

## Endoscopic Experience Improves Interobserver Agreement in the Grading of Esophagitis by Los Angeles Classification: Conventional Endoscopy and Optimal Band Image System

Si Hyung Lee\*, Byung Ik Jang\*, Kyeong Ok Kim\*, Seong Woo Jeon<sup>†</sup>, Joong Goo Kwon<sup>‡</sup>, Eun Young Kim<sup>‡</sup>, Jin Tae Jung<sup>‡</sup>, Kyung Sik Park<sup>§</sup>, Kwnag Bum Cho<sup>§</sup>, Eun Soo Kim<sup>§</sup>, Chang Geun Park<sup>||</sup>, Chang Heon Yang<sup>¶</sup>, and DeaguGyeongbook Gastrointestinal Study Group

\*Department of Internal Medicine, Yeungnam University School of Medicine, <sup>†</sup>Department of Internal Medicine, Kyungpook National University School of Medicine, <sup>‡</sup>Department of Internal Medicine, Catholic University of Daegu School of Medicine, <sup>§</sup>Department of Internal Medicine, Keimyung University School of Medicine, <sup>||</sup>Department of Internal Medicine, Daegu Fatima Hospital, and <sup>¶</sup>Department of Internal Medicine, Dongguk University School of Medicine, Daegu, Korea

**Background/Aims:** Interobserver variation by experience was documented for the diagnosis of esophagitis using the Los Angeles classification. The aim of this study was to evaluate whether interobserver agreement can be improved by higher levels of endoscopic experience in the diagnosis of erosive esophagitis. **Methods:** Endoscopic images of 51 patients with gastroesophageal reflux disease (GERD) symptoms were obtained with conventional endoscopy and optimal band imaging (OBI). Endoscopists were divided into an expert group (16 gastroenterologic endoscopic specialists guaranteed by the Korean Society of Gastrointestinal Endoscopy) and a trainee group (individuals with fellowships, first year of specialty training in gastroenterology). All endoscopists had no or minimal experience with OBI. GERD was diagnosed using the Los Angeles classification with or without OBI. **Results:** The mean weighted paired  $\kappa$  statistics for interobserver agreement in grading erosive esophagitis by conventional endoscopy in the expert group was better than that in the trainee group (0.51 vs 0.42,  $p < 0.05$ ). The mean weighted paired  $\kappa$  statistics in the expert group and in the trainee group based on conventional endoscopy with OBI did not differ (0.42, 0.42). **Conclusions:** Interobserver agreement in the expert group using conventional endoscopy was better than that in the trainee group. Endoscopic experience can improve the interobserver agreement in the grading of esophagitis using the Los Angeles classification. (**Gut Liver 2014;8:154-159**)

**Key Words:** Gastroesophageal reflux; Agreement; Experience

### INTRODUCTION

Gastroesophageal reflux disease (GERD) is a condition characterized by heartburn or regurgitation caused by esophageal exposure to acidic gastric content.<sup>1</sup> Due to the increasing prevalence of GERD in Asia, the social and medical importance of GERD is increasing.

Several diagnostic tools have been developed for diagnosis of GERD; however, endoscopy is the major tool for assessing esophagitis. Detection of erosion or ulcer at the gastroesophageal junction by endoscopy can allow for estimation of patients' symptoms and can be used in design of a treatment plan. Several classification systems have been introduced for improvement of the consistency of diagnosis of GERD.<sup>2,3</sup> Los Angeles (LA) classification was recently introduced and is used as a mainstay for diagnosis of erosive esophagitis, because of good correlation between esophagitis grade and acid exposure, heartburn severity, and risk of symptom relapse.<sup>2,4</sup> However, this system also has several limitations, including low interobserver agreement in trainees and low symptom correlation in endoscopy-negative reflux disease (NERD) or minimal change on the gastroesophageal junction.<sup>5,6</sup>

The optimal band imaging (OBI) system was developed as a new type of computed virtual chromoendoscopy and was introduced as flexible spectral-imaging color enhancement (FICE) by Fujinon Corporation (Saitama, Japan). This system can enhance

Correspondence to: Byung Ik Jang

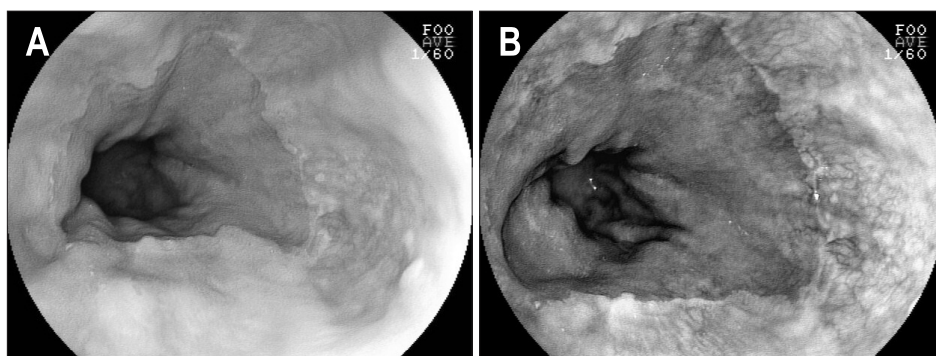
Department of Internal Medicine, Yeungnam University College of Medicine, 170 Hyeonchung-ro, Nam-gu, Daegu 705-717, Korea

Tel: +82-53-620-3310, Fax: +82-53-654-8386, E-mail: jbi@med.ac.kr

Received on February 7, 2013. Revised on March 25, 2013. Accepted on April 19, 2013. Published online on November 21, 2013

pISSN 1976-2283 eISSN 2005-1212 <http://dx.doi.org/10.5009/gnl.2014.8.2.154>

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Fig. 1.** (A) A still image obtained using conventional endoscopy in a patient without erosive esophagitis. (B) A still image of optimal band imaging in the same patient.

the mucosal surface and mucosal microvasculature without use of dye. Several studies have reported that this system was very useful for diagnosis of various GI diseases, such as detection of depressed-type early gastric cancer and differential diagnosis of colonic polyps.<sup>7-11</sup> Recently, a narrow band imaging (NBI) system was found to increase the interobserver consistency in diagnosis of GERD.<sup>12</sup> Several previous studies reported that the level of experience influenced the interobserver agreement for diagnosis GERD by LA classification. But, a study with many endoscopists to confirm interobserver agreement to diagnose GERD by LA classification is limited.

In this study, we evaluated that interobserver agreement in diagnosis of erosive esophagitis using conventional endoscopy with or without OBI can be improved by the level of conventional endoscopic and OBI system experience.

## MATERIALS AND METHODS

We retrospectively recruited patients with symptoms suggestive of GERD between January 2009 and December 2009. All subjects underwent conventional endoscopy with the OBI system using video endoscopy (EG-590WR; Fujinon Corp.). The FICE process was performed with R 520 nm, G 500 nm, and B 405 nm. Fifty-one static photographs of the esophagogastric junction were selected according to their resolution. All names or dates were removed from the images.

Twenty-six endoscopists participated in this study for evaluation of interobserver agreement for diagnosis of GERD. We divided 26 endoscopists into two group, expert group and trainee group. Expert group was composed of 16 endoscopists (gastroenterologic endoscopic specialist guaranteed by Korean Society of Gastrointestinal Endoscopy) with more than 4 years conventional endoscopy experience and no or minimal OBI system experience. Trainee group was composed of 10 endoscopists with minimal conventional and OBI system experience. Sixteen expert endoscopists were gastroenterologist specialists and 10 trainees were in their first year of specialty training in gastroenterology. The definition of minimal experience for OBI is defined when endoscopist has experience for OBI without experience for diagnosis of GERD by OBI.

**Table 1.** Experience with OBI for the Diagnosis of GERD and Opinions on the Efficacy of OBI in the Diagnosis of GERD according to the Los Angeles Classification

	Expert group	Trainee group
Experience with OBI		
No	14	6
Minimal	2	4
Total	16	10
Opinion for efficacy of OBI for diagnosis of GERD		
Good	4	1
Mild to moderate	8	7
No	2	2
Total	14	10

Two subjects declined to provide their opinion on the efficacy of OBI in the diagnosis of GERD.

OBI, optimal band imaging; GERD, gastroesophageal reflux disease.

To enhance the consistency in diagnosing GERD, introduction for OBI system and examples of GERD by OBI system was introduced to investigators before test. The endoscopic grade of GERD was determined using LA classification, which was presented at a symposium World Congress of Gastroenterology in Los Angeles in 1994. Mucosal breaks were graded as normal, A, B, C, and D according to the extent. The still image of the esophagogastric junction with a ruler was provided in order to minimize interobserver variation on examination of the extent of the mucosal break. The 51 still images were shown to 26 investigators using a video projector for 10 seconds. First, investigators recorded the grade of GERD by only conventional endoscopy image for 51 still images. Second, investigators recorded the grade of GERD by conventional endoscopy image and the OBI system image. Representative images on conventional and OBI endoscopy images are shown in Fig. 1.

### 1. Statistical analysis

Our primary interest was to compare interobserver agreement in the interpretation of conventional images with or without the combination of OBI system image. The index we used to as-



from 220 (16.6%) to 291 (21.9%) after using the OBI system. The number of patients diagnosed as normal by conventional endoscopy, but diagnosed as grade A esophagitis using OBI was 128 (43.9%, 128 of 291).

The mean weighted paired  $\kappa$  statistics for interobserver agreement in grading of erosive esophagitis in the expert group and trainee group by conventional endoscopy was 0.51 and 0.42, respectively ( $p < 0.05$ ) (Tables 3 and 4). The proportion of fair agreement and good agreement in the expert group by conventional endoscopy was 48.3% (58 of 120) and 27.5% (33 of 120). The proportion of above fair agreement in the expert group

by conventional endoscopy only was 75.8% (91 of 120). The proportion of fair agreement and good agreement in the trainee group by conventional endoscopy only was 53.3% (24 of 45) and 8.9% (4 of 45). The proportion of above fair agreement in the trainee group by conventional endoscopy only was 62.2% (28 of 45).

The mean weighted paired  $\kappa$  statistics in the expert group and trainee group by conventional endoscopy with OBI was 0.42 and 0.42, respectively (Tables 5 and 6). The proportion of fair agreement and good agreement in the expert group by conventional endoscopy using OBI was 34.1% (41 of 120) and

**Table 4.** Weighted Paired  $\kappa$  Statistics for Interobserver Consistency in the Grading of Esophagitis in the Trainee Group Based on Conventional Imaging

	1	2	3	4	5	6	7	8	9	10
1	1	0.579	0.575	0.524	0.417	0.375	0.318	0.477	0.311	0.463
2		1	0.609	0.332	0.563	0.402	0.468	0.56	0.437	0.33
3			1	0.328	0.452	0.452	0.457	0.609	0.421	0.491
4				1	0.362	0.321	0.198	0.25	0.383	0.559
5					1	0.48	0.286	0.617	0.485	0.493
6						1	0.338	0.462	0.409	0.362
7							1	0.32	0.252	0.229
8								1	0.568	0.453
9									1	0.616
10										1

**Table 5.** Weighted Paired  $\kappa$  Statistics for Interobserver Consistency in the Grading of Esophagitis in the Expert Group Based on Conventional Imaging with OBI

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
A	1	0.755	0.325	0.763	0.654	0.508	0.42	0.669	0.368	0.637	0.333	0.645	0.676	0.286	0.343	0.653
B		1	0.39	0.791	0.477	0.35	0.323	0.496	0.364	0.505	0.199	0.589	0.625	0.234	0.342	0.649
C			1	0.429	0.224	0.25	0.38	0.282	0.111	0.268	0.132	0.409	0.207	0.213	0.205	0.363
D				1	0.57	0.44	0.366	0.53	0.351	0.511	0.283	0.7	0.593	0.317	0.373	0.691
E					1	0.444	0.343	0.631	0.412	0.607	0.406	0.545	0.635	0.366	0.468	0.503
F						1	0.284	0.481	0.375	0.498	0.448	0.414	0.481	0.377	0.326	0.293
G							1	0.362	0.192	0.372	0.186	0.434	0.344	0.176	0.233	0.545
H								1	0.486	0.802	0.526	0.62	0.586	0.453	0.465	0.409
I									1	0.5	0.353	0.375	0.537	0.368	0.328	0.21
J										1	0.488	0.538	0.61	0.498	0.398	0.368
K											1	0.249	0.322	0.477	0.328	0.154
L												1	0.561	0.329	0.477	0.561
M													1	0.417	0.372	0.463
N														1	0.299	0.177
O															1	0.282
P																1

OBI, optimal band imaging.

**Table 6.** Weighted Paired  $\kappa$  Statistics for Interobserver Consistency in the Grading of Esophagitis in the Trainee Group Based on Conventional Imaging with OBI

	1	2	3	4	5	6	7	8	9	10
1	1	0.415	0.605	0.529	0.321	0.466	0.418	0.319	0.598	0.438
2		1	0.428	0.422	0.409	0.368	0.474	0.238	0.493	0.472
3			1	0.394	0.389	0.455	0.522	0.311	0.601	0.447
4				1	0.313	0.476	0.37	0.259	0.512	0.504
5					1	0.246	0.561	0.247	0.417	0.433
6						1	0.343	0.275	0.578	0.417
7							1	0.364	0.585	0.424
8								1	0.45	0.345
9									1	0.67
10										1

OBI, optimal band imaging.

15% (18 of 120). The proportion of above fair agreement in the expert group by conventional endoscopy only was 49.1% (59 of 120). The proportion of fair agreement and good agreement in the trainee group by conventional endoscopy using OBI was 57.5% (26 of 45) and 6.6% (3 of 45). The proportion of above fair agreement in the trainee group by conventional endoscopy only was 64.4% (29 of 45).

## DISCUSSION

GERD can be diagnosed by several diagnostic modalities, including traditional endoscopy, pH testing, and questionnaire. So far, the diagnosis of GERD using endoscopy is a mainstay in the clinical field; LA classification system is a popular system for grading the esophagitis. Interobserver variation for diagnosis of GERD using LA classification and virtual chromoendoscopy (NBI) was recently introduced by Lee *et al.*<sup>12</sup> But, no study regarding diagnosis of GERD using traditional endoscopy with OBI has been reported so far.

We conducted this study in order to confirm whether endoscopic experience could influence interobserver consistency for diagnosis of GERD by LA classification and for evaluation of whether OBI could improve interobserver consistency for diagnosis of GERD in each expert and trainee endoscopist group. Interobserver consistency with conventional endoscopy showed fair agreement (0.41 to 0.60) in both groups. However, consistency in the expert group was better than that in the trainee group (0.51 vs 0.42,  $p < 0.05$ ). In addition, the proportion of the above fair agreement with conventional endoscopy in the expert group was better than that in the trainee group (75.8% vs 62.2%). This result is similar to those of several previous studies.<sup>3,6</sup>

Interobserver consistency for diagnosis of GERD using conventional endoscopy with OBI showed fair agreement in both groups. The consistency was 0.42 in both groups. The propor-

tion of the above fair agreement in the expert group and the trainee group was 49.1% and 64.4%, respectively.

We think that OBI cannot improve the consistency of GERD diagnosis because of the following problems. First, most observers have no or minimal experience in use of OBI system for diagnosing GERD. The proportion of those who have experience in use of OBI system was 12.5% and 40% in the expert group and the trainee group, respectively. The lack of experience in use of OBI system in the expert group may influence the reduction of interobserver consistency after adding OBI images.

A previous study on experience in performance of virtual chromoendoscopy concluded that diagnostic accuracy could improve for less-experienced endoscopists after participation in an expanded training program.<sup>13</sup> Thus, study to determine whether a training program for OBI can improve diagnostic accuracy and interobserver consistency is needed.

Secondly, we used a still image instead of video image in grading of esophagitis by LA classification. In addition, time duration of the slide was only 10 seconds. This method may influence the negative effect for interobserver consistency.

In conclusion, experience in diagnosis of GERD by endoscopy improves interobserver agreement in grading of esophagitis by LA classification.

## CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

## REFERENCES

1. Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones R; Global Consensus Group. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. *Am J Gastroenterol* 2006;101:1900-1920.

2. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut* 1999;45:172-180.
3. Rath HC, Timmer A, Kunkel C, et al. Comparison of interobserver agreement for different scoring systems for reflux esophagitis: impact of level of experience. *Gastrointest Endosc* 2004;60:44-49.
4. Armstrong D, Bennett JR, Blum AL, et al. The endoscopic assessment of esophagitis: a progress report on observer agreement. *Gastroenterology* 1996;111:85-92.
5. Kusano M, Ino K, Yamada T, et al. Interobserver and intraobserver variation in endoscopic assessment of GERD using the "Los Angeles" classification. *Gastrointest Endosc* 1999;49:700-704.
6. Pandolfino JE, Vakil NB, Kahrilas PJ. Comparison of inter- and intraobserver consistency for grading of esophagitis by expert and trainee endoscopists. *Gastrointest Endosc* 2002;56:639-643.
7. Mouri R, Yoshida S, Tanaka S, Oka S, Yoshihara M, Chayama K. Evaluation and validation of computed virtual chromoendoscopy in early gastric cancer. *Gastrointest Endosc* 2009;69:1052-1058.
8. Togashi K, Osawa H, Koinuma K, et al. A comparison of conventional endoscopy, chromoendoscopy, and the optimal-band imaging system for the differentiation of neoplastic and non-neoplastic colonic polyps. *Gastrointest Endosc* 2009;69(3 Pt 2):734-741.
9. Osawa H, Yoshizawa M, Yamamoto H, et al. Optimal band imaging system can facilitate detection of changes in depressed-type early gastric cancer. *Gastrointest Endosc* 2008;67:226-234.
10. Cha JM, Lee JI, Joo KR, Jung SW, Shin HP. A prospective randomized study on computed virtual chromoendoscopy versus conventional colonoscopy for the detection of small colorectal adenomas. *Dig Dis Sci* 2010;55:2357-2364.
11. Pohl J, Lotterer E, Balzer C, et al. Computed virtual chromoendoscopy versus standard colonoscopy with targeted indigocarmine chromoscopy: a randomised multicentre trial. *Gut* 2009;58:73-78.
12. Lee YC, Lin JT, Chiu HM, et al. Intraobserver and interobserver consistency for grading esophagitis with narrow-band imaging. *Gastrointest Endosc* 2007;66:230-236.
13. Higashi R, Uraoka T, Kato J, et al. Diagnostic accuracy of narrow-band imaging and pit pattern analysis significantly improved for less-experienced endoscopists after an expanded training program. *Gastrointest Endosc* 2010;72:127-135.