Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos)

Yuichi Mori, MD, PhD,1 Shin-ei Kudo, MD, PhD,1 Kunihiko Wakamura, MD,1 Masashi Misawa, MD,1 Yushi Ogawa, MD,1 Makoto Kutsukawa, MD,1 Toyoki Kudo, MD,1 Takemasa Hayashi, MD,1 Hideyuki Miyachi, MD, PhD,1 Fumio Ishida, MD, PhD,1 Haruhiko Inoue, MD, PhD2

Yokohama, Japan

Background: Endocytoscopy enables in vivo observation of nuclei at 450× magnification during GI endoscopy, thus allowing precise prediction of lesion pathology. However, because it requires training and experience, it may be beneficial only when performed by expert endoscopists.

Objective: To develop and evaluate a novel computer-aided diagnosis system for endocytoscopic imaging (EC-CAD) of colorectal lesions.

Design: Pilot study.

Setting: University hospital.

Patients: One hundred fifty-two patients with small colorectal polyps (≤10 mm) who had undergone endocytoscopy.

Intervention: Test sets of white-light endoscopic images and endocytoscopic images from 176 small colorectal polyps (137 neoplastic and 39 non-neoplastic polyps) were assessed by EC-CAD, 2 expert endoscopists, and 2 trainee endoscopists.

Main Outcome Measurement: Sensitivity, specificity, and accuracy in predicting neoplastic change by EC-CAD comparing expert and trainee endoscopists.

Results: EC-CAD had a sensitivity of 92.0% and an accuracy of 89.2%; these were comparable to those achieved by expert endoscopists (92.7% and 92.3%; P = .868 and .256, respectively) and significantly higher than those achieved by trainee endoscopists (81.8% and 80.4%; P < .001 and .002, respectively). EC-CAD achieved a specificity of 79.5%; this did not differ significantly from that achieved by the experts and trainees. EC-CAD also enabled instant diagnosis, taking only 0.3 seconds for each lesion with perfect reproducibility.

Limitations: No sample size calculation.

Conclusions: EC-CAD provides fully automated instant classification of colorectal polyps with excellent sensitivity, accuracy, and objectivity. Thus, it can be a powerful tool for facilitating decision making during routine colonoscopy. (Gastrointest Endosc 2015;81:621-9.)

(footnotes appear on last page of article)
Computer-aided diagnosis by using endocytoscopy

Mori et al

Both the endoscopic examinations and the test set evaluation took place at Showa University Northern Yokohama Hospital, a tertiary referral center in Japan. The patients had undergone bowel preparation with 2 to 3 L of polyethylene glycol solution before colonoscopy. Diazepam and butylscopolamine had been administered intravenously for sedation and prevention of peristalsis, respectively. The endoscopic images had been taken by experienced endoscopists who were not the test evaluators in this study. The study protocol was approved by the Ethics Committee of Showa University Northern Yokohama Hospital (no. 1312-02; December 27, 2013) and registered as a UMIN clinical trial (UMIN000012797). All participants gave written informed consent, and the study was conducted according to the Declaration of Helsinki.

Endocytoscopy

An integrated-type EC scope (CF Y-0020-I, prototype from Olympus) was used to acquire the endoscopic images in all cases in the study. The integrated-type EC has 2 separate observation modes: standard videoendoscopy and EC modes. By using a hand-operated lever, endoscopists can consecutively perform EC observations in addition to standard videoendoscopy without changing the endoscope. EC observation is based on the principles of contact light microscopy: the CF Y-0020-I has a 380× magnification with a focusing depth of 50 μm and a field of view of 700 × 600 μm. EC images were obtained after staining with 1% methylene blue and 0.05% crystal violet, thus providing morphological images of nuclei and gland duct lumens in the superficial layer that are similar to micrographic images of fixed specimens stained with hematoxylin and eosin (Fig. 1). The EC equipment does not differ in its main specifications and appearance from a normal colonoscope (eg, CF Y-0020-I is 13.6 mm in diameter and 1330 mm in length and has a 3.2-mm instrumental channel in the 5 o’clock position with a water jet channel in the 7 o’clock position). Thus, any form of treatment (even a large polypectomy) can be performed with EC in the same manner as with a normal colonoscope.

Computer-aided diagnostic system for EC imaging

In this study, we evaluated a newly developed fully automated diagnostic system named EC-CAD. EC-CAD is directly connected to an endoscopy unit (EVIS LUCERA ELITE; Olympus). Thus, instant output of real-time automated diagnoses is available by simply pushing the release button of the endoscope to acquire an image. Total time from pushing the release button of the endoscope to output of a diagnosis is only 0.3 seconds. Receiving real-time histology predictions from EC-CAD is expected to assist endoscopists (especially trainee endoscopists) in diagnosing lesions more precisely. This system is now used routinely for screening colonoscopies at Showa University Northern Yokohama Hospital with the approval of
the local Ethics Committee (no. 1401-09; February 20, 2014). Illustrative examples of performing colonoscopy by using EC-CAD are shown in Videos 1 and 2 (available online at www.giejournal.org).

The image processing and calculation algorithm of EC-CAD comprises the following 3 steps: nuclear segmentation, feature extraction, and output of predicted pathological classification.

Nuclear segmentation is automatically performed after image acquisition from the original EC image. Image processing for nuclei segmentation has 2 phases. First, Otsu's adaptive thresholding based on the red ingredient is performed to label the nucleus-like spots. In the second phase, the labeled nuclear-like spots are filtered by using many types of threshold, such as size, diameter, and shape, to remove artifacts, thus extracting the nuclei. The sequence of automated image processing for nuclear segmentation and examples of extracted nuclei are shown in Figures 2 and 3, respectively.

After nuclear segmentation, features are automatically extracted from each nucleus. EC-CAD extracts the following 6 features of the nuclei: area, standard deviation of area, circularity, circularity of the 20 largest nuclei, shortest diameter, and longest diameter (Fig. 4). These 6 features were identified by multivariate logistic regression analysis as independently relevant to pathological classification.

After these features of the nuclei have been extracted, the predicted pathological classification is calculated according to an algorithm, and the result is displayed. The algorithm for EC-CAD has been programmed based on regression equations for each of the pathological classifications defined by the 6 assessed features of nuclei. The output of the
predicted pathological classification is expressed as 1 of the following 3 categories: non-neoplastic, adenoma, and cancer. If fewer than 30 nuclei have been extracted, the display shows “unable to diagnose” rather than a pathological classification. We select 30 nuclei as the minimum number for sampling because identification of too few nuclei (<30) significantly decreases the diagnostic accuracy of EC-CAD. This is due to the inherent strong bias caused by small samples and the relatively greater significance of artifacts created by mucus.

**Image evaluation by EC-CAD and endoscopists**

The test sets were evaluated by EC-CAD, 2 expert endoscopists (K.W. and M.M.), and 2 trainee endoscopists.
The EC classiﬁcation was either neoplastic or non-neoplastic according to the experts and trainees diagnosed each EC image in the test sets, which were displayed on the monitor. Both the endoscopists evaluated the test image was automatically transferred from the test set to EC-CAD, whereas the endoscopists evaluated the test images. For assessment of EC-CAD, each EC image was automatically transferred from the test set to EC-CAD without sample size calculation. The sensitivity, speciﬁcity, and accuracy for diagnosing neoplastic change were compared with the McNemar test. In this analysis, diagnostic outputs of “unable to diagnose” from EC-CAD were provisionally treated as misdiagnoses. The average times to make diagnoses were compared by applying the Student t test. Two-sided P values < .017 after applying Bonferroni’s correction were considered statistically signiﬁcant for comparisons among the 3 groups (EC-CAD, the experts, and the trainees). For validation of the classiﬁcations, the proportion of agreement and unweighted Cohen κ coefﬁcient were determined (strength of agreement considered as follows: 0.01–0.2, slight; 0.21–0.4, fair; 0.41–0.6, moderate; 0.61–0.8, substantial; and 0.81–1.0, almost perfect). All statistical analyses were performed by using R Project for Statistical Computing version 2.10.0 (Vienna, Austria).

**RESULTS**

**Relevant data concerning patients and tumors in the test set**

The test set, comprising 176 polyps from 152 patients (107 male and 45 female; average age, 64.2 ± 12.1 years), was extracted from the EC database. The characteristics of the 176 polyps are presented in Table 1. The mean size of the lesions was 6.3 ± 2.4 mm. The polyps were located in the right colon in 77 cases, in the left colon in 66, and in the rectum in 33. Ninety-seven of the lesions were of the protruded type (Paris types Is and Ip), 71 of the slightly elevated type (Paris type IIa), and 8 of the slightly depressed type (Paris types Iic and Iic+Iic).19 Histopathological evaluation conﬁrmed that 137 of the polyps were neoplastic (130 adenomas, 7 invasive cancers) and 39 were non-neoplastic.

**EC-CAD versus EC image evaluation by the endoscopists**

EC-CAD automatically output the predicted pathology of all except 8 EC-images immediately after their input from the test set. These 8 images (2 of lesions diagnosed pathologically as non-neoplastic and 6 as neoplastic) were assessed by EC-CAD as “unable to diagnose” because too few nuclei were extracted; these were treated as misdiagnoses in our analysis.

Table 2 shows the diagnostic performance of EC-CAD compared with EC-image evaluation by the endoscopists. EC-CAD had a sensitivity of 92.0% (95% conﬁdence interval [CI], 86.1%-95.9%) and an accuracy of 89.2% (95% CI, 83.7%-93.4%), which were comparable to those achieved by the experts (92.7%; 95% CI, 89.0%-95.5% and 92.3%; 95% CI, 89.0%-94.9%; P = .868 and .256, respectively) and signiﬁcantly higher than those achieved by the trainees (81.8%; 95% CI, 76.6%-86.1% and 81.8%; 95% CI, 76.6%-86.1% and respectively). Statistical analysis

This study was an initial pilot study for validation of EC-CAD without sample size calculation. The sensitivity, speciﬁcity, and accuracy for diagnosing neoplastic change were compared by the McNemar test. In this analysis, diagnostic outputs of “unable to diagnose” from EC-CAD were provisionally treated as misdiagnoses. The average times to make diagnoses were compared by applying the Student t test. Two-sided P values < .017 after applying Bonferroni’s correction were considered statistically signiﬁcant for comparisons among the 3 groups (EC-CAD, the experts, and the trainees). For validation of the classiﬁcations, the proportion of agreement and unweighted Cohen κ coefﬁcient were determined (strength of agreement considered as follows: 0.01–0.2, slight; 0.21–0.4, fair; 0.41–0.6, moderate; 0.61–0.8, substantial; and 0.81–1.0, almost perfect). All statistical analyses were performed by using R Project for Statistical Computing version 2.10.0 (Vienna, Austria).

**TABLE 1. Characteristics of the study lesions**

<table>
<thead>
<tr>
<th>Summary (N = 176)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesion size, mm</td>
<td>6.3 (2.4)</td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Right colon</td>
<td>77 (43.7)</td>
</tr>
<tr>
<td>Left colon</td>
<td>66 (37.5)</td>
</tr>
<tr>
<td>Rectum</td>
<td>33 (18.8)</td>
</tr>
<tr>
<td>Shape (Paris classification)</td>
<td></td>
</tr>
<tr>
<td>Polypoid (Is, Ip)</td>
<td>97 (55.2)</td>
</tr>
<tr>
<td>Slightly elevated (Ila)</td>
<td>71 (40.3)</td>
</tr>
<tr>
<td>Slightly depressed (Iic, Iic+Iic)</td>
<td>8 (4.5)</td>
</tr>
<tr>
<td>Histopathology of resected specimens</td>
<td></td>
</tr>
<tr>
<td>Non-neoplastic</td>
<td></td>
</tr>
<tr>
<td>Hyperplastic polyp</td>
<td>30 (17.0)</td>
</tr>
<tr>
<td>Inflammatory polyp</td>
<td>5 (2.8)</td>
</tr>
<tr>
<td>Juvenile polyp</td>
<td>4 (2.3)</td>
</tr>
<tr>
<td>Neoplastic</td>
<td></td>
</tr>
<tr>
<td>Low-grade adenoma</td>
<td>104 (59.1)</td>
</tr>
<tr>
<td>High-grade adenoma</td>
<td>26 (14.8)</td>
</tr>
<tr>
<td>Invasive cancer</td>
<td>7 (4.0)</td>
</tr>
</tbody>
</table>

Data are presented as mean (standard deviation) or number (%).

(F.U. and K.S.). For assessment of EC-CAD, each EC image was automatically transferred from the test set to EC-CAD, whereas the endoscopists evaluated the test sets, which were displayed on the monitor. Both the experts and trainees diagnosed each EC image in the test set as either neoplastic or non-neoplastic according to the EC classification proposed by Kudo et al14 after assessing a WLE image of the same lesion. Both the experts and trainees were blinded to the histological findings on the subject lesions. Each expert had experience of more than 2000 colonoscopies and 50 EC procedures, whereas each trainee had experience of fewer than 500 colonoscopies and had never used EC. The trainees had been trained to make diagnoses from EC images by studying relevant lecture slides for 1 hour before assessing the test set. The ability of EC-CAD to predict neoplastic change was compared with that of the experts and trainees with reference to the pathology of the resected specimens. The average time to make a diagnosis from each image and inter- and intraobserver agreement (by using 20 WLE and EC images 4 weeks after the first test) were also measured and compared among the 3 groups.
80.4%; 95% CI, 75.9%-84.4%; P < .001 and .002, respectively). EC-CAD achieved a specificity of 79.5% (95% CI, 63.5%-90.7%); this was not significantly different from that achieved by the experts and trainees. EC-CAD also made instant diagnoses (0.3 seconds) from each image, which was significantly shorter than the time taken by both the experts and trainees (P < .001 and .001, respectively).

Intraobserver agreement was perfect for EC-CAD (κ score = 1), whereas for the experts inter- and intraobserver agreement was substantial (κ score = 0.68 and 0.79, respectively). For the trainees, interobserver agreement was moderate, and intraobserver agreement was substantial (κ score = 0.58 and 0.71, respectively).

**EC-CAD versus WLE image evaluation by the endoscopists**

We also compared the diagnostic performance of EC-CAD with that of WLE image evaluation by the endoscopists. Table 3 shows the details. EC-CAD achieved significantly higher sensitivity, specificity, and accuracy than the trainees (92.0% [95% CI, 86.1%-95.9%] versus 83.2% [95% CI, 78.2%-87.4%] in sensitivity, 79.5% [95% CI, 63.5%-90.7%] versus 56.4% [95% CI, 44.7%-67.6%] in specificity, and 89.2% [95% CI, 83.7%-95.4%] versus 77.3% [95% CI, 72.5%-81.5%] in accuracy; P = .002, .009, and <.001, respectively). In contrast, there were no significant differences between EC-CAD and the experts in sensitivity, specificity, and accuracy.

Regarding inter- and intraobserver agreement of WLE image evaluation by endoscopists, the experts achieved substantial inter- and intraobserver agreement (κ score = 0.64 and 0.74, respectively), whereas the trainees achieved fair interobserver and substantial intraobserver agreement (κ score = 0.40 and 0.74, respectively).

**Subanalysis based on location and histology**

We performed a subgroup analysis of the accuracy of diagnosis of polyps by EC-CAD and endoscopists based on location: right colon, left colon, and rectum. We found no significant differences between these locations in diagnostic accuracy. The diagnostic accuracies in the right side of the colon, left side of the colon, and rectum were 90.9, 89.4, and 84.8% by EC-CAD; 90.3, 94.7, and 92.4% by the
By expert endoscopists.10-12,21-23 The same is true of EC, performance, most of them achieving accuracies of greater than 90%. These have been proven to have excellent diagnostic performance. Many endoscopic technologies, such as NBI, FICE, and confocal laser endomicroscopy, have recently been developed to enable more precise diagnosis of small polyps. These have been proven to have excellent diagnostic performance, most of them achieving accuracies of greater than 90%; however, such accuracy can only be achieved by expert endoscopists.10-12,21-23 The same is true of EC, an emerging endoscopic technology that allows real-time cellular imaging of alimentary mucosa and provides very accurate predictions of pathology (eg, accuracies of 94.1%17 and 96.5%).25 However, with this technology as well, such accuracy is achieved only by trained and experienced expert endoscopists. We were therefore determined to develop automated diagnostic software for EC that could support nonexpert endoscopists to diagnose lesions precisely without any training.

In this study, EC-CAD achieved a sensitivity of 92.0%, specificity of 79.5%, and accuracy of 89.2% for identifying neoplastic change in small polyps (≤10 mm). The excellent sensitivity and accuracy were comparable to those achieved by the EC image evaluation by experts and significantly higher than those achieved by EC image evaluation by the trainees. We consider that EC-CAD’s specificity of 79.5% makes it a feasible means of screening. In addition, EC-CAD provides diagnoses extremely rapidly (0.3 seconds for analysis of 1 EC image) and with perfect reproducibility (κ score = 1). Such high sensitivity, rapidity, and perfect reproducibility are very important characteristics for a screening modality. Thus, EC-CAD would be an attractive real-time aid in decision making by nonexpert endoscopists assessing small colorectal polyps. In this study, we also compared the diagnostic performance of EC-CAD with that of WLE-image evaluation by endoscopists. We found that EC-CAD achieved significantly higher diagnostic performance than the trainees in terms of specificity, sensitivity, and accuracy, which means that trainees can receive much more benefit from using EC-CAD than diagnosing lesions manually based on WLE image assessment.

As for the practicability of EC-CAD, there are no significant barriers that would hinder its installation in clinical practice; endoscopists need no expertise in technique or computer skills, they only need to push the endoscope’s release button to receive real-time diagnosis output. We also consider that acquisition of EC images by nonexperts would have no influence on the accuracy of EC-CAD diagnoses because it is relatively easy, even for trainees. It requires no focus adjustment; putting the endoscope in contact with the lesion with the hand lever fully pulled is all that is required. We believe that a precise and objective fully automated diagnostic system for endoscopy is necessary for improving the management of colonic lesions, although as yet no automated diagnostic system is commercially available. Practically, with the “polyp leave in” protocol of our institution, given that we have great confidence in our endoscopic diagnoses based on the diagnostic output from EC-CAD, all polyps in the proximal colon and all adenomas are resected, whereas rectosigmoid hyperplastic polyps 5 mm or smaller are left in situ.

As a future prospect, EC-CAD could potentially detect dysplasia in patients with ulcerative colitis by assessing nuclear abnormalities. These dysplasias are still difficult to detect with normal endoscopic technologies; thus, future application of EC-CAD to this area would provide a major benefit in routine medical practice.

On the other hand, EC-CAD has some drawbacks that could influence its practicability. First, it is relatively cumbersome, taking longer than a conventional procedure because of the time required to place the endoscope in contact with the lesion, press the release button of the endoscope, and check the computer diagnosis, in addition to the time for staining. However, the Preservation and Incorporation of Valuable endoscopic Innovations initiative stresses the importance of an accurate diagnosis based on high-quality endoscopic technology for safe management of diminutive polyps for both “resect and discard” and “polyp leave in” protocols.27 Thus, given that EC-CAD provides excellent accuracy and objectivity regardless of the skills of the endoscopist, the relatively long time required for this procedure could be acceptable, especially for trainee endoscopists. Second, in this study, no diagnosis was made by EC-CAD in 4.5% (8/176) of the subjects’ EC images because too few nuclei were extracted as a result of insufficient staining. Investigation of a more effective staining method is needed.
There are some previous studies in the field of automated diagnosis of endoscopy, most of which have been based on NBI. For computer analysis of vascularization features on NBI, Tischendorf et al. reported a sensitivity of 93.8% and a specificity of 61.2%, Gross et al. reported a sensitivity of 93.0% and a specificity of 90.3%, and Takeiura et al. a sensitivity of 97.8% and a specificity of 97.9%.

All of these are very high diagnostic performances; however, these diagnostic systems were not fully automated but required manual operation during image acquisition and processing by the system. In this regard, our newly developed EC-CAD provides fully automated diagnoses, requires no manual operation, and facilitates real-time decision making during endoscopy. Most previously reported systems have used support vector machines with linear kernels as classifiers for their algorithm because support vector machines are one of the optimal classifiers for separating highly complex distributions. In contrast, multivariate analysis of relevant features was performed to construct EC-CAD’s algorithm, achieving satisfactory results. The high diagnostic ability of EC-CAD may be attributable to the fact that EC-CAD focuses on shapes of nuclei, which are the most relevant factors in conventional pathological diagnosis. However, we are investigating using support vector machines with the classification algorithm of EC-CAD with the aim of achieving better performance from our next version of our model.

This study has several limitations. First, it was an initial pilot study without sample size calculation, which has possibly caused some methodological bias, compromising generalization of the study results. Second, a miss rate for diagnosis of neoplastic polyps of 8% in a “polyp leave in” protocol would be a concern. However, in clinical practice, diagnoses of polyps are based not only on the diagnostic output from EC-CAD, but also on many other factors, such as their location, color, and size. With this additional information, the miss rate would therefore be less than 8% to provide an acceptable diagnostic performance for a “polyp leave in” protocol. Third, diagnosis from still images may not be as accurate as real-time analysis: this may have skewed the data of our study in favor of EC-CAD. To compensate for these 3 limitations, we are now conducting a prospective study to evaluate the clinical effectiveness of EC-CAD (UMIN000013917); this study takes into account other endoscopically provided information in addition to EC-CAD’s diagnostic output. Fourth, heterogeneity of cytoplasm and reversed nucleus/cytoplasm ratios were not evaluated when developing EC-CAD, which were both important features for in vivo assessment of neoplastic lesions. This is because extraction of cytoplasm from EC images is considerably difficult, whereas extraction of nuclei is relatively easy. EC can recognize clusters of cytoplasm but cannot identify cytoplasm of individual cells because it is unable to recognize cell membranes. To compensate for this drawback, we evaluated variability in size and average size of nuclei instead. To some extent, these nuclear features indirectly reflect the heterogeneity of cytoplasm and nucleus/cytoplasm ratio. Finally, excluding SSAs/Ps was a major limitation of this study, considering the importance of SSAs/Ps that have malignant potential and are managed clinically in a similar fashion to adenomatous polyps. Given that the algorithm for EC-CAD so far focuses exclusively on morphology of nuclei, it is currently very difficult to diagnose SSAs/Ps, and there is a risk of EC-CAD misdiagnosing them as non-neoplastic; SSAs/Ps usually have no nuclear abnormalities. Adoption of a technique for assessing the morphology of crypts will be necessary to enable diagnosis of SSAs/Ps by EC-CAD in the future.

In conclusion, the fully automated EC-CAD system described here provides excellent sensitivity and accuracy for identifying neoplastic change in small colorectal polyps, its sensitivity and accuracy being comparable with those achieved by expert endoscopists and significantly higher than those achieved by trainees. In addition, its diagnostic output occurs ultrarapidly and with perfect objectivity. These superiorities of EC-CAD can help nonexpert endoscopists to identify adenomas during routine colonoscopy.

ACKNOWLEDGMENTS

We express our gratitude to Olympus (Tokyo, Japan) and Cybernet Systems (Tokyo, Japan) for their invaluable support in developing EC-CAD.

REFERENCES


