

Artificial intelligence for identification and characterization of colonic polyps

Nasim Parsa  and Michael F. Byrne

Ther Adv Gastrointest Endosc

2021, Vol. 14: 1–12

DOI: 10.1177/

26317745211014698

© The Author(s), 2021.
Article reuse guidelines:
sagepub.com/journals-
permissions

Abstract: Colonoscopy remains the gold standard exam for colorectal cancer screening due to its ability to detect and resect pre-cancerous lesions in the colon. However, its performance is greatly operator dependent. Studies have shown that up to one-quarter of colorectal polyps can be missed on a single colonoscopy, leading to high rates of interval colorectal cancer. In addition, the American Society for Gastrointestinal Endoscopy has proposed the “resect-and-discard” and “diagnose-and-leave” strategies for diminutive colorectal polyps to reduce the costs of unnecessary polyp resection and pathology evaluation. However, the performance of optical biopsy has been suboptimal in community practice. With recent improvements in machine-learning techniques, artificial intelligence–assisted computer-aided detection and diagnosis have been increasingly utilized by endoscopists. The application of computer-aided design on real-time colonoscopy has been shown to increase the adenoma detection rate while decreasing the withdrawal time and improve endoscopists’ optical biopsy accuracy, while reducing the time to make the diagnosis. These are promising steps toward standardization and improvement of colonoscopy quality, and implementation of “resect-and-discard” and “diagnose-and-leave” strategies. Yet, issues such as real-world applications and regulatory approval need to be addressed before artificial intelligence models can be successfully implemented in clinical practice. In this review, we summarize the recent literature on the application of artificial intelligence for detection and characterization of colorectal polyps and review the limitation of existing artificial intelligence technologies and future directions for this field.

Keywords: artificial intelligence, computer-aided detection, computer-aided diagnosis, convolutional neural network, deep learning

Received: 20 January 2021; revised manuscript accepted: 7 April 2021.

Introduction

Colorectal cancer (CRC) is the third most common cancer and second most common cause of cancer deaths worldwide.¹ Colonoscopy reduces the risk of CRC through detection and resection of pre-cancerous lesions such as adenomas.² The ability to detect adenomas during colonoscopy (ADR) is greatly operator dependent, with studies reporting a wide ADR range of 7% to 53% among different endoscopists.³ Failure to detect and remove neoplastic lesions is associated with the development of interval CRC, which accounts for nearly 10% of all diagnosed CRC.⁴ In addition, most of the detected polyps during colonoscopy are diminutive in size (1–5 mm), with a

negligible risk of progression to cancer.⁵ Unnecessary resection and pathology evaluation of these non-neoplastic lesions are associated with increased costs and adverse events. The American Society for Gastrointestinal Endoscopy (ASGE) has published a Preservation and Incorporation of Valuable endoscopic Innovations (PIVI) statement for optical biopsy of diminutive polyps. The “resect and discard” paradigm is recommended when the optical biopsy provides 90% agreement with histologic assessment for post-polypectomy surveillance intervals, and the “diagnose and leave” strategy is recommended for hyperplastic polyps when the negative predictive value (NPV) for diminutive rectosigmoid adenomas is 90% or

Correspondence to:

Nasim Parsa
Division of
Gastroenterology and
Hepatology, Department
of Medicine, University of
Missouri, Columbia, MO
65211, USA.
parsan@health.missouri.
edu

Michael F. Byrne
Division of
Gastroenterology,
Department of Medicine,
The University of British
Columbia, Vancouver, BC,
Canada; Satisfai Health,
Vancouver, BC, Canada

more.⁶ However, the performance of endoscopists' optical biopsy has not consistently reached these thresholds in community practice.

To overcome these challenges, artificial intelligence (AI) has been introduced to the field of endoscopy. AI-assisted computer-aided detection (CADe) and diagnosis (CADx) systems, especially deep-learning techniques, are promising options to improve detection and optical biopsy and decrease human variation through the ability to process high-dimensional endoscopic data and to self-identify trainable parameters not appreciable to humans. The application of computer-aided design (CAD) on real-time colonoscopy has been shown to increase the ADR, reduce the withdrawal time, improve endoscopists' optical biopsy, while reducing the time to make a diagnosis. Yet, issues such as real-world applications and regulatory approval need to be addressed before AI models can be successfully implemented in clinical practice. In this review, we summarize the recent literature on the application of AI for detection and characterization of colorectal polyps, and review the clinical implementation, current limitation of existing AI technologies, and future directions for this field.

AI for detection of colorectal polyps (CADe)

Table 1 summarizes the important studies on CADe for detection of colorectal polyps.

The initial CADe systems were reported in the early 2000s.^{7,8,24} These systems were designed with a handcrafted algorithm, based on certain polyp features, and provided accuracy more than 90%. Several other groups designed and evaluated different handcrafted CADe solutions, using small numbers of static images. While these systems typically showed high accuracy on carefully chosen data sets, they were limited in real-world application due to low sensitivity, high false-positive rates, and long processing time. More recently, deep-learning algorithms such as convolutional neural networks (CNNs) have been utilized for the development of CADe systems, enabling the continuous recognition of abnormal lesions without the need for external input. Using 50 polyp and 85 non-polyp videos, Misawa and colleagues¹¹ developed a three-dimensional CNN-based CADe with a sensitivity and specificity of 90% and 63%, respectively. Urban and colleagues reported the first real-time application of CNN-based CADe, trained on more than 8,000

images from 2,000 patients. Their CADe showed 97% sensitivity, 95% specificity, and 96% accuracy for detection of colorectal polyps, which was superior to the performance of the endoscopist (45% vs 36%). The unique feature of this study was that of the 73 polyps missed by endoscopist, 67 were detected by CADe, with a false-positive rate of 5%.¹² Klare and colleagues prospectively studied CADe during live colonoscopy performed by a trained endoscopist while a second observer monitored the CADe output. The system analyzed with an average delay of only 50 ms and achieved a polyp detection rate (PDR) of 51% and ADR of 29%, comparable to the endoscopist's PDR of 56% and ADR of 31%. The first commercially available CADe (GI-Genius, Medtronic) was recently studied in a retrospective validation trial which showed an excellent performance with a per-lesion sensitivity rate of 99.7%.¹⁵

To date, eight randomized controlled trials (RCTs) have compared CADe to standard colonoscopy, all demonstrating a significantly higher ADR by CADe. Wang and colleagues reported the first RCT (non-blinded) on 1,058 patients (536 with CADe, 522 without CADe) and reported a significantly higher ADR (29.1% vs 20.3%, $p < 0.001$) and increased number of adenomas per patient (0.53 vs 0.31) in the CADe group. However, the increased ADR was limited to an increase in detection of diminutive adenomas, and there was no difference in detection of polyps more than 10 mm between the two groups. Moreover, a higher proportion of polyps detected by CADe were hyperplastic (43.6% vs 34.9%) and there was no difference in the proportion of detected advanced adenomas or sessile serrated lesions (SSL) between the two groups.¹⁷ The same authors performed a double-blind RCT using sham-AI and showed significantly greater ADR in the CADe than the sham group (34% vs 28%, $p = 0.03$).¹⁸ Su and colleagues designed a CADe that was able to evaluate the quality of bowel preparation and measure the withdrawal time. In their study, 308 and 315 patients were analyzed in the CADe and control groups. The CADe group had a significantly higher ADR (29% vs 17%, $p < 0.001$) with prolonged exposure time (7.0 vs 5.6 min, $p < 0.001$) and adequate bowel preparation.²² Liu and colleagues²¹ conducted an RCT on 1,026 patients and found that the CADe group had a significantly higher ADR (39% vs 24%, $p < 0.001$). Repici and colleagues conducted a multicenter RCT for the GI Genius CADe system

Table 1. Clinical studies on computer-aided detection (CADe) for colorectal polyps.

Author	Study design	CADe system	AI method	Study subjects, n	Outcomes, %
Park and Sargent ⁷	Retrospective ex vivo	-	CNN	Training and testing sets: 562 images	Sensitivity 86 Specificity 85
Billah and colleagues ⁸	Retrospective ex vivo	-	CNN	Combined training and testing sets: 14,000 still images	Sensitivity 99 Specificity 99
Yu, 2017 ⁹	Retrospective ex vivo	-	CNN	Training set: ASU-Mayo 20 videos Testing set: ASU-Mayo 18 videos	Sensitivity 71 PPV 88
Zhang, 2017 ¹⁰	Retrospective ex vivo	-	CNN	Training set: 2262 still images Testing set: 150 random + 30 NBI still images	Sensitivity 98 PPV 99 Accuracy 86
Misawa and colleagues ¹¹	Retrospective ex vivo	-	CNN	Training set: 411 videos Testing set: 135 videos	Per-polyp sensitivity 94 Per-frame sensitivity 90 Specificity 63 False Positive 60 Accuracy 76
Urban and colleagues ¹²	Retrospective ex vivo	-	CNN	Multiple training sets: 8641 images, 1330 images, 9 videos, 53,588 images from videos, 11 challenging videos Testing set: 20 colonoscopy videos	Sensitivity 90 AUC 0.991 Accuracy 96 False positive 7
Yamada, 2018 ¹³	Retrospective ex vivo	-	CNN	Training set: 139,983 video images Testing set: 4840 video images	Sensitivity 97 specificity 99 AUC 0.975
Misawa and colleagues ¹⁴	Ex vivo	-	CNN	Training set: 3,017,088 frames Testing set: 64 videos	Sensitivity 86 False positive 26

(Continued)

Table 1. (Continued)

Author	Study design	CADe system	AI method	Study subjects, n	Outcomes, %
Klare and colleagues ¹⁵	Prospective in vivo	–	CNN	Training set: not reported Testing set: 55 live colonoscopies	Per-polyp sensitivity 75 ADR 29 (31 in endoscopist)
Ozawa and colleagues ¹⁶	Ex vivo	–	CNN	Training set: 16,418 images Testing set: 7,077 still images	Sensitivity 92 PPV 93 Accuracy 85
Wang and colleagues ¹⁷	Single-center RCT in vivo	EndoScreener	CNN	CADe patients with adenoma 151, total patients 522 Control patients with adenoma 109, total patients 536	ADR: WLI 20 vs CADe 29
Wang and colleagues ¹⁸	Single-center RCT in vivo	EndoScreener	CNN	CADe patients with adenoma 165, total patients 484 Control patients with adenoma 134, total patients 478	ADR: WLI 28 vs CADe 34
Gong and colleagues ¹⁹	Single-center RCT in vivo	ENDOANGEL	CNN	CADe patients with adenoma 54, total patients 324 Control patients with adenoma 26, total patients 318	ADR: WLI 8 vs CADe 16
Repici and colleagues ²⁰	Multicenter RCT in vivo	GI-Genius	CNN	CADe group patients with adenoma 187, total patients 341 Control group patients with adenoma 139, total patients 348	ADR: WLI 40, 40.4 vs CADe 54
Liu and colleagues ²¹	Single-center RCT in vivo	HenanTongyu	CNN	CADe patients with adenoma 199, total patients 508 WLI patients with adenoma 124, total patients 518	ADR: WLI 24 vs CADe 39
Su and colleagues ²²	Single-center RCT in vivo	AQCS	CNN	CADe group patients with adenoma 89, total patients 308 Control group patients with adenoma 52, total patients 315	ADR: WLI 16 vs CADe 28
Wang and colleagues ²³	Single-center RCT in vivo	EndoScreener	CNN	CADe group patients with adenoma 124, total patients 184 Control group patients with adenoma 72, total patients 185	ADR: WLI 39 vs CADe 67

ADR, adenomas during colonoscopy; AI, artificial intelligence; AQCS, Automatic Quality Control System; AUC, area under the curve; CADe, computer-aided detection; CNN, convolutional neural network; NBI, narrow band imaging; PPV, positive predictive value; RCT, randomized controlled trials; WLI, white light imaging.

on 685 patients and identified a significantly higher ADR in the CADE group (54.8% vs 40.4%). It is important to note that this study showed higher ADR for both diminutive (33.7% vs 26.5%) and small (6–9 mm) size adenomas (10.6% vs 5.8%) which was irrespective of the polyp shape or location. Another unique feature of this study was its high baseline ADR, as opposed to the aforementioned studies.²⁰ Gong and colleagues developed a CADE with the ability to recognize cecal intubation. In addition to showing a significantly higher ADR in the CADE group (16% vs 8%, $p = 0.001$), they demonstrated a significantly higher detection rate for advanced polyps as well (3% vs 1%).¹⁹ Wang and colleagues conducted the first randomized tandem trial comparing CADE with standard colonoscopy. The adenoma miss rate was significantly lower in the CADE group (13.8% vs 40.0%, $p < 0.001$) and was significant for diminutive (39.6% vs 13.1%, $p < 0.001$), and small polyps (46.9% vs 13.7%, $p < 0.0001$), but not for the polyps bigger than 10 mm in size (15.3% vs 33.3%).²³ They further evaluated the miss rate among visible polyps (exposed but not recognized by endoscopists) and invisible (not exposed) and reported that CADE rarely misses that polyp if the mucosa is exposed by the operator (visible in the CADE: adenoma miss rate 1.5%, polyp miss rate 2.3%). Regarding sessile serrated polyps, serrated miss rate was found not to be significantly different between the two groups.

AI for characterization of colorectal polyps (CADx)

Table 2 summarizes the studies on CADx for characterization of colorectal polyps.

CADx for digital image-enhanced endoscopy

Narrow band imaging (NBI; Olympus Corp., Tokyo, Japan)-based CADx systems are the most extensively studied modality to date. The initial CADx systems utilized a support vector machine (SVM) and were made for magnifying NBI, which limited the widespread use of these systems in clinical practice.^{25,26,28} Recent integration of CNN with CADx has resulted in systems with higher diagnostic accuracy and faster processing times.^{31,40,41} Using standard non-magnified NBI, Chen and colleagues³¹ developed a CNN-based CADx that had sensitivity, specificity, positive predictive value (PPV), NPV, and accuracy of 96.3%,

78.1%, 89.6%, 91.5%, and 91%, respectively. Byrne and colleagues developed the first CADx that reached the ASGE optical biopsy thresholds in real-time clinical practice.³⁴ Using standard NBI, they trained the CADx with 223 polyp videos (60,089 frames) and tested their system on 125 diminutive polyp videos, of which credibility score did not reach more than 50% for 19 polyps. Of the remaining 106 polyp videos, the sensitivity, specificity, PPV, NPV, and accuracy for identifying diminutive adenomas and hyperplastic polyps were 98%, 83%, 90%, 97%, and 94%, respectively. Zachariah and colleagues³⁷ designed a CNN-based CADx with both white-light imaging (WLI) and NBI that exceeded the ASGE PIVI thresholds with NPV and accuracy of 93% and 94%, respectively. This study resulted in accurate automatic classification of diminutive polyps, irrespective of endoscopists' experience and NBI usage, which could potentially be a positive factor for the community endoscopists. Using both NBI and blue light imaging (BLI), Zorron Cheng Tao Pu developed a CADx based on the modified Sano (MS) classification and validated it with two internal and external polyp image data sets.^{39,42} The CADx had a mean area under the curve (AUC) of 94.3% for the internal set, and 84.5% and 90.3% for the external sets (NBI and BLI, respectively). A unique feature of this study was to show an equal highly accurate CADx prediction across two different imaging technologies (NBI and BLI), suggesting the potential to have a CADx trained and used with two different technologies, even when the predicted endoscopy imaging technology is not part of the training set. Moreover, the CADx AUC was comparable with experts and similar with both NBI and BLI. Song and colleagues developed and compared their CNN-based CADx model with both trainees and NBI expert endoscopists. The CADx system had a significantly higher diagnostic accuracy (81%–82%) compared with the trainees (63.8%–71.8%, $p < 0.01$), and comparable to the experts (82.4%–87.3%, $p = 0.72$).³⁵ Importantly, the addition of CADx as a support tool resulted in significant improvement in trainees' diagnostic accuracy (63.8%–72% vs 82.7%–84.2%, $p < 0.001$). Similar results were also noted by Jin and colleagues, who showed that the addition of CADx as a support tool resulted in improvement of endoscopists' diagnostic accuracy (82.5% to 88.5%, $p < 0.05$). The greatest improvement was noted in novice endoscopists (73.8% to 85.6%, $p < 0.05$), almost reaching the accuracy of experts (89.0%, $p = 0.10$).³⁸

Table 2. Clinical studies on computer-aided diagnosis (CADx) for characterization of colorectal polyps.

Author	Study design	AI method	Study subjects	Imaging modality	Study comparison	Outcomes, %
Tischendorf and colleagues ²⁵	Retrospective pilot	SVM	Training set: not reported Testing set: 209 polyps (160 neoplastic, 49 non-neoplastic)	Magnifying NBI	Adenomas vs non-adenomas	Sensitivity 90 Specificity 70 Accuracy 85
Gross and colleagues ²⁶	Prospective	SVM	Training set: not reported Testing set: 434 small (<10 mm) polyps (258 neoplastic, 176 non-neoplastic)	Magnifying NBI	Small adenomas vs non-adenomas	Sensitivity 95 Specificity 90 PPV 93 NPV 92 Accuracy 93
Takemura and colleagues ²⁷	Retrospective	SVM	Training set: 1,519 polyps Testing set: 371 images	Magnification chromoendoscopy	Neoplastic or non-neoplastic polyps	Sensitivity 97 Specificity 97 Accuracy 97
Komimami and colleagues ²⁸	Prospective	SVM	Training set: 1,262 polyps Testing set: 118 images (73 neoplastic, 45 non-neoplastic)	Magnifying NBI	Small adenomas vs non-adenomas	Sensitivity 93 Specificity 95 PPV 95 NPV 93 Accuracy 97
Mori and colleagues ²⁹	Retrospective	SVM	Training set: 6,051 images Testing set: 205 polyps (147 neoplastic and 58 non-neoplastic)	Endocytoscopy	Adenomas vs non-adenomas	Sensitivity 87 Specificity 91 PPV 98 NPV 84 Accuracy 89
Misawa and colleagues ²⁹	Retrospective	SVM	Training set: 1,079 images (431 benign, 648 malignant) Testing set: 100 images (50 benign, 50 malignant)	Endocytoscopy with NBI	Predict histology Preclinical	Sensitivity 84 Specificity 97 Accuracy 90 PPV 98 NPV 82
Komeda and colleagues ³⁰	Retrospective	CNN	Training set: 1,800 polyp images (1,200 adenomatous, 600 non-adenomatous), cross-validation 10 videos	WLI, NBI	Adenomatous vs non-adenomatous polyps	Accuracy 75
Chen and colleagues ³¹	Retrospective	CNN	Training set: 2,157 images (1,476 neoplastic, 681 hyperplastic) Testing set: 284 images (188 neoplastic, 96 hyperplastic)	Standard NBI	Diminutive (<5 mm) neoplastic vs hyperplastic polyps	Sensitivity 96 Specificity 78 PPV 89 NPV 90 Accuracy 90

(Continued)

Table 2. (Continued)

Author	Study design	AI method	Study subjects	Imaging modality	Study comparison	Outcomes, %
Mori and colleagues ³²	Prospective	SVM	Training set: 61,952 images Testing set: 466 polyps (250 rectosigmoid)	Endocytoscopy with NBI	Diminutive adenomas vs non-adenomas	Sensitivity 95 Specificity 92 PPV 96 NPV 96 Accuracy 98
Sánchez-Montes and colleagues ³³	Retrospective	SVM	Testing set: 225 polyps (142 dysplastic and 83 nondysplastic)	WLI	Adenomatous vs non-adenomatous polyps	Sensitivity 92 NPV 91 Accuracy 90 Diminutive polyps: NPV 96, accuracy 87
Byrne and colleagues ³⁴	Retrospective	CNN	Training set: 223 videos, 60,089 frames Validation set: 40 videos Testing set: 125 videos, 51 hyperplastic, 74 adenomatous polyps	NBI	Diminutive adenomas vs non-adenomas	Sensitivity 98 Specificity 83 PPV 90 NPV 97 Accuracy 92
Song and colleagues ³⁵	Retrospective	CNN	Training set: 12,480 images from 624 polyps Testing set: Set 1: 15 hyperplastic, 24 SSLs, 106 adenomas. Set 2: 30 hyperplastic, 70 SSLs, 206 adenomas	NBI	Adenomatous polyps vs SSLs	Sensitivity 82, 84 Specificity 93, 88 Accuracy 81, 82
Kudo and colleagues ³⁶	Retrospective	EndoBRAIN	69, 142 endocytoscopic images, taken at 520-fold magnification from 2,000 polyps	Endocytoscopy, NBI	Neoplastic vs non-neoplastic	Sensitivity 96 Specificity 94 PPV 96 NPV 94 Accuracy 96
Zachariah and colleagues ³⁷	Retrospective	CNN	Training set: 5,278 images, 3,310 adenomatous, 1,968 SSLs Testing set: 634 images	NBI, WLI	Diminutive adenomatous vs SSLs/hyperplastic polyps	Sensitivity 96 Specificity 90 PPV 94 NPV 93 Accuracy 94
Jin and colleagues ³⁸	Retrospective	CNN	Training set: 1,100 adenomatous and 1,050 hyperplastic polyps Testing set: 180 adenomatous, 120 hyperplastic polyps	NBI	Diminutive adenomas vs non-adenomas	Sensitivity 83 Specificity 91 Accuracy 87
Ozawa and colleagues ¹⁶	Retrospective	CNN	Training set: 16,418 images of 4,752 polyps, 4,013 images of normal colorectums Validation set: 7,077 images including 1,172 polyp images (309 polyps)	NBI, WLI	Diminutive adenomas vs non-adenomas	WLI NPV 85 NBI NPV 91
Zorron Cheng Tao Pu and colleagues ³⁹	Retrospective	CNN	Training, testing, and internal validation set: 1,235 polyp images External validation set: 69 polyps (20 NBI, 49 BLI)	NBI, BLI	Differentiate lesions into 5 subtypes, including SSLs	Internal set AUC 94 External set AUC 84 External set AUC 90

AI, artificial intelligence; AUC, area under the curve; BLI, blue light imaging; CADx, computer-aided diagnosis; CNN, convolutional neural network; NBI, narrow band imaging; NPV, negative predictive value; PPV, positive predictive value; SSL, sessile serrated lesion; SVM, support vector machine; WLI, white light imaging.

CADx for chromoendoscopy

There are a few older studies on CADx for chromoendoscopy. Takemura and colleagues developed a software that enabled computer-aided prediction of pit pattern by extracting six features (e.g. area, perimeter, circularity) from crystal violet-stained images. Their CADx performed surprisingly well, with 98.5% accuracy.²⁷ Pit pattern classification requires crystal violet staining by endoscopist, and the depth of color depends on how much dye is sprayed. Therefore, it is difficult to obtain uniform image quality and as a result, to obtain robust CADx for chromoendoscopy.

CADx for white-light imaging

Studies on CADx for WLI have failed to report high diagnostic accuracy, likely because optical diagnosis using WLI is usually less informative than by NBI or chromoendoscopy. Komeda and colleagues developed a WLI-based CADx model with a reported diagnostic accuracy rate of only 75.1%. Sánchez-Montes WLI-based CADx reached 95.0% sensitivity, 87.9% specificity, 82.6% PPV, 96.7% NPV, and 91.1% accuracy for differentiating diminutive rectosigmoid adenomas.^{33,30}

CADx for endocytoscopy

Endocytoscopy (H290ECI, Olympus, Tokyo, Japan) is a novel in vivo microscopic imaging technique that allows real-time visualization of cellular and microvascular patterns of colorectal polyps.⁴³ Endocytoscopy is considered ideal for pairing with CAD systems because it consistently provides focused, fixed-size images, thus facilitating easier image analysis. In 2015, Mori and colleagues developed a CAD system which used stained feature extraction to predict neoplastic polyps in 152 patients. Polyps less than 10 mm were analyzed in real-time and the system was able to achieve a sensitivity of 92.0% and specificity of 79.5%, with an accuracy of 89.2% for identifying neoplastic changes, comparable to those of expert endoscopists.⁴⁴ In a prospective trial on 791 patients and 466 diminutive rectosigmoid polyps, the NPV was 93.7%, reaching the performance level required for the ASGE diagnose-and-leave strategy.³² Misawa and colleagues²⁹ developed an NBI-based CADx for endocytoscopy that achieved more impressive results with overall sensitivity of 84.5%, specificity of 97.6%, and accuracy of 90.0% using the existing training images. When the resulting probability of diagnosis was greater than 90%, the result was considered a

“high-confidence” diagnosis. These diagnoses carried an overall sensitivity of 97.6%, specificity of 95.8%, and accuracy of 96.9%, surpassing the proposed cutoffs for the diagnose-and-leave strategy.²⁹ In a retrospective comparison of 30 endoscopists (trainee and expert) of both stained endocytoscopy and NBI images versus endocytoscopy, endocytoscopy identified colon lesions with 96.9% sensitivity, 100% specificity, 98% accuracy, 100% PPV, and 94.6% NPV, which were all significantly greater than those of the endoscopy trainees and experts. For NBI, endocytoscopy distinguished neoplastic from non-neoplastic lesions with 96.9% sensitivity, 94.3%, 96.0% accuracy, 96.9% PPV, and a 94.3% NPV, all significantly higher than those of the endoscopy trainees. Sensitivity and NPV were significantly higher, but the other values are comparable to those of the experts.³⁶ A recent cost-effectiveness analysis on the use of AI for implementing the diagnose-and-leave strategy showed that through AI, 145 rectosigmoid diminutive polyps were not resected, which suggested that one could reduce the average colonoscopy cost and the gross annual reimbursement for colonoscopies by 18.9% and US\$149.2 million in Japan, 6.9% and US\$12.3 million in England, 7.6% and US\$1.1 million in Norway, and 10.9% and US\$85.2 million in the United States, respectively.⁴⁵ However, endocytoscopy is not widely used in clinical practice. Given its cost-efficient potential, more attention should be paid toward regulation, accessibility, and effective implementation of this powerful technology.

Full workflow systems (CADe + CADx)

To enhance the integration of CAD systems into clinical practice, full workflow systems with the ability to perform both polyp detection and characterization have been developed. Mori and colleagues¹⁷ designed a novel CAD that included two algorithms, a deep learning-based CAD for polyp detection with WLI, and an algorithm for optical biopsy by endocytoscopic images. Guizard and colleagues⁴⁶ developed a full workflow system using both WL and NBI, which was also able to tag polyps with unique identifiers that could be tracked throughout the procedure. Ozawa and colleagues designed a CNN-based CAD for both WLI and NBI, using a single-shot MultiBox detector that could detect and characterize a target object simultaneously. For WLI, the sensitivity and PPV were 90% and 83%, and for NBI, the sensitivity and PPV were 97% and 98%, respectively. Among those lesions that were accurately

identified as polyps, 83% were correctly classified through images and 97% of adenomas were precisely identified under the WLI.¹⁷

Limitations and future directions

While AI technologies have shown impressive results for detection and histologic prediction of colorectal polyps, there are still several points that need to be addressed before the use of CAD can be implemented in routine clinical practice. To improve the reliability and minimize bias, the performance of CAD systems should be evaluated in prospective RCTs, conducted in both community and academic centers, and among endoscopists with different levels of experience. The preferred study endpoint would be those of ASGE PIVI strategies, for example, the design of the CAD models should use widely available technology (such as standard NBI), with the ability to process raw videos taken during real-time colonoscopy. Moreover, training should be performed with a large number of standardized high-quality data sets, and testing should be done with several data sets and diverse contents. Recently, Misawa and colleagues launched a publicly accessible colonoscopy video database (SUN-database) that contains 49,799 polyp frames annotated with bounding boxes and 102,761 frames without polyps, making a total of 152,560 frames.⁴⁷ It is important to note that the pathology is not always the gold standard for diagnosis, especially regarding the ≤ 3 mm colorectal lesions. In a recent study on 644 colon polyps ≤ 3 mm in size, there was a 28.9% (13.2% HPs, 0.3% SSLs, and 15.4% normal mucosa; respectively) discrepancy between expert endoscopic and histologic opinion, of which 15.4% were diagnosed as normal by the pathologist. Following a blinded optical evaluation by two expert endoscopists, agreement with the endoscopic diagnosis was made in 94% and 100% of cases, respectively.⁴⁸ Based on these data, Shahidi and colleagues evaluated the application of AI as the arbitration between endoscopist and pathologist when discordant diagnoses occur. They used an established real-time AI clinical decision support solution (CDSS), which agreed with the endoscopic diagnosis in 89.6% lesions. In discordant cases, CDSS agreed with the endoscopic diagnosis in 90.3% lesions. Interestingly, of those lesions identified on pathology as normal mucosa, CDSS agreed with the endoscopic diagnosis in 90.9% of cases.⁴⁹ In addition to adenomas, the CAD designs should also focus on detecting the proximal colon lesions, specifically SSLs.

Obtaining regulatory approval is an essential factor for using CAD systems in clinical practice. Currently, the CAD EYE™ (Fujifilm Corp, Tokyo, Japan), DISCOVERY™ (Pentax Corp, Tokyo, Japan), Endo-AID (Olympus Corp), and GI-Genius (Medtronic Corp, Minneapolis, MN) have successfully obtained the regulatory approval, which hopefully will open doors for more platforms. Medico-legal issues are important topics to be discussed. As AI systems do not always provide accurate information, negative results due to the use of AI can possibly happen, which could lead to medico-legal challenges. We should recognize the strengths and weaknesses of AI and avoid over-relying on the results of AI. However, with wide spread of the AI tools in medical fields, we will have to reconsider the medico-legal issues in the near future.

Summary

In recent years, the application of AI has significantly expanded in the field of gastrointestinal endoscopy. Multiple studies have shown that integration of CAD with colonoscopy can improve the endoscopists' performance in detection and characterization of colorectal polyps, which are promising steps toward improving and standardizing colonoscopy quality and implementing the ASGE PIVI paradigm, among others. However, the majority of these data are based on small studies at tertiary care centers, with relatively small number of images used for the AI model's training set, with possible selection bias and no randomization. There is a substantial need for large, multicenter clinical trials to establish the diagnostic accuracy of AI technology in real-time clinical practice, which will be an essential step for obtaining regulatory approval and widespread use of AI technologies.

Conflict of interest statement

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: M.F.B.: CEO and shareholder: Satisfai Health; founder of AI4GI joint venture. Co-development agreement between Olympus America and AI4GI in artificial intelligence and colorectal polyps. N.P. has no conflicts to declare.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Nasim Parsa  <https://orcid.org/0000-0003-3882-266X>

References

1. Bray F, Ferlay J, Soerjomataram I, *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68: 394–424.
2. Brenner H and Stock CHM. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 2014; 348: g2467.
3. Corley DA, Jensen CD, Marks AR, *et al.* Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; 370: 1298–1306.
4. Robertson DJ, Lieberman DA, Winawer SJ, *et al.* Colorectal cancers soon after colonoscopy: a pooled multicohort analysis. *Gut* 2014; 63: 949–956.
5. Ponugoti PL, Cummings OW, Rex DK. Risk of cancer in small and diminutive colorectal polyps. *Dig Liver Dis* 2017; 49: 34–37.
6. 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; 73: 419–422.
7. Park SY and Sargent D. Colonoscopic polyp detection using convolutional neural networks. In: Tourassi GD and Armato SG III (eds) *Medical imaging 2016: computer-aided diagnosis*. Bellingham WA: International Society for Optics and Photonics, 2016, p. 978528.
8. Billah M, Waheed S and Rahman MM. An automatic gastrointestinal polyp detection system in video endoscopy using fusion of color wavelet and convolutional neural network features. *Int J Biomed Imaging* 2017; 2017: 9545920.
9. Yu L, Chen H, Dou Q, Qin J, *et al.* Integrating Online and Offline Three-Dimensional Deep Learning for Automated Polyp Detection in Colonoscopy Videos. *IEEE J Biomed Health Inform* 2017; 21(1): 65–75.
10. Zhang R, Zheng Y, Mak TW, *et al.* Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE J Biomed Health Inform* 2017; 21(1): 41–47.
11. Misawa M, Kudo SE, Mori Y, *et al.* Artificial intelligence assisted polyp detection for colonoscopy. *Gastroenterology* 2018; 154: 2027–2029.
12. Urban G, Tripathi P, Alkayali T, *et al.* Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 2018; 155: 1069–1078.
13. Yamada M, Saito Y, Imaoka H, *et al.* Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Sci Rep* 2019; 9: 1–9.
14. Mori Y, Kudo SE, Misawa M, *et al.* Simultaneous detection and characterization of diminutive polyps with the use of artificial intelligence during colonoscopy. *VideoGIE* 2019; 4: 7–10.
15. Klare P, Sander C, Prinzen M, *et al.* Automated polyp detection in the colorectum: a prospective study (with videos). *Gastrointest Endosc* 2019; 89: 576–582.
16. Ozawa T, Ishihara S, Fujishiro M, *et al.* Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. *Therap Adv Gastroenterol* 2020; 13: 1756284820910659.
17. Wang P, Berzin TM, Glissen Brown JR, *et al.* Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; 68: 1813–1819.
18. Wang P, Liu X, Berzin TM, *et al.* Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADE-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; 5: 343–351.
19. Gong D, Wu L, Zhang J, *et al.* Detection of colorectal adenomas with a realtime computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020; 5: 352–361.
20. Repici A, Badalamenti M, Maselli R, *et al.* Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. *Gastroenterology* 2020; 159: 512–520.

21. Liu WN, Zhang YY, Bian XQ, *et al.* Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; 26: 13–19.
22. Su JR, Li Z, Shao XJ, *et al.* Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc* 2020; 91: 415–424.
23. Wang P, Liu P, Glissen Brown JR, *et al.* Lower adenoma miss rate of computer-aided detection-assisted colonoscopy vs routine white-light colonoscopy in a prospective tandem study. *Gastroenterology* 2020; 159: 1252–1261.
24. Karkanis SA, Iakovidis DK, Maroulis DE, *et al.* Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Trans Inf Technol Biomed* 2003; 7: 141–152.
25. Tischendorf JJ, Gross S, Winograd R, *et al.* Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. *Endoscopy* 2010; 42: 203–207.
26. Gross S, Trautwein C, Behrens A, *et al.* Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc* 2011; 74: 1354–1359.
27. Takemura Y, Yoshida S, Tanaka S, *et al.* Quantitative analysis and development of a computer-aided system for identification of regular pit patterns of colorectal lesions. *Gastrointest Endosc* 2010; 72: 1047–1051.
28. Kominami Y, Yoshida S, Tanaka S, *et al.* Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; 83: 643–649.
29. Misawa M, Kudo S-E, Mori Y, *et al.* Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy. *Gastroenterology* 2016; 150: 1531–1532.
30. Komeda Y, Handa H, Watanabe T, *et al.* Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience. *Oncology* 2017; 93(Suppl. 1): 30–34.
31. Chen PJ, Lin MC, Lai MJ, *et al.* Accurate classification of diminutive colorectal polyps using computer-aided analysis. *Gastroenterology* 2018; 154: 568–575.
32. Mori Y, Kudo SE, Misawa M, *et al.* Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Ann Intern Med* 2018; 169: 357–366.
33. Sánchez-Montes C, Sánchez FJ, Bernal J, *et al.* Computer-aided prediction of polyp histology on white-light colonoscopy using surface pattern analysis. *Endoscopy* 2019; 51: 261–265.
34. Byrne MF, Chapados N, Soudan F, *et al.* Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2019; 68: 94–100.
35. Song E, Park B, Ha C, *et al.* Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model. *Sci Rep* 2020; 10: 30.
36. Kudo SE, Misawa M, Mori Y, *et al.* Artificial intelligence-assisted system improves endoscopic identification of colorectal neoplasms. *Clin Gastroenterol Hepatol* 2020; 18: 1874–1881.
37. Zachariah R, Samarasekera J, Luba D, *et al.* Prediction of polyp pathology using convolutional neural networks achieves “resect and discard” thresholds. *Am J Gastroenterol* 2020; 115: 138–144.
38. Jin EH, Lee D, Bae JH, *et al.* Improved accuracy in optical diagnosis of colorectal polyps using convolutional neural networks with visual explanations. *Gastroenterology* 2020; 158: 2169–2179.e8.
39. Zorron Cheng Tao Pu L, Maicas G, Tian Y, *et al.* Computer-aided diagnosis for characterization of colorectal lesions : comprehensive software that includes differentiation of serrated lesions. *Gastrointest Endosc* 2020; 92: 891–899.
40. Alagappan M, Brown JRG and Mori YBT. Artificial intelligence in gastrointestinal endoscopy: the future is almost here. *World J Gastrointest Endosc* 2018; 10: 239–249.
41. Djinbachian R, Dube AJ and von Renteln D. Optical diagnosis of colorectal polyps: recent developments. *Curr Treat Options Gastroenterol* 2019; 17: 99–114.
42. Singh R, Jayanna M, Navadgi S, *et al.* Narrow-band imaging with dual focus magnification in differentiating colorectal neoplasia. *Dig Endosc* 2013; 25(Suppl. 2): 16–20.
43. Mori Y, Kudo S, Ikehara N, *et al.* Comprehensive diagnostic ability of endocytoscopy compared with biopsy for colorectal neoplasms: a prospective randomized noninferiority trial. *Endoscopy* 2013; 45: 98–105.

44. Mori Y, Kudo S-e, Wakamura K, *et al.* Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos). *Gastrointest Endosc* 2015; 81: 621–629.
45. Mori Y, Kudo SE, East JE, *et al.* Cost savings in colonoscopy with artificial intelligence-aided polyp diagnosis : an add-on analysis of a clinical trial (with video). *Gastrointest Endosc* 2020; 92: 905–911.e1.
46. Guizard N, Ghalehjegh SH, Henkel M, *et al.* Artificial intelligence for realtime multiple polyp detection with identification, tracking, and optical biopsy during colonoscopy. *Gastroenterology* 2019; 156: S48–S49.
47. Misawa M, Kudo S-E, Mori Y, *et al.* Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). *Gastrointest Endosc* 2021; 93: 960–967.e3.
48. Ponugoti P, Rastogi A, Kaltenbach T, *et al.* Disagreement between high confidence endoscopic adenoma prediction and histopathological diagnosis in colonic lesions \leq 3 mm in size. *Endoscopy* 2019; 51: 221–226.
49. Shahidi N, Rex DK, Kaltenbach T, *et al.* Use of endoscopic impression, artificial intelligence, and pathologist interpretation to resolve discrepancies from endoscopy and pathology analyses of diminutive colorectal polyps. *Gastroenterology* 2020; 158: 783–785.

Visit SAGE journals online
[journals.sagepub.com/
home/cmj](https://journals.sagepub.com/home/cmj)

 SAGE journals