ORIGINAL ARTICLE



Artificial Intelligence-Assisted Colonoscopy for Detection of Colon Polyps: a Prospective, Randomized Cohort Study

Yuchen Luo¹ · Yi Zhang¹ · Ming Liu¹ · Yihong Lai¹ · Panpan Liu¹ · Zhen Wang¹ · Tongyin Xing¹ · Ying Huang¹ · Yue Li¹ · Aiming Li¹ · Yadong Wang¹ · Xiaobei Luo¹ · Side Liu¹ · Zelong Han¹

Received: 20 May 2020 / Accepted: 9 September 2020 / Published online: 23 September 2020 \odot 2020 The Author(s)

Abstract

Background and aims Improving the rate of polyp detection is an important measure to prevent colorectal cancer (CRC). Realtime automatic polyp detection systems, through deep learning methods, can learn and perform specific endoscopic tasks previously performed by endoscopists. The purpose of this study was to explore whether a high-performance, real-time automatic polyp detection system could improve the polyp detection rate (PDR) in the actual clinical environment.

Methods The selected patients underwent same-day, back-to-back colonoscopies in a random order, with either traditional colonoscopy or artificial intelligence (AI)-assisted colonoscopy performed first by different experienced endoscopists (> 3000 colonoscopies). The primary outcome was the PDR. It was registered with clinicaltrials.gov. (NCT047126265).

Results In this study, we randomized 150 patients. The AI system significantly increased the PDR (34.0% vs 38.7%, p < 0.001). In addition, AI-assisted colonoscopy increased the detection of polyps smaller than 6 mm (69 vs 91, p < 0.001), but no difference was found with regard to larger lesions.

Conclusions A real-time automatic polyp detection system can increase the PDR, primarily for diminutive polyps. However, a larger sample size is still needed in the follow-up study to further verify this conclusion.

Trial Registration clinicaltrials.gov Identifier: NCT047126265

Keywords Colonoscopy · Artificial intelligence · Computer-aided diagnose

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed malignancy and one of the leading causes of cancerrelated death.¹ Colonoscopy is the primary method for detecting and removing polyps. The detection and resection of tumor lesions by colonoscopy have been shown to be effective for the prevention of CRC.² There is evidence that for every 1.0% increase in the rate of adenoma detection, the risk of

Yuchen Luo, Yi Zhang and Ming Liu contributed equally to this paper.

Xiaobei Luo luoxiaobei63@126.com

Side Liu liuside2011@163.comZelong Han

hzl198886@163.com

¹ Department of Gastroenterology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China CRC decreases by 3.0%.³ However, colonoscopy is not perfect, and occasionally, interval cancer is detected in patients with a recent normal colonoscopy.⁴ Due to the characteristics of polyps and operators, polyps are prone to missed diagnosis, and the missed diagnosis rate can be as high as 27%.^{5,6} Two factors are considered to affect the missed diagnosis rate: blind spots and human error. The first factor can be addressed by using a wide-angle range or wide-angle remote attachment, but human error is difficult to overcome. A physician's workload may influence his/her level of performance. Even when the colonic mucosa is fully exposed, polyps may be missed because they are small or flat or because the color difference between the polyp and the normal mucosa is very small. In a 2006 systematic review and meta-analysis, six tandem colonoscopies showed a cumulative missed diagnosis rate of 22% for all polyps. This condition varied according to the size of the adenoma. The total missed diagnosis rate for adenomas \geq 10 mm was 2.1%, that of adenomas 5-10 mm was 13%, and that of adenomas 1-5 mm was 26%.8 These findings emphasize that small polyps are prone to missed diagnosis regardless

of experience. Furthermore, some studies have shown that the rate of polyp detection can be improved with the help of a second observer.^{9,10} Ideally, a real-time automatic polyp detection system with a performance similar to that of an expert endoscopist can be applied to help endoscopists detect polyp lesions; additionally, artificial intelligence (AI) has attracted attention in this area.¹¹ To further reduce the missed diagnosis of colonic polyps, further technical progress is needed to optimize the detection and endoscopic evaluation of colonic polyps. Computer-aided diagnosis (CAD) can take advantage of progress in the field of AI, especially deep learning technology, and play an auxiliary role in colonoscopy, thus providing a promising solution for human performance changes. Deep learning patterns depend on artificial neural networks, which are inspired by the concept of the network of neurons and synapses in the human brain. For image analysis, the best results so far have been obtained using a model based on convolutional neural networks (CNNs), which consist of several simple computing nodes and complex connections to simulate the human visual cortex.¹² CAD is a system that encompasses the ability of a computer to learn and perform specific tasks. Several automatic polyp detection systems have been developed in the past decade^{13,14}; however, there is little evidence demonstrating the ability of this technique to locate and track polyps in real time during colonoscopy in clinical practice.

The purpose of this study was to explore whether AIassisted colonoscopy could improve the polyp detection rate (PDR) in the actual clinical environment.

Methods

Patients

This study was conducted at the Endoscopy Center of Nanfang Hospital, China. Consecutive patients who underwent colonoscopy between April 2019 and September 2019 were eligible for enrollment. The specific inclusion criteria were as follows: (I) Chinese patients between the ages of 18 and 70 years; (II) voluntary signature of an informed consent form; and (III) no colonoscopy or anesthesia-related contraindications. The exclusion criteria were as follows: (I) history of inflammatory bowel disease (IBD); (II) history of colorectal surgery; (III) previously failed colonoscopy; (IV) polyposis syndrome; and (V) highly suspected CRC. In the course of the study, the following circumstances allowed patients to withdraw: (I) inadequate intestinal preparation; (II) withdrawal of informed consent by the patient; and (III) obvious adverse events that interrupted or affected the continuation of the examination. Basic demographic characteristics, including age, sex, and the Boston Bowel Preparation Scale (BBPS) score, were recorded.

Study Design

This was a prospective cohort trial. The conventional colonoscopy group was the control group, and the AI-assisted colonoscopy group was the research group. All included patients were required to sign an informed consent form before the screening. Routine bowel preparation consisted of 4 L of

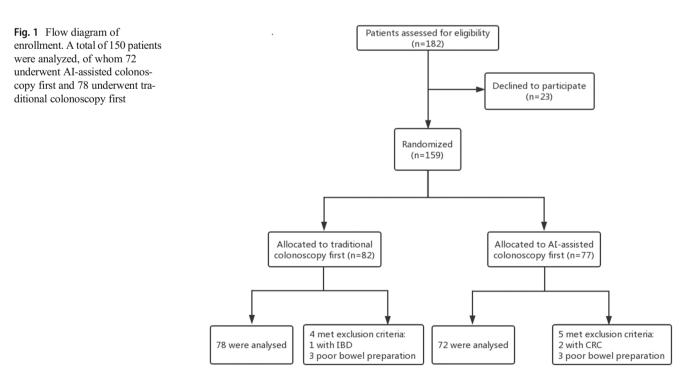




Fig. 2 Host device of the real-time polyp detection system

polyethylene glycol given in split doses. Colonoscopies were performed with high-definition colonoscopes (Olympus 290, CF-HQ290ZI) and high-definition monitors. The real-time automatic polyp detection system adopted was obtained from Xiamen Innovision Co., Ltd. (Fig. 2). All procedures were carried out by two experienced endoscopists, each having performed more than 3000 standard colonoscopies. The patients were assigned, by digital random number generators, to undergo back-to-back tandem colonoscopies with either conventional colonoscopy or AI-assisted colonoscopy first, followed immediately by the other procedure performed by the other endoscopist. Two endoscopists performed each colonoscopy independently, and the patient's final colonoscopy report was combined with the endoscopy results of the two doctors. Patients in the control group underwent routine colonoscopy (Figs. 1 and 2). In the research group, the real-time automatic polyp detection system was used for endoscopic assistance.

The system, turned on during withdrawal only, was connected to the endoscope generator and synchronously captured the video stream. The system displayed the position of the detected polyp with a hollow blue tracking box on the adjacent monitor (Fig. 3). Polyps detected during the first procedure were left in situ to be removed at the end of the second procedure. The research was registered with clinicaltrials.gov (NCT047126265).

The primary outcome was the PDR. The secondary outcomes were the number of polyps detected, the number of diminutive polyps (diameter < 6 mm), the number of polyps of each Paris type detected, and the number of false-positive results.

Statistical Analysis

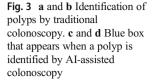
We prospectively designed this study to allow 90% power or more to detect a 20% difference (45% vs 25%) between the colonoscopy procedures with a two-group χ^2 test and a twosided α level of 0.05. A total of 118 samples were required based on the paired χ^2 difference test. A sample size of 118 participants was needed, and the overall participant enrollment goal was set at 157 to allow for potential exclusions or dropouts.

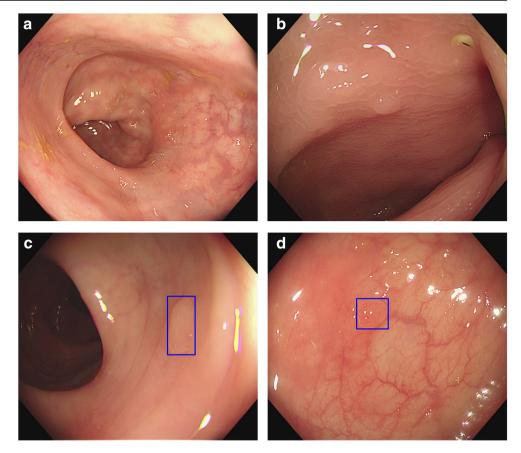
Measurement data are described as the mean and standard deviation. Count data are described as the number and percentage of patients. Comparisons of baseline and demographic characteristics between the research group and the control group were performed using the paired χ^2 test (McNemar's test) or Fisher's exact test for categorical variables and using the paired *t* test for continuous variables. A *p* value less than or equal to 0.05 was considered to indicate a statistically significant difference between the experimental and control groups. The observational indexes of the two groups were statistically analyzed by SPSS 23.0 statistical software.

The AI-Assisted System

The AI-assisted system (Xiamen Innovision Co., Ltd.) was developed by employing a CNN algorithm, specifically, a YOLO network architecture for object detection (Fig. 4). The model was trained on 112,199 colonoscopy images (including 64,134 images with 69,716 polyps and 48,065 images without polyps). All the images were recorded in the white light model and labeled by colonoscopists with more than 5 years of experience.

To enable the trained model for real-time detection, we captured each frame of the colonoscopy video during a colonoscopy and sequentially sent the frames to a PC. The model was then called to analyze the colonoscopy frame by frame in the PC and display the results on the adjacent monitor. The hardware configuration of the PC contains an NVIDIA GeForce GTX 2080 Ti graphics card, an i7-6600 CPU, and an 8-GB memory. The whole system, including the software and hardware, can process at least





50 frames per second with a latency of 33.20 ± 10.13 ms in real-time video analysis (Fig. 5).

Results

Clinical Characteristics

A total of 150 people (76 males and 74 females) met the inclusion criteria and were therefore considered for enrollment in the study. Baseline characteristics are presented in Table 1. The average age of the patients was 41.3 years, with a maximum age of 68 years and a minimum age of 18 years. Each patient was evaluated for bowel preparation at the first withdrawal. According to the BBPS score, 84% of the patients

scored greater than or equal to 6, and 16% of the patients scored less than 6. In these patients, a good examination could still be performed after water flushing and full suction, so they were not excluded. The average withdrawal time was 370.15 \pm 31.44 s in the traditional enteroscopy group and 373.17 \pm 33.37 s in the AI-assisted enteroscopy group (p = 0.102). Overall, there were no statistically significant differences between the two groups in terms of demographic data.

Polyp Characteristics

There was no significant difference in the location, size, or classification of polyps between the two groups (Table 2). The number of polyps detected in the control group and the research group was 80 and 105, respectively (p = 0.020). The

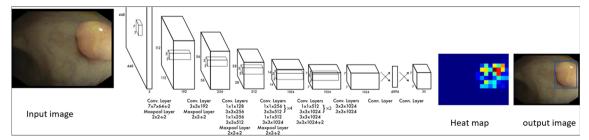
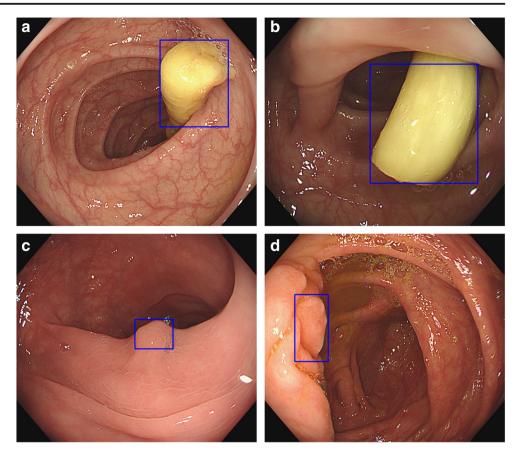


Fig. 4 Artificial intelligence (AI) system. The detection algorithm is a deep convolutional neural network (CNN) based on the YOLO network architecture

Fig. 5 a and b Feces mistakenly identified by AI. c Mucosal fold mistakenly identified by AI. d Ileocecal lobe misrecognized by AI



PDRs of the control group and the research group were 34.0%and 38.7%, respectively (p < 0.001) (Table 3). Significant differences were observed in the number of polyps and the PDR between the two groups. Furthermore, we classified statistics and compared them according to the size and the Paris classification of the polyps. According to the size of the polyps, 69 and 91 diminutive polyps were found in the control group and in the research group, respectively. The positive rates of diminutive polyp detection in the control group and the research group were 30.0% and 34.7%, respectively (p < 0.001). In contrast, there was no significant difference in the number of polyps larger than 6 mm (11 vs 14, p = 0.319). By comparing polyp morphology according to the Paris classification system, we observed a significant difference in the number of 0-IIa polyps between the two groups (61 vs 87, p = 0.010). AI-assisted colonoscopy is more sensitive to the detection of Paris type 0-IIa polyps.

False Positives with the Automatic Polyp Detection System

There was a total of 52 false positives in the AI-assisted colonoscopy group, averaging 0.35 false positives per colonoscopy (Table 4), mostly due to feces, mucosal folds, and so on.

 Table 1 Baseline characteristics

 of patients

Patient characteristic	Traditional colonoscopy $(n = 150)$	AI-assisted colonoscopy $(n = 150)$	p value
Withdrawal time, seconds, mean ± SD	370.15 ± 31.44	373.17 ± 33.37	0.102
Age, years, mean \pm SD (range)	41.3 ± 10.4 (18-68)	$41.3 \pm 10.4 \ (18-68)$	1.000
Sex			
Male, <i>n</i> (%)	76 (51)	76 (51)	1.000
Female, n (%)	74 (49)	74 (49)	1.000
Bowel preparation score			
< 6, <i>n</i> (%)	24 (16)	24 (16)	1.000
\geq 6, <i>n</i> (%)	126 (84)	126 (84)	1.000

Table 2 Polyp characteristics

Variable	Traditional colonoscopy ($n = 80$)	AI-assisted colonoscopy ($n = 105$)	p value
Location of polyp			0.764
Left colon, n (%)	44 (55.00)	57 (54.29)	0.923
Transverse colon, n (%)	22 (27.50)	33 (31.43)	0.562
Right colon, n (%)	14 (17.50)	15 (14.28)	0.551
Polyp size			0.935
<6 mm, <i>n</i> (%)	69 (86.25)	91 (86.67)	
$\geq 6 \text{ mm}, n (\%)$	11 (13.75)	14 (13.33)	
Polyp type ^a			0.352
0-IIa, <i>n</i> (%)	61 (76.25)	87 (82.86)	0.266
0-Is, n (%)	8 (10.00)	5 (4.76)	0.167

^a Paris classification

11 (13.75)

0-Ip, n (%)

Discussion

This was a prospective cohort study conducted to investigate the advantages of AI-assisted colonoscopy. Compared with the study by Wang et al.,¹⁵ our study ruled out the interference caused by intestinal differences between patients. Our study compared the number and positive detection rate of colorectal polyps between traditional colonoscopy and AI-assisted colonoscopy by observing withdrawal by different doctors in the same patient. Furthermore, because the order of the two procedures was completely random and the two operators were experienced endoscopists with the same level of seniority, the influence of the order of the procedure and the experience level of the operators on the results were minimized.

In this study, we found that AI-assisted colonoscopy could significantly increase the number of polyps detected compared with traditional colonoscopy, and the PDR of AI-assisted colonoscopy was also significantly higher than that of traditional colonoscopy. Moreover, we found that AI-assisted colonoscopy could significantly improve the detection of diminutive

polyps (diameter < 6 mm). This may be because small polyps			
in the field of view are more likely to be missed by			
endoscopists, while larger polyps are more likely to be detected.			
Although smaller polyps have a lower risk of malignancy than			
larger polyps, an increase in the overall PDR may ultimately			
help reduce the risk of interval CRC. Further research can also			
focus on the role of CAD in decreasing the risk of interval			
CRC. ³ However, the improved detection of diminutive polyps			
may lead to additional unnecessary polypectomies and add to			
the workload. In our study, there was no significant difference			
in the average withdrawal time between the two groups. It			
shows that the detection of more diminutive polyps by AI-			
assisted colonoscopy will not significantly increase the with-			
drawal time but also shows the real-time performance of CAD.			
Another study demonstrated that real-time CAD methods can			
achieve the performance level required for a diagnosis-and-			
leave strategy for diminutive, non-neoplastic rectosigmoid			
polyps, ¹⁶ which could potentially reduce costs.			

13 (12.38)

At the same time, we also classified and statistically analyzed the polyps according to the Paris classification system.

Variable	Traditional colonoscopy ($n = 150$)	AI-assisted colonoscopy ($n = 150$)	p value
Patients with at least one polyp, <i>n</i> (PDR)	51 (34.0%)	58 (38.7%)	< 0.001
Patients with at least one diminutive polyp, n (%)	45 (30.0%)	52 (34.7%)	< 0.001
Patients with at least one Paris type 0-IIa polyp, n (%)	39 (26.0%)	48 (32.0%)	< 0.001
Polyps detected, n	80	105	0.020
Polyps detected, by dimension			
< 6 mm, <i>n</i>	69	91	< 0.001
\geq 6 mm, <i>n</i>	11	14	0.319
Polyps detected, by type ^a			
0-IIa, <i>n</i>	61	87	0.010
0-Is, <i>n</i>	8	5	0.181
0-Ip, <i>n</i>	11	13	0.319

^a Paris classification

0.784

Table 4 Statistical results of false positives

	AI-assisted colonoscopy, n (%)
False positive	52 (100.00)
Submucosal tumor	4 (7.69)
Cyst	1 (1.92)
Feces	29 (55.77)
Ulcer	4 (7.69)
Mucosal fold	10 (19.23)
Other (bubble, circular blood vessel, and so on)	4 (7.69)

Compared with traditional colonoscopy, AI-assisted colonoscopy showed a higher detection rate for 0-IIa polyps. This could be because most of the polyps were type 0-IIa and because the sample size was not large enough.

We found that the real-time polyp detection system yielded a total of 52 false positives, most of which were due to feces and mucosal folds. There were approximately 0.35 instances per colonoscopy. This may suggest that the AI system is highly sensitive. Moreover, the recognition of color, shape, and property by AI still needs to be improved. The AI-assisted colonoscopy group did not demonstrate a longer withdrawal time. These false alarm lesions could be eliminated by the identification of endoscopists.

Some limitations to this study need to be mentioned. First, this was a single-center study with a small sample size. Second, AI has different effects on improving the PDR among different doctors. Third, the trial did not compare the adenoma detection rate (ADR) between the two groups. In the future, we will conduct a trial to improve the limitations listed above and to explore the effect of AI-assisted colonoscopy on the PDR using doctors with different seniorities and with different ADRs. Improvements based on the above aspects will improve the AI algorithm and its application in the field of medicine.

In conclusion, this study shows that an AI system based on deep learning and its real-time performance led to significant increases in colorectal PDR. The study is the early stages of AI for polyp detection that may be clinically relevant and is justification for further investigation.

Acknowledgments We want to express our gratitude to the Guangdong Gastrointestinal Disease Research Center (no. 2017B020209003).

Authors' Contributions Yuchen Luo, Yi Zhang, and Ming Liu: conception and design; analysis and interpretation of the data; drafting of the article; critical revision of the article for important intellectual content; final approval of the article.

Panpan Liu, Yihong Lai, Tongyin Xing, Zhen Wang, Ying Huang, Yue Li, Aiming Li, and Yadong Wang: acquisition of data; drafting of the article; final approval of the article. Xiaobei Luo and Side Liu: analysis and interpretation of the data; critical revision of the article for important intellectual content; final approval of the article.

Zelong Han: conception and design; critical revision of the article for important intellectual content; final approval of the article

Funding This project is supported by the "Guangdong Gastrointestinal Disease Research Center" (No.2017B020209003).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflicts of interest.

Ethical Statement This trial was approved by the Nanfang Hospital Institutional Review Board (internal reference number: NFEC-2019-160) and was registered with clinicaltrials.gov (NCT047126265). All participants will be enrolled only after providing written informed consent. No clinical data or bio-samples will be collected without the Participants' consent. The trial is conducted per national laws, Good Clinical Practice guidelines, and the Declaration of Helsinki as revised in 2013. The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethical Approval The study protocol was approved by the Institutional Review Board of Nanfang Hospital, Southern Medical University (NO. NFEC-2019-160).

Informed Consent All the informed consent has been obtained from the study patients.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- International Agency for Research on Cancer. Globocan 2018: Cancer Fact Sheets — Colorectal Cancer. IARC http://gco.iarc.fr/ today/data/factsheets/cancers/10_8_9-Colorectum-fact-sheet.pdf (2018)
- Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. N Engl J Med 2000;343:162-8.
- Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;370: 1298-306.

- Bressler B, Paszat LF, Chen Z, Rothwell DM, Vinden C, Rabeneck L. Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis. Gastroenterology 2007;132:96-102.
- Ahn SB, Han DS, Bae JH, Byun TJ, Kim JP, Eun CS. The Miss Rate for Colorectal Adenoma Determined by Quality-Adjusted, Back-to-Back Colonoscopies. Gut Liver 2012;6:64-70.
- Mahmud N, Cohen J, Tsourides K, Berzin TM. Computer vision and augmented reality in gastrointestinal endoscopy. Gastroenterol Rep (Oxf) 2015;3:179-84.
- Zheng B, Rieder E, Cassera MA, Martinec DV, Lee G, Panton ON, Park A, Swanstrom LL. Quantifying mental workloads of surgeons performing natural orifice transluminal endoscopic surgery (NOTES) procedures. Surg Endosc 2012;26:1352-8.
- van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol 2006;101:343-50.
- Aslanian HR, Shieh FK, Chan FW, Ciarleglio MM, Deng Y, Rogart JN, Jamidar PA, Siddiqui UD. Nurse observation during colonoscopy increases polyp detection: a randomized prospective study. Am J Gastroenterol 2013;108:166-72.
- Lee CK, Park DI, Lee SH, Hwangbo Y, Eun CS, Han DS, Cha JM, Lee BI, Shin JE. Participation by experienced endoscopy nurses increases the detection rate of colon polyps during a screening colonoscopy: a multicenter, prospective, randomized study. Gastrointest Endosc 2011;74:1094-102.
- Byrne MF, Shahidi N, Rex DK. Will computer-aided detection and diagnosis revolutionize colonoscopy? Gastroenterology 2017;153: 1460-1464. e1.

- Shin H, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Transactions on Medical Imaging 2016;35:1285-1298.
- Tajbakhsh N, Gurudu SR, Liang J. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information. IEEE Trans Med Imaging 2016;35:630-44.
- 14. Misawa M, Kudo SE, Mori Y, Cho T, Kataoka S, Yamauchi A, Ogawa Y, Maeda Y, Takeda K, Ichimasa K, Nakamura H, Yagawa Y, Toyoshima N, Ogata N, Kudo T, Hisayuki T, Hayashi T, Wakamura K, Baba T, Ishida F, Itoh H, Roth H, Oda M, Mori K. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. Gastroenterology 2018;154:2027-2029. e3.
- Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019:gutjnl-2018-317500.
- 16. Mori Y, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. Ann Intern Med 2018;169:357-366.

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