JOURNAL OF THE
ANUS, RECTUM AND COLON
dx.doi.org/10.23922/jarc.2023-070
http://journal-arc.jp

Original Research Article

Non-polypoid Colorectal Lesions Detection and False Positive Detection by Artificial Intelligence under Blue Laser Imaging and Linked Color Imaging

Satoshi Sugino[1)], Naohisa Yoshida[2)], Zhe Guo[3)], Ruiyao Zhang[3)], Ken Inoue[2)], Ryohei Hirose[2)], Osamu Dohi[2)], Yoshito Itoh[2)], Daiki Nemoto[4)], Kazutomo Togashi[4)], Hironori Yamamoto[5)] and Xin Zhu[3)]

1) Department of Gastroenterology, Asahi University Hospital, Gifu, Japan
2) Department of Gastroenterology, Kyoto Prefectural University of Medicine, Kyoto, Japan
3) Biomedical Information Engineering Lab, The University of Aizu, Fukushima, Japan
4) Department of Coloproctology, Aizu Medical Center, Fukushima Medical University, Fukushima, Japan
5) Department of Gastroenterology, Jichi Medical University, Tochigi, Japan

Abstract

Objectives: Artificial intelligence (AI) with white light imaging (WLI) is not enough for detecting non-polypoid colorectal polyps and it still has high false positive rate (FPR). We developed AIs using blue laser imaging (BLI) and linked color imaging (LCI) to detect them with specific learning sets (LS).

Methods: The contents of LS were as follows, LS (WLI): 1991 WLI images of lesion of 2-10 mm, LS (IEE): 5920 WLI, BLI, and LCI images of non-polypoid and small lesions of 2-20 mm. LS (IEE) was extracted from videos and included both in-focus and out-of-focus images. We designed three AIs as follows: AI (WLI) finetuned by LS (WLI), AI (IEE) finetuned by LS (WLI)+LS (IEE), and AI (HQ) finetuned by LS (WLI)+LS (IEE) only with images in focus. Polyp detection using a test set of WLI, BLI, and LCI videos of 100 non-polypoid or non-reddish lesions of 2-20 mm and FPR using movies of 15 total colonoscopy were analyzed, compared to 2 experts and 2 trainees.

Results: The sensitivity for LCI in AI (IEE) (83%) was compared to that for WLI in AI (IEE) (76%: p= 0.02), WLI in AI (WLI) (57%: p<0.01), BLI in AI (IEE) (78%: p=0.14), and LCI in trainees (74%: p <0.01). The sensitivity for LCI in AI (IEE) (83%) was significantly higher than that in AI (HQ) (78%: p <0.01). The FPR for LCI (6.5%) in AI (IEE) were significantly lower than that in AI (HQ) (17.3%: p <0.01).

Conclusions: AI finetuned by appropriate LS detected non-reddish and non-polypoid polyps under LCI.

Keywords

artificial intelligence, colorectal polyps, linked color imaging, blue laser imaging

J Anus Rectum Colon 2024; 8(3): 212-220

Introduction

Colonoscopy is important for reducing the mortality rate of colorectal cancer (CRC), and resection of colorectal polyps contributes to the prevention of CRC[1]. However, the rate of polyp miss in colonoscopy has been reported to be 22%[2]. The risk factors for polyp miss are small polyps, non-polypoid morphology, poor preparation, and non-reddish lesions such as sessile serrated lesions (SSL)[3,4]. Previous reports, including randomized control trials (RCTs) and systematic reviews, have discussed the prevention of polyp miss and the improvement of polyp detection under

Corresponding author: Naohisa Yoshida, naohisa@koto.kpu-m.ac.jp
Received: November 28, 2023, Accepted: March 28, 2024

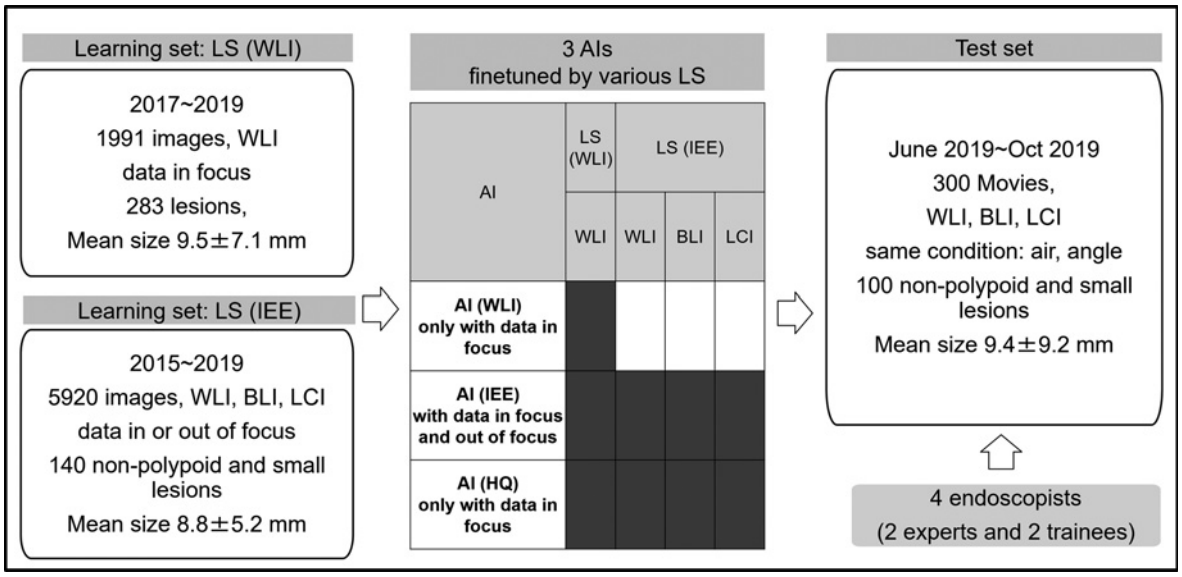

Figure 1. Study flow.

image-enhanced endoscopy (IEE), such as blue laser imaging (BLI) and linked color imaging (LCI)[5-7].

On the other hand, image recognition by an artificial intelligence (AI) has been dramatically improved by deep learning such as a convolutional neural network (CNN) and it has been reported that the detection of colorectal polyps is improved with the help of AI under white light imaging (WLI)[8-10]. AI using IEE is promising to assist endoscopists to detect more polyps than AI using WLI, and various AI for polyp detection in colonoscopy are now marketed[11]. However, non-reddish and non-polypoid lesions that are poorly visible on WLI are still not well detected by both endoscopists and marketed AIs[10,12,13]. A marketed AI, CAD EYE (Fujifilm Co., Tokyo, Japan) used both WLI and LCI for lesion detection and reported the superiority of LCI for polyp detection rate compared with WLI[14-16]. On the other hand, false positive rate (FPR) is one of the problems in AI, because it can affect prolong the observation time and make endoscopists tired[12,17,18].

Learning content is so important to improving AI, and in-focus endoscopic images of colorectal polyps are regularly used to finetune AI. However, endoscopic images in a monitor are sometimes out of focus during observation, and expert endoscopists can detect polyps even in these situations. From this point of view, we hypothesized that the use of out-of-focus images for finetuning AI might be useful to improve polyp detection and FPR of AI.

In the current study, we investigated the efficacy of AI finetuned with IEE images such as BLI or LCI for the detection of non-reddish or non-polypoid lesions. In addition, we evaluated whether out of focus images as learning content could improve the AI.

Methods

We used two learning set (LS) contents for finetuning original AIs. LS (WLI) consisted of 1991 WLI non-magnified endoscopic still WLI images for colorectal lesions of 2-10 mm, obtained at a single institution (Fukushima Medical University Aizu Medical Center) from 2017 to 2019, which was made in a previous study (Figure 1)[19]. All images of LS (WLI) were in-focus. LS (IEE) consisted of 5920 still images extracted from WLI, BLI, and LCI videos of 140 non-reddish, non-polypoid, or small lesions of 2-20 mm obtained at Kyoto Prefectural University of Medicine between June 2015 and October 2019. Extraction of these still images from the videos was performed by two technicians (Z.G. and R.Z.) with the assistance of an expert endoscopist (N.Y.). LS (IEE) included both in-focus and out-of-focus images and it included images with both lesions in the center and in the edges. Out-of-focus images were defined as images in which a polyp was not sharply depicted but was confirmed to be present by a movie of the polyp. Using these two LSs, we designed AI (WLI), AI (IEE), and AI (high quality; HQ) programs as follows: AI (WLI) finetuned by LS (WLI), AI (IEE) finetuned by LS (WLI)+LS (IEE), and AI (HQ) finetuned by only excellent images in focus of LS (WLI)+LS (IEE) (2490 images). The test set included 300 videos under the same condition of WLI, BLI, and LCI of 100 non-reddish or non-polypoid lesions of 2-20 mm from June 2019 to September 2019 at Kyoto Prefectural University of Medicine.

All lesions of LS (IEE) and test set were selected and recorded according to each lesion definition by 2 experienced endoscopists (N.Y., K.I.) in the period. The area of 5 cm on the oral and anal sides of a lesion was recorded under the

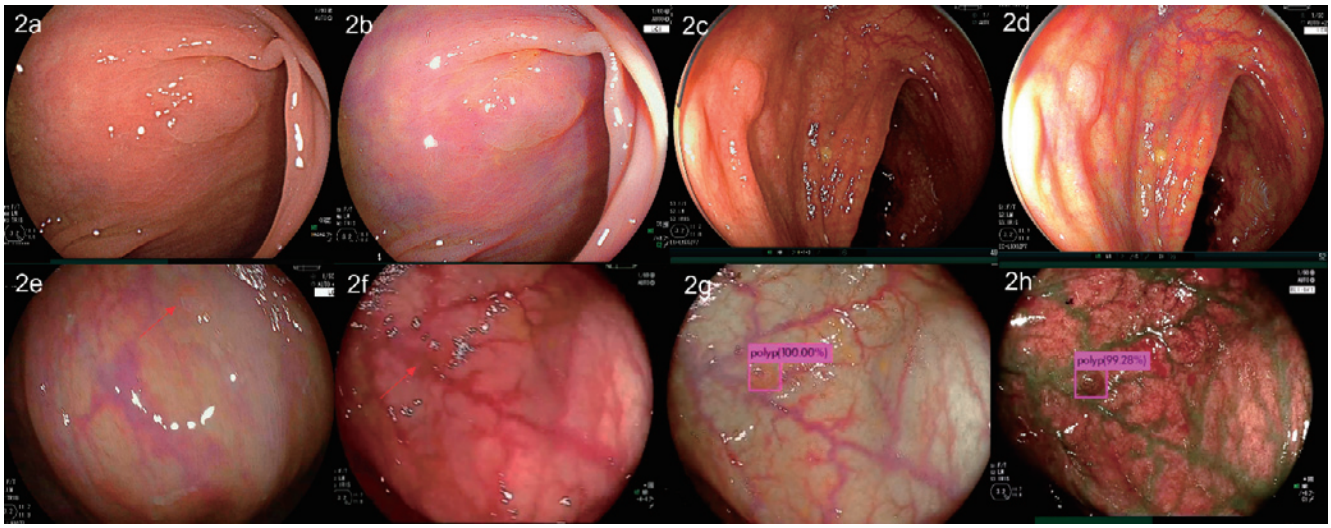

Figure 2. Learning contents for finetuning artificial intelligence (AI) and evaluations of test sets by AI under WLI, BLI, and LCI. 2a. A clear WLI image for a non-polypoid lesion of 12 mm (SSL) in the leaning set (LS) 2. 2b. A clear LCI image for the same lesion in LS (IEE). 2c. A clear WLI image for a non-reddish polypoid lesion of 8 mm (adenoma) in LS (IEE). 2d. A clear LCI image for the same lesion in LS (IEE). 2e. A LCI image which was not in-focus for a non-polypoid lesion of 4 mm (adenoma) in LS (IEE) (red arrow). 2f. A non-polypoid lesion of 2 mm (adenoma) on the descending colon (red arrow). The AI (IEE) was not able to detect a lesion under WLI. 2g. The AI (IEE) was able to detect the same lesion with an annotation box under LCI, showing a 100% confidence level in the detection of the lesion. 2h. The AI (IEE) was able to detect the lesion under BLI, showing a 99% confidence level.

same conditions (air, residual fluid, and observation speed) whenever possible over approximately 10 seconds with WLI, BLI, and LCI. Cases with inflammatory bowel disease and melanosis were excluded. Lesions recorded without the same conditions of all WLI, BLI, and LCI were excluded.

The test set was evaluated by both AIs (AI (WLI), AI (IEE), and AI (HQ)) and endoscopists (2 experts and 2 trainees). The four endoscopists were different from those who prepared the image data set. Polyp visibility was evaluated by endoscopists using a polyp visibility score according to a previous report[20]. Score 4 indicates excellent visibility, where it is easy to detect a polyp; Score 3 indicates good visibility, where if an endoscopist looked in the direction of the polyp on the monitor, it would be easy to detect the polyp; Score 2 indicates fair visibility, where it would be difficult to detect the polyp without careful observation; Score 1 indicates poor visibility. For the assessment of polyp detection by endoscopists, lesions with polyp visibility score 3 and 4 were defined as detected lesions.

Polyp detection and false positive rate of AI

The automatic polyp detection algorithm was based on the computer-aided detection (CADe) algorithm developed by the Biomedical Information Engineering Lab, Aizu University. This CADe algorithm was based on YOLOv3, a CNN architecture for real-time object detection from videos[19]. Each endoscopic image was resized from 1,280×1,024 pixels to 416×416 pixels to conform to YOLOv3, and data augmentation was performed to improve the performance without overfitting. Finally, different combinations of two LSs were used to finetune each AI. In addition, the false alarm detection for bubbles, feces, wrinkles, etc. was finetuned using the parts of movies without polyps. For polyp detection from videos, an annotation box was programmed to appear on the screen when a polyp was detected (Figure 2). If this annotation box contained polyps for at least 1 frame (1/30 sec), we defined the AI as being able to accurately detect polyps and calculated the proportion of polyp videos from which the AI was able to accurately detect polyps as the per-video sensitivity[19]. For the analysis of FPR, 15 cases of total colonoscopy movies without any polyps under observation of the whole colon in WLI, BLI, and LCI modes during the same period (5 videos each for BLI, LCI, and WLI) were selected by an expert endoscopist (N.Y.). The definition of false positive detection was decided as the AI highlighting an annotation box of at least 45 frames (1.5 secs) to an area without polyps in this study. We calculated the FPR as the rate of false positive frames in whole frames of these videos. In addition, we thought that multiple simultaneous false-positive detections on one monitor could be problematic for endoscopists. Therefore, if there were multiple false positive detections at the same time, we calculated them all separately. The AI confidence level threshold could be changed from ≥10% to ≥90% every 10%. A confidence level ≥30% was used in the comparison for each AI because a previous study on AI (WLI) showed that

a confidence level ≥30% had the best balance between sensitivity and FPR[19]. In the previous study, AI (WLI) with confidence level ≥30% had a sensitivity of 88% in a test set of 100 adenomas ≤10 mm, compared with the sensitivities of two experts (88% and 88%) and two trainees (84% and 76%).

The primary outcome of the current study was the improvement in sensitivity of LCI and BLI in AI (IEE) for non-reddish or non-polypoid lesions compared to that of WLI in AI (WLI), AI (IEE), and AI (HQ). As a secondary outcome, the sensitivity of AI (IEE) was also compared with that of trainees and experts. In addition, polyp detection and FPR of WLI, BLI, and LCI in AI (IEE) finetuned by both in-focus and out-of-focus images as LS were compared to those in AI (HQ) finetuned by only in-focus images. Different confidence level from ≥10% to ≥90% in AI were also investigated in terms of sensitivity and FPR of WLI, BLI, and LCI. In addition, risk factors for polyp miss regarding non-reddish or non-polypoid lesions were analyzed among AI, experts, and trainees regarding morphology, location, size, and histopathology. In the risk factor analysis, lesions with scores of 2-4 were analyzed as detected lesions compared to lesions with score 1.

The location of each polyp was divided into the proximal side from the cecum to the transverse colon and the distal side from the descending colon to the rectum. The macroscopic type was classified as non-polypoid and polypoid, and the size was calculated using forceps and a snare. Histopathological diagnosis was made according to the WHO classification from specimens obtained by cold snare polypectomy, endoscopic mucosal resection (EMR), and biopsy[21]. All lesions were classified as hyperplastic polyp (HP), SSL, adenoma (ADE), and T1 cancer. Bowel preparation was assessed using the Aronchick bowel scale, where a score of 3 or 4 was defined as poor preparation status[22]. Regarding endoscopist experience, expert endoscopists were defined as those with experience of more than 5000 colonoscopies including BLI and LCI observation. Trainees were defined as those with experience of 100-300 colonoscopies including BLI and LCI observations.

The endoscopic systems and light sources used in this study were the VP-7000 and LL-7000 (Fujifilm), and the colonoscopes were the EC-L600ZP and EC-L600ZP7 (Fujifilm). These were high-resolution complementary metal oxide semiconductor (CMOS) colonoscopes.

All patients' personal information was removed from the clinical data to maintain anonymity. This study was approved by our ethics review board (ERB-C-1704-3) and conducted in accordance with the World Medical Association's Declaration of Helsinki.

Statistical analyses

The results were analyzed using Pearson's Chi Square Test, Mann-Whitney U test, McNemar test, and logistic regression analysis. Statistical analyses were performed using the Stata 16 (Stata Corp., College Station, TX, USA) and SPSS (Version 22.0: IBM Japan, Ltd., Tokyo, Japan) software programs for Windows. Significant differences were set at a p value of <0.05.

Results

Regarding lesion characteristics, the LS (WLI), LS (IEE), and test set included 1991 images of 283 lesions, 5920 images of 140 lesions, and 300 movies of 100 lesions (Table 1). The mean polyp sizes (mm) were 9.5±7.1, 8.8±5.2, and 9.4±9.2. The rates of polypoid were 62.5%, 73.6%, and 62.0%, respectively.

The sensitivity for LCI in AI (IEE) (83%) was significantly higher than that for WLI (76%, p=0.02), but not significantly higher than that for BLI (78%, p=0.14) (Table 2). The sensitivity of WLI, BLI and LCI of AI (IEE) was significantly higher than that of WLI in AI (WLI) (57%) (WLI: p<0.01, BLI: p<0.01, LCI: p<0.01). They were also significantly higher than those of AI (HQ) (WLI: 67%, p<0.01, BLI: 68%, p<0.01, LCI: 78%, p<0.01).

Regarding the comparison between AI (IEE) and endoscopists, the sensitivity of WLI (76%) and LCI (83%) in AI (IEE) was significantly higher than in trainees (52%, p <0.01, 74%, p<0.01) (Table 3). However, that of LCI was significantly lower than that of experts (89%, p=0.03). Those of BLI (74%) and LCI (74%) were better than that of WLI (52%) for both trainees (BLI: p<0.01, LCI: p<0.01) and experts (BLI: p<0.01, LCI: p<0.01). In addition, the LCI was better than the BLI for the expert (p=0.04).

The FPR for BLI (2.5%) and LCI (6.5%) in AI (IEE) were significantly lower than those in AI (HQ) (17.3%, p <0.01, 18.3%, p=0.01). In AI (IEE), there was no significant difference about the FPR between LCI and WLI (6.5% vs. 4.7%, p=0.73) (Table 4). The FPR of BLI (2.5%) was significantly lower than that of WLI (p=0.01). In AI (HQ), the FPR of LCI (18.3%) was significantly higher than that of WLI (6.1%, p=0.02).

The sensitivity and FPR of AI (IEE) with different confidence levels were examined (Supplemental Table 1). At different confidence levels of ≥10-90% in AI (IEE), the overall sensitivities of LCI were significantly higher than those of WLI (p=0.04). There were no significant differences of them between WLI and BLI (p=0.53) and between LCI and BLI (p=0.19). On the other hand, the FPR of BLI in AI (IEE) was significantly lower than that of LCI (p<0.01) and WLI (p=0.01). There was no significant difference of them between LCI and WLI (p=0.73).

Regarding risk factors of polyp miss for non-reddish or non-polypoid lesions in AI (IEE), there were significant differences between proximal colon and distal colon for WLI

Table 1. The Lesion Characteristics of Learning Sets and Test Sets.

	LS (WLI)	LS (IEE)	Test set
Number of images	1991	5920	300
Source of data	Image	Movie	Movie
Extracted data	Image	Image	Movie
Modality	WLI	WLI, BLI, LCI	WLI, BLI, LCI
Lesion number	283	140	100
Location, n (%)			
Proximal	192 (68.0)	58 (41.4)	70 (70.0)
Distal	91 (32.0)	82 (58.6)	30 (30.0)
Polyp size (mm), mean±SD, range	9.5±7.1 2-10	8.8±5.2 2-20	9.4±9.2 2-20
2-9 mm, n (%)	177 (62.5)	103 (73.6)	62 (62.0)
≥10 mm, n (%)	106 (37.5)	37 (26.4)	38 (38.0)
Macroscopic type, n (%)			
Polypoid type	93 (33)	36 (25.7)	36 (36.0)
Non-polypoid type	190 (67)	104 (74.3)	64 (64.0)
Lesion color, n (%)			
Reddish	n.e.	45 (32.1)	35 (35.0)
Non-reddish	n.e.	95 (67.9)	65 (65.0)
Histopathology, n (%)			
HP	20 (7)	9 (6.4)	8 (8.0)
SSL	57 (20)	32 (22.9)	18 (18.0)
ADE	198 (70)	95 (67.9)	83 (61.0)
T1	8 (3)	4 (2.9)	1 (1.0)

LS: leaning set, Proximal: cecum to transverse colon, Distal; descending colon to rectum, SD: standard deviation, ADE: adenoma, HP: hyperplastic polyp, SSL: sessile serrated lesion

Table 2. The Sensitivity of WLI, BLI, and LCI among 3 Finetuned AIs.

	Sensitivity, %			p value		
AI model ≥30% confidence level	WLI	BLI	LCI	WL vs. BLI	WL vs. LCI	BLI vs. LCI
AI (WLI)	57	-	-	-	-	-
AI (IEE)	76	78	83	0.53	0.02	0.14
AI (HQ)	67	68	78	0.21	<0.01	<0.01
p value WLI of AI (WLI) vs. AI (IEE)	<0.01	<0.01	<0.01			
p value AI (IEE) vs. AI (HQ)	<0.01	<0.01	<0.01			

AI: artificial intelligence, WLI: white light imaging, BLI: blue laser imaging, LCI: linked color imaging

(p<0.01) and LCI (p=0.01) (Table 5). For experts and trainees, there was a significant difference between proximal colon and distal colon for WLI (experts and trainees) and BLI (trainees). Additionally, there were significant differences between HP+SSL and ADE+T1 for BLI (trainees) and LCI (trainees).

Discussion

Several RCTs have described the improvement of polyp detection with LCI better than WLI, in addition to previous papers showing the efficacy of BLI and LCI for polyp detection[5-7,23-25]. Our recent RCT showed that the adenoma detection rate of LCI was significantly better than that of WLI (58.7% vs. 46.7%, p<0.01)[6]. In addition, the SSL detection rate of LCI was also better than that of WLI (4.8% vs. 2.8%, p<0.01). These studies led us to hypothesize that AI could detect more polyps with BLI and LCI than with WLI. In the current study, we demonstrated that AI finetuned with appropriate learning contents of WLI, BLI

Table 3. The Sensitivity of WLI, BLI, and LCI between AI and Endoscopists.

	Sensitivity, %			p value		
AI model ≥30% confidence level	WLI	BLI	LCI	WL vs. BLI	WL vs. LCI	BLI vs. LCI
AI (IEE)	76	78	83	0.53	0.02	0.14
Trainee	52	74	74	<0.01	<0.01	0.50
Expert	71	82	89	<0.01	<0.01	0.04
p value AI (IEE) vs. Trainee	<0.01	0.22	<0.01			
p value AI (IEE) vs. Expert	0.15	0.22	0.03			

AI: artificial intelligence, WLI: white light imaging, BLI: blue laser imaging, LCI: linked color imaging

Table 4. The False Positive Rate of WLI, BLI, and LCI among 3 Finetuned AIs.

	FPR, %			p value		
AI model ≥30% confidence level	WLI	BLI	LCI	WLI vs. BLI	WLI vs. LCI	BLI vs. LCI
AI (WLI)	2.1	-	-	-	-	-
AI (IEE)	4.7	2.5	6.5	0.01	0.73	<0.01
AI (HQ)	6.1	17.3	18.3	0.08	0.02	0.54
p value AI (IEE) vs. AI (HQ)	0.66	<0.01	0.01			

AI: artificial intelligence, WLI: white light imaging, BLI: blue laser imaging, LCI: linked color imaging, FPR: false positive rate

Table 5. The Sensitivity of Lesion Detection for Various Characteristics for Risk Factors of Polyp Miss about Non-Reddish and Non-Polypoid Lesions for AI and Endoscopists.

	Mode	Non-polypoid, (%)	Polypoid, (%)	p	Proximal, (%)	Distal, (%)	p	2-9 mm, (%)	10-20 mm, (%)	p	HP SSL, (%)	ADE, T1, (%)	p
AI (IEE)	WLI	73.4	80.5	0.42	68.6	93.3	<0.01	74.2	78.9	0.58	80.8	74.3	0.50
	BLI	75.0	80.5	0.52	72.9	90.0	0.05	72.6	86.8	0.09	80.8	77.0	0.69
	LCI	79.7	88.9	0.23	77.1	96.7	0.01	77.4	92.1	0.05	84.6	82.4	0.79
Expert	WLI	85.9	88.8	0.67	81.4	100.0	0.01	82.3	94.7	0.07	84.6	87.8	0.67
	BLI	93.6	100.0	0.12	94.3	100.0	0.18	93.5	100.0	0.11	92.3	97.3	0.24
	LCI	98.4	100.0	0.45	98.6	100.0	0.51	98.4	100.0	0.43	96.2	100	0.08
Trainee	WLI	73.4	77.8	0.41	65.7	90.0	0.01	67.7	81.6	0.13	84.6	68.9	0.12
	BLI	75.0	88.9	0.31	77.1	100.0	<0.01	80.6	89.5	0.24	96.2	79.7	0.04
	LCI	79.7	91.7	0.52	85.7	96.7	0.10	87.1	92.1	0.43	100.0	86.5	0.04

AI: artificial intelligence, WLI: white light imaging, BLI: blue laser imaging, LCI: linked color imaging, HP: hyperplastic polyp, SSL: sessile serrated lesions

and LCI improved to detect more lesions under LCI and BLI than WLI. However, the efficacy was lower in BLI than in LCI, suggesting that there is a difference between LCI and BLI detection in AI. A marketed AI for polyp detection is available under not BLI but either WLI or LCI (CAD EYE, Fujifilm, Tokyo, Japan)[1]. However, in the current study, FPR was significantly less in BLI though the number of images in LSs was limited and the sensitivity of lesions between BLI and LCI was not significant. Thus, we suggested the use of both BLI and LCI for marketed AI is expected in the future. Additionally, the sensitivity of LCI in AI (IEE) was superior to trainees and was inferior to experts. Because some lesions which could be detected by experts could not be detected by AI (IEE) in the current study. This might be a limitation of our AI. Or the increase of leaning contents to AI (IEE) may improve AI more. Further analysis should be performed.

The Preservation and Incorporation of Valuable endo-

scopic Innovations (PIVI)-1 threshold, proposed by the American Society of Gastrointestinal Endoscopy recommended the assessment whether AI-assisted optical diagnosis with a high degree of confidence achieved ≥ 90% negative predictive value for adenomatous histology of diminutive rectosigmoid polyps, having histopathology as the reference standard[26]. We could not analyze the nagetive predictive value due to the setting. Regarding the quality of AI, the sensitivity of AI for colorectal polyps becomes >90% after the advent of deep learning[11,27]. In the current study, > 90% sensitivity was not achieved in any AI because we adopted only non-polypoid and non-reddish lesions as the test set in the current study, which were difficult to detect even by experts. In fact, the sensitivity of WLI in AI (WLI) and experts was low (57% and 71%). However, the same AI (WLI) in the previous study showed that the sensitivity of WLI was high (88%) using 100 adenomas of ≤10 mm as a test set[19]. The result was better than that of the current study (the sensitivity of AI (WLI): 76%) though the same AI (WLI) was used. Thus, the sensitivity of WLI is affected by the content of a test set. According to these things, we suggest that the sensitivities of studies on AI need to be discussed considering the contents of a test set and the comparison with the sensitivity of endoscopists using the same test set.

The content and quantity of LS are important for finetuning AI. In this study, we specifically used images of non-reddish and non-polypoid lesions under WLI, BLI, and LCI as LS to improve polyp detection around them. The number of images was slightly smaller than other recent studies on AI for colorectal polyp detection (more than about 10000 images for learning contents)[28,29]. However, we could achieve the higher sensitivity of WLI, BLI and LCI in AI (IEE) compared to that of WLI in AI (WLI). Thus, we demonstrated that specific LS was useful to overcome the general weakness of AI in polyp detection. In addition, we used not only in-focus images but also out-of-focus images as LS in this study. Images in focus improved the sensitivity of WLI, BLI, and LCI in AI (HQ), and images out of focus could improve the sensitivity in AI (IEE) more than those in AI (HQ). Additionally, in AI (HQ) finetuned with 2490 images, the contents of BLI and LCI might be not enough for decreasing FPR of BLI and LCI. On the other hand, in AI (IEE) finetuned with 5920 images could decrease both of FPR. The number of images was enough for decreasing FPR and images out of focus could decrease the FPR of WLI, BLI, and LCI in AI (IEE) more than those in AI (HQ). From this point of view, out-of-focus images could improve AI. However, there was still a significant difference between LCI and BLI in AI (IEE). We thought this suggested LCI with AI needed more learning contents for decreasing FPR compared to BLI.

The definition of FPR has not been formally established and has varied in previous limited studies. In one study, FPR was defined as the number of frames showing a false annotation with AI per all frames in a test set using a short video of each polyp, and it was calculated as 0.001%[28]. In another study, it was calculated the number of false annotation with AI in a test set with a whole withdrawal procedure of colonoscopy and it was calculated 2.2 times per colonoscopy[29]. However, these studies didn't mention the definition of the frame length of an incorrect annotation. In a clinical study, a false positive was defined as a false annotation that persisted for ≥2 seconds of events to improve specificity compared to that for ≥0.5 seconds (99.8% and 93.2%, respectively)[30]. In the current study, FPR was defined as the rate of flames highlighting an annotation box of at least 45 frames (1.5 secs) to an area without polyps with AI in full frames of colorectal observation videos under each WLI, BLI, and LCI. FPR should be investigated more in AI, although the sensitivity of polyp detection was mainly discussed in the studies of AI[12]. We also suggest that the definition of FPR should be unified to improve the rate of FPR.

Regarding the risk factors of polyp miss, there was a significant difference between proximal and distal colon for WLI and LCI in AI (IEE). It might be related to SSL in the proximal colon, which were difficult to detect in both WLI and LCI and poor preparation, although detailed analysis was not performed. Interestingly, the risk factors of AI and endoscopists were different. These results may lead to the improvement of AI in the future.

We were able to show improvement in AI in the current study. However, this may lead to the increase of non-neoplastic lesions such as HP. In addition, it may prolong the observation time and affect the fatigue of endoscopists. Some papers have reported these disadvantages for AI[31]. So far, the recently marketed AI can't differentiate HP from SSL. However, recent research has shown a possibility to differentiate HP from SSL[31,32]. Further analysis is expected to reduce the disadvantages of AI.

Several limitations associated with the present study warrant mention. This was a retrospective examination using recorded images and videos. All of lesions for LSs and the test set were selected only by two experts. The number of test set was small. This study was in vivo analysis of AI using recorded small amount of images so that the improved sensitivity of AI under LCI may not necessarily lead to improved polyp detection in the clinical setting, as polyp detection is influenced by a number of factors. Each LS was taken in different hospital. LS (WLI) was made in a previous study[19]. In the study, all images were taken in Fukushima University Aizu Medical center. In the previous study, Kyoto Prefectural University of Medicine was not included. In the current study, Kyoto Prefectural University of Medicine joined and all of images in LS (IEE) were taken

there. The images of LS (WLI) were taken in a previous system (Processor: VP-4450HD, Light source: LL-4450) and endoscopes (EC-L590ZP). The bright of the images in the previous system were less than those in the recent system. It might affect the results. Polyp detection was a gold standard for evaluating AI. However, we adopted polyp visibility by endoscopists for evaluating test set though it was subjective. The detection of non-reddish lesion was the important endpoint of this study. However, we did not analyze the rate of reddish lesions in LS (WLI) because it was made in a previous study[19]. We did not include large lesions ≥20 mm in the current study. Our previous study showed that polyp visibility was significantly better in lesions ≥10 mm than in lesions <10 mm, although the study included both polypoid and non-polypoid lesions[33]. However, further studies should be expected for examining the detectability of AI for non-polypoid lesions of ≥20 mm.

In conclusion, our original AI revealed better sensitivity under LCI than WLI for detecting non-reddish or non-polypoid lesions. This efficacy could be affected with LS. Additionally, images out of focus might be useful for increasing sensitivity and decreasing FPR in AI. Our study suggested a marketed AI can be improved by images out of focus.

Acknowledgements

We thank Dr. Yoshikazu Hayashi for making an initial AI and we also thank all members of the Department of Molecular Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, for helping with this study.

Conflicts of Interest

All authors have had access to the data and have control of the decision to publish. There were no financial supports for this study. Naohisa Yoshida and Osamu Dohi received a research grant from Fujifilm Co. Naohisa Yoshida received a payment for lectures from Fujifilm Co. Hironori Yamamoto had a consultant relationship with Fujifilm Co. and had received honoraria, grants and royalties from the company. The other authors have no conflicts of interest to declare.

Authors Contributions

Regarding each author's special contribution, Satoshi Sugino and Naohisa Yoshida organized this study. Ken Inoue, Ryohei Hirose, and Osamu Dohi collected the data, Yoshito Itoh helped to organize this study. Zhe Guo, Ruiyao Zhang, Daiki Nemoto, Kazutomo Togashi, and Hironori Yamamoto helped to design AI algorithms and software. Xin Zhu organized the team for R&A of AI algorithms and software.

Approval by Institutional Review Board (IRB)

ERB-C-1704-3 in the Ethics Committee of Kyoto Prefectural University of Medicine

Disclaimer

Hironori Yamamoto is one of the Associate Editors of Journal of the Anus, Rectum and Colon and on the journal's Editorial Board. He was not involved in the editorial evaluation or decision to accept this article for publication at all.

References

1. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. N Engl J Med. 2013 Sep; 369(12): 1095-105.
2. van Rijn JC, Reitsma JB, Stoker J, et al. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol. 2006 Feb; 101(2): 343-50.
3. Zimmermann-Fraedrich K, Sehner S, Rex DK, et al. Right-Sided Location Not Associated With Missed Colorectal Adenomas in an Individual-Level Reanalysis of Tandem Colonoscopy Studies. Gastroenterology. 2019 Sep; 157(3): 660-71.
4. Heresbach D, Barrioz T, Lapalus MG, et al. Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies. Endoscopy. 2008 Apr; 40(4): 284-90.
5. Ikematsu H, Sakamoto T, Togashi K, et al. Detectability of colorectal neoplastic lesions using a novel endoscopic system with blue laser imaging: a multicenter randomized controlled trial. Gastointest Endosc. 2017 Aug; 86(2): 386-94.
6. Suzuki S, Aniwan S, Chiu HM, et al. Linked-Color Imaging Detects More Colorectal Adenoma and Serrated Lesions: An International Randomized Controlled Trial. Clin Gastroenterol Hepatol. 2023 Jun; 21(6): 1493-502.
7. Shinozaki S, Kobayashi Y, Hayashi Y, et al. Colon polyp detection using linked color imaging compared to white light imaging: Systematic review and meta-analysis. Dig Endosc. 2020 Sep; 32(6): 874-81.
8. Kudo SE, Misawa M, Mori Y, et al. Artificial Intelligence-assisted System Improves Endoscopic Identification of Colorectal Neoplasms. Clin Gastroenterol Hepatol. 2020 Jul; 18(8): 1874-81.e2.
9. Barua I, Vinsard DG, Jodal HC, et al. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. Endoscopy. 2021 Mar; 53(3): 277-84.
10. Hassan C, Spadaccini M, Mori Y, et al. Real-Time Computer-Aided Detection of Colorectal Neoplasia During Colonoscopy: A Systematic Review and Meta-analysis. Ann Intern Med. 2023 Sep; 176(9): 1209-20.
11. Mori Y, Neumann H, Misawa M, et al. Artificial intelligence in colonoscopy - Now on the market. What's next? J Gastroenterol Hepatol. 2021 Jan; 36(1): 7-11.
12. Huang D, Shen J, Hong J, et al. Effect of artificial intelligence-aided colonoscopy for adenoma and polyp detection: a meta-analysis of randomized clinical trials. Int J Colorectal Dis. 2022 Mar; 37(3): 495-506.
13. Kobayashi R, Yoshida N, Tomita Y, et al. Detailed superiority of the CAD EYE artificial intelligence system over endoscopists for lesion detection and characterization using unique movie sets. J Anus Rectum Colon. 2024 Apr; 8(2): 61-9.
14. Rondonotti E, Di Paolo D, Rizzotto ER, et al. Efficacy of a computer-aided detection system in a fecal immunochemical test-based organized colorectal cancer screening program: a randomized controlled trial (AIFIT study). Endoscopy. 2022 Dec; 54(12): 1171-9.

15. Sakamoto T, Nakashima H, Nakamura K, et al. Performance of Computer-Aided Detection and Diagnosis of Colorectal Polyps Compares to That of Experienced Endoscopists. Dig Dis Sci. 2022 Aug; 67(8): 3976-83.

16. Yoshida N, Inoue K, Tomita Y, et al. An analysis of the function of a new artificial intelligence, CAD EYE with the lesion recognition and diagnosis for colorectal polyps in clinical practice. Int J Colorectal Dis. 2021 Oct; 36(10): 2237-45.

17. Ahmad OF, Mori Y, Misawa M, et al. Establishing key research questions for the implementation of artificial intelligence in colonoscopy: a modified Delphi method. Endoscopy. 2021 Sep; 53 (9): 893-901.

18. Holzwanger EA, Bilal M, Glissen Brown JR, et al. Benchmarking definitions of false-positive alerts during computer-aided polyp detection in colonoscopy. Endoscopy. 2021 Sep; 53(9): 937-40.

19. Guo Z, Nemoto D, Zhu X, et al. A polyp detection algorithm can detect small polyps: An ex vivo reading test compared with endoscopists. Dig Endosc. 2021 Jan; 33(1): 162-9.

20. Yoshida N, Hayashi Y, Kashida H, et al. Images of laser and light-emitting diode colonoscopy for comparing large colorectal lesion visibility with linked color imaging and white-light imaging. Dig Endosc. 2022 Nov; 34(7): 1413-21.

21. WHO Classification of Tumours Editorial Board. WHO Classification of Tumours. Digestive system tumours. 5th ed. Lyon: International Agency for Research on Cancer; 2019. p. 532-34.

22. Aronchick CA. Bowel preparation scale. Gastrointest Endosc. 2004 Dec; 60(6): 1037-8.

23. Aniwan S, Vanduangden K, Kerr S, et al. Linked color imaging, mucosal exposure device, their combination, and standard colonoscopy for adenoma detection: a randomized trial. Gastrointest Endosc. 2021 Nov; 94(5): 969-77.

24. Dos Santos CEO, Malaman D, Arciniegas Sanmartin ID, et al. Effect of Linked-color Imaging on the Detection of Adenomas in Screening Colonoscopies. J Clin Gastroenterol. 2022 Apr; 56(4): e268-72.

25. Miyaguchi K, Takabayashi K, Saito D, et al. Linked color imaging versus white light imaging colonoscopy for colorectal adenoma detection: A randomized controlled trial. J Gastroenterol Hepatol. 2021 Oct; 36(10): 2778-84.

26. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc. 2011 Mar; 73(3): 419-22.

27. Misawa M, Kudo SE, Mori Y, et al. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. Gastroenterology. 2018 Jun; 154(8): 2027-29.e3.

28. Neumann H, Kreft A, Sivanathan V, et al. Evaluation of novel LCI CAD EYE system for real time detection of colon polyps. PLoS One. 2021 Aug; 16(8): e0255955.

29. Zhao SB, Yang W, Wang SL, et al. Establishment and validation of a computer-assisted colonic polyp localization system based on deep learning. World J Gastroenterol. 2021 Aug; 27(31): 5232-46.

30. Holzwanger EA, Bilal M, Brown JRG, et al. Benchmarking definitions of false-positive alerts during computer-aided polyp detection in colonoscopy. Endoscopy. 2021 Sep; 53(9): 937-40.

31. Mori Y, East JE, Hassan C, et al. Benefits and challenges in implementation of artificial intelligence in colonoscopy: World Endoscopy Organization position statement. Dig Endosc. 2023 May; 35 (4): 422-9.

32. Kader R, Cid-Mejias A, Brandao P, et al. Polyp characterization using deep learning and a publicly accessible polyp video database. Dig Endosc. 2023 Jul; 35(5): 645-55.

33. Yoshida N, Hisabe T, Ikematsu H, et al. Comparison between linked color imaging and blue laser imaging for improving the visibility of flat colorectal polyps -A multicenter pilot study-. Dig Dis Sci. 2020 Jul; 65(7): 2054-62.

Supplementary Files

Supplemental Table 1.

Video 1. A lesion detected by artificial intelligence (AI (IEE)) not with white light imaging (WLI) but with linked color imaging (LCI) and blue light imaging (BLI).

The AI could not detect the lesion (polypoid type, adenoma) of 2vmm under WLI, but could detect it under LCI and BLI.

Please find supplementary file(s);
http://dx.doi.org/10.23922/jarc.2023-070