External Validation of the Endoscopic Features of Sessile Serrated Adenomas in Expert and Trainee Colonoscopists

Hyo-Joon Yang¹, Jeong In Lee¹, Soo-Kyung Park¹, Yoon Suk Jung¹, Jin Hee Sohn², Kyu Yong Choi¹ and Dong Il Park¹

¹Division of Gastroenterology, Department of Internal Medicine, ²Department of Pathology, Gastrointestinal Cancer Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Background/Aims: It is unclear whether the endoscopic features of sessile serrated adenomas (SSAs) would be useful to trainee colonoscopists to predict SSA. Therefore, the present study aimed to identify features that expert and trainee colonoscopists can use to independently and reliably predict SSA by using high-resolution white-light endoscopy.

Methods: Endoscopic images of 81 polyps (39 SSAs, 22 hyperplastic polyps, and 20 tubular adenomas) from 43 patients were retrospectively evaluated by 10 colonoscopists (four experts and six trainees). Eight endoscopic features of SSAs were assessed for each polyp.

Results: According to multivariable analysis, a mucous cap (odds ratio [OR], 10.44; 95% confidence interval [CI], 5.72 to 19.07), indistinctive borders (OR, 4.21; 95% CI, 2.74 to 7.16), dark spots (OR, 3.64; 95% CI, 1.89 to 7.00), and cloud-like surface (OR, 2.43; 95% CI, 1.27 to 4.668) were independent predictors of SSAs. Among these, a mucous cap, indistinctive borders, and cloud-like surface showed moderate interobserver agreement (mean κ >0.40) among experts and trainees. When ≥1 of the three predictors was observed, the sensitivity and specificity for diagnosing SSAs were 79.0% and 81.4%, respectively.

Conclusions: Colonoscopy trainees and experts can use several specific endoscopic features to independently and reliably predict SSAs.

Key Words: Colorectal neoplasms; Colonoscopy; Sensitivity and specificity; Observer variation; Education, medical, graduate

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer worldwide and the fourth most common cause of cancer-related deaths.⁰ Coloscopy is favored as a primary screening modality for CRC in many countries, because it is the most accurate method for detecting cancer early, and it enables colonoscopists to remove precancerous lesions or adenomas.¹ Data from large observational studies and systematic reviews have indicated that screening colonoscopies were associated with a decreased incidence and mortality of CRC.²,³ However, the magnitude of risk reduction was smaller for proximal colon cancer than for distal colon cancer.²,³,⁴ In addition, about 6% to 8% of newly diagnosed CRCs develop in patients who have recently received a colonoscopy;⁵,⁶ these postcolonoscopy CRCs also have a disproportionately proximal location.⁷ A serrated pathway, recently recognized as the colorectal carcinogenesis pathway, was suggested as a probable explanation for these observations, because in the serrated pathway, cancer may arise more frequently than that through the conventional adenoma-carcinoma pathway.⁸,⁹ Moreover, a precancerous lesion in the pathway or sessile serrated adenoma (SSA) is difficult to identify during colonoscopy owing to its sessile feature with minimal change in the vascular surface.¹⁰ Thus, great efforts have been made to better understand the serrated.
pathway and improve the endoscopic identification of SSAs.

Various endoscopic features of SSAs on high-resolution white-light endoscopy (HR-WLE),14-16 narrow-band imaging (NBI),15,16 or NBI with magnification17,18 have been suggested. However, only highly experienced endoscopists identified and validated these findings, and it has not been determined whether colonoscopists with training could use these features to accurately and reliably identify SSAs. Some features may be easier for trainee colonoscopists to identify than other features. Although prospective studies have demonstrated that a short NBI training session improves the discrimination accuracy between tubular adenomas (TAs) and hyperplastic polyps (HPs),19,20 many colonoscopists are still unfamiliar with using NBI. Moreover, despite an increased availability of non-magnifying NBI, HR-WLE is currently more widely used in routine colonoscopy than NBI. Thus, it would be practical to focus on the findings of HR-WLE.

Therefore, in the present study, we aimed to (1) externally validate the endoscopic features of SSAs, (2) identify features that expert and trainee colonoscopists can independently and reliably use to predict SSA, and (3) examine the diagnostic accuracy of endoscopically predicting SSAs by using HR-WLE.

MATERIALS AND METHODS

Patients and polyps

In this retrospective image analysis study, we screened consecutive polyp images from 115 adult patients with at least one SSA diagnosed during their colonoscopic polypectomy between March 2005 and April 2014 at Kangbuk Samsung Hospital, Seoul, Korea. Among them, 51 SSAs in 43 patients were randomly selected using a simple random number list, and their images were reviewed by one expert colonoscopist (DIP) who was not involved in the image analysis process and had >15 years of experience. After excluding 12 SSAs in 11 patients with poor image quality, 39 SSAs in 32 patients were finally included in the study. For comparison, images of 22 HPs and 20 TAs in 33 patients were randomly selected from the sample population. The images were randomly divided into two groups: a learning set comprising 30 polyps (10 SSAs, 10 HPs, and 10 TAs) and a validation set comprising 51 polyps (29 SSAs, 12 HPs, and 10 TAs). All images were obtained by using HR-WLEs (CF-Q260AI or CF-H260AI; Olympus Inc., Tokyo, Japan). Histologic diagnoses of SSA were made based on the World Health Organization criteria. The diagnostic criteria included a serrated lesion with irregular dilated crypts such as boot-shaped, L-shaped, or inverted T-shaped dilatation of the crypt base. All pathological results of the included polyps were reviewed by one pathologist (JHS) who was blinded to the endoscopic appearances of the polyps and had >30 years of experience.

Image evaluation

On the basis of previous studies, one expert colonoscopist (DIP) selected eight endoscopic features: indistinctive borders, irregular shape, rim of debris, cloud-like surface, mucous cap, nodular surface, absence of surface vessels, and dark spots. The detailed definitions and endoscopic images of the features are shown in Table 1, Fig. 1, respectively.

The images were evaluated in three stages: a training session, consensus meeting, and validation session. Four expert colonoscopists, including H.J.Y., S.K.P., and Y.S.J. with experience with at least 3,000 colonoscopy cases, and six trainee

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition</th>
<th>Image examples from Fig. 1</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indistinctive borders</td>
<td>Vague demarcation of a border of a lesion</td>
<td>B, C, D</td>
<td>Hazewinkel et al. (2013)15</td>
</tr>
<tr>
<td>Irregular shape</td>
<td>An asymmetric shape in contrast to the oval, circular shape of small hyperplastic polyps and conventional adenomas</td>
<td>B, C, E, F</td>
<td>Hazewinkel et al. (2013)15</td>
</tr>
<tr>
<td>Rim of debris</td>
<td>A conspicuous ring of debris/bubbles encircling ≥25% of the lesion</td>
<td>A, C</td>
<td>Tadepalli et al. (2011)14</td>
</tr>
<tr>
<td>Cloud-like surface</td>
<td>A bumpy, soft-looking nodular surface resembling a cumulus cloud</td>
<td>B, F</td>
<td>Hazewinkel et al. (2013)15</td>
</tr>
<tr>
<td>Mucous cap</td>
<td>Focal collection of mucus (clear, bile stained, or debris stained) on the mucosal surface that can be washed off with irrigation</td>
<td>A, E</td>
<td>Tadepalli et al. (2011)14</td>
</tr>
<tr>
<td>Nodular surface</td>
<td>Focal, subtle irregularity or bumpiness of the mucosal surface</td>
<td>E</td>
<td>Tadepalli et al. (2011)14</td>
</tr>
<tr>
<td>Absence of surface vessels</td>
<td>Small superficial vessels or telangiectasias occasionally seen on the surface of particular distal hyperplastic polyps</td>
<td>A, B, D, G</td>
<td>Hazewinkel et al. (2013)15</td>
</tr>
<tr>
<td>Dark spots</td>
<td>Small red dots inside the open crypts on high-resolution white-light endoscopy</td>
<td>F, G</td>
<td>Hazewinkel et al. (2013)15</td>
</tr>
</tbody>
</table>
Colonoscopists served as image assessors during the entire process. Before the training session, the assessors were provided with descriptive and visual examples of the endoscopic features of SSAs (Table 1, Fig. 1). During the training session, 30 polyps from the learning set were presented with the location and size in a random order, and the assessors independently determined the presence or absence of each endoscopic feature for each polyp. They were blinded to the histology of the polyps. After the training session, all assessors were invited to a consensus meeting where they were unblinded to the histopathologic results of the polyps from the learning set; these assessors discussed the presence of endoscopic features for each polyp with the colonoscopist who prepared the images and features (DIP), and they reached a consensus. During the validation session, the assessors independently evaluated 51 polyps from the validation set in the same manner as they did for the learning set. The study protocol was approved by the Institutional Review Board of Kangbuk Samsung Hospital (KBSMC 2014-01-069), and the requirement to obtain informed consent was waived.

**Statistical analysis**

The results from the validation set were used for the data analysis in the form of pooled data: 10 assessments for each feature of each polyp. Initially, we assessed the proportion of endoscopic features according to the types of polyps. To identify the independent predictive features of SSA, a multiple logistic regression model was fitted to the statistically significant (p<0.1) features determined by univariate analysis. Then interobserver agreement of independent predictors among

---

**Fig. 1.** Endoscopic features of sessile serrated adenoma (SSA). (A) A 6-mm SSA located in the proximal transverse colon showing a rim of debris, mucous cap, and absence of surface vessels. (B) A 15-mm SSA located in the cecum showing indistinctive borders (arrows), an irregular shape, and a cloud-like surface without surface vessels. (C) A 12-mm SSA located in the mid-ascending colon showing a rim of debris, indistinctive borders, and an irregular shape. (D) A 12-mm SSA located in the splenic flexure demonstrating the absence of surface vessels with indistinctive borders. (E) A 12-mm SSA with an irregular shape located in the cecum covered with a mucous cap on the nodular surface. (F) A 6-mm SSA located in the cecum showing dark spots inside the crypts in addition to an irregular shape and cloud-like surface. (G) A 15-mm SSA located in the proximal ascending colon with dark spots inside the crypts and the absence of surface vessels.
the expert and trainee colonoscopists was assessed by using $\kappa$ statistics. Agreement was considered very good for $\kappa$ values of 0.81 to 1.00, good for $\kappa$ values of 0.61 to 0.80, moderate for $\kappa$ values of 0.41 to 0.60, fair for $\kappa$ values of 0.21 to 0.40, and poor for $\kappa$ values of 0 to 0.20.\textsuperscript{23} Finally, the area under the receiver-operating characteristic curve (AUROC) was measured to evaluate the diagnostic accuracy of the selected endoscopic features. All statistical analyses were performed using SPSS version 19.0 (IBM Co., Armonk, NY, USA), and $p<0.05$ was considered significant.

RESULTS

Prevalence of individual endoscopic features

During the validation session, the images of 51 polyps (29 SSAs, 12 HPs, and 10 TAs) were assessed by 10 colonoscopists (four experts and six trainees). The size and location of the polyps are summarized in Table 2. Although SSAs were more likely to be located proximal than the non-SSAs ($p<0.001$), the proportions of large ($\geq 10$ mm) polyps were not different between SSAs and non-SSAs ($p=0.155$).

Overall, the median number of endoscopic features assessed to be present was 2 (interquartile range [IQR], 0 to 4). SSAs had a significantly more features (median 3 features; IQR, 2 to 5) than HPs (median 0; IQR, 0 to 2) and TAs (median 0; IQR, 0 to 1) ($p<0.001$). The prevalence of individual features according to the histology of the polyps among the colonoscopy experts and trainees are presented in and Supplementary Table 1, Fig. 2. The most frequently reported feature of SSAs observed by experts (57.8%) and trainees (51.7%) was indistinctive borders, followed by a mucous cap (experts 53.4%, trainees 48.9%), the absence of surface vessels (experts 55.2%, trainees 44.3%), and an irregular shape (experts 51.7%, trainees 42.5%). All features were more frequently found in SSAs than in non-SSAs, according to experts and trainees ($p<0.01$, all).

Independent predictors and interobserver agreement

In the multivariable logistic regression model, a mucous cap (odds ratio [OR], 10.44; 95% confidence interval [CI], 5.72 to 19.07; $p<0.001$), indistinctive borders (OR, 4.21; 95% CI, 2.74 to 7.16; $p<0.001$), cloud-like surface (OR, 2.43; 95% CI, 1.27 to 4.68; $p=0.008$), and dark spots (OR, 3.64; 95% CI, 1.89 to 7.00; $p<0.001$) were independent predictive features associated with SSAs (Table 3). In the comparison between SSAs and HPs, a mucous cap (OR, 7.05; 95% CI, 3.64 to 13.69; $p<0.001$), indistinctive borders (OR, 2.60; 95% CI, 1.46 to 4.63; $p=0.001$), and dark spots (OR, 2.29; 95% CI, 1.14 to 4.59; $p=0.020$) were still independent predictors, whereas a cloud-like surface (OR, 1.81; 95% CI, 0.91 to 3.62; $p=0.092$) was not.

When the interobserver agreement was evaluated, three of the independent features showed moderate interobserver agreement both among the experts and trainees: a mucous cap (mean $\kappa$ values: experts 0.41, trainees 0.46), indistinctive borders (mean $\kappa$ values: experts 0.43, trainees 0.52), and cloud-like surface (mean $\kappa$ values: experts 0.54, trainees 0.42) (Table 3). Although the presence of dark spots was also an independent predictor, it had poor interobserver agreement (mean $\kappa$ val-

---

### Table 2. The Location and Size of Polyps According to Histological Analysis in the Validation Set

<table>
<thead>
<tr>
<th>Variable</th>
<th>SSA ($n=29$)</th>
<th>HP ($n=12$)</th>
<th>TA ($n=10$)</th>
<th>$p$-value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>28 (96.6)</td>
<td>4 (33.3)</td>
<td>7 (70.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cecum</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Ascending colon</td>
<td>17</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Hepatic flexure</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Transverse colon</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Distal</td>
<td>1 (3.4)</td>
<td>8 (66.7)</td>
<td>3 (30.0)</td>
<td></td>
</tr>
<tr>
<td>Descending colon</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Sigmoid colon</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Rectum</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Size, mm</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.155</td>
</tr>
<tr>
<td>&lt;10</td>
<td>14 (48.3)</td>
<td>9 (75.0)</td>
<td>6 (60.0)</td>
<td></td>
</tr>
<tr>
<td>$\geq 10$</td>
<td>15 (51.7)</td>
<td>3 (25.0)</td>
<td>4 (40.0)</td>
<td></td>
</tr>
</tbody>
</table>

Values are presented as number (%).

SSA, sessile serrated adenoma; HP, hyperplastic polyp; TA, tubular adenoma.

$^a$SSAs vs. non-SSAs.
Endoscopic prediction of SSAs

The sensitivity and specificity of endoscopically predicting SSAs were evaluated based on three independent predictors with moderate interobserver agreement (a mucous cap, indistinctive borders, and cloud-like surface). When there was at least one of three features, SSAs could be differentiated from HPs and TAs with a sensitivity of 79.0% and a specificity of 81.4% in the validation set (Fig. 3). The AUROC of the number of features that predicted SSAs was 0.815 (95% CI, 0.777 to 0.854), indicating excellent discrimination. However, regarding the discrimination between SSAs and HPs using these features, the specificity decreased to 72.5%, although the sensitivity was maintained at 79.0% (AUROC, 0.771; 95% CI, 0.720 to 0.822).

As 21.0% (61 of 290) of SSAs in the validation set were considered to have none of the three features, the prevalence
of the other five features was estimated in these false-negative cases. The most prevalent additional features of SSAs were dark spots (26.2%, 16 of 61) and the absence of surface vessels (26.2%, 16 of 61), followed by rim of debris (11.5%, 7 of 61), nodular surface (3.3%, 2 of 61), and irregular shape (3.3%, 2 of 61).

DISCUSSION

We examined eight previously reported characteristic features of SSAs using HR-WLE to determine their usefulness in the independent and reliable endoscopic prediction of SSAs among colonoscopy trainees and experts. We found that a mucous cap, indistinctive borders, and cloud-like surface were independent predictive features for SSA histology with moderate interobserver agreement both in expert and trainee colonoscopists. Moreover, we showed that the endoscopic prediction of SSAs was feasible based on these features.

With increased awareness of the serrated pathway, there have been attempts to characterize the morphology of SSA systematically using HR-WLE, NBI, and NBI with magnification. However, it may be too complicated and difficult for colonoscopists after training to be well acquainted with all these findings, especially those that can be found using NBI with or without magnification. Thus, it is necessary to identify endoscopic features on HR-WLE that are simple, specific, and reliable for colonoscopy training. We found three features that meet those requirements: a mucous cap, indistinctive borders, and cloud-like surface. These findings are comparable to those of previous reports. A mucous cap is composed of adherent bile salts and fecal debris that accumulate from excessively secreted mucus. This can be a specific characteristic of SSAs, because SSAs have abundant mucinous material and mature goblet cells in the crypt base, which cause the hypersecretion of mucin. Tadepalli et al. reported that a mucous cap was the most frequently observed characteristic of SSA. Similarly, in our validation set, more than half of cases were assessed to have it. Indistinctive borders and a cloud-like surface are also well-known characteristics of SSAs. In a systematic evaluation reported by Hazewinkel et al., these were the only two features associated with SSAs on HR-WLE. In another Japanese study by Yamada et al., a granular surface and vague margin were reported to be endoscopic findings more frequently observed for SSAs than for HPs. These three predictors also showed moderate interobserver agreement among experts and trainees, rendering their use feasible during colonoscopy. On the basis of the three characteristics, we could predict SSAs with a sensitivity of 79.0% and a specificity of 81.4%. As it is important to not overlook SSAs in clinical practice, the sensitivity is more important than the specificity. From this perspective, a sensitivity of <80% would be unsatisfactory, although our result is similar to that in a previous study (75%) in which HR-WLE was used; additionally, our result was even higher than that reported in another study where the sensitivity of a single finding on NBI magnification was 65%. To determine how to improve the sensitivity of identifying SSAs, we identified additional characteristics that should be assessed when all the main features are absent. Dark spots can explain one-fourth of the false-negative identifications of SSAs. Dark spots were also one of the independent predictors
associated with SSAs. Histologically, dark spots also represent the characteristic mucin hypersecretion of SSAs. However, they had a poor interobserver agreement in our study. This was probably because it may be more suitable to screen for this feature using NBI than HR-WLE, as shown in previous studies where dark spots were useful for differentiating SSAs from HPs or TAs using NBI with or without magnification, but not with HR-WLE. Therefore, NBI may be able to improve the identification of dark spots, which may then increase the sensitivity of predicting SSAs. It is also noteworthy that the interobserver agreement among trainee colonoscopists was not inferior to that among expert colonoscopists, although only a short image test and consensus meeting were performed in our study. Nevertheless, the overall interobserver agreement was suboptimal, and it needs to be improved. Thus, further study is necessary to evaluate the usefulness of NBI for colonoscopy trainees to make a diagnosis of SSAs with special emphasis on the overall and dark spot-specific interobserver agreements.

The strength of our study is that we systematically validated known SSA features among expert colonoscopists and trainee colonoscopists. Consequently, we suggest a set of features that can be used in colonoscopy training to endoscopically identify SSAs. However, this study has several limitations. First, polyps were evaluated retrospectively by using images rather than videos. Second, the numbers of different polyp types did not represent those seen in clinical practice. In clinical practice, TAs and HPs are observed more frequently than SSAs. However, in our image set, SSAs were two to three times more prevalent than non-SSAs. Third, the polyp images were presented to the assessors with the location and size during the test. This information could have influenced the assessors' decisions considering that SSAs are more likely to be proximally located than other polyps; however, this study can be considered a simulation of clinical practice. Fourth, we only evaluated the interobserver agreement once after the training session and consensus meeting. If changes in interobserver agreement before and after the training session and consensus meeting had been evaluated, more information could have been provided. Lastly, although our results may be helpful to trainee colonoscopists for differentiating between SSAs and non-SSAs, this does not necessarily mean that the SSA detection rate will be improved during routine colonoscopy, as this needs to be evaluated in further studies.

In conclusion, our study's results indicated that trainee and expert colonoscopists can use a mucous cap, indistinctive borders, and cloud-like surface on HR-WLE to independently and reliably predict SSAs. A combination of these features may enable an acceptable on-site prediction of SSA histology during colonoscopy. However, the interobserver agreement of the predictors and sensitivity of SSA prediction using HR-WLE remain to be improved.

Conflicts of Interest
The authors have no financial conflicts of interest.

Supplementary Material
Supplementary materials are available at https://doi.org/10.5946/ce.2016.107 or via http://e-ce.org/.

REFERENCES
17. Yamashina T, Takeuchi Y, Uedo N, et al. Diagnostic features of ses-