Artificial Intelligence–Assisted Colonoscopy for Colorectal Cancer Screening: A Multicenter Randomized Controlled Trial

Q1 Q8

Q9 Hong Xu,[1] Raymond S. Y. Tang,[2,3] Thomas Y. T. Lam,[3,4,5] Guijun Zhao,[6] James Y. W. Lau,[3,7] Yunpeng Liu,[8] Qi Wu,[9] Long Rong,[10] Weiran Xu,[1] Xue Li,[6] Sunny H. Wong,[2,3,12] Shuntian Cai,[8] Jing Wang,[9] Guanyi Liu,[10] Tantan Ma,[1] Xiong Liang,[6] Joyce W. Y. Mak,[2,3] Hongzhi Xu,[8] Peng Yuan,[9] Tingting Cao,[1] Fudong Li,[1] Zhenshi Ye,[8] Zhang Shutian,[11] and Joseph J. Y. Sung[2,3,5,12]

[1]Department of Gastroenterology and Endoscopy Center, First Hospital of Jilin University, Jilin, China; [2]Department of Medicine and Therapeutics, Chinese University of Hong Kong, Hong Kong SAR, China; [3]Institute of Digestive Disease, Chinese University of Hong Kong, Hong Kong SAR, China; [4]JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong SAR, China; [5]Stanley Ho Big Data Decision Analytics Research Centre, Chinese University of Hong Kong, Hong Kong SAR, China; [6]Department of Endoscopy Center, Inner Mongolia Key Laboratory of Endoscopic Digestive Diseases, Inner Mongolia People's Hospital, Hohhot, China; [7]Department of Surgery, Chinese University of Hong Kong, Hong Kong SAR, China; [8]Department of Gastroenterology, Zhongshan Hospital Affiliated to Xiamen University, Xiamen, China; [9]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Endoscopy Center, Peking University Cancer Hospital and Institute, Beijing, China; [10]Endoscopy Center, Peking University First Hospital, Beijing, China; [11]Department of Gastroenterology and Hepatology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Diseases, Beijing, China; and [12]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

BACKGROUND AND AIMS: Artificial intelligence (AI)–assisted colonoscopy improves polyp detection and characterization in colonoscopy. However, data from large-scale multicenter randomized controlled trials (RCT) in an asymptomatic population are lacking.

METHODS: This multicenter RCT aimed to compare AI-assisted colonoscopy with conventional colonoscopy for adenoma detection in an asymptomatic population. Asymptomatic subjects 45–75 years of age undergoing colorectal cancer screening by direct colonoscopy or fecal immunochemical test were recruited in 6 referral centers in Hong Kong, Jilin, Inner Mongolia, Xiamen, and Beijing. In the AI-assisted colonoscopy, an AI polyp detection system (Eagle-Eye) with real-time notification on the same monitor of the endoscopy system was used. The primary outcome was overall adenoma detection rate (ADR). Secondary outcomes were mean number of adenomas per colonoscopy, ADR according to endoscopist's experience, and colonoscopy withdrawal time. This study received Institutional Review Board approval (CRE-2019.393).

RESULTS: From November 2019 to August 2021, 3059 subjects were randomized to AI-assisted colonoscopy (n = 1519) and conventional colonoscopy (n = 1540). Baseline characteristics and bowel preparation quality between the 2 groups were similar. The overall ADR (39.9% vs 32.4%; $P < .001$), advanced ADR (6.6% vs 4.9%; $P = .041$), ADR of expert (42.3% vs 32.8%; $P < .001$) and nonexpert endoscopists (37.5% vs 32.1%; $P = .023$), and adenomas per colonoscopy (0.59 ± 0.97 vs 0.45 ± 0.81; $P < .001$) were all significantly higher in the AI-assisted colonoscopy. The median withdrawal time (8.3 minutes vs 7.8 minutes; $P = .004$) was slightly longer in the AI-assisted colonoscopy group.

CONCLUSIONS: In this multicenter RCT in asymptomatic patients, AI-assisted colonoscopy improved overall ADR, advanced ADR, and ADR of both expert and nonexpert attending endoscopists. (ClinicalTrials.gov, Number: NCT04422548).

Keywords: Artificial Intelligence; Colonoscopy; Colorectal Cancer Screening.

Abbreviations used in this paper: **ADR, adenoma detection rate; AI, artificial intelligence; AMR, adenoma miss rate; aOR, adjusted odds ratio; APC, adenoma per colonoscopy; BBPS, Boston Bowel Preparation Scale; BMI, body mass index; CC, conventional colonoscopy; CI, confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; ITT, intention to treat; IQR, interquartile range; PDR, polyp detection rate; PP, per protocol; SSL, sessile serrated lesion.**

1542-3565
https://doi.org/10.1016/j.cgh.2022.07.006

While screening colonoscopy with polypectomy has been shown to reduce colorectal cancer (CRC) incidence and mortality, missed colorectal neoplasia remains a significant problem and is associated with postcolonoscopy CRC development.[1,2] In a recent meta-analysis including more than 15,000 tandem colonoscopies, missed rates for adenomas and advanced adenomas were up to 26% and 9%, respectively.[3] Small and flat lesions, inadequate bowel preparation, and variable endoscopist performance in polyp detection are important factors contributing to missed colorectal neoplasia.[4,5]

Because a high adenoma detection rate (ADR) is associated with a lower risk of postcolonoscopy CRC, significant efforts have been devoted to improve ADR.[4,6] Various colonoscopy-based strategies, including image enhanced endoscopy, mucosal fold flattening devices, colonoscopes with a wide field of view, and second right colon examination by retroflexion or forward view have been used and reported to improve polyp detection rate or ADR.[7–10] While some of the previous strategies can improve detection of lesions hidden in blind spots, their impact on polyp detection failure by endoscopists remains uncertain.

In recent years, artificial intelligence (AI)–assisted colonoscopy has gained significant clinical interest because it is expected to reduce adenoma miss rate by overcoming human error in polyp detection.[11] Multiple AI systems using deep learning and convolutional neural network with visual or audio notices for endoscopists have been developed and evaluated in clinical setting for real-time polyp detection or monitoring of colonoscopy withdrawal quality (eg, withdrawal speed and time, bowel cleansing quality).[12–17] While AI-assisted colonoscopy was reported to have a higher ADR and adenomas per colonoscopy (APC) in RCTs and a lower adenoma miss rate (AMR) in tandem studies when compared with conventional colonoscopy (CC), none of these studies targeted a purely screening population. In addition, the benefit of AI-assisted colonoscopy for endoscopists of different level of experience remains uncertain. Because data from large scale multicenter RCT assessing the efficacy of real-time AI assistance for adenoma detection in an asymptomatic population eligible for CRC screening are lacking, we performed a multicenter RCT comparing AI-assisted colonoscopy with standard colonoscopy for adenoma detection in an asymptomatic population.

What You Need to Know

Background

Artificial intelligence (AI)–assisted colonoscopy has been reported to improve the adenoma detection rate (ADR) in symptomatic and asymptomatic patients undergoing diagnostic and screening colonoscopy. However, there are no large randomized controlled trials assessing the efficacy of AI-assisted colonoscopy in an asymptomatic population eligible for colorectal cancer screening, and involving endoscopists with different levels of experience.

Findings

In this multicenter randomized controlled trial involving an asymptomatic population of 3059 patients and 24 endoscopists, AI-assisted colonoscopy significantly increased the overall ADR by 7.5% when compared with conventional colonoscopy (39.9% vs 32.4%; $P < .001$). ADR for advanced adenomas, adenomas $\geq$10 mm, nonpedunculated adenomas, and both expert and nonexpert attending endoscopists were also significantly improved in the AI-assisted colonoscopy group.

Implications for patient care

Routine use of AI assistance for real-time adenoma detection in colonoscopy should be considered.

Materials and Methods

Study Design

This was a prospective, multicenter, single-blinded, parallel RCT conducted in 6 university-affiliated endoscopy centers in Hong Kong, Beijing (2 centers), Inner Mongolia, Jilin, and Xiamen, China. The study was conducted in accordance with the Principles of Good Clinical Practice and the Declaration of Helsinki. The study protocol was approved by the Joint Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee (CRE-2019.393) and the institutional review boards at all participating sites, and was registered at ClinicalTrials.gov (NCT04422548). All authors had access to the study data and reviewed and approved the final manuscript.

Study Population and AI Polyp Detection System

Asymptomatic subjects between 45 and 75 years of age undergoing CRC screening by direct screening colonoscopy or fecal immunochemical test (FIT)–based screening program were invited to participate between November 2019 and August 2021. In China, different strategies for CRC screening exist across the country. Some provinces and cities allow direct colonoscopy, whereas others require FIT prior to colonoscopy. In this study, all 3059 individuals included were asymptomatic subjects eligible for CRC screening. Among the 6 participating centers, 2 offered only FIT-based screening programs, and did not offer direct screening colonoscopy to the general public. In order for these centers to capture the average-risk population for CRC screening, asymptomatic subjects have to go through the government-subsidized FIT-based screening program first. For enrollment into this study, all asymptomatic subjects were offered colonoscopy irrespective of their FIT results. Written informed consent was obtained from all enrolled patients. Patients were excluded from the study if they were unable to provide informed consent, had contraindications for endoscopy, overt symptoms suggestive of colorectal disorders, history of inflammatory bowel disease, CRC, polyposis syndrome, prior colorectal surgery, or colonoscopy within 10 years. Intention-to-treat (ITT) analysis included all recruited patients. Per-protocol (PP) analysis included only those who have successful cecal intubation and adequate bowel preparation (defined by Boston Bowel Preparation Scale (BBPS) score $\geq$6 and all segmental BBPS score $\geq$2).[18] The artificial intelligence system for polyp detection used in this study is detailed in the Supplementary Methods (Artificial Intelligence Polyp Detection System Section).

Randomization and Blinding

Recruited asymptomatic subjects were randomized in a 1:1 ratio to undergo either AI-assisted colonoscopy or CC. T.Y.T.L generated the randomization sequence in a concealed allocation fashion in block sizes of 10 for the participating centers. Subjects enrollment and group assignment were done independently by the study team members of each participating center. Randomization was stratified by endoscopist's colonoscopy experience (nonexpert attending vs expert attending). Group assignments were contained in sealed, opaque envelopes. This is a single-blinded randomization with enrolled subjects being blinded to the result of their randomization while endoscopists were not blinded to the group assignment. Colonic polyp specimens were evaluated by pathologists who were also blinded to the study group allocation.

Colonoscopy Procedures and Data Collection

Study colonoscopies were performed by 12 nonexpert attending (colonoscopy experience <5000) and 12 expert attending endoscopists (colonoscopy experience $\geq$5000), with 2 nonexpert and 2 expert attending endoscopists from each of the 6 centers. All procedures were performed with high-definition colonoscopes, endoscopy processors, and monitors (EVIS Lucera Elite, 290 series [Olympus, Tokyo, Japan] or 4450 series [Fujifilm, Tokyo, Japan]). In order to standardize the endoscopic procedures, electronic image enhancing function was not allowed for polyp detection but could be used for polyp characterization at the discretion of the endoscopists. Mucosal exposure devices were not allowed, and antispasmodic agents were not used during the procedures.

All enrolled patients received a low-fiber diet 3 days prior to colonoscopy and underwent bowel cleansing with polyethylene glycol solution in split dose based on institutional protocol. Baseline demographics and clinical history relevant to CRC were obtained from patients before colonoscopy. Colonoscopies were performed under conscious sedation. Bowel preparation quality was rated by the BBPS, with adequate bowel preparation being defined as BBPS score $\geq$6 and any segmental BBPS score $\geq$2.

For patients randomized to the AI-assisted colonoscopy group, the colonoscopy procedure was the same as that of the CC group except that the AI system was turned on for real-time polyp detection during both the intubation and withdrawal phases of the procedure. Lesions detected by the AI system were marked by a blue tracking box and were shown on the same high definition monitor of the endoscopy system (Figure 1). After confirmation by the endoscopist, polyps detected were removed. In both arms of the study, while polyp detection and removal were intended to be performed during withdrawal phase per usual practice, removal of polyp with size <10 mm during intubation phase immediately upon detection was allowed to avoid missing the smaller lesion (already detected on intubation) for polypectomy during scope withdrawal. Data collection of procedure time and characteristics of polyps resected are described in the Supplementary Methods.

Outcome Measures

The primary outcome was overall ADR, which was defined as the proportion of patients with at least 1 adenoma detected during colonoscopy. Secondary outcomes were the mean number of APC, ADR according to the experience of endoscopists, adenoma size, morphology, histology, colonoscopy intubation time, and withdrawal time. Advanced adenomas were defined as adenomas with size $\geq$10 mm, villous component, or high grade dysplasia. Tubular adenomas, tubulovillous adenomas, villous adenomas, and CRC were included in the ADR calculation.

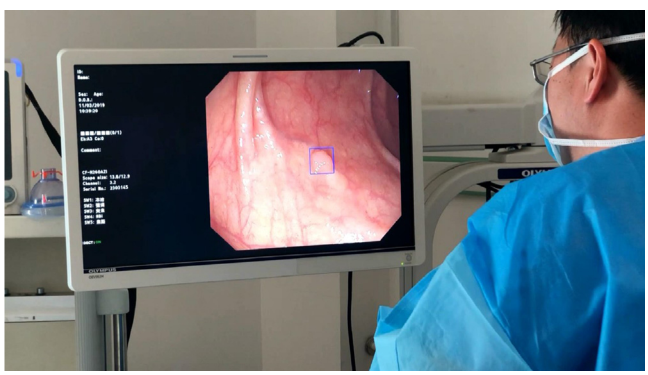


Figure 1. The AI polyp detection system.

Sample Size Estimation and Statistical Analysis

An overall ADR of 37% was reported in subjects who received direct screening colonoscopy in a recently published international multicenter study conducted in the Asia-Pacific region.[19] Assuming that AI-assisted colonoscopy would yield a 5% increase in overall ADR, at least 1497 subjects per study group (ie, at least 2994 subjects in total) would be required to achieve a statistical power of $\geq$80% to show superiority of AI-assisted colonoscopy against CC. Type I error of $\leq$5% is 2-sided in this study and is powered for superiority. Statistical analyses are described in the Supplementary Methods (Statistical Analysis section).

Results

From November 2019 to August 2021, 3227 patients were recruited. A total of 115 patients were excluded from the study; 53 patients did not meet inclusion criteria and 62 declined to participate. A total of 3059 patients were enrolled and randomized, of whom 1519 patients and 1540 patients were assigned to the AI-assisted colonoscopy group and CC group, respectively, and included in the ITT analysis. After excluding subjects with incomplete colonoscopy and inadequate bowel preparation, 1238 subjects in the AI-assisted colonoscopy group and 1289 subjects in the CC group were included in the PP analysis (Figure 2). The baseline characteristics of patients were similar between the 2 groups. Among the 3059 patients, 1684 (55.1%) patients underwent direct screening colonoscopy. Among the 1375 patients who underwent FIT-based screening initially, 127 (9.2%) were FIT positive and 1248 (90.8%) were FIT negative. In order to capture asymptomatic patients who would otherwise be ineligible for direct screening colonoscopy in centers in which only government-run FIT-based screening program was available, both FIT-positive and FIT-negative patients were invited for colonoscopy and randomized to AI or CC groups (Table 1).

In ITT analysis, the cecal intubation rate was the same between the 2 groups (98.7% vs 98.7%; $P = .965$). The median overall BBPS score, the median segmental BBPS score of the left, transverse, and right colon, and the proportion of patients having adequate bowel preparation quality were similar in both groups (81.5% vs 83.7%; $P = .108$). There were no immediate adverse events during colonoscopy in this study. In ITT analysis, both the median intubation time (4.48 [interquartile range (IQR), 3.33–6.30] minutes vs 4.17 [IQR, 3.17–5.76] minutes; $P < .001$) and the median withdrawal time (8.25 [IQR, 6.67–12.00] minutes vs 7.78 [IQR, Q4

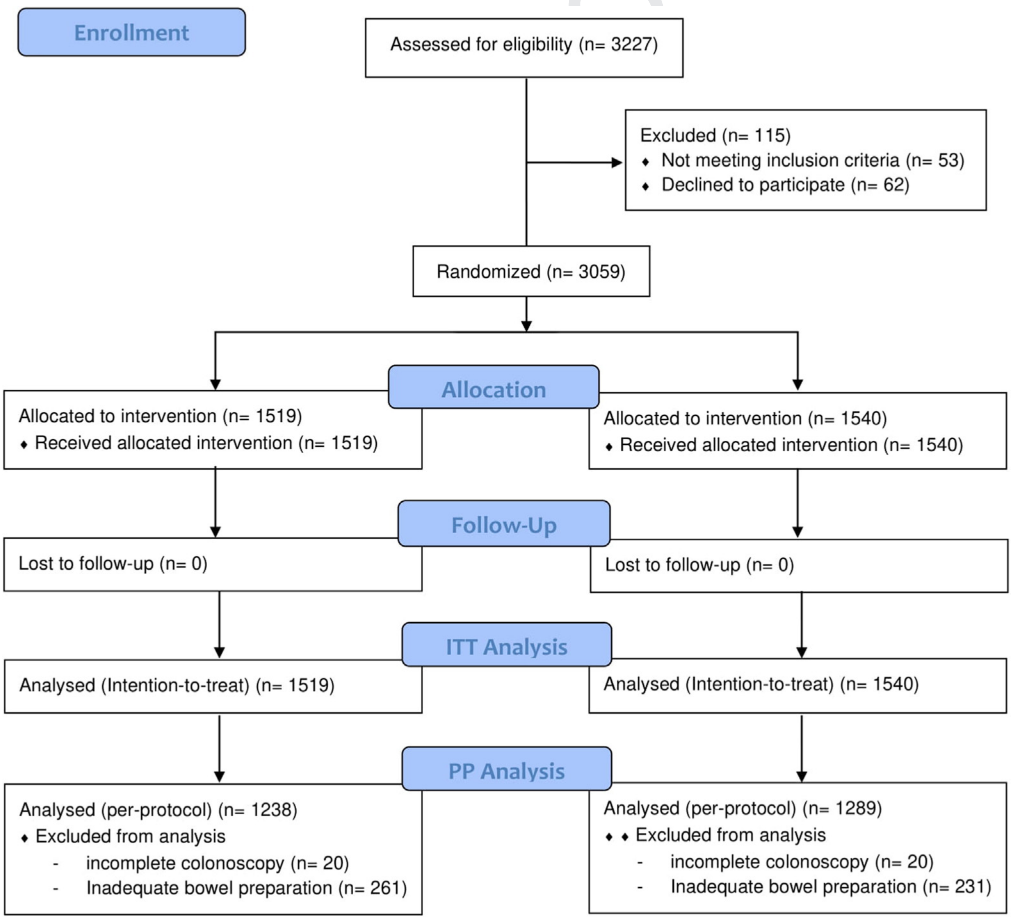


Figure 2. CONSORT flow diagram.

web 4C/FPO

6.45–11.00] minutes; $P = .004$) were significantly longer in the AI-assisted colonoscopy group when compared with CC group (Table 2). In PP analysis, significant prolongation of intubation time and withdrawal time of similar magnitude was also observed in the AI-assisted colonoscopy group (Table 2).

Adenoma Detection

In ITT analysis, the overall ADR of the AI-assisted colonoscopy group was significantly higher than that of the CC group (39.9% vs 32.4%; $P < .001$). The APC in the AI-assisted colonoscopy group was also significantly higher (0.59 ± 0.97 vs 0.45 ± 0.81; $P < .001$). There was also significant improvement in overall ADR (39.7% vs 33.5%; $P = .001$) and APC (0.65 ± 1.02 vs 0.52 ± 0.85; $P < .001$) in the AI-assisted colonoscopy group compared with the CC group in PP analysis (Table 2).

In subgroup analysis, AI-assisted colonoscopy improved ADR for both expert (42.3% vs 32.8%; $P < .001$) and nonexpert (37.5% vs 32.1%; $P = .023$) attending endoscopists in the ITT analysis. The PP analysis produced similar results, but the improvement of ADR of nonexpert attendings by AI did not achieve statistical significance (39.9% vs 36.5%; $P = .201$). Among patients with inadequate bowel preparation (BBPS <6 or any segmental BBPS <2), the use of AI also improved ADR (40.7% vs 29.5%; $P = .039$) when compared with CC (Table 2).

Characteristics of polyps detected and removed in both groups are listed in Table 3. Detection of nonadvanced adenoma in the AI-assisted colonoscopy group was significantly higher than in the CC group (nonadvanced ADR, 32.3% vs 26.7%; $P = .001$). Similarly, detection of advanced adenoma was also significantly higher in the AI-assisted colonoscopy group than the CC group (advanced ADR, 6.6% vs 4.9%; $P = .041$). On the other hand, CRC and sessile serrated lesion (SSL) detection rates were not different between 2 groups (0.9% vs 0.8%; $P = .819$; and 1.1% vs 1.3%; $P = .65$, respectively). In subgroup analysis according to size and morphology of polyps, AI-assisted colonoscopy also significantly increased detection of adenoma <5 mm (16.5% vs 11.5%; $P < .001$), adenoma ≥10 mm (6.5% vs 4.7%; $P = .033$), and nonpedunculated adenomas (27.6% vs 21.8%; $P < .001$) when compared with CC (Table 3). AI-assisted colonoscopy improved the detection of adenoma in both proximal and distal colon when compared with CC (28.4% vs 23.8%; $P = .004$; and 10.6% vs 7.7%; $P = .006$, respectively). The results show that there was an increased resection of nonadvanced adenoma (<5 mm) as well as advanced adenoma. There was also increased resection of nonpedunculated adenomas. No increase in hyperplastic polyp resection was shown.

Table 1. Patient Baseline Characteristics

	AI-Assisted Colonoscopy Group (n = 1519)	CC Group (n = 1540)	*P* Value
Age, y	57.49 ± 7.55	57.03 ± 7.43	.086
Male	707 (46.5)	728 (47.3)	.686
BMI, kg/m^2	23.91 ± 3.24	23.91 ± 3.19	.961
Family history of CRC	50 (3.3)	46 (3.0)	.629
Ever smoking	222 (14.6)	221 (14.4)	.836
Alcohol consumption	139 (9.2)	139 (9.0)	.904
Hypertension	168 (11.1)	189 (12.3)	.296
Diabetes	70 (4.6)	75 (4.9)	.733
Ischemic heart disease	38 (2.5)	30 (1.9)	.299
Stroke	6 (0.4)	6 (0.4)	.981
Liver cirrhosis	3 (0.2)	3 (0.2)	.987
Fatty liver	122 (8.0)	116 (7.5)	.606
GERD	34 (2.2)	34 (2.2)	.954
Asymptomatic subjects			
Direct screening colonoscopy	840 (55.3)	844 (54.8)	.758
Colonoscopy for FIT-positive subjects	59 (3.9)	68 (4.4)	
Colonoscopy for FIT-negative subjects	620 (40.8)	628 (40.8)	

Values are mean ± SD or n (%).
AI, artificial intelligence; BMI, body mass index; CC, conventional colonoscopy; CRC, colorectal cancer; FIT, fecal immunochemical test; GERD, gastroesophageal reflux disease.

Table 2. Findings and Outcomes in AI-Assisted Colonoscopy and CC

	Intention-to-Treat Analysis (n = 3059)			Per-Protocol Analysis (n = 2527)		
	AI-Assisted Colonoscopy Group (n = 1519)	CC Group (n = 1540)	*P* Value	AI-Assisted Colonoscopy Group (n = 1238)	CC Group (n = 1289)	*P* Value
Cecal intubation rate	1499 (98.7)	1520 (98.7)	.965	1238 (100.0)	1289 (100.0)	NA
BBPS	7 (6–8)	7 (6–8)	.023	7 (6–8)	7 (6–8)	.037
Left colon BBPS score	2 (2–3)	2 (2–3)	.416	2 (2–3)	2 (2–3)	.888
Transverse colon BBPS score	2 (2–3)	2 (2–3)	.032	2 (2–3)	2 (2–3)	.104
Right colon BBPS score	2 (2–3)	2 (2–3)	.008	2 (2–3)	2 (2–3)	.050
Adequate bowel preparation	1238 (81.5)	1289 (83.7)	.108	1238 (100.0)	1289 (100.0)	NA
Immediate adverse event	0	0	NA	0	0	NA
Intubation time, min	4.48 (3.33–6.30)	4.17 (3.17–5.76)	<.001	4.50 (3.38–6.33)	4.17 (3.21–5.83)	<.001
Withdrawal time, min	8.25 (6.67–12.00)	7.78 (6.45–11.00)	.004	8.33 (6.67–12.09)	7.83 (6.5–11.00)	.008
Overall ADR	606 (39.9)	499 (32.4)	<.001	492 (39.7)	432 (33.5)	.001
Number of adenomas per colonoscopy	0.59 ± 0.97	0.45 ± 0.81	<.001	0.63 ± 1.01	0.49 ± 0.83	<.001
ADR of expert attending endoscopist	319/754 (42.3)	236/720 (32.8)	<.001	258/594 (43.4)	197/585 (33.7)	.001
ADR of nonexpert attending endoscopist	287/765 (37.5)	263/820 (32.1)	.023	234/644 (36.3)	235/704 (33.4)	.255
ADR for direct screening colonoscopy	322/840 (38.3)	252/844 (29.8)	<.001	267/700 (38.1)	222/726 (30.6)	.003
ADR for colonoscopy after positive FIT	39/59 (66.1)	35/68 (51.5)	.095	37/55 (67.3)	35/65 (53.8)	.135
ADR for colonoscopy after negative FIT	245/620 (39.5)	212/628 (33.8)	.035	188/483 (38.9)	175/498 (35.1)	.220
ADR of patients with inadequate bowel preparation[a]	114/281 (40.6)	67/251 (26.7)	.001	NA	NA	NA

Values are n (%), median (interquartile range), mean ± SD, or n/n (%).
ADR, adenoma detection rate; AI, artificial intelligence; BBPS, Boston Bowel Prep Score; CC, conventional colonoscopy; FIT, fecal immunochemical test; NA, ●●●. Q6
[a]BBPS <6 or segmental BBPS <2.

Table 3. Characteristics of Polyps Resected

	Intention-to-Treat Analysis (n = 3059)			Per-Protocol Analysis (n = 2527)		
	AI-Assisted Colonoscopy Group (n = 1519)	CC Group (n = 1540)	*P* Value	AI-Assisted Colonoscopy Group (n = 1238)	CC Group (n = 1289)	*P* Value
Nonadvanced adenoma	492 (32.3)	411 (26.7)	<.001	407 (32.9)	359 (27.9)	.006
Advanced adenoma	100 (6.6)	75 (4.9)	.041	83 (6.7)	66 (5.1)	.091
Colorectal cancer	14 (0.9)	13 (0.8)	.819	2 (0.2)	8 (0.6)	.109
Sessile serrated lesion	17 (1.1)	20 (1.3)	.650	16 (1.3)	19 (1.5)	.696
Hyperplastic polyp	234 (15.2)	244 (15.8)	.627	186 (15)	201 (15.6)	.691
Other polyps[a]	22 (1.4)	22 (1.4)	.963	21 (1.7)	19 (1.5)	.655
Adenoma size						
<5 mm	250 (16.5)	177 (11.5)	<.001	216 (17.4)	168 (13)	.002
5–9mm	243 (16.0)	236 (15.3)	.609	204 (16.5)	210 (16.3)	.899
≥10 mm[b]	99 (6.5)	73 (4.7)	.033	85 (6.9)	65 (5.0)	.053
Pedunculated adenoma	173 (11.4)	150 (9.7)	.138	139 (11.2)	136 (10.6)	.585
Nonpedunculated adenoma[c]	419 (27.6)	336 (21.8)	<.001	365 (29.5)	307 (23.8)	.001
Adenoma in proximal colon[d]	431 (28.4)	367 (23.8)	.004	359 (29.0)	317 (24.6)	.012
Adenoma in distal colon[e]	161 (10.6)	119 (7.7)	.006	129 (10.4)	107 (8.3)	.067

Values are n (%).
AI, artificial intelligence; CC, conventional colonoscopy.
[a]Inflammatory polyps, mucosal polyps, and lymphoid polyps.
[b]A total of 59 (59.6%) of the 99 adenomas with size ≥ 10 mm were nonpedunculated lesions.
[c]Included flat, depressed, sessile adenomas.
[d]From transverse colon to cecum.
[e]From rectum to splenic flexure.

Univariate and Multivariate Regression Analysis

Regression analysis was performed to identify predictors of adenoma detection in the AI-assisted colonoscopy and CC groups in this study. AI assistance, gender, age of patients, lifestyle (smoking, drinking, BMI), comorbid condition (diabetes, ischemic heart disease, fatty liver, use of aspirin), experience of endoscopist, and withdrawal time of colonoscopes were included in univariate and multivariate analyses (Table 4). Based on ITT data, multivariate analysis showed that AI assistance (adjusted odds ratio [aOR], 1.302; 95% confidence interval [CI], 1.106–1.532; *P* = .001), male sex (aOR, 1.648; 95% CI, 1.394–1.948; *P* < .001), older age groups (60–64 years of age: aOR, 1.527; 95% CI, 1.156–2.017; *P* = .003, 65–69 years of age: aOR, 2.214; 95% CI, 1.653–2.964; *P* < .001, 70–75 years of age: aOR, 2.008; 95% CI, 1.393–2.894; *P* < .001), alcohol drinking (aOR, 1.399; 95% CI, 1.061–1.845; *P* = .017), body mass index (aOR, 1.026; 95% CI, 1.000–1.052; *P* = .049), expert endoscopist (aOR, 1.213; 95% CI, 1.030–1.428; *P* = .020), and withdrawal time (aOR, 1.179; 95% CI, 1.156–1.202; *P* < .001) were independently associated with higher ADR. Similarly, in PP analysis, AI assistance (aOR, 1.219; 95% CI, 1.019–1.458; *P* = .031), male sex (aOR, 1.599; 95% CI, 1.330–1.921; *P* < .001), elder age groups (aged 60–64: aOR, 1.521; 95% CI, 1.117–2.071; *P* = .008; 65–69 years of age: aOR, 2.281; 95% CI, 1.650–3.154; *P* < .001; 70–75 years of age: aOR, 1.775; 95% CI, 1.182–2.664; *P* = .006), drinking (aOR, 1.368; 95% CI, 1.011–1.851; *P* = .042), expert endoscopist (aOR, 1.279; 95% CI, 1.068–1.532; *P* = .007), and withdrawal time (aOR, 1.193; 95% CI, 1.167–1.220; *P* < .001) were independently associated with higher ADR (Table 4). On the other hand, positive FIT result was not found to be an independent predictor of higher ADR on multivariate regression analysis.

Discussion

This is the first RCT to investigate the benefit of AI-assisted colonoscopy in adenoma detection in a large, asymptomatic population. The study involved 6 centers and 24 attending endoscopists of 2 levels of experience in a single-blinded design. The findings confirm that AI-assisted colonoscopy significantly improves ADR for colonoscopy in asymptomatic subjects, for both advanced and nonadvanced adenomas, for adenomas <5 mm and ≥10 mm, for nonpedunculated adenoma, and in both the proximal and distal colon. There were more adenoma (advanced and nonadvanced, pedunculated and

Table 4. Univariate and Multivariate Regression Analysis of Predictors of Adenoma Detection

	Intention-to-Treat Analysis (n = 3059)						Per-Protocol Analysis (n = 2527)					
	Univariate Analysis			Multivariate Analysis			Univariate Analysis			Multivariate Analysis		
Variable	OR	95% CI	*P* Value	aOR	95% CI	*P* Value	OR	95% CI	*P* Value	aOR	95% CI	*P* Value
Study intervention—AI assistance[a]	1.385	1.194–1.606	<.001[b]	1.302	1.106–1.532	.001[b]	1.308	1.112–1.539	.001[b]	1.219	1.019–1.458	.031[b]
Male[a]	1.893	1.630–2.198	<.001[b]	1.648	1.394–1.948	<.001[b]	1.830	1.553–2.155	<.001[b]	1.599	1.330–1.921	<.001[b]
45–49 years of age		Reference			Reference			Reference			Reference	
50–54 years of age	1.071	0.832–1.378	.596	0.997	0.758–1.310	.982	1.063	0.805–1.403	.667	0.981	0.726–1.326	.901
55–59 years of age	1.349	1.064–1.710	.013[b]	1.265	0.978–1.637	.073	1.278	0.984–1.661	.066[b]	1.188	0.894–1.579	.235
60–64 years of age	1.662	1.287–2.145	<.001[b]	1.527	1.156–2.017	.003[b]	1.664	1.255–2.208	<.001[b]	1.521	1.117–2.071	.008[b]
65–69 years of age	2.392	1.830–3.127	<.001[b]	2.214	1.653–2.964	<.001[b]	2.517	1.871–3.386	<.001[b]	2.281	1.650–3.154	<.001[b]
70–75 years of age	2.287	1.641–3.187	<.001[b]	2.008	1.393–2.894	<.001[b]	2.286	1.584–3.299	<.001[b]	1.775	1.182–2.664	.006[b]
Smoker[a]	1.474	1.202–1.808	<.001				1.528	1.221–1.911	<.001			
Alcohol drinker[a]	1.576	1.229–2.020	<.001[b]	1.399	1.061–1.845	.017[b]	1.538	1.174–2.016	.002[b]	1.368	1.011–1.851	.042[b]
Aspirin user	1.257	0.815–1.938	.301				1.176	0.729–1.896	.507			
Diabetes	1.119	0.794–1.576	.521				1.156	0.793–1.687	.451			
Ischemic heart disease	1.169	0.715–1.910	.534				0.963	0.554–1.674	.894			
Fatty liver	0.905	0.685–1.197	.485				1.001	0.744–1.346	.995			
BMI[a]	1.041	1.017–1.065	.001[b]	1.026	1.000–1.052	.049[b]	1.039	1.012–1.065	.004			
Experience of endoscopist – expert[a]	1.136	0.980–1.317	.089[b]	1.213	1.030–1.428	.020[b]	1.178	1.001–1.385	.048[b]	1.279	1.068–1.532	.007[b]
Withdrawal time[a]	1.184	1.162–1.207	<.001[b]	1.179	1.156–1.202	<.001[b]	1.198	1.172–1.224	<.001[b]	1.193	1.167–1.220	<.001[b]
Colonoscopy indication[a]												
FIT Positive	2.700	1.871–3.896	<.001				2.874	1.963–4.208	<.001			
FIT Negative	1.117	0.959–1.302	.156				1.126	0.950–1.3333	.172			
Direct screening		Reference						Reference				

aOR, adjusted odds ratio; CI, confidence interval; AI, artificial intelligence; BMI, body mass index; FIT, fecal immunochemical test; OR, odds ratio.
[a]Factors considered in the multivariate regression analysis.
[b]●●●.

Q7

nonpedunculated) resections, which could translate into better cancer prevention in the long term. The positive results were at the expense of slightly prolonged intubation time (~0.3 minutes) and withdrawal time (~0.5 minutes). The benefit of AI-assisted colonoscopy in adenoma detection is confirmed in regression analysis, among other determining factors such as male sex, ≥60 years of age, and withdrawal time.

Suboptimal quality of colonoscopy is well known to be related to postcolonoscopy CRC. Studies have shown that quality varies between individual screening examinations, hospital and healthcare institutions, and endoscopists, in terms of the rate of complete examination, detection rates of neoplastic lesions, and polypectomy technique.[20,21] ADR is one of the most important quality indicators of colonoscopy and is inversely associated with the risk of postcolonoscopy CRC, advanced stage of CRC, and CRC mortality.[5,22] In this study, AI-assisted colonoscopy has shown the clear advantage of increasing the detection of adenoma irrespective of location (proximal vs distal colon) and with the experience of endoscopist, and could even attenuate the effect of suboptimal bowel preparation.

While prior studies of real-time AI assistance have reported improved ADR for adenoma <10 mm,[12–15] our study is the first to report improvement in detection for advanced adenomas, and this was driven by increased detection of adenomas ≥10 mm. This is unlikely to be explained by a difference in bowel preparation quality, given that the BBPS was higher in the CC group. This finding was unexpected and warrants additional study in other settings and with different AI platforms. We also found that AI-assisted colonoscopy improved the detection of nonpedunculated adenomas (27.6% vs 21.8%; $P < .001$), and 59 (59.6%) of the 99 adenomas ≥10 mm were nonpedunculated lesions.

Real-time polyp detection by AI on the same endoscopy display monitor is designed to enhance prompt removal of adenomas and in turn reduce adenoma miss rate. However, this potential benefit often comes at the expense of slightly longer intubation time and withdrawal time in AI-assisted colonoscopy, as was the case in our study. This is likely due to more time to confirm suspected polyps detected by AI when the system is turned on during insertion and withdrawal. Given the benefit of higher overall ADR observed in this large-scale randomized study, we believe that the relatively small prolongation of intubation and withdrawal times are an acceptable drawback.

The real-time polyp detection system used in this study is different from other studies that have reported high reliability in detecting the image of polypoid structures, better positioning of the lesions, and shorter time lag between machine detection and real-time imaging. The algorithm of the YOLOv3 system is a deep convolution neural network–based on the Darknet 53 structure, and it is beyond the scope of this article to discuss the engineering details of this AI system. While there is no direct comparison of different AI-assisted tools to date, we believe that the reported differences between AI systems are not major. To date, there is also a lack of direct comparison between AI-assisted colonoscopy against various image-enhanced endoscopy technologies such as narrow-band imaging, i-SCAN, and linked color imaging.

Studies have shown that SSLs are associated with CRC, especially those on the right colon and in the elderly age group, and hence should be detected and removed.[23] In our study, despite the improved ADR for nonpedunculated lesions observed in the AI-assisted colonoscopy, the detection of SSL was not improved by AI. Our AI system used in this study did not undergo dedicated training with a large dataset of SSLs; thus, the sensitivity and specificity for SSLs of the AI system remains unclear and constitutes a limitation of this study. While the prevalence of SSLs was reported to range from 3% to 8% in the West, [24] the prevalence of SSL is much lower in Asia (around 1%).[25] The could be related to low awareness of the lesions and hence underdiagnosis in Asian settings. Nevertheless, better AI tools should be developed for SSL detection because this may further enhance the benefit of AI-assisted colonoscopy.

In conclusion, this large-scale multicenter randomized study showed the benefits of AI-assisted colonoscopy in asymptomatic subjects undergoing CRC screening. AI-assisted image analysis has already been applied in mammography for the screening of breast cancer, as well as in 3-dimensional low-dose computed tomography for the screening of lung cancer.[26,27] It is time for us to consider generalizing the use of AI-assisted endoscopy in the gastrointestinal tract.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at https://doi.org/10.1016/j.cgh.2022.07.006.

References

1. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;366:687–696.
2. Singh H, Turner D, Xue L, et al. Risk of developing colorectal cancer following a negative colonoscopy examination: evidence for a 10-year interval between colonoscopies. JAMA 2006;295:2366–2373.
3. Zhao S, Wang S, Pan P, et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. Gastroenterology 2019;156:1661–1674.e11.
4. 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.
5. le Clercq CM, Bouwens MW, Rondagh EJ, et al. Postcolonoscopy colorectal cancers are preventable: a population-based study. Gut 2014;63:957–963.

6. Wieszczy P, Waldmann E, Løberg M, et al. Colonoscopist performance and colorectal cancer risk after adenoma removal to stratify surveillance: two nationwide observational studies. Gastroenterology 2021;160:1067–1074.e6.
7. Bisschops R, East JE, Hassan C, et al. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy 2019;51:1155–1179.
8. Ng SC, Tsoi KK, Hirai HW, et al. The efficacy of cap-assisted colonoscopy in polyp detection and cecal intubation: a meta-analysis of randomized controlled trials. Am J Gastroenterol 2012;107:1165–1173.
9. Hassan C, Senore C, Radaelli F, et al. Full-spectrum (FUSE) versus standard forward-viewing colonoscopy in an organised colorectal cancer screening programme. Gut 2017;66:1949–1955.
10. Tang RSY, Lee JWJ, Chang LC, et al. Two vs one forward view examination of right colon on adenoma detection: an international multicenter randomized trial. Clin Gastroenterol Hepatol 2022;20:372–380.e2.
11. Misawa M, Kudo SE, Mori Y, et al. Artificial intelligence-assisted polyp detection for colonoscopy: initial experience. Gastroenterology 2018;154:2027–2029.e3.
12. 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.
13. Su J-R, Li Z, Shao X-J, 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.e4.
14. Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol 2020;5:352–361.
15. Repici A, Spadaccini M, Antonelli G, et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. Gut 2022;71:757–765.
16. Luo Y, Zhang Y, Liu M, et al. Artificial intelligence-assisted colonoscopy for detection of colon polyps: a prospective, randomized cohort study. J Gastrointest Surg 2021;25:2011–2018.
17. Kamba S, Tamai N, Horiuchi H, et al. A multicentre randomized controlled trial to verify the reducibility of adenoma miss rate of colonoscopy assisted with artificial intelligence based software. Gastrointest Endosc 2021;93:AB195.
18. Lai EJ, Calderwood AH, Doros G, et al. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc 2009;69:620–625.
19. Wong JCT, Chiu HM, Kim HS, et al. Adenoma detection rates in colonoscopies for positive fecal immunochemical tests versus direct screening colonoscopies. Gastrointest Endosc 2019;89:607–613.e1.
20. Morris EJA, Rutter MD, Finan PJ, et al. Post-colonoscopy colorectal cancer (PCCRC) rates vary considerably depending on the method used to calculate them: a retrospective observational population-based study of PCCRC in the English National Health Service. Gut 2015;64:1248–1256.
21. Baxter NN, Goldwasser MA, Paszat LF, et al. Association of colonoscopy and death from colorectal cancer. Ann Intern Med 2009;150:1–8.
22. Pohl H, Anderson JC, Aguilera-Fish A, et al. Recurrence of colorectal neoplastic polyps after incomplete resection. Ann Intern Med 2021;174:1377–1384.
23. Song M, Emilsson L, Bozorg SR, et al. Risk of colorectal cancer incidence and mortality after polypectomy: a Swedish record-linkage study. Lancet Gastroenterol Hepatol 2020;5:537–547.
24. Meester RGS, van Herk M, Lansdorp-Vogelaar I, et al. Prevalence and clinical features of sessile serrated polyps: a systematic review. Gastroenterology 2020;159:105–118.e25.
25. Lui RN, Kyaw MH, Lam TYT, et al. Prevalence and risk factors for sessile serrated lesions in an average risk colorectal cancer screening population. J Gastroenterol Hepatol 2021;36:1656–1662.
26. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med 2019;25:954–961.
27. Rodriguez-Ruiz A, Lang K, Gubern-Merida A, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst 2019;111:916–922.

Correspondence
Address correspondence to: Joseph J.Y. Sung, MD, PhD, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore. e-mail: josephsung@ntu.edu.sg; fax: ●●●. Q2

CRediT Authorship Contributions

Hong Xu (Conceptualization: Equal; Data curation: Equal; Writing – original draft: Lead)
Raymond S.Y. Tang (Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Project administration: Equal; Validation: Equal; Writing – review & editing: Equal)
Thomas Y.T. Lam (Data curation: Equal; Formal analysis: Equal; Project administration: Equal; Validation: Equal; Writing – review & editing: Equal)
Guijun Zhao (Data curation: Equal; Investigation: Equal)
James Y.W. Lau (Data curation: Equal; Investigation: Equal)
Yunpeng Liu (Data curation: Equal; Funding acquisition: Equal)
Qi Wu (Data curation: Equal; Investigation: Equal)
Rong Long (Data curation: Equal; Investigation: Equal)
Weiran Xu (Data curation: Equal; Investigation: Equal)
Xue Li (Data curation: Equal; Investigation: Equal)
Sunny H. Wong (Data curation: Equal; Investigation: Equal)
Shuntian Cai (Data curation: Equal; Funding acquisition: Equal)
Jing Wang (Data curation: Equal; Investigation: Equal)
Guanyi Liu (Data curation: Equal; Investigation: Equal)
Tantan Ma (Data curation: Equal; Investigation: Equal)
Xiong Liang (Data curation: Equal; Investigation: Equal)
Joyce W.Y. Mak (Data curation: Equal; Investigation: Equal)
Hongzhi Xu (Data curation: Equal; Investigation: Equal)
Peng Yuan (Data curation: Equal; Investigation: Equal)
Tingting Cao (Data curation: Equal; Investigation: Equal)
Fudong Li (Data curation: Equal; Investigation: Equal)
Zhenshi Ye (Data curation: Equal; Investigation: Equal)
Zhang Shutian (Conceptualization: Equal)
Joseph J.Y. Sung (Conceptualization: Lead; Data curation: Equal; Investigation: Equal; Project administration: Lead; Supervision: Lead; Validation: Lead; Writing – review & editing: Lead)

Conflicts of interest
The authors disclose no conflicts. Q3

Supplementary Methods

Artificial Intelligence Polyp Detection System

The artificial intelligence (AI) polyp detection system (Eagle Eye [version 5.1]; Xiamen Innovision, Xiamen, China) used in this study has been described previously and validated in a single-center randomized controlled trial of 150 patients.[1] Briefly, it is developed on the basis of the deep learning architecture with the assistance of endoscopists and modelers. The detection algorithm is a YOLOv3 deep convolution neural network, which allows image analysis from multiple scales (Supplementary Figure 1).[2] Lesions of different sizes can be detected via features from different scales which in turn improves the positioning accuracy of the tracking boxes. The AI platform adopts a multithreaded processing system with a delay time of 17 ms. The algorithm was trained by 112,199 still images (including 64,134 images with histopathologically confirmed polyps and 48,065 images without polyps) of colonoscopy from Zhongshan Hospital affiliated with Xiamen University. All the images were recorded in white light mode and labeled by colonoscopists with more than 5 years of experience. A total of 90% and 10% of the images in the dataset were randomly selected for training and validation, respectively. A separate dataset of 21 colonoscopy videos was employed to fine-tune the model in which 7 videos were recorded by Olympus 290 system (Olympus, Tokyo, Japan), while 14 videos were recorded by Fujifilm 4450 system (Fujifilm, Tokyo, Japan). Afterward, the AI model was validated by an independent dataset of 40,000 colonoscopy images (20,000 with polyp and 20,000 without polyp) extracted from colonoscopy videos that were recorded by the Olympus endoscopy system and Fujifilm endoscopy system. The version of the AI system used in this study has a sensitivity of >98% and specificity of >89% for colonic polyps.

The AI polyp detection system has the size of an ordinary desktop computer tower and is installed between the endoscopy processor and monitor (Supplementary Figure 2). The AI system receives the colonoscopy video from the endoscopy system and analyzes the video frame by frame with the AI polyp detection algorithm. The processed results are then displayed on the endoscopy monitor with a hollow blue tracking box over suspected polyp (Figure 2) in real-time. This AI polyp detection system supports various endoscopy systems, including Olympus CV260/290, Fujifilm EPX-4450 HD /7000 HD, and PENTAX EPK-i5000/7000 (Pentax, Tokyo, Japan).

Data Collection of Procedure Time and Characteristics of Polyps Resected Section

Procedure time was recorded at the commencement of rectal intubation. In the AI group, when lesion was detected by the AI system, confirmation of the lesion identified in the blue tracking box was included in the time recording. In both the AI and conventional colonoscopy groups, time recording was stopped during polypectomy. Therefore polypectomy time would not count into the procedure time.

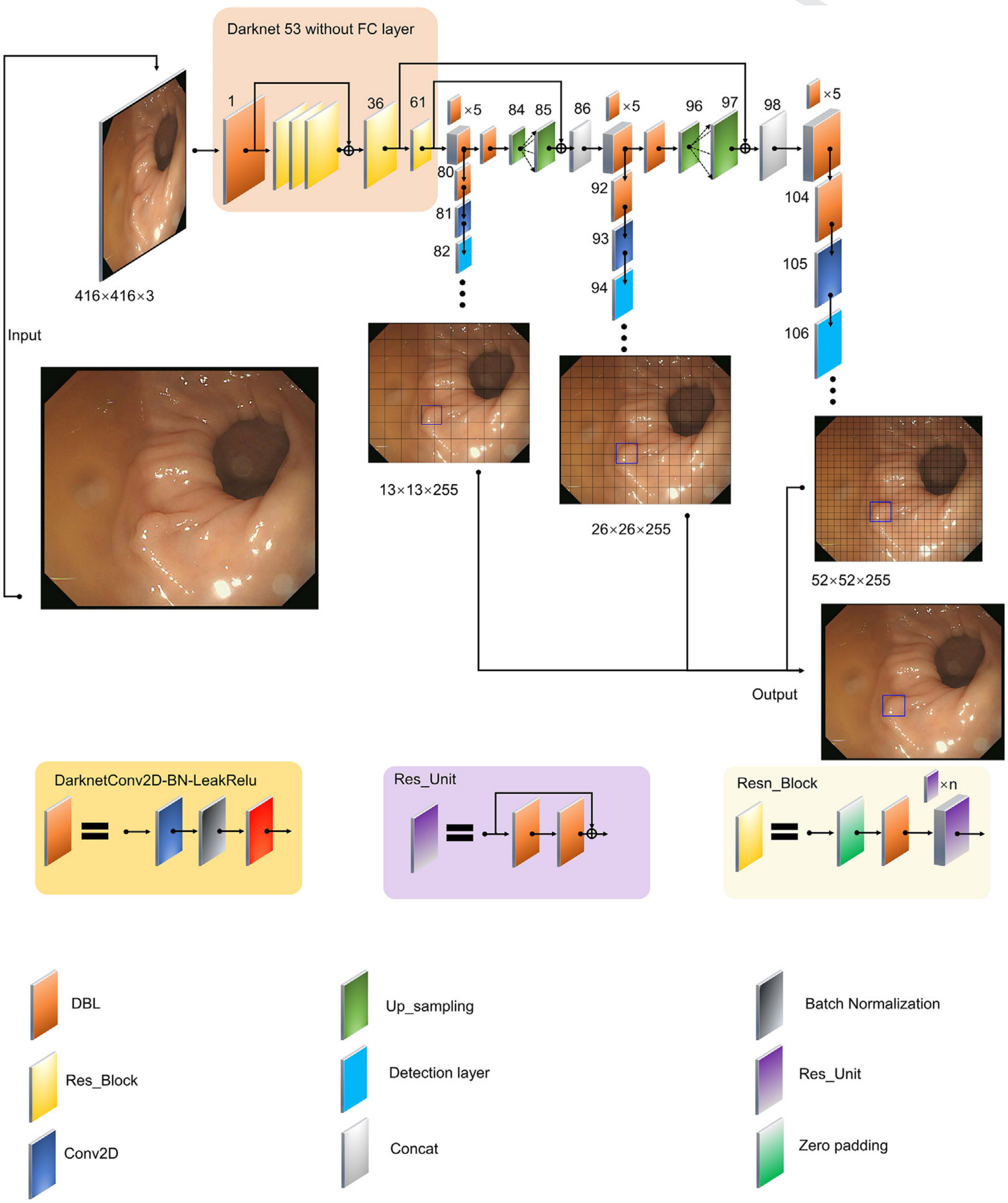
Location, size and morphology of all removed polyps were documented. Nonresectable lesions were biopsied. The proximal colon was defined as the segment from cecum to transverse colon, while distal colon was defined as segment from splenic flexure to rectum. Polyps were classified into pedunculated or nonpedunculated morphology, and nonpedunculated lesions were further characterized as sessile, flat, or depressed lesions. Polyps removed were placed in separate specimen bottles based on anatomical location of the polyps (eg, cecum, ascending colon). Tissues were stained by hematoxylin and eosin stain and Periodic acid–Schiff stain when necessary. Colonic polyp specimens were evaluated by pathologists of individual study center who were blinded to study group assignment. Colonic polyp specimens were evaluated according to the 2019 World Health Organization classification of tumors of the digestive system.[3] Advanced adenomas were defined as adenomas with size ≥10 mm, villous component, or high-grade dysplasia. Patients were monitored for any immediate adverse event postcolonoscopy and were discharged after recovery from sedation.

Statistical Analysis

Intention-to-treat and per-protocol analysis were used to evaluate the primary outcome and secondary outcomes. Intention-to-treat analysis included all recruited subjects, while per-protocol analysis excluded patients who had incomplete colonoscopy or inadequate bowel preparation (Boston Bowel Preparation Scale <6 or any segmental Boston Bowel Preparation Scale <2). Statistical analyses were performed using SPSS software (version 26.0; IBM, Armonk, NY). Descriptive statistics were reported as mean with standard deviation, median with interquartile range, or frequency with percentage as appropriate. The mean and median value between groups were compared by Student's t test and independent samples median test, respectively. Differences in categoric variables were analyzed using chi-square tests. P values <.05 were considered statistically significant. Univariate and multivariate logistic regression were performed to identify predictors of adenoma detection during colonoscopy for the following covariates: AI assistance, patient's age, sex, smoking status, drinking status, aspirin usage, existing diabetes, existing ischemic heart diseases, existing fatty liver, body mass index, experience of endoscopist, withdrawal time, and colonoscopy indication. Variables with P values <.1 in the univariate model were included in the multivariate logistic regression model (backward stepwise selection) to determine the independent predictive effect of the variables of interest. *P* values <.05 were considered statistically significant.

Supplementary References

1. Luo Y, Zhang Y, Liu M, et al. Artificial intelligence-assisted colonoscopy for detection of colon polyps: a prospective, randomized cohort study. J Gastrointest Surg 2021; 25:2011–2018.
2. Redmon J, Farhadi A. YOLOv3: an incremental improvement. Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 18-23, 2018; Salt Lake City, UT.
3. Dixon MF. Gastrointestinal epithelial neoplasia: Vienna revisited. Gut 2002;51:130–131.



Supplementary Figure 1. The YOLOv3 architecture.

web 4C/FPO

web 4C/FPO

Supplementary Figure 2. The installation of the AI polyp detection system.