



Endoscopic three-categorical diagnosis of *Helicobacter pylori* infection using linked color imaging and deep learning: a single-center prospective study (with video)

Hiroataka Nakashima¹ · Hiroshi Kawahira² · Hiroshi Kawachi³ · Nobuhiro Sakaki¹

Received: 5 March 2020 / Accepted: 23 April 2020 / Published online: 7 May 2020
© The International Gastric Cancer Association and The Japanese Gastric Cancer Association 2020

Abstract

Background *Helicobacter pylori* (*H. pylori*) eradication is required to reduce incidence related to gastric cancer. Recently, it was found that even after the successful eradication of *H. pylori*, an increased, i.e., moderate, risk of gastric cancer persists in patients with advanced mucosal atrophy and/or intestinal metaplasia. This study aimed to develop a computer-aided diagnosis (CAD) system to classify the status of *H. pylori* infection of patients into three categories: uninfected (with no history of *H. pylori* infection), currently infected, and post-eradication.

Methods The CAD system was based on linked color imaging (LCI) combined with deep learning (DL). First, a validation dataset was formed for the CAD systems by recording endoscopic movies of 120 subjects. Next, a training dataset of 395 subjects was prepared to enable DL. All endoscopic examinations were recorded using both LCI and white-light imaging (WLI). These endoscopic data were used to develop two different CAD systems, one for LCI (LCI-CAD) and one for WLI (WLI-CAD) images.

Results The diagnostic accuracy of the LCI-CAD system was 84.2% for uninfected, 82.5% for currently infected, and 79.2% for post-eradication status. Comparisons revealed superior accuracy of diagnoses based on LCI-CAD data relative based on WLI-CAD for uninfected, currently infected, and post-eradication cases. Furthermore, the LCI-CAD system demonstrated comparable diagnostic accuracy to that of experienced endoscopists with the validation data set of LCI.

Conclusions The results of this study suggest the feasibility of an innovative gastric cancer screening program to determine cancer risk in individual subjects based on LCI-CAD.

Keywords Deep convolutional neural network · *Helicobacter pylori* infection · Image enhanced endoscopy · Risk of gastric cancer

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10120-020-01077-1>) contains supplementary material, which is available to authorized users.

✉ Hiroataka Nakashima
nakashima@soiken.or.jp

¹ Foundation for Detection of Early Gastric Carcinoma, 2-6-12 Nihombashikayabacho Chuo Ward, Tokyo 103-0025, Japan

² Medical Simulation Center, Jichi Medical University, Tochigi, Japan

³ Department of Pathology, Cancer Institute Hospital Japanese Foundation for Cancer Research, Tokyo, Japan

Introduction

Patients with *Helicobacter pylori* (*H. Pylori*)-associated gastritis have an increased risk of gastric carcinogenesis. Moreover, *H. pylori* eradication therapy is a highly effective approach to reduce incidence related to gastric cancer [1–4]. In Japan, 1.5 million people receive this treatment annually [5–7]. As a consequence, the following three categories of the *H. pylori* infection status coexist among the Japanese: persons having no history of *H. pylori* infection (hereinafter referred to as “uninfected”); those with ongoing gastritis associated with *H. pylori* infection (hereinafter referred to as “currently infected”), and those with a history of successful *H. pylori* eradication therapy (hereinafter referred to as “post-eradication”) [8, 9].

Previous studies revealed that patients with a currently infected status have an increased risk of gastric carcinogenesis, whereas those with an uninfected status have a low risk [1–7]. Furthermore, recent studies have found that, even in patients with a post-eradication status, an increased, i.e., moderate, risk to develop gastric cancer persists in patients with advanced mucosal atrophy and/or intestinal metaplasia [8–12]. Therefore, because the infection status is indicative of the risk of developing gastric cancer, the effectiveness of cancer screening programs can be improved by a system to classify individuals according to their infection status. The latest knowledge is consistent with reports from western countries [3, 4, 13–15].

Recently, our study group has developed computer-aided diagnosis (CAD) systems to distinguish *H. pylori*-infected and non-infected subjects using endoscopic images and deep learning (DL) [16–18]. DL is a machine-learning approach which mimics the functioning of neural networks in the brain and is capable of extracting features from images for automatic classification. The present study aimed to progress toward the development of a CAD system to classify the *H. pylori* infection status of cases into three categories using DL combined with linked color imaging (LCI), a new technology of image-enhanced endoscopy (IEE) [4, 19–21]. Another aim was to validate the diagnostic ability of the CAD system using it to classify images from endoscopic movies.

Methods

Subjects and collection of endoscopic images

This study was a prospective, single-center clinical study. The cohort consisted of health checking examinees, which did not include any patients with severe liver, renal, or cardiopulmonary dysfunction or any with a history of gastric cancer. Subjects who received a blood test for *H. pylori* (serum *H. pylori* antibodies: E-plate Eiken *Helicobacter pylori* antibody II; Eiken Chemical, Japan) or those who received a urea breath test (UBT, i.e., a ^{13}C -urea breath test using a UBT tablet; Otsuka Pharmaceutical Co., Ltd., Japan), as well as a screening esophagogastroduodenoscopy (EGD) following *H. pylori* eradication at the Foundation for Detection of Early Gastric Carcinoma between November 2015 and March 2019, were selected. The definition of three *H. pylori* infection status was as follows: subjects who had never experienced *H. pylori* eradication before and those with *H. pylori* antibodies at levels lower than 3.0 U/ml were classified as “uninfected,” whereas those with *H. pylori* antibody concentrations equal to or higher than 10.0 U/ml were classified as “currently infected.” Since the diagnosis of *H. pylori* infection based solely on blood data may be

inaccurate, subjects with *H. pylori* antibodies equal to or higher than 3.0 U/ml and less than 10.0 U/ml were excluded from this study as a serological gray zone of *H. pylori* infection [22]. Finally, subjects with a UBT less than 2.5‰ following *H. pylori* eradication were defined as having a “post-eradication” status. The large majority of cases received *H. pylori* eradication therapy in our clinic. Figure 1 presents a flowchart of 515 subjects registered in this study. From the start of this study, 120 consecutive subjects were registered for the recording of endoscopic movies for the validation dataset, to obtain an estimate of the diagnostic accuracy of the CAD systems. The endoscopist recorded endoscopic movies of the lesser gastric curvature using white-light imaging (WLI) and LCI simultaneously in the same area. The breakdown of subjects according to their infection status was as follows: 40 uninfected, 40 currently infected, and 40 post-eradication cases. The authors believe that continuously collecting the validation data set first might suppressed a confirmation bias, i.e., occurred from the accumulation of the learning effect of the endoscopist, on this study. Next, 395 subjects (141 of which were classified as being uninfected, 138 as being currently infected, and 116 as having the post-eradication status) were selected for the capture of endoscopic still images by WLI and LCI. The images were used as the training dataset to build deep-learning models. The validation and training datasets were obtained from entirely different subsets of subjects.

The endoscopic devices used were the EG-L580NW and EG-6400 N (Fujifilm Co., Japan). The EG-L580NW features

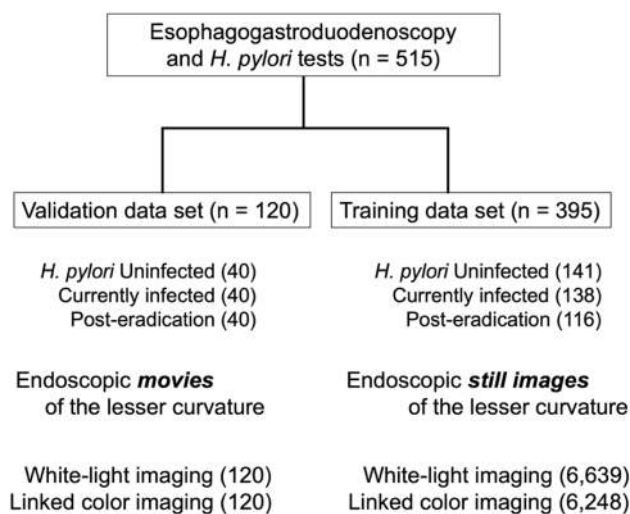


Fig. 1 Flowchart of subject enrollment and image generation. The subjects were examinees who underwent esophagogastroduodenoscopy and who were tested for serum *Helicobacter pylori* antibodies or by a ^{13}C -urea breath test. All 515 subjects were allocated to a validation group ($n=120$), in which the diagnostic accuracy of the computer-aided diagnosis was evaluated, and a training group ($n=395$), to which computer-aided diagnosis system was constructed

a laser light source, whereas in the EG-6400 N, a light-emitting diode is used. Both devices have LCI functionality. The endoscopic images of the lesser gastric curvature were serially captured from all registered subjects in retroflexed view and both in WLI and LCI modes. *H. pylori* infection status was diagnosed based on the endoscopic images of the lesser gastric curvature, i.e., the area where atrophy and intestinal metaplasia are commonly observed [23–25]. Importantly, no real-time magnification was used during the endoscopic recording. All EGD procedures were performed by one physician (H N), a Board-Certified Trainer of the Japan Gastroenterological Endoscopy Society.

This protocol for the present study was designed in accordance with the Helsinki Declaration update 2013. For the use of endoscopic images and clinical information, written consent was obtained from all 512 subjects, and the study protocol was approved by the institutional ethics committee (Approval Number: 15–02 in 2015).

Two deep convolutional neural networks and computer specifications

The training dataset obtained from 395 subjects contained 6639 WLI and 6248 LCI images, which were linked to the clinically determined *H. pylori* infection status. Two deep convolutional neural networks (DCNNs), each equipped with a 22-layer skip-connection architecture, were constructed, and each was used to develop a model adapted to either WLI or LCI imaging. Finally, two different CAD systems were developed using the DCNNs with the same architecture. They were adapted to either CAD for WLI (WLI-CAD) or CAD for LCI (LCI-CAD) imaging.

Computer specifications used in this study were as follows: Ubuntu 18.04 LTS (Canonical Ltd., UK) as the operating system, TensorFlow (Google Co. USA) as the framework, and Nvidia GeForce GTX 1080 (Nvidia Co., USA)

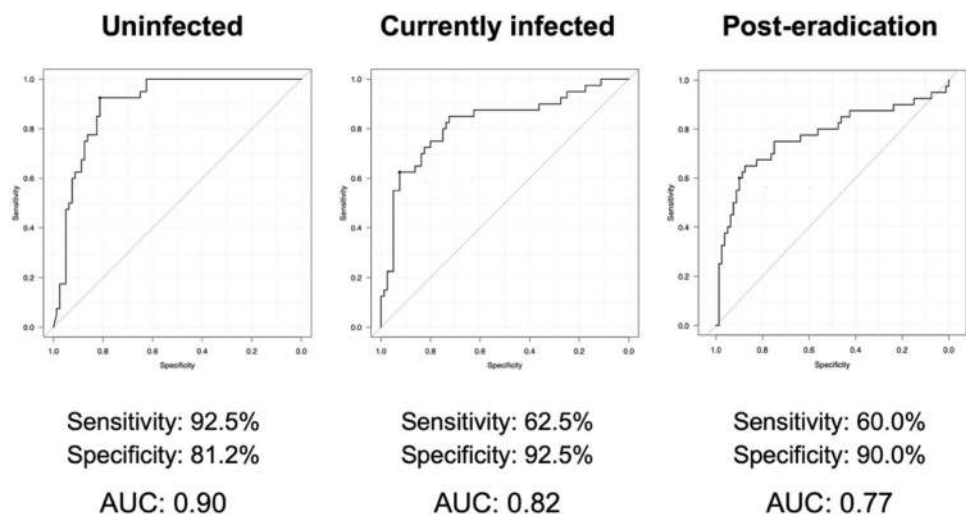
as the graphics processing unit. The implementation of the DCNNs was achieved by programs written in Python 3.6.

Assessment of the diagnostic accuracy of CAD systems using clinical endoscopic movies

The ultimate goal of the two CAD systems was to diagnose *H. pylori* infection status based solely on endoscopic movies. Diagnostic procedures using these CAD systems were as follows. First, endoscopic movies of 120 cases in which the lesser gastric curvature was recorded using LCI were included in the validation dataset. The recordings were entered into the LCI-CAD system, one case at a time. In response, the LCI-CAD system returned outputs as continuous variables ranging from 0 to 1 corresponding to the predicted *H. pylori* infection status, i.e., uninfected, currently infected, and post-eradication. Consequently, a receiver operating characteristic curve was generated for each *H. pylori* infection status based on the predictions provided by the LCI-CAD system (Fig. 2).

Next, CAD-generated diagnoses were made for all validation cases, corresponding to the *H. pylori* infection status with the highest CAD-returned value. Based on the obtained diagnoses, the sensitivity, specificity, positive predictive value (PPV), and diagnostic accuracy were calculated separately for the WLI-CAD and LCI-CAD systems (Table 2). Sensitivity is a true positive value, i.e., (number of correctly diagnosed with a certain *H. pylori* infection status by CAD)/(actual number of the corresponding *H. pylori* infection category: $n = 40$). Specificity is true negative values, i.e., (number of correctly diagnosed *H. pylori* infection statuses by CAD with excepting a certain category)/(actual number of the corresponding two infection categories: $n = 80$). PPV is (actual number of the corresponding *H. pylori* infection status)/(number of a certain *H. pylori* infection status diagnosed by CAD system). Accuracy is (number of correct

Fig. 2 Diagnostic ability of LCI-CAD with receiver operating characteristic curve. The area under receiver operating characteristic curve for the uninfected 0.90 (95% CI 0.842–0.953), currently infected 0.82 (95% CI 0.74–0.91), and post-eradication 0.77 (95% CI 0.67–0.87) categories



diagnoses for *H. pylori* infection status by CAD system, i.e., true positive plus true negative)/(number of all validation data: $n = 120$).

Diagnostic accuracy of experienced endoscopists using the validation data set of LCI

The same validation data set of LCI ($n = 120$) was diagnosed for *H. pylori* infection status by four experienced endoscopists. Three of the four endoscopists were Board-Certified Fellows of the Japan Gastroenterological Endoscopy Society. They were referred to as the expert. The fourth is an advanced trainee who has five years of gastrointestinal endoscopic experience. These were not the endoscopist responsible for collecting the endoscopic images for this study. They viewed the LCI of the lesser gastric curvature and diagnosed *H. pylori* infection status, one case at a time. Accuracy was calculated from (number of correct diagnoses for *H. pylori* infection status by endoscopists, i.e., true positive plus true negative)/(number of all validation data: $n = 120$) in this evaluation.

Sample size calculation and statistical analysis

Previously, we have reported areas under the curves (AUCs) of 0.6–0.9 for endoscopic diagnoses of *H. pylori* infection status using DL [17, 18]. Based on this information, we assumed the AUC to be 0.8, whereas the detection power was set at 0.90. The threshold level for the significance of one-tailed statistical tests was set as 0.05. Deriving from the three different *H. pylori* infection status and setting the proportion of positive infection (currently infected) as 1 and that of negative infection (i.e., uninfected and post-eradication) as 2, the required sample size per category was calculated to be 10.3, implying a required total sample size of 31 cases, or more. We conservatively selected a higher sample size of 40 cases per *H. pylori* infection status category, i.e., leading to a total of 120 cases in the validation dataset. The software environment and language, R (version 3.3.2), was used for all statistical analyses, including sample size calculation.

Results

Table 1 presents the clinical characteristics of subjects included in the validation and training datasets, respectively. The degree of gastric atrophy was determined according to the Kimura–Takemoto Classification [26]. Specifically, the proportion of subjects who presented with the entire lesser gastric curvature being covered with atrophic mucosa, the so-called open type,” was referred to as the “ratio of marked gastric atrophy.”

Table 1 Clinical features of the subjects

Characteristics	Validation data set	Training data set
Number of cases	120	395
Age	57.2 (12.9)	59.0 (14.0)
<i>H. pylori</i> infection status		
Uninfected	40	141
Currently infected	40	138
Post-eradication	40	116
Ratio of marked gastric atrophy		
Uninfected (%)	1/40 (2.5)	10/141 (7.1)
Currently infected (%)	28/40 (70.0)	97/138 (70.3)
Post-eradication (%)	25/40 (62.5)	86/116 (74.1)

The degree of gastric atrophy was determined according to the Kimura–Takemoto Classification. The proportion of subjects who presented with the entire lesser gastric curvature being covered with atrophic mucosa, the so-called open type,” was referred to as the “ratio of marked gastric atrophy”. Data of age is presented as mean (standard deviation)

Figure 2 presents the diagnostic ability of the LCI-CAD system when applied to the endoscopic validation dataset. After setting a value on the receiver operating characteristic curve as a threshold to maximize the sum of sensitivity and specificity, the following AUCs emerged: 0.90 [95% confidence interval (CI) 0.842–0.953] for uninfected; 0.82 [95% CI 0.74–0.91] for currently infected; and 0.77 [95% CI 0.67–0.87] for post-eradication cases.

Figure 3 presents images from endoscopic examination movies, obtained by two CAD systems from a case classified with the post-eradication status. The two-screen images were from the same subject. The computer-aided diagnosis system for white-light imaging (WLI-CAD, upper side) returned a prediction value of 0.492 for a post-eradication status, which turned out to be an incorrect prediction. Whereas, linked color imaging (LCI-CAD, lower side) returned a prediction value of 0.985 for a post-eradication status, which turned out to be a correct prediction. In addition, we prepared heat maps to visualize the areas that the CAD systems considered important to diagnose. The lower heat map demonstrated that hot spots were drawn against the contrast between a pale reddish tone and a whitish tone on the gastric mucosa in the captured LCI image.

Video 1 (Electronic Supplementary Material) shows EGD scenes based on LCI-CAD. The first subject was uninfected, the next was currently infected, and the last one had a post-eradication status. The numbers on the lower right of the screens are the predicted values of *H. pylori* infection status indicated by LCI-CAD.

Table 2 presents the results of sensitivity, specificity, PPV, and accuracy of the two different CAD systems. Although there was no statistically significant difference between the diagnostic results of the two CAD systems,

Fig. 3 Diagnosis with CAD systems for the endoscopic video of the post-eradication subject. The computer-aided diagnosis system for white-light imaging (WLI-CAD, upper side) returned a prediction value of 0.492 for a post-eradication status, which turned out to be an incorrect prediction. Whereas, linked color imaging (LCI-CAD, lower side) returned a prediction value of 0.985 for a post-eradication status, which turned out to be a correct prediction. The lower heat map demonstrated that hot spots were drawn against the contrast between a pale reddish tone and a whitish tone on the gastric mucosa in the captured LCI image

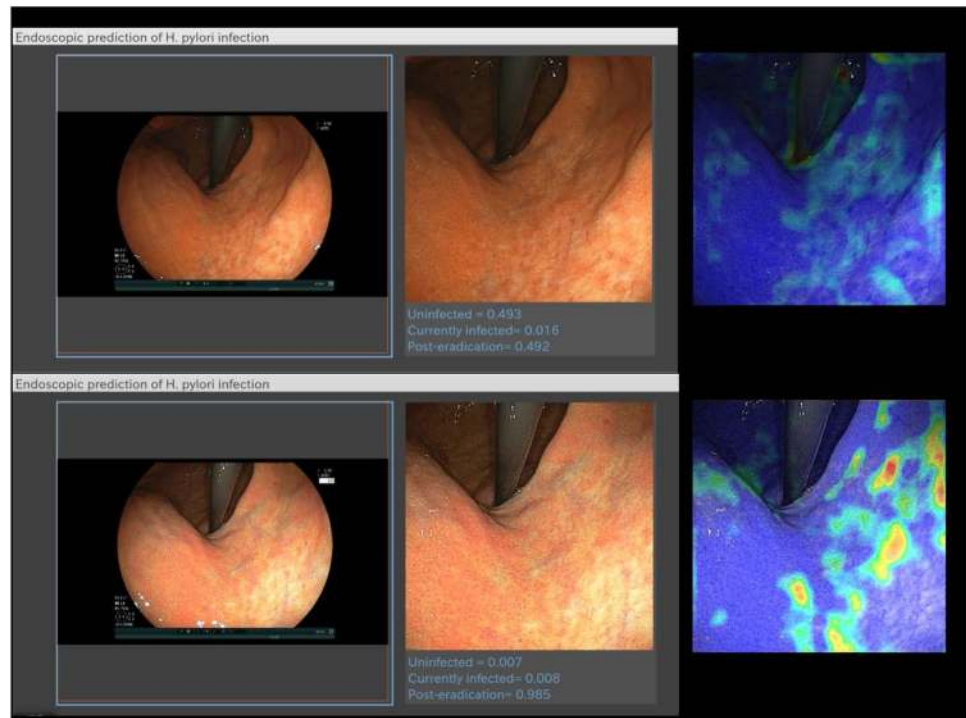


Table 2 Diagnostic ability of WLI-CAD and LCI-CAD with sensitivity, specificity, PPV, and overall accuracy

	Uninfected		Currently infected		Post-eradication	
	WLI-CAD	LCI-CAD	WLI-CAD	LCI-CAD	WLI-CAD	LCI-CAD
Sensitivity, % (95% CI)	95.0 (83.1–99.4)	92.5 (79.6–98.4)	60.0 (44.3–75.1)	62.5 (45.8–77.3)	35.0 (20.6–51.7)	65.0 (48.3–79.4)
Specificity, % (95% CI)	65.0 (53.5–75.3)	80.0 (69.6–88.1)	86.2 (76.7–92.9)	92.5 (84.4–97.2)	93.8 (86.0–97.9)	86.2 (76.7–92.9)
PPV, % (95% CI)	57.0 (44.8–69.7)	69.8 (55.7–81.7)	68.8 (44.3–75.1)	80.6 (62.5–92.5)	73.7 (48.8–90.9)	70.3 (53.0–84.1)
Accuracy, % (95% CI)	75.0 (66.3–82.5)	84.2 (76.4–90.2)	77.5 (69.0–84.6)	82.5 (74.5–88.8)	74.2 (65.4–81.7)	79.2 (70.8–86.0)
<i>P</i> value	<i>P</i> =0.126	<i>P</i> =0.126	<i>P</i> =0.559	<i>P</i> =0.559	<i>P</i> =0.556	<i>P</i> =0.556

Comparing the accuracies of the two developed computer-aided diagnosis systems, the computer-aided diagnosis system for linked color imaging was superior to the computer-aided diagnosis system for white-light imaging by 9.2% for uninfected, 5.0% for currently infected, and 5.0% for post-eradication cases. *P* values were calculated from ROCs of WLI-CAD and LCI-CAD

PPV positive predictive value

the LCI-CAD system exhibited a higher diagnostic ability of the *H. pylori* infection status than the WLI-CAD system. The accuracy of the LCI-CAD system was 84.2% for the uninfected, 82.5% for the currently infected, and 79.2% for the post-eradication status.

Table 3 shows the diagnostic accuracy of experienced endoscopists. The average accuracy of each *H. pylori* infection status by four endoscopists, using the data set of LCI was calculated at 91.2% for uninfected, 79.4% for currently infected, and 78.1% for post-eradication cases.

Discussion

Machine-learning-based CAD system is an advanced, hot topic in the field of gastroenterological endoscopy. In this field, input information includes WLIs obtained during endoscopy, magnified images obtained through the combined use of pigment staining and IEE, and endocytoscopic images enabling in vivo observation of living cells [27]. Regarding endoscopic CAD systems to classify the

Table 3 Diagnostic accuracy of experienced endoscopists for *H. pylori* infection status using the validation data set of LCI

	Uninfected	Currently infected	Post-eradication
Expert A, % (95% CI)	87.5 (80.2–92.8)	77.5 (69.0–84.6)	73.3 (64.5–81.0)
Expert B, % (95% CI)	92.5 (86.2–96.5)	83.3 (75.4–89.5)	84.2 (76.4–90.2)
Expert C, % (95% CI)	95.0 (89.4–98.1)	81.7 (73.6–88.1)	81.7 (73.6–88.1)
Advanced trainee D, % (95% CI)	90.0 (83.2–94.7)	75.0 (66.3–82.5)	73.3 (64.5–81.0)
Mean, % (95% CI)	91.2 (88.4–93.6)	79.4 (75.5–82.9)	78.1 (74.2–81.7)

The average diagnostic accuracy of each *H. pylori* infection status by four experienced endoscopists was calculated 91.2% for uninfected, 79.4% for currently infected, and 78.1% for post-eradication cases

H. pylori infection status, previous reports focused on the use of DL procedures on WLI images without real-time magnification [17, 28]. These reports, which attempted to classify *H. pylori* infection status into two categories, i.e., uninfected and currently infected, exhibited excellent performances of the CAD systems, with sensitivities and specificities exceeding 85%. However, when subjects with a clinical history of eradication therapy against *H. pylori* were added as the post-eradication category, the accuracy of CAD systems remains potential for improvement [29]. Meanwhile, diagnostic difficulties of using a three-category classification scheme were also evident in some studies on the diagnostic accuracy of *H. pylori* infection by experienced endoscopists, because endoscopically obtained images of mucosal surfaces are often altered in complex ways by *H. pylori* eradication treatments [30]. An article stated that a long-term, i.e., 2-year, training was required to attain adequate diagnostic ability of newcomers in endoscopy [8]. Indeed, classifying *H. pylori* infection status into three categories based on endoscopic images is a challenging task, both for CAD systems and endoscopists.

Accordingly, two novel approaches were adopted to improve the diagnostic ability of our CAD systems that were developed in the current study. The first approach was to use images generated by LCI as input. LCI, a new type of IEE technology, can enhance slight differences in mucosal color, thereby exhibiting excellent performance in endoscopic diagnosis of gastrointestinal inflammation [19–21]. In this study, when comparing the accuracies of the two developed CAD systems, LCI-CAD was superior to WLI-CAD by 9.2% for uninfected, 5.0% for currently infected, and 5.0% for post-eradication cases (Table 2). The second approach was to restrict endoscopic images, which were used as input for the CAD systems, in terms of their origin. Particularly, we focused on the lesser gastric curvature, which can display clear characteristics of inflammation, atrophy, and intestinal metaplasia. Moreover, the endoscopic images of the lesser curvature exhibit less diversity among subjects in terms of the appearance of the stomach mucosa, i.e., the distribution of mucosal folds and the amount of mucus secretion. Therefore, we speculate that the CAD systems may respond

better to complex changes in the mucosal color tone of the lesser gastric curvature than to endoscopic images of other parts [25, 26].

Since those two novel approaches improved the diagnostic ability of the LCI-CAD system, the authors have inferred the procedure of our CAD systems for which diagnose *H. pylori* infection status by endoscopic images, as described below. This study used simultaneously captured WLI and LCI images in the same view of the lesser curvature for the training data set, so the CAD systems likely to be affected better by mucosal color tone than morphological features. The endoscopic findings of *H. pylori* gastritis in the lesser curvature comprise two visual components, namely, a reddish tone due to inflammation and a whitish tone when atrophy and/or intestinal metaplasia is present. Hence, the color tone of the mucosa uninfected by *H. pylori* as a benchmark, our CAD systems diagnose subjects whose mucosa appears redder due to inflammation as being currently infected and those with a more whitish mucosa as having a post-eradication status with atrophy and/or intestinal metaplasia in the lesser curvature. The heat maps in Fig. 3 were generated to visualize the endoscopic findings that the CAD systems strongly responded. The computer program Grad-CAM was used for this demonstration [31]. In the heatmap of the LCI-CAD system, hot spots were drawn against the contrast between a pale reddish tone and a whitish tone on the gastric mucosa in the lesser curvature of the post-eradicated case. These heat maps are a piece of evidence for our hypothesis.

Another result of this study revealed accuracies of the LCI-CAD system when used on endoscopic movies to be 84.2% for uninfected, 82.5% for currently infected, and 79.2% for post-eradication cases in Table 2. Generally, if a three-category classification scheme is set, the basic value of the accuracy rate is 33.3%; therefore, we believe that the LCI-CAD system has improved the diagnostic accuracies of the *H. pylori* infection status from the basic value by 50.9% for uninfected, 49.2% for currently infected, and 45.9% for post-eradication cases.

In this study, a comparison of diagnostic accuracy was also examined between endoscopists and the CAD system. The LCI-CAD system was superior to human endoscopists using WLI images in a previous report, in which the

obtained accuracy was 88.9% for uninfected cases, 62.1% for currently infected cases, and 55.8% for cases with the post-eradication status [8]. Meanwhile, the LCI-CAD system demonstrated comparable diagnostic accuracy to that of experienced endoscopists with the same validation data set of LCI, in which the obtained mean accuracy was 91.2% for uninfected cases, 79.4% for currently infected cases, and 78.1% for cases with the post-eradication status (Table 3).

Several limitations of this study should be noted. The first one relates to the single-center character of our study design. When attempting to implement CAD systems in clinical practice, their accuracy should be determined using validation datasets containing endoscopic images collected from multiple facilities. A second limitation concerns the used gold standard for *H. pylori* infection diagnosis. This study used blood testing with *H. pylori* antibodies as the sole means to distinguish uninfected and currently infected status. Since the sensitivity and specificity of this test are not 100%, we addressed this issue by excluding subjects who exhibited *H. pylori* antibody titers equal to or higher than 3.0 U/ml and less than 10.0 U/ml from the study as a serological gray zone of *H. pylori* infection. However, excluding these subjects does not guarantee that all serological false-negative and false-positive diagnoses of *H. pylori* infection were completely excluded from the datasets.

The final limitation of this study was that the CAD systems are still fundamentally restricted to serving as an aid for diagnosis. In other words, we believe that the diagnoses made by the CAD system should be considered as mere predictions of *H. pylori* infection status based on mucosal color and that the final diagnosis should be made by human visual inspection and judgment, i.e., by experienced endoscopists.

Nevertheless, the observed strong diagnostic ability of LCI-CAD to classify the status of *H. pylori* infection of cases into three categories may have considerable significance for gastric cancer screening. The application of LCI-CAD procedures to endoscopic inspections may facilitate risk stratification of examinees and improved detection of gastric cancer. More specifically, it seems plausible that a high-quality gastric cancer screening program is achievable by consolidating the resources of endoscopic examinations to subjects in high-risk categories (i.e., currently infected and post-eradication) after excluding subjects categorized as uninfected who are at an extremely low risk of gastric cancer. Furthermore, the authors speculate that using the LCI-CAD system to classify the infection status into three-categories may not only be indicative of the risk of developing gastric cancer but may also prevent accidental re-treatment of patients whom it is already eradicated.

Overall, we strongly advocate the future implementation of LCI-CAD system similar to the one tested in our clinical practice.

Author contributions HN: conception and design, analysis and interpretation of the data and drafting of the article. HK: interpretation of the data and critical revision of the article for important intellectual content. HK: critical revision of the article for important intellectual content. NS: final approval of the article.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Ethical approval All procedures involving human participants were performed in accordance with the ethical standards of the institutional and national research committee and with the Helsinki Declaration of 1964 and later versions. The study was approved by the institutional ethics committee (Approval Number: 15–02 in 2015).

Informed consent The authors obtained written agreement to participation in this study from all 512 subjects.

References

1. Chey WD, Leontiadis GI, Howden CW, et al. ACG clinical guideline: treatment of *Helicobacter pylori* Infection. *Am J Gastroenterol*. 2017;112:212–39. <https://doi.org/10.1038/ajg.2016.563>.
2. IARC *Helicobacter pylori* Working Group. *Helicobacter pylori* eradication as a strategy for preventing gastric cancer. IARC Working Group Reports, No. 8. Lyon, France: International agency for research on cancer; 2014. Available from: https://www.iarc.fr/en/publications/pdfs-online/wrk/wrk8/Helicobacter_pylori_Eradication.pdf. Accessed 18 Apr 2020.
3. Pimentel-Nunes P, Libânio D, Marcos-Pinto R, et al. Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European Society of Gastrointestinal Endoscopy (ESGE), European *Helicobacter* and Microbiota Study Group (EHMSG), European Society of Pathology (ESP), and Sociedade Portuguesa de Endoscopia Digestiva (SPED) guideline update 2019. *Endoscopy*. 2019;51:365–88. <https://doi.org/10.1055/a-0859-1883>.
4. Banks M, Graham D, Jansen M, et al. British Society of Gastroenterology guidelines on the diagnosis and management of patients at risk of gastric adenocarcinoma. *Gut*. 2019;68:1545–75. <https://doi.org/10.1136/gutjnl-2018-318126>.
5. Sugano K. Effect of *Helicobacter pylori* eradication on the incidence of gastric cancer: a systematic review and meta-analysis. *Gastric Cancer*. 2019;22:435–45.
6. Tsuda M, Asaka M, Kato M, et al. Effect on *Helicobacter pylori* eradication therapy against gastric cancer in Japan. *Helicobacter*. 2017;22:e12415. <https://doi.org/10.1111/hel.12415>.
7. Kato M, Ota H, Okuda M, et al. Guidelines for the management of *Helicobacter pylori* infection in Japan: 2016 Revised Edition. *Helicobacter*. 2019;24:e12597. <https://doi.org/10.1111/hel.12597> **Epub 2019 May 20**.
8. Watanabe K, Nagata N, Shimbo T, et al. Accuracy of endoscopic diagnosis of *Helicobacter pylori* infection according to level of endoscopic experience and the effect of training. *BMC Gastroenterol*. 2013;13:128.
9. Tsukamoto T, Nakagawa M, Kiriya Y, et al. prevention of gastric cancer: eradication of *Helicobacter pylori* and beyond. *Int J Mol Sci*. 2017;18:1699. <https://doi.org/10.3390/ijms18081699>.
10. Kobayashi M, Sato Y, Terai S. Endoscopic surveillance of gastric cancers after *Helicobacter pylori* eradication. *World J Gastroenterol*. 2015;21:10553–62.

11. Mori G, Nakajima T, Asada K, et al. Incidence of and risk factors for metachronous gastric cancer after endoscopic resection and successful *Helicobacter pylori* eradication results of a large-scale, multicenter cohort study in Japan. *Gastric Cancer*. 2016;19:911–8. <https://doi.org/10.1007/s10120-015-0544-6>.
12. Sugano K, Tack J, Kuipers E, et al. Kyoto global consensus report on *Helicobacter pylori* gastritis. *Gut*. 2015;64:1353–67. <https://doi.org/10.1136/gutjnl-2015-309252>.
13. Rokkas T, Rokka A, Portincasa P. A systematic review and meta-analysis of the role of *Helicobacter pylori* eradication in preventing gastric cancer. *Ann Gastroenterol*. 2017;30:414–23.
14. Rugge M, Meggio A, Pennelli G, et al. Gastritis staging in clinical practice: the OLGA staging system. *Gut*. 2007;56:631–6. <https://doi.org/10.1136/gut.2006.106666>.
15. Esposito G, Pimentel-Nunes P, Angeletti S, et al. Endoscopic grading of gastric intestinal metaplasia (EGGIM): a multicenter validation study. *Endoscopy*. 2019;51:515–21. <https://doi.org/10.1055/a-0808-3186>.
16. Nakashima H, Itoh T, Kawahira H, et al. Computer-aided endoscopic diagnosis of *Helicobacter pylori* infection. *J Gastroenterol Hepatol*. 2016;31:274.
17. Itoh T, Kawahira H, Nakashima H, et al. Deep learning analyzes *Helicobacter pylori* infection by upper gastrointestinal endoscopy images. *Endosc Int Open*. 2018;06:E139–E144. <https://doi.org/10.1055/s-0043-120830>.
18. Nakashima H, Kawahira H, Kawachi H et al. Artificial intelligence diagnosis of *Helicobacter pylori* infection using blue laser imaging-bright and linked color imaging: a single-center prospective study. *Ann Gastroenterol*. 2018;31: 462–468. <https://doi.org/10.20524/aog.2018.0269>
19. Dohi O, Yagi N, Onozawa Y, et al. Linked color imaging improves endoscopic diagnosis of active *Helicobacter pylori* infection. *Endosc Int Open*. 2016;04:E800–E805. <https://doi.org/10.1055/s-0042-109049>.
20. Fukuda H, Miura Y, Osawa H, et al. Linked color imaging can enhance recognition of early gastric cancer by high color contrast to surrounding gastric intestinal metaplasia. *J Gastroenterol*. 2019;54:396–406. <https://doi.org/10.1007/s00535-018-1515-6>.
21. Sun X, Bi Y, Dong T, et al. Linked colour imaging benefits the endoscopic diagnosis of distal gastric diseases. *Sci Rep*. 2017;17:5638. <https://doi.org/10.1038/s41598-017-05847-3>.
22. Kotachi T, Ito M, Yoshihara M, et al. Serological evaluation of gastric cancer risk based on pepsinogen and *Helicobacter pylori* antibody: relationship to endoscopic findings. *Digestion*. 2017;95:314–8.
23. Saka A, Yagi K, Nimura S. OLGA- and OLGIM-based staging of gastritis using narrow-band imaging magnifying endoscopy. *Dig Endosc*. 2015;27:735–42. <https://doi.org/10.1111/den.12483>.
24. Nakashima H, Kawahira H, Kawachi H, et al. Artificial intelligence diagnosis of *Helicobacter pylori* infection using linked color imaging. *UEG J*. 2017;8:A215.
25. Yasuda T, Hiroyasu T, Hiwa S, et al. Potential of automatic diagnosis system with linked color imaging for diagnosis of *Helicobacter pylori* infection. *Dig Endosc*. 2019. <https://doi.org/10.1111/den.13509>.
26. Kimura K, Takemoto T. An endoscopic recognition of the atrophic border and its significance in chronic gastritis. *Endoscopy*. 1969;1:87–97.
27. Mori Y, Kudo S, Hussein E, et al. Artificial intelligence and upper gastrointestinal endoscopy: current status and future perspective. *Dig Endosc*. 2019;31:378–88. <https://doi.org/10.1111/den.13317>.
28. Shichijo S, Nomura S, Aoyama K, et al. Application of convolutional neural networks in the diagnosis of *Helicobacter pylori* infection based on endoscopic images. *EBioMedicine*. 2017;25:106–11.
29. Shichijo S, Endo Y, Aoyama K, et al. Application of convolutional neural networks for evaluating *Helicobacter pylori* infection status on the basis of endoscopic images. *Scand J Gastroenterol*. 2019;54:158–63. <https://doi.org/10.1080/00365521.2019.1577486>.
30. Yoshii S, Mabe K, Watano K, et al. Validity of endoscopic features for the diagnosis of *Helicobacter pylori* infection status based on the Kyoto classification of gastritis. *Dig Endosc*. 2019. <https://doi.org/10.1111/den.13486>.
31. Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vis*. 2020;128:336–59. <https://doi.org/10.1007/s11263-019-01228-7>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.