration in 12 (19.0%), 32 (50.8%), 19 (30.2%), and 0 case, respectively. Among the 342 cases of LG, 13 (3.8%), 31 (9.1%), 295 (86.3%), and three (0.9%) were originally diagnosed as CA, HG, LG, and regeneration, respectively. Among the 41 other cases, two (4.9%), three (7.3%), 10 (24.4%), and 26 (63.4%) cases were originally diagnosed as CA, HG, LG, and regeneration, respectively. Therefore, the highest accuracy was seen in the CA (91.5%) group, followed by the LG (86.3%), other (63.4%), and HG (50.8%) groups. The comparison of the original and final diagnoses is shown in Table 1.

The disease composition varied among the institutions that participated in the study. For example, institution A had the highest CA incidence (42/43, 97.7%), with only one case of LG (1/43, 2.3%). The lowest rate of CA (45.5%) was found at institution L. The composition of the original and final diagnoses at each institution is shown in Table 2.

2. Assessment of kappa value

The total weighted kappa value including all the institutions was 0.83, indicating excellent agreement. To evaluate the gray zone effect of CA vs. HG and LG vs. HG, regrouping of disease were performed. When HG was combined with CA or LG, the kappa values were 0.82 and 0.83, respectively (Table 3). The kappa value between the CA group and the non-CA group was 0.84. All these kappa values were similarly high, so the diagnostic agreement was excellent in general.

In contrast, the kappa values of the institutions varied from 0.74 to 1, and this was related to the cancer incidence (Table 2). For institution A, which had the highest CA ratio (42/43, 97.7%), the kappa value was 1. In contrast, institution L, which had the lowest CA rate (46/121, 45.5%), had a much

lower kappa value of 0.75. The weighted kappa value of each institution correlated statistically with the cancer incidence (Spearman's correlation analysis, $p=0.02$), but not with the HG, LG and other groups. The kappa value of institution A differed significantly ($p < 0.05$) from those of five other institutions (D, G, H, I, and L), but did not differ from those of the remaining institutions in a Bonferroni examination. Although the kappa value was different at each institution, these differences were considered to be minor.

3. Assessment of pathologic findings in discrepant cases

1) Cases revised as CAs

Sixty-four cases (8.5%) originally diagnosed as non-CA were finally diagnosed as CA based on pathological characteristics. On the other hand, 27 of the 719 cases (3.8%) originally diagnosed as CA were finally diagnosed as non-CA; thus, there was a higher frequency of under-diagnosis than over-diagnosis. In some cases, the specimen condition was poor and the correct diagnosis was difficult to make. The pathologic characteristics observed in CA were structural and cellular atypia. Structural abnormalities could be observed in low-, medium-, and high-power views. The characteristics of CA observed at low magnification were irregular gland shapes, sizes and distributions and a loss of gland polarity. In some instances, the fusion of glands that maintained polarity was observed (Fig. 1A). At low and medium power, glandular fusion, irregular branching and sometimes small glands at the base of the lamina propria were observed (Fig. 1B and D). At high magnification, invasion could be observed (Fig. 1C and F). The outside of the basement membrane was irregular and angular glands were present (Fig. 1B, C, and F).

If cytological atypia such as hyperchromasia and pleomorphism was observed to be very severe, even though without great architectural abnormality, these cases could be diagnosed as CA (Fig. 1G-I). In addition, the glandular epithelial cells were observed to have a single-layer arrangement,

Table 2. Comparison between the original and final pathologic diagnoses in individual institution

No.	Institution	Original	Final				Total	Kappa	Weighted kappa
			CA	HG	LG	RA			
1	A	CA	42	0	0	0	43	1	1
		HG	0	0	0	0			
		LG	0	0	1	0			
		RA	0	0	0	0			
2	B	CA	98	3	0	0	110	0.68	0.80
		HG	1	2	0	0			
		LG	0	2	3	0			
		RA	0	0	0	1			
3	C	CA	90	0	3	0	124	0.88	0.89
		HG	0	2	1	0			
		LG	1	1	26	0			
		RA	0	0	0	0			
4	D	CA	75	2	2	0	123	0.67	0.74
		HG	4	2	3	0			
		LG	6	2	24	1			
		RA	0	0	0	2			
5	E	CA	27	0	0	0	49	0.68	0.82
		HG	6	2	1	0			
		LG	0	0	8	2			
		RA	0	0	0	3			
6	F	CA	29	0	0	0	54	0.73	0.81
		HG	5	3	0	0			
		LG	1	2	14	0			
		RA	0	0	0	0			
7	G	CA	64	3	2	0	119	0.68	0.78
		HG	7	4	6	0			
		LG	2	1	28	0			
		RA	0	0	0	2			
8	H	CA	37	0	0	0	74	0.72	0.79
		HG	3	2	1	0			
		LG	4	2	22	2			
		RA	0	0	0	1			
9	I	CA	129	2	1	2	266	0.80	0.86
		HG	9	8	7	0			
		LG	2	6	89	3			
		RA	0	0	0	8			
10	J	CA	27	0	1	0	59	0.71	0.81
		HG	2	0	6	0			
		LG	1	0	20	0			
		RA	0	0	0	2			
11	K	CA	28	0	2	0	60	0.77	0.83
		HG	0	0	4	0			
		LG	1	1	22	0			
		RA	0	0	0	2			
12	L	CA	46	2	2	0	121	0.69	0.75
		HG	6	7	2	3			
		LG	2	2	38	2			
		RA	1	0	3	5			
Total			756	63	342	41	1,202		

CA, carcinoma; HG, high-grade dysplasia; LG, low-grade dysplasia; RA, others.

Table 3. Kappa value according to the disease categorization

Disease category	Kappa	Weighted kappa
CA/HG/LG/Others	0.76	0.83
CA+HG/LG/Others	0.81	0.82
CA/HG+LG/Others	0.82	0.83
CA/HG/LG+Others	0.83	0.83
CA+HG/LG+Others	0.83	0.83
CA/Non-CA	0.84	0.84

CA, carcinoma; HG, high-grade dysplasia; LG, low-grade dysplasia.

round nuclei, vesicular chromatin, nuclear polarity loss, surface epithelial crowding and stratification in CA cases (Fig. 1E, H, and I). The observation of surface epithelial atypia was helpful for diagnosis. In particular, the difference between regenerating change and CA was that surface atypia and invasion existed in CA (Fig. 1E, 2G, and H). During the observation of surface cells, areas where the cells are well maintained should be selected. The pathologic characteristics of CA were summarized in Table 4.

2) Cases revised as high- and low-grade dysplasia

Among the 719 cases originally diagnosed as CA, 12 cases were revised to HG and 13 cases were revised to LG. In addi-

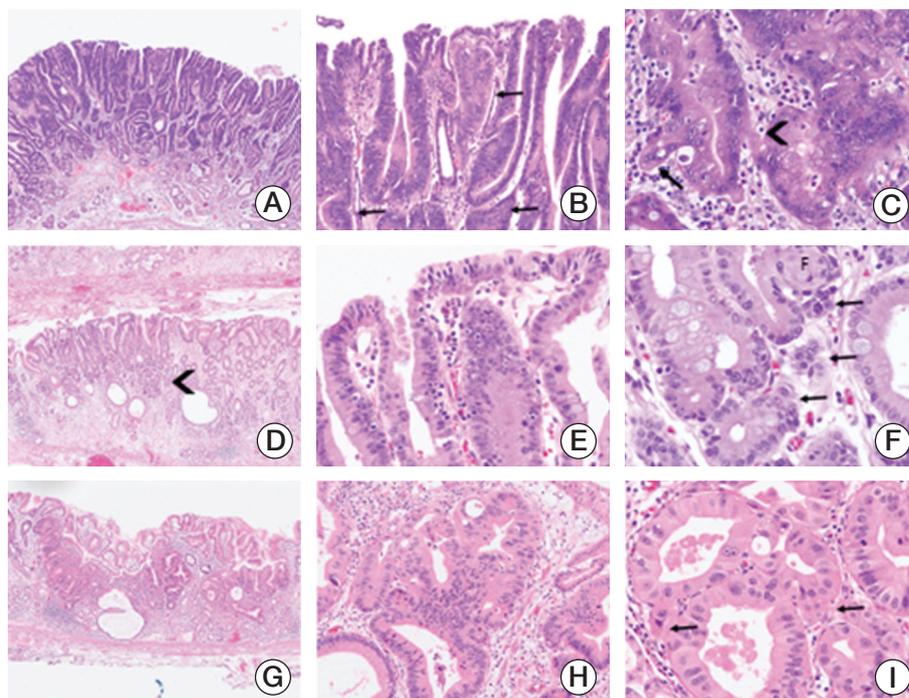


Fig. 1. Examples of underdiagnosed cases. (A-C) Carcinoma originally diagnosed as high-grade dysplasia. (A) Although the glandular polarity is maintained, irregular branching and fusion of glands are observed ($\times 40$). (B) Irregular branching of glands, papillary growth into the lumen, and focal surface maturation is present. A few angular glands are found (arrows, $\times 200$). (C) In high power view, invasion of tumor cells into the lamina propria is present (arrow). Blurring of the outside of basement membrane is also present (arrow head). The nucleus shows vesicular chromatin and prominent nucleoli ($\times 400$). (D-F) Carcinoma originally diagnosed as low-grade dysplasia. (D) Low power view shows maintained glandular polarity but small glandular proliferation in basal portion (arrowhead) of the lesion ($\times 40$). (E) Crowding, stratification and atypia of surface epithelia are present ($\times 400$). (F) Invasions and blurring of basement membrane are identified in high power view (arrows, $\times 400$). (G-I) Carcinoma showing obvious cytologic atypia originally diagnosed as high-grade dysplasia. (G) Hyperchromasia is easily noted in low power view ($\times 40$). (H) Not only marked cytologic atypia but also structural abnormality such as glandular branching and fusion is present. Definite invasion is not found ($\times 200$). (I) Marked cytologic atypia is present showing single-layer arrangement, prominent nucleoli and vesicular chromatin. In careful examination, minute invasions are identified (arrows, $\times 400$).

Table 4. Pathologic characteristics of carcinoma

Magnification	Characteristic	Example (Figures)	
Low	1. Irregular size and shape of glands	1G	
	2. Irregular distribution (density) of glands	1G	
	3. Loss of polarity of glands	1G	
	4. Fusion of glands	1A	
Medium	1. Fusion of glands	1B, H	
	2. Irregular branching	1B, H	
	3. Small glands in lamina propria	1D	
High	Structure	1. Invasion	1C, F
		2. Irregular basement membrane	1C, F, I
		3. Angular glands	1B
	Cytology	4. Extreme atypia (hyperchromasia, pleomorphism)	1I
		5. Round nuclei	1C, F, I
		6. Vesicular chromatin	1C, F, I
		7. Nuclear polarity loss	1C, F, I
		8. Surface epithelial crowding and stratification	1E

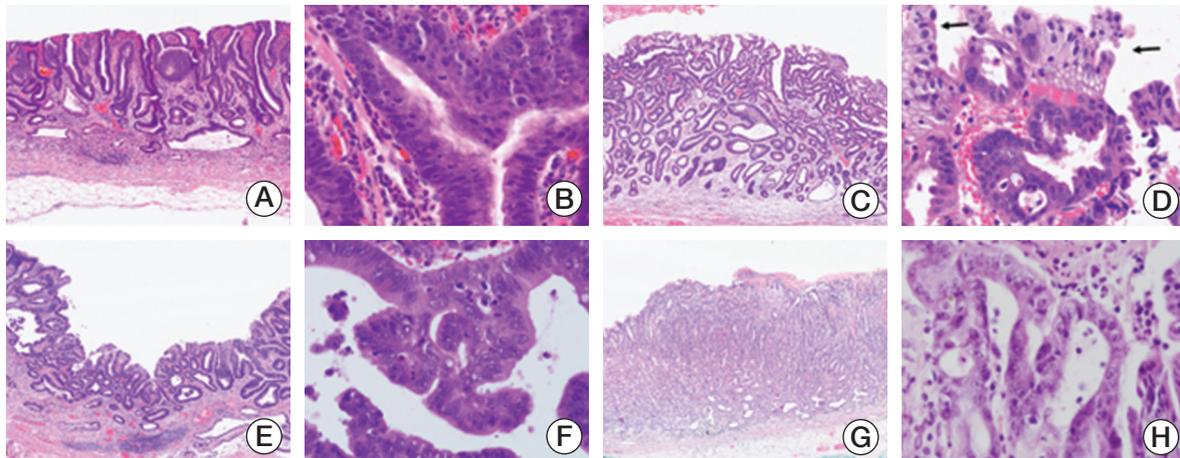


Fig. 2. Examples of over-diagnosed cases. (A, B) Low-grade dysplasia originally diagnosed as high-grade dysplasia. (A) At low power, the glands are evenly distributed and not crowded. The size and shape of glands are relatively uniform without branching or papillary projection. However, polarity loss of glandular axis is present ($\times 40$). (B) Intraepithelial neutrophilic infiltration may cause marked nuclear stratification in low-grade dysplasia ($\times 400$). (C, D) Another example of low-grade dysplasia originally diagnosed as high-grade dysplasia. (C) The superficial portion of glands are irregularly proliferated due to regenerative change, mimicking high-grade dysplasia. However, the glands in deeper portion are evenly distributed and not crowded ($\times 40$). (D) The glandular epithelia show intraglandular proliferation, tufting and papillary features but surface maturation is evident (arrows). Atypical cytologic features such as nuclear pleomorphism, vesicular chromatin pattern, and prominent nucleoli are occasionally found ($\times 400$). (E, F) High-grade dysplasia originally diagnosed as carcinoma. (E) Low power view demonstrates uneven distribution of glands with loss of polarity in the left side ($\times 40$). (F) Intraglandular proliferation, tufting and papillary features with cellular atypia is present. Atypical nuclear features mimicking carcinoma such as polarity loss and vesicular chromatin are focally present ($\times 400$). (G, H) Regenerative change originally diagnosed as carcinoma. (G) Glandular branching and irregularity are mimicking carcinomas ($\times 40$). (H) Stratification or hyperchromasia of the surface epithelium is not evident in regenerative change ($\times 400$).

Table 5. Differential points of high- and low-grade dysplasia

Characteristic	LG	HG
Structure		
Gland distribution	Even	Crowded
Gland shape	Regular	Irregular/intraglandular proliferation
Gland axis polarity	Present/loss	Loss
Surface maturation	Present	Absent
Nuclear stratification	Less than half	More than half
Cytology		
Nuclear polarity loss	Absent	Present
Brush border	Present	Absent/present
Pitfall		
Tufting and papillae in regeneration		
Polarity loss of glandular axis in LG		
Severe nuclear stratification in LG		
Brush border in HG		
Caution when erosion, inflammation, regeneration, and/or intraglandular inflammatory cell infiltration are associated		

LG, low-grade dysplasia; HG, high-grade dysplasia.

tion, 31 of the 109 cases originally diagnosed as HG were revised to LG, and conversely, 19 of the 344 cases originally diagnosed as LG were revised to HG. After reviewing these low- and high-grade dysplasia cases, we identified several characteristics important for differentiation (Table 4): (1) at low magnification, the glandular distribution was even in LG (Fig. 2A-D), but was irregularly crowded in HG (Fig. 2E). (2) LG had an evenly distributed and regular glandular shape (Fig. 2A-D), but HG displayed structural abnormalities such as irregular shapes, intraglandular proliferation, tufting and papillary features with cellular atypia (Fig. 2F). However, tufting and papillary features could be seen in LG cases associated with regeneration, which need careful concern (Fig. 2C and D). (3) The overall axis of the gland was not important for distinguishing HG from LG (Fig. 2A and B). Polarity loss of the glandular axis was observed primarily in HG (Fig. 2E), but not necessarily. (4) Surface maturation was helpful for distinguishing LG from HG (Fig. 2D). Surface maturation of glands could be seen in both LG and HG, even in CA, but it was usually focal in HG. If surface maturation was widely present, LG was more favorable. (5) If the nuclear stratification was less than half of the epithelial cell height, LG was favored, but if the nuclear stratification was severe, the possibility of HG was high. However, severe cell proliferation did not necessarily indicate HG, so it was necessary to judge structural abnormalities together. In addition to the architecture, nuclear features such as polarity loss and cellular atypia were important for differentiating between LG and HG. If the polarity loss of the nucleus was localized, the case could be diagnosed as HG rather than CA (Fig. 2F). (6) In

cases of LG associated with erosion, inflammation and regeneration, there was a high possibility of misdiagnosing LG as HG or CA due to nuclear atypia (such as vesicular chromatin and prominent nucleoli), especially at high magnification (Fig. 2D). In addition, intraglandular lymphocyte or neutrophil infiltration could increase the possibility of over-diagnosing HG (Fig. 2B), so careful examination is important at higher magnifications. The presence of a brush border could favor LG, but can also be seen in HG. The pathologic characteristics of LG and HG were summarized in Table 5.

Diagnostic pitfalls were present in non-neoplastic lesions originally diagnosed as IEN. For the differentiation between regenerating change and CA, cellular atypia of glands was extending into the surface epithelium in CA (Fig. 1E) but not in regeneration. The identification of definite invasion was important for cancer diagnosis, also. Non-neoplastic lesions exhibiting artifacts, improper/bad staining, erosion or regeneration and focality could be misdiagnosed as neoplastic lesions.

Discussion

The pathologic diagnosis of gastric IEN and EGC is important for the selection of the appropriate therapeutic options. However, there has been controversy among pathologists regarding the classification of these lesions, especially bet-

between the East and West. In the West, the terms gastric adenoma (for raised lesions) and gastric dysplasia (for flat/depressed lesions) are used, whereas in Japan, the term borderline lesion (groups 3 or 4) is used instead [5-7]. Not only the terminology, but also the criteria for CA differ between Western and Japanese pathologists. In Japan, the diagnosis of gastric CA is based on cytologic criteria, while in the West, gastric CA is diagnosed if invasion is detected [8]. International efforts have been made to reduce this confusion such as the Vienna classification [1] and Padova classification [9]. The diagnostic concordance rate for gastric lesions increased to 80% following the introduction of the revised Vienna classification. However, this classification system is not perfect, and diagnostic discrepancies still exist not only between the East and West but also among the pathologists in the same nation [10-12].

Since Korea is geographically close to Japan, academic exchanges have been frequent. The endoscopic mucosal resection/ESD technology in Korea was mainly introduced by Japan. There has been much discussion and cooperation between the two countries, not only among endoscopists, but also among pathologists. However, since Korean medicine developed mainly under Western influence, doctors in Korea are familiar with Western-style medical terminology. Thus, the terms, definitions and diagnostic criteria for gastric IEN and early CA in Korea are highly heterogeneous due to the combination of Western and Japanese influences.

On the other hand, this situation has been helpful in highlighting the differences in diagnostic criteria between Japan and the West and in demonstrating their strengths and weaknesses. As a result, more scientific and meaningful diagnostic criteria have been established, and efforts have been made to standardize the diagnosis of gastric IEN and early CA in Korea after a period of confusion. Invasion is the diagnostic criterion for gastric cancer in Korea, as in the West, but both cytological and structural atypia should be considered. Because the incidence of gastric IEN and CA is high, Korean pathologists are more familiar with gastric biopsy and can better identify invasion. Thus, it may appear that the diagnosis is based on cytology in Korea as it is in Japan, but this is likely because even the smallest invasion is recognized and diagnosed as CA. Korean pathologists have gradually standardized their diagnostic criteria after a season of chaos [13]. Nevertheless, there has not yet been a large-scale study on the diagnostic consistency of Korean pathologists.

To promote diagnostic consensus, NECA planned a nation-wide study to evaluate both clinical and pathological features in ESD-performed patients. In our study, we evaluated inter-observer differences, diagnostic features of IEN and early CA, and pathological characteristics contributing to discordance in ESD specimens. The diagnostic agreement rate of Korean gastrointestinal pathologists was excellent

($k=0.83$). This indicates that the existing criteria for cancer are working well and the inter-observer discrepancy rate in judging invasion is low. In the review of the literature, we could find two reports concerning the inter-observer reproducibility of histologic classification of gastric cancer [14,15]. Palli et al. [14], the study in Italy, reported agreement in histologic classification for about 70%-80% of 100 gastric cancers. The kappa was 0.34-0.64 (median, 0.51) among six pathologists when they applied WHO system. In our study, the kappa value of 12 institutions was 0.67-1, although the study design was different. Another study about inter-observer variation in the diagnosis of gastric epithelial dysplasia and CA between two pathologists in Japan and Korea [15] revealed that the agreement rate of diagnosis was 73.8% (31/42 cases). The most common disagreement occurred in the diagnosis of adenoma with high-grade dysplasia (9/17 cases, 52.9%): eight cases diagnosed as adenocarcinoma by the Japanese pathologist were diagnosed as high-grade dysplasia by the Korean pathologist.

The kappa value correlated statistically with the cancer proportion at each institution. In the case of institution A, almost all cases were CA, and the kappa value was 1. The highest kappa value was achieved when cases were grouped as cancer vs. non-cancer. On the other hand, the higher inter-observer discrepancy was present in diagnoses of LG and HG. The discrimination between low- and high-grade dysplasia is subjective because it depends on the quantitative degree of structural and cytologic abnormalities. So, the chance of intervention of pathologist's subjective judgment is more likely.

The histological evidence of invasion includes infiltration of the stroma by single cells or small clusters of cells, the presence of a stromal response such as desmoplasia, and evidence of lymphovascular invasion [16]. Stromal desmoplasia and lymphovascular invasion are difficult to identify in cases of very early invasion. In our study, we found that careful observation should be performed to identify invasion if there are irregularities outside the glandular basement membrane, severe structural or cytological atypia, angular glands or atypical single-layered glandular epithelial cells having vesicular nuclei. Although the presence of invasion is the essential criterion for differentiating CA from HG, cytologic atypia are also very important. If the extent of atypia is too great for dysplasia but there is no definite invasion, more careful examination and serial sectioning are necessary. Minute invasions are very difficult to identify, so under-diagnosis was more common than over-diagnosis: 5.7% (43/756) and 2.6% (20/756) of original HG and LG cases were revised to CA, while only 1.7% (12/719) and 0.8% (13/719) of original CA cases were revised to HG and LG, respectively. In terms of the total number of cases, the rate of under-diagnosing CA as HG was 3.6% (43/1,202), and this

was mostly due to unrecognized invasions. Especially in cases of well-differentiated adenocarcinoma, the detection of single or small clusters of cells suggesting lamina propria invasion was very difficult due to the usual features of well-formed glands.

The total rate of over-diagnosing HG as CA was only 1.0% (12/1,202). In most cases, there was inter-observer discrepancy regarding the invasion. When the lesion displayed irregularly-sized/shaped or branching glands with more cytological atypia than a usual adenoma, CA was suspected. Most over-diagnosed cases did not exhibit invasion, but displayed cytologic atypia. A noteworthy point is that tangential cutting or portions of glandular epithelial cells resembling small clusters of invasion led to the incorrect diagnosis of invasion. At high magnification, nuclear features such as nucleoli may be exaggerated, increasing the possibility of over-diagnosis. Thus, microscopic diagnosis should be performed at low magnification to evaluate the general distribution, variability, size and shape of the glands, as well as polarity and nuclear features.

Most gastric dysplasias are of the intestinal phenotype and contain crowded, tubular glands displaying nuclear stratification, pencil-like hyperchromatic nuclei, inconspicuous nucleoli, and mucin depletion, without surface maturation [17]. In the case of low-grade dysplasia, architectural abnormalities are minimal and the severity of cytologic atypia is only mild to moderate. The columnar cells contain stratified but polarized nuclei, most of which are basally located, below the half of the cytoplasm with mild to moderate mitotic activity. In high-grade dysplasia, the architectural disarray is more pronounced. The glandular cells become cuboidal in shape, with a high nucleus-to-cytoplasm ratio, nuclear stratification over half of the cytoplasm, loss of nuclear polarity, prominent amphophilic nucleoli and numerous mitoses, which can be atypical [16]. In spite of these criteria, HG is sometimes very difficult to distinguish. In our study, there was a high rate of revising HG to LG (31/109 cases, 28.4%), but a rather low rate of revising LG to HG (19/344 cases, 4.8%). This illustrates the high possibility of over-diagnosing LG as HG in cases with focal or more atypical features that nevertheless fall short of the diagnosis of HG. We also suspect that pathologists tend to be conservative in diagnosing CA but aggressive in diagnosing the dysplasia grade in ambiguous cases, because there is less responsibility of treatment option and less harm to patients leading to the close follow-up. Based on this study, we can suggest the following as useful points for distinguishing LG from HG: the degree of glandular crowding, the irregularity of the glands, the amount of inter-glandular lamina propria and the presence or absence of glandular axis polarity loss. In addition, nuclear stratification extending into the luminal aspect more than half of the cell height in a few contiguous glands [18] seems

to be helpful in diagnosing HG.

Regenerative atypia is a diagnostic pitfall in gastric biopsies because it is difficult to distinguish from adenocarcinoma. However, this misdiagnosis was very rare for our ESD specimens. This was presumably due to sufficient specimen volumes and careful diagnoses.

Additionally, the differential diagnosis of CA was difficult in cases of gastritis cystica profunda with dysplastic glands in the muscularis mucosa or submucosa. In such cases, overtly malignant cytologic atypia such as marked pleomorphism and vesicular chromatin were important for differential diagnosis. If desmoplastic reactions, single cells or small clusters of cells were present, the diagnosis of CA was easier.

In summary, the overall reproducibility of the diagnosis of gastric IEN and early CA in ESD specimens was excellent. The concordance correlated with the proportion of CA cases, demonstrating that the diagnostic criteria for CA are more reproducible than those for dysplasia. Further study is needed to establish the pathologic characteristics of CA and dysplasia so that diagnosis can be standardized. Caution should be exercised when secondary changes such as inflammation, erosion and regeneration are observed and the lesion is very focal.

Conflicts of Interest

Woo Ho Kim, a contributing editor of the Cancer Research and Treatment, was not involved in the editorial evaluation or decision to publish this article. All remaining authors have declared no conflicts of interest.

Acknowledgments

This study was conducted as part of project number NA2014-001 funded by the National Evidence-based Healthcare Collaborating Agency (NECA) in Korea.

Author Details

¹Department of Pathology, Inha University School of Medicine, Incheon, ²Department of Pathology, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, ³Department of Pathology, Yonsei University Wonju Severance Christian Hospital, Wonju, ⁴Department of Pathology, Seoul National University College of Medicine, Seoul, ⁵Department of Pathology, Kosin University College of Medicine, Busan, ⁶Department of Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, ⁷Department of Pathology, National Cancer Center, Goyang, ⁸Department of Pathology, Soonchunhyang University Seoul Hospital, Seoul, ⁹Department of Pathology, Korea University College of Medicine, Seoul, ¹⁰Department of Pathology, Asan Med-

ical Center, University of Ulsan College of Medicine, Seoul, ¹¹Department of Pathology, Inje University Busan Paik Hospital, Busan, ¹²Department of Pathology, Yonsei University College of Medicine, Seoul, ¹³Department of Pathology, Chonnam National University Hwasun Hospital, Hwasun, ¹⁴Department of Pathology, Pusan National University College of Medicine, Busan, ¹⁵Department of Pathology, Samsung Medical Center, Sungkyunkwan University

School of Medicine, Seoul, ¹⁶Department of Biomedical Sciences, Inha University School of Medicine, Incheon, ¹⁷Department of Internal Medicine, Inje University Busan Paik Hospital, Busan, ¹⁸Department of Gastroenterology, Asan Medical Center, University of Ulsan College of Medicine, Asan Digestive Disease Research Institute, Seoul, ¹⁹Department of Statistics, Inha University, Incheon, ²⁰National Evidence-based Healthcare Collaborating Agency, Seoul, Korea

References

- Schlemper RJ, Riddell RH, Kato Y, Borchard F, Cooper HS, Dawsey SM, et al. The Vienna classification of gastrointestinal epithelial neoplasia. *Gut*. 2000;47:251-5.
- Dixon MF. Gastrointestinal epithelial neoplasia: Vienna revisited. *Gut*. 2002;51:130-1.
- Kim JM, Sohn JH, Cho MY, Kim WH, Chang HK, Jung ES, et al. Pre- and post-ESD discrepancies in clinicopathologic criteria in early gastric cancer: the NECA-Korea ESD for Early Gastric Cancer Prospective Study (N-Keep). *Gastric Cancer*. 2016;19:1104-13.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B*. 1995;57:289-300.
- Goldstein NS, Lewin KJ. Gastric epithelial dysplasia and adenoma: historical review and histological criteria for grading. *Hum Pathol*. 1997;28:127-33.
- Riddell RH, Iwafuchi M. Problems arising from eastern and western classification systems for gastrointestinal dysplasia and carcinoma: are they resolvable? *Histopathology*. 1998;33:197-202.
- Schlemper RJ, Kato Y, Stolte M. Review of histological classifications of gastrointestinal epithelial neoplasia: differences in diagnosis of early carcinomas between Japanese and Western pathologists. *J Gastroenterol*. 2001;36:445-56.
- Schlemper RJ, Itabashi M, Kato Y, Lewin KJ, Riddell RH, Shimoda T, et al. Differences in diagnostic criteria for gastric carcinoma between Japanese and western pathologists. *Lancet*. 1997;349:1725-9.
- Rugge M, Correa P, Dixon MF, Hattori T, Leandro G, Lewin K, et al. Gastric dysplasia: the Padova international classification. *Am J Surg Pathol*. 2000;24:167-76.
- Kasamatsu E, Bravo LE, Bravo JC, Aguirre-Garcia J, Flores-Luna L, Nunes-Velloso Mdel C, et al. Reproducibility of histopathologic diagnosis of precursor lesions of gastric carcinoma in three Latin American countries. *Salud Publica Mex*. 2010;52:386-90.
- Ahn S, Park DY. Practical points in gastric pathology. *Arch Pathol Lab Med*. 2016;140:397-405.
- Zhao G, Xue M, Hu Y, Lai S, Chen S, Wang L. How commonly is the diagnosis of gastric low-grade dysplasia upgraded following endoscopic resection? A meta-analysis. *PLoS One*. 2015;10:e0132699.
- Kim JM, Cho MY, Sohn JH, Kang DY, Park CK, Kim WH, et al. Diagnosis of gastric epithelial neoplasia: Dilemma for Korean pathologists. *World J Gastroenterol*. 2011;17:2602-10.
- Palli D, Bianchi S, Cipriani F, Duca P, Amorosi A, Avellini C, et al. Reproducibility of histologic classification of gastric cancer. *Br J Cancer*. 1991;63:765-8.
- Kushima R, Kim KM. Interobserver variation in the diagnosis of gastric epithelial dysplasia and carcinoma between two pathologists in Japan and Korea. *J Gastric Cancer*. 2011;11:141-5.
- Bosman FT, Carneiro F, Hruban RH, Theise ND. WHO classification of tumours of the digestive system. 4th ed. Lyon: IARC Press; 2010. p. 10-4, 56.
- Jass JR. A classification of gastric dysplasia. *Histopathology*. 1983;7:181-93.
- Kim WH, Park CK, Kim YB, Kim YW, Kim HG, Bae HI, et al. A standardized pathology report for gastric cancer. *Korean J Pathol*. 2005;39:106-13.