A computer-aided polyp detection system in screening and surveillance colonoscopy: an international, multicentre, randomised, tandem trial

Michiel H J Maas, Helmut Neumann, Haim Shirin, Lior H Katz, Ariel A Benson, Arslan Kahloon, Elsa Soons, Rawi Hazzan, Marc J Landsman, Benjamin Lebwohl, Suzanne K Lewis, Visvakanth Sivanathan, Saowanee Ngamruengphong, Harold Jacob, Peter D Siersema

Summary

Background Studies on the effect of computer-aided detection (CAD) in a daily clinical screening and surveillance colonoscopy population practice are scarce. The aim of this study was to evaluate a novel CAD system in a screening and surveillance colonoscopy population.

Methods This multicentre, randomised, controlled trial was done in ten hospitals in Europe, the USA, and Israel by 31 endoscopists. Patients referred for non-immunochemical faecal occult blood test (iFOBT) screening or surveillance colonoscopy were included. Patients were randomomly assigned to CAD-assisted colonoscopy or conventional colonoscopy; a subset was further randomly assigned to undergo tandem colonoscopy: CAD followed by conventional colonoscopy or conventional colonoscopy followed by CAD. Primary objectives included adenoma per colonoscopy (APC) and adenoma per extraction (APE). Secondary objectives included adenoma miss rate (AMR) in the tandem colonoscopies. The study was registered at ClinicalTrials.gov, NCT04640792.

Findings A total of 916 patients were included in the modified intention-to-treat analysis: 449 in the CAD group and 467 in the conventional colonoscopy group. APC was higher with CAD compared with conventional colonoscopy (0·70 *vs* 0·51, p=0·015; 314 adenomas per 449 colonoscopies *vs* 238 adenomas per 467 colonoscopies; poisson effect ratio 1·372 [95% CI 1·068–1·769]), while showing non-inferiority of APE compared with conventional colonoscopy (0·59 *vs* 0·66; p<0·001 for non-inferiority; 314 of 536 extractions *vs* 238 of 360 extractions). AMR in the 127 (61 with CAD first, 66 with conventional colonoscopy first) patients completing tandem colonoscopy was 19% (11 of 59 detected during the second pass) in the CAD first group and 36% (16 of 45 detected during the second pass) in the conventional colonoscopy first group (p=0·024).

Interpretation CAD increased adenoma detection in non-iFOBT screening and surveillance colonoscopies and reduced adenoma miss rates compared with conventional colonoscopy, without an increase in the resection of non-adenomatous lesions.

Funding Magentiq Eye.

Lancet Digit Health 2024; 6: e157–65

Department of Gastroenterology & Hepatology, Radboud University Medical Center, Nijmegen, Netherlands (M H J Maas MD, E Soons MD PhD, Prof P D Siersema MD PhD); **University Medical Center Mainz, Interventional Endoscopy Center, I Medizinische Klinik und Poliklinik, Mainz, Germany** (Prof H Neumann MD PhD, V Sivanathan MD); **Institute of Gastroenterology and Liver Diseases, Shamir (Assaf Harofeh) Medical Center, Zerifin, Israel** (Prof H Shirin MD); **Hadassah Medical Center and Faculty of Medicine, Hebrew University of Jerusalem, Institute of Gastroenterology and Liver Diseases, Jerusalem, Israel** (Prof L H Katz MD, A A Benson MD, H Jacob MD); **College of Medicine, Division of Gastroenterology, University of Tennessee, Chattanooga, TN, USA** (A Kahloon MD); **Assuta Centers, Haifa Gastroenterology Institute, Haifa, Israel** (R Hazzan MD); **Department of Gastroenterology, MetroHealth Medical Center, Cleveland, OH, USA** (M J Landsman MD); **Department of Gastroenterology, Columbia University Irving Medical Center, New York, NY, USA** (B Lebwohl MD, S K Lewis MD); **Department of Gastroenterology, Johns Hopkins University, Baltimore, MD, USA** (S Ngamruengphong MD); **Department of Gastroenterology & Hepatology, Erasmus MC, University Medical Center, Rotterdam, Netherlands** (Prof P D Siersema)

Introduction

Colonoscopy is the gold standard for early detection and removal of premalignant colorectal polyps. However, a substantial number of polyps are missed during colonoscopy. A meta-analysis of tandem colonoscopy studies reported a pooled adenoma miss rate of 26%.[1] Several factors can result in missing lesions during colonoscopy, such as incomplete visualisation of the colonic mucosa, operator distraction, and fatigue.[2,3] Missed lesions could potentially develop into colorectal cancer, and it is hypothesised that at least 50% of all post-colonoscopy colorectal cancer, defined as a colorectal cancer diagnosed after a colonoscopy in which no cancer was found, develop from missed lesions.[4]

The use of machine-learning or deep-learning systems, also known as computer-aided detection (CAD) systems, in colonoscopy has been shown to increase detection rates and reduce adenoma miss rates.[5–7] However, international studies evaluating the effect of CAD on adenoma detection and miss rates in non-immunochemical faecal occult blood test (iFOBT) screening and surveillance colonoscopy populations are scarce.[8,9] A novel CAD device (MAGENTIQ-COLO, Magentiq Eye, Haifa, Israel) has been developed to detect colorectal polyps in real time during colonoscopy.

The aim of this multicentre, randomised, tandem trial was to evaluate the efficacy and safety of colonoscopy assisted by a novel CAD device compared with conventional colonoscopy in a daily clinical practice non-iFOBT screening and surveillance population.

Methods

Study design

This prospective, multicentre, international, randomised controlled trial (RCT) was done in ten hospitals in Europe (n=3), Israel (n=3), and the USA (n=4) by 31 endoscopists.

Correspondence to:
Michiel H J Maas, Department of Gastroenterology & Hepatology, Radboud University Medical Center, 6500 HB Nijmegen, Netherlands
michiel.maas@radboudumc.nl

Research in context

Evidence before this study

We searched PubMed, with no language restrictions, from inception to Oct 1, 2023, for randomised controlled trials (RCTs) and meta-analyses using the following search terms: "artificial intelligence", "computer-aided detection", "adenoma per colonoscopy", "adenoma miss rate", "randomised controlled trial", "meta-analysis", and "systematic review". During conventional colonoscopy, up to 26% of all adenomas are missed as reported by a 2019 meta-analysis of tandem studies. These missed lesions could potentially develop into post-colonoscopy colorectal cancer. Computer-aided detection (CAD) systems, using artificial intelligence, have been developed to assist endoscopists in detecting lesions during colonoscopy by highlighting polyps in real-time on the monitor. A 2023 meta-analysis, which pooled the results of 21 RCTs across China, Europe, Japan, and the USA, reported an approximately absolute significant increase of 0·20 in adenoma per colonoscopy (APC), and an absolute significant reduction of 19% in adenoma miss rate (AMR) for CAD tandem studies.

Added value of this study

Our tandem RCT strongly supports the benefit of CAD-assisted colonoscopy compared with conventional colonoscopy in terms of APC and AMR. The results of our study are similar to the pooled results of the aforementioned meta-analysis of 21 CAD RCTs. Notably, our study contributes to the current literature through our novel parallel tandem design, which supports the hypothesis that CAD detects visually neglected adenomas. Moreover, detection of both proximal and 6–9 mm adenomas was increased in our study, suggesting that this CAD system might be able to detect these clinically relevant lesions.

Implications of all the available evidence

CAD has shown its efficacy in improving detection rates and reducing miss rates across various colonoscopy indications and settings, supporting the routine integration of CAD during colonoscopy to increase detection rates. Nonetheless, the focus of future studies should shift towards the effect of CAD on long-term outcomes, including incidence of post-colonoscopy colorectal cancer, and colorectal cancer-related and overall mortality.

The study was approved by the independent institutional review boards at each of the study sites in accordance with the Declaration of Helsinki and applicable guidelines for Good Clinical Practice. This study was reported according to the CONSORT-AI guidelines for RCTs. Participant records underwent source data verification and monitoring.

Participants

Patients aged 18–90 years and scheduled for non-iFOBT screening or surveillance colonoscopy (at least 3 years after the last examination) were recruited and prospectively enrolled. Patients were excluded in case of an already known or suspected colorectal tumour or polyp, therapeutic colonoscopy, inadequately corrected anticoagulation use or disorder, pregnancy or potential pregnancy, inadequate bowel preparation, defined as an overall Boston Bowel Preparation Score (BBPS)[10] of less than 6 or a score of less than 2 in any segment, or known inflammatory bowel disease. Patients with a new diagnosis of active colitis, polyposis syndrome, colonic stricture, or obstructing colorectal cancer not allowing complete colonoscopy were withdrawn from the study. All study participants provided written informed consent.

Randomisation

Participants were randomly assigned in a 6:6:1:1 ratio to four different study groups: single CAD-assisted colonoscopy, single conventional colonoscopy, CAD first followed by conventional colonoscopy, and conventional colonoscopy first followed by CAD. Tandem colonoscopies were done by the same endoscopist, back-to-back, on the same day. Randomisation was stratified according to the study site and colonoscopy indication, and was done on-site within 24 h before colonoscopy, using a central (cloud-based) service (Sealed Envelope, London, UK). The study site coordinator revealed the randomly assigned allocation to the endoscopist just before the first procedure.

Procedures

Participants randomly assigned to conventional colonoscopy underwent a routine colonoscopy per standard of care. In the CAD-assisted colonoscopy study group, the investigational device was switched on at the start of the procedure. The endoscopist completed the study procedure per standard of care with the assistance of the investigational device. Conventional high-definition scopes were used during all procedures. Use of distal attachments to improve field of view was not allowed. In each study group, participating endoscopists were instructed to aim for a withdrawal time of 6–10 min, excluding interventions. Bowel preparation was evaluated with the BBPS. Each colonic lesion detected and removed during the procedure was separately sent for histopathological examination. Diminutive polyps (ie, polyps of 1–5 mm in diameter) located in the rectum and determined to be hyperplastic could be left in-situ. Experienced pathologists masked to the endoscopic diagnosis or intervention determined the histopathological diagnosis with the Vienna criteria.[11] The use of sedatives was according to local protocol and the assessment of the physician. Patient records were reviewed for adverse events occurring for up to 30 days after the study procedure.

The investigational CAD device, MAGENTIQ-COLO, is based on a computing device that acquires digital video output signals from the colonoscope during colonoscopy. The CAD system runs a dedicated convolutional neural network in real time on the video output of the endoscopic camera. It then feeds a video output with the detected lesion in an overlay to the physician to assist in polyp detection. The CAD device is to be used as an adjunct to the colonoscopy procedure, aiming to assist the endoscopist in highlighting regions with visual characteristics consistent with different types of mucosal abnormalities. The CAD device has been internally tested on 172 full colonoscopy videos, composed of 4656632 video frames featuring 263 different polyps verified by pathology, resulting in 99·6% polyp-wise sensitivity and 98% specificity. The MAGENTIQ-COLO was clinically tested and found compatible with processors and endoscope models of all current main manufacturers. During the study, MAGENTIQ-COLO version 1.0, with software version 1.7.2, was used.

All participating endoscopists were trained in using the CAD device. CAD was installed at each individual study site. Each endoscopist had at least one training colonoscopy with the CAD system, and each endoscopist confirmed that they were familiar with the system before enrolling patients in the study. The 31 participating endoscopists were required to have a registered adenoma detection rate (ADR) in the range of 25–40% before participation in the study.

Outcomes

This study had two primary endpoints: adenoma per colonoscopy (APC), defined as the total number of histologically confirmed adenomas divided by the total number of colonoscopies during the first examination, and adenoma per extraction (APE), defined as the total number of histologically confirmed adenomas divided by the total number of extractions (polypectomies or biopsies) during the first examination. Mean APE was reported on a per-patient basis. Secondary endpoints included ADR, defined as the number of patients with one or more adenomas detected and removed divided by the total number of patients, and adenoma miss rate (AMR), defined as the number of adenomas detected during the second examination divided by the total number of adenomas detected during the first and second colonoscopy. Additional endpoints included sessile serrated lesions (SSLs) detection rate, defined as the number of patients with one or more histologically confirmed SSLs divided by the total number of colonoscopies, adenoma size, adenoma location (subdivided in proximal [ie, cecum, ascending colon, hepatic flexure, or transverse colon] or distal [ie, splenic flexure, descending colon, sigmoid, or rectum]), adenoma morphology according to the Paris classification,[12] and withdrawal time excluding intervention time. In secondary analyses, outcomes were also evaluated for each colonoscopy indication.

Statistical analysis

The study was powered to detect a significant effect in both coprimary endpoints, APC, and APE. Additionally, sample size calculation was done to determine the required number of participants in the tandem colonoscopy study groups. Sample size calculations were done with the PASS software (version 19.0.4). The minimum required sample size for APC was calculated with a superiority test following a Poisson distribution with a superiority margin of 5% (superiority margin ratio of 1·05). APC for the conventional colonoscopy group was assumed to be 0·30. Based on previous system testing and consultations, we expected an increase to 0·45 with CAD-assisted colonoscopy. At a one-sided alpha of 0·025 with 80% power, a minimum of 343 patients were required per group, totalling 686 patients. The minimum required sample size for APE was calculated with a non-inferiority design. Two previous studies reported a decrease in APE with CAD compared with conventional colonoscopy, from 0·79 to 0·58,[13] and from 0·65 to 0·64, respectively.[14] The average APE from both studies for each study group (average of 0·72 for the conventional colonoscopy group and 0·61 for the intervention group) resulted in an expected difference in favour of conventional colonoscopy of 0·11. Since it was expected that CAD would detect more polyps, it was reasonable to expect that more polyps would be extracted, with some being negative for adenoma. Therefore, a lower APE was expected for CAD-assisted colonoscopy compared with conventional colonoscopy, and the non-inferiority margin was set to 0·20 as an expected value. With a non-inferiority margin set at 0·20, a one-sided alpha of 0·025, a statistical power of 80%, and a 35% prevalence of examinations with no extractions, a minimum of 448 patients per group were required for a total of 896 participants. Accounting for a 5% dropout rate, the minimum required sample size was 944. This sample size was increased to 952 participants to account for a reasonable randomisation ratio. The minimum required sample size for the tandem colonoscopy study groups was calculated using binominal proportions. We assumed the proportion of AMR in the conventional colonoscopy first group to be 0·400 and in the CAD first group to be 0·138. Assuming a two-sided alpha of 5%, with 80% power, a sample size of 104 was required. Assuming a dropout rate of 5%, the sample size was increased to 110. To obtain a reasonable randomisation ratio, the sample size was further increased to 136 participants, 68 in each tandem study group.

The primary analysis was a per-patient analysis done according to the modified intention-to-treat (mITT) of participants not excluded due to an inadequate bowel preparation score or inability to examine the colon. The statistical analysis was done with the JMP Pro Statistical

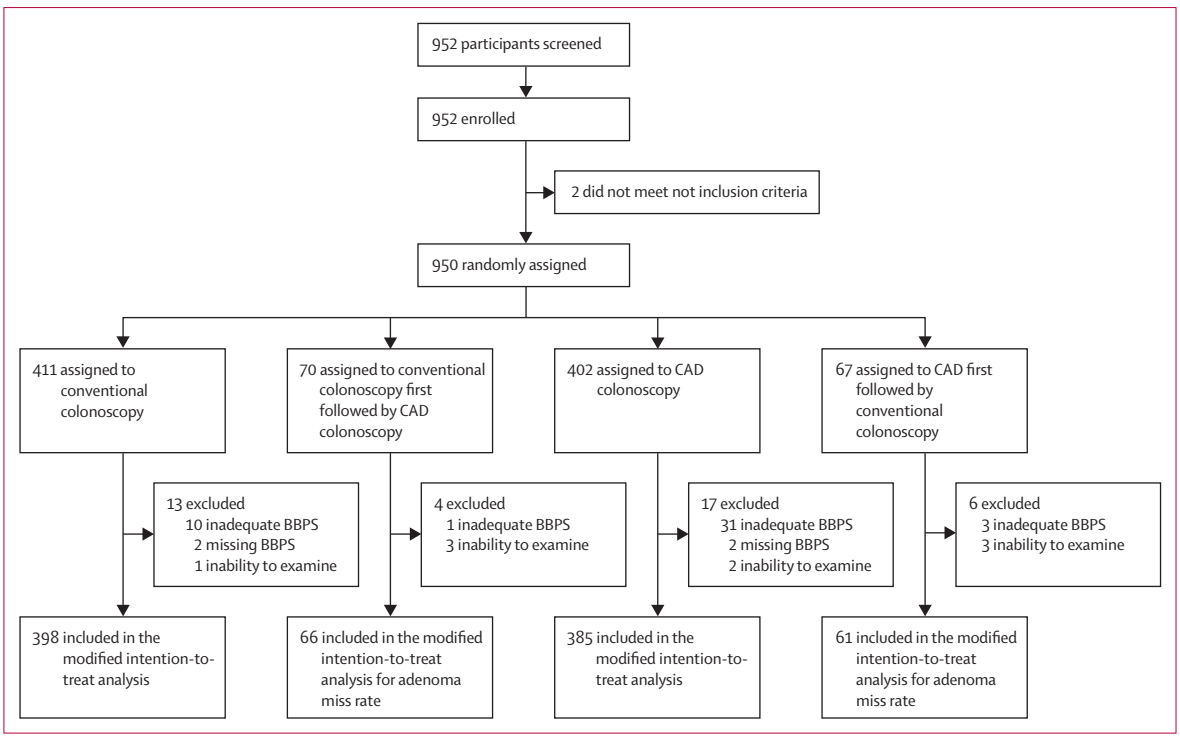

Figure: **Trial profile**
BBPS=Boston Bowel Preparation Score. CAD=computer-aided detection.

Discovery software, version 16.2.0 (SAS Institute, Cary, NC, USA).

Participant demographics by age, gender (per the participants' electronic health record registration), BMI, colonoscopy indication, race, and BBPS were summarised with descriptive statistics. The difference in APC between study groups was assessed with a two-sided *t*-test, with an alpha set at 5%. This analysis deviated from the sample size calculation method to assess the difference between means, reasonably assuming normal approximation due to the large sample size. Additionally, following the original sample size calculations, a generalised linear Poisson regression analysis was done as sensitivity analysis with a log link function. This analysis allowed for obtaining the ratio of CAD to conventional colonoscopy ratio in APC and its 95% CI, with the requirement that the lower limit of the 95% CI was above 1·05 to show superiority of CAD. Additionally, a post-hoc Poisson regression analysis was done to address potential stratification effects. This model included study site, colonoscopy, and treatment group as independent variables, and APC as the dependent variable. Only the first colonoscopy was considered in the analysis of APC for participants undergoing the tandem procedure.

For APE, a two-sided (alpha 5%) Wilcoxon test with a non-inferiority margin of 20% was performed after subtracting 0·2 from the values of APE in the conventional colonoscopy cohort to test for non-inferiority. A Wilcoxon test was used since APE, being a ratio on an examination level, might not adhere to a normal distribution. Additionally, the 95% CI of the mean per-patient APE of CAD and conventional colonoscopy was calculated using normal approximation. Examinations with no extractions were considered null. To maintain an overall type I error of 5%, secondary endpoints were only analysed after both coprimary endpoints (APE and APC) met their criteria. This hierarchy of statistical tests maintained the required overall type I error, addressing the multiplicity issue.

The AMR analysis was done with the non-parametric Wilcoxon Rank-Sum test, with the success criterion defined as $p<0·025$. Normal approximation was used to identify the 95% CI of the mean participant-level AMR. To simulate a worst-case scenario, participants with no adenomas in both examinations were included in the analysis, assuming that one adenoma was missed by both techniques. AMR was calculated in patients who completed both assigned procedures. ADR, assumed to follow a binomial distribution as a binary variable, was assessed with a Pearson χ^2 test. The 95% CIs of ADR and SSL detection rate for conventional colonoscopy and CAD were calculated with the Wilson Score Method. For all analyses, $p\leq0·05$ was considered statistically significant. The analysis of the counts of the secondary endpoints was identified after obtaining the properties of the participant-level data. Therefore, the analysis deviated from the study protocol, which specified a non-parametric approach for handling counts. Data analysis was

independently done by DataSights, Haifa, Israel. The study was registered at ClinicalTrials.gov, NCT04640792.

Role of the funding source

The funder of the study had no role in data collection, data interpretation, and writing of the report. The funder provided feedback during study design and contracted a statistical analyst to assist in data analyses.

Results

A total of 952 patients were screened and enrolled from Nov 8, 2020, to Jan 11, 2022. After the exclusion of two patients due to not meeting inclusion criteria, 950 were randomly assigned to four study groups: 402 to CAD, 411 to conventional colonoscopy, 67 to CAD first followed by conventional colonoscopy, and 70 to conventional colonoscopy first followed by CAD. 34 patients were excluded from the final analysis due to inadequate bowel preparation, missing BBPS data, or inability to undergo colonoscopy (figure).

In total, 916 patients were included in the final mITT analysis (385 in CAD, 398 in conventional colonoscopy, 64 in CAD first followed by conventional colonoscopy, and 69 in conventional colonoscopy first followed by CAD), with 514 having a screening indication, of whom 512 underwent a first-time colonoscopy. 402 had a surveillance indication. Cecal intubation and complete colonoscopy was done in 914 (>99%) of 916 patients. Colonoscopy indication was distributed between the index conventional colonoscopy group (n=467; 259 screening, 208 surveillance) and the index CAD group (n=449; 255 screening, 194 surveillance). Total of BBPS scores was distributed between the index conventional colonoscopy group (n=467; 77 BBPS score of 7, 40 BBPS score of 75, 275 BBPS score of 9) and the index CAD group (n=449; 86 BBPS score 6, 33 BBPS score 7, 65 BBPS score 8, 265 BBPS score 9). A total of 127 patients (61 CAD first, 66 conventional colonoscopy first) completed both assigned tandem procedures. The median age of all patients was 60 years (IQR 12·0), and 493 (54%) of 916 were men (table 1). No missing values were observed for the calculation of the primary and secondary endpoints.

APC was higher in the CAD group than in the conventional colonoscopy group (0·70 [95% CI 0·58 to 0·82] *vs* 0·51 [0·42 to 0·60], p=0·015; total detected adenomas per colonoscopy, 314 of 449 *vs* 238 of 467). Poisson analysis showed a 37% relative increase in APC with CAD compared with conventional colonoscopy (relative risk [RR] 1·372 [95% CI 1·068 to 1·769]), exceeding the superiority margin requirement of 1·05. In the post-hoc Poisson regression analysis accounting for study site and colonoscopy indication, the observed increase in APC retained its statistical significance (p=0·005), with an RR of 1·56 (1·16 to 2·11). APE was 0·59 (314 adenomas in 536 extractions) in the CAD group and 0·66 (238 adenomas in 360 extractions) in the conventional colonoscopy group (p<0·001 for non-inferiority). Mean per-patient APE was higher in the CAD group compared with the conventional colonoscopy group (0·27 [95% CI 0·23 to 0·31] *vs* 0·31 [0·27 to 0·35]), corresponding to a point-estimate difference of 0·04 (95% CI –0·018 to 0·093). Furthermore, ADR was higher in the CAD group than in the conventional colonoscopy group (37% [95% CI 33 to 42] *vs* 30% [26 to 34], p=0·014; colonoscopies with at least one adenoma, 167 of 449 *vs* 138 of 467) and SSL detection rate was higher in the CAD group than in the conventional colonoscopy group, although not statistically significant (6·0% [95% CI 4·2 to 8·6] *vs* 3·9% [2·5 to 6·0], p=0·087; colonoscopies with at least one SSL, 27 of 449 *vs* 18 of 467, respectively).

In the CAD first tandem colonoscopy group, 48 adenomas were detected during the index colonoscopy, whereas 11 adenomas were detected during the subsequent conventional colonoscopy, corresponding to an AMR of 19% (11 of 59 total adenomas). In the conventional colonoscopy first tandem colonoscopy group, 29 adenomas were detected, and an additional 16 adenomas were detected during the subsequent CAD colonoscopy, corresponding to an AMR of 36% (16 of 45 total adenomas; p=0·024). Mean per-patient AMR was higher in the conventional colonoscopy first group than in the CAD first group (0·81 [95% CI 0·72–0·90] *vs* 0·64 [0·52–0·76]). Withdrawal times without interventions were similar in the CAD group compared with the conventional colonoscopy group (withdrawal time without interventions 6·5 [IQR 4·0] min *vs* 6·5 [4·1] min, respectively; p=0·98). Withdrawal time without inventions in the tandem colonoscopy group was similar in the CAD first group compared with the conventional colonoscopy first group (6·0 [1·0] min *vs* 6·0 [0·8] min, respectively; p=0·58).

	Conventional colonoscopy (n=398)	Conventional colonoscopy followed by CAD (n=69)	CAD (n=385)	CAD followed by conventional colonoscopy (n=64)
Age, years	60·0 (13·0)	60·1 (10·1)	60·0 (12·0)	59·3 (9·0)
Sex				
Male	219 (55%)	38 (55%)	202 (52%)	34 (53%)
Female	179 (45%)	31 (45%)	183 (48%)	30 (47%)
BMI, kg/m²	26·4 (5·8)	25·9 (4·8)	26·1 (4·7)	27·2 (4·8)
Indication for colonoscopy				
non-iFOBT screening	222 (56%)	37 (54%)	219 (57%)	36 (56%)
Surveillance	176 (44%)	32 (46%)	166 (43%)	28 (44%)
Race or ethnicity*				
White	374 (94%)	67 (97%)	370 (96%)	63 (98%)
African American	18 (5%)	2 (3%)	12 (3%)	1 (2%)
Asian	1 (<1%)	..	1 (<1%)	..
Other	4 (1%)	..	2 (1%)	..

Data are n (%) or median (IQR). CAD=computer-aided detection. iFOBT=immunochemical faecal occult blood test.
*Race was not reported for one participant in the conventional colonoscopy group.

Table 1: **Baseline characteristics**

	Conventional colonoscopy (n=467)	CAD (n=449)	p value
Adenomas per colonoscopy*			
Overall†	238/467 (0·51); 0·42–0·60	314/449 (0·70); 0·58–0·82	0·015
Poisson model effect ratio, CAD to conventional colonoscopy†	..	1·37 (1·07–1·77)	..
Non-iFOBT screening	106/259 (0·41); 0·31–0·51	149/255 (0·58); 0·46–0·71	0·039
Surveillance	132/208 (0·63); 0·47–0·80	165/194 (0·85); 0·63–1·07	0·12
Adenomas per extraction*‡§			
Overall	238/360 (0·66)	314/536 (0·59)	<0·001
Mean per-patient adenomas per extraction	0·27 (0·23–0·31)	0·31 (0·27–0·35)	..
non-iFOBT screening	106/171 (0·62)	149/239 (0·62)	<0·001
Surveillance	132/189 (0·70)	165/279 (0·56)	<0·001
Adenoma detection rate*			
Overall	138/467 (30%); 26–34	167/449 (37%); 33–42	0·014
non-iFOBT screening	69/259 (27%); 22–32	88/255 (35%); 29–41	0·014
Surveillance	69/208 (33%); 27–40	79/194 (41%); 34–48	0·001
Adenoma miss rate¶			
Overall	16/45 (36%)	11/59 (19%)	0·024
non-iFOBT screening	2/11 (18%)	4/29 (14%)	0·080
Surveillance	14/34 (41%)	7/30 (23%)	0·16
Sessile serrated lesion detection rate*			
Overall	18/449 (4%); 2·5–6·0	27/467 (6%); 4·2–8·6	0·087
non-iFOBT screening	7/259 (3%); 1·3–5·5	18/255 (7%); 4·5–10·9	0·017
Surveillance	11/208 (5%); 3·0–9·2	9/194 (5%); 2·5–8·6	0·70
Withdrawal time without interventions during first examination, minutes (IQR)*	6·5 (4·0)	6·5 (4·1)	0·86
Withdrawal time without interventions during second examination, minutes (IQR)	6·0 (1·0)	6·0 (0·8)	0·58

Data are n/N (% or proportion) or median (IQR), with 95% CI, unless otherwise stated. CAD=computer-aided detection. iFOBT=immunochemical faecal occult blood test. *Calculated on the regular parallel study groups, and including the first examination of tandem procedures. †Ratio calculated with a Poisson regression model with treatment group and adenomas per colonoscopy as model variables. ‡Calculated as the total number of adenomas divided by the total number of extractions. §p value calculated for a non-inferiority margin of 0·20 using a Wilcoxon test after subtracting 0·20 of the conventional colonoscopy group. ¶Calculated on the tandem parallel study groups, the number of adenomas detected during the second colonoscopy divided by the total number of adenomas detected during the first and second colonoscopy. 95% CIs are missing because the adenoma miss rate is reported not on subject-level, but as the total across all participants.

Table 2: **Primary and secondary endpoints according to each study group in the modified intention-to-treat population**

When stratified for colonoscopy indication, the results were similar. For screening colonoscopy (n=514), APC was higher in the CAD group than in the conventional colonoscopy group (0·58 [95% CI 0·46–0·71] *vs* 0·41 [0·31–0·51], p=0·039; total detected adenomas per colonoscopy, 149 of 255 *vs* 106 of 259). For surveillance colonoscopy (n=402), APC was higher in the CAD group than in the conventional colonoscopy group, although not statistically significant (0·85 [0·63–1·07] *vs* 0·63 [0·47–0·80], p=0·115; total detected adenomas, 165 of 194 *vs* 132 of 208). Additional endpoints are reported in table 2.

In the CAD group, adenoma detection was significantly increased on an adenoma-per-colonoscopy basis in the proximal colon compared with the conventional colonoscopy group (47% *vs* 31%, p=0·006; detected adenomas, 209 *vs* 145). Adenoma detection was not significantly increased in the distal colon in the CAD group compared with the conventional colonoscopy group (23% *vs* 20%, p=0·85; detected adenomas, 105 *vs* 93). A significant increase in diminutive adenomas was seen in the CAD group compared with the conventional colonoscopy group (45% *vs* 33%, p=0·035; detected adenomas, 203 *vs* 153), and in adenomas sized 6–9 mm (16% *vs* 10%, p=0·036; detected adenomas, 70 *vs* 47). Adenoma characteristics are reported in table 3.

Discussion

In this multicentre, randomised, controlled, tandem trial, the use of this novel CAD system resulted in an absolute increase of 0·19 (relative increase 37%) in APC and 7% (relative increase 23%) in ADR, while showing non-inferiority of APE compared with conventional colonoscopy. Correspondingly, AMR was decreased by 17% (relative decrease 47%). Adenoma detection was higher for both non-iFOBT screening and surveillance colonoscopy indications and this was primarily driven by an increased detection of both proximal adenomas and adenomas smaller than 10 mm.

Our study is in concordance with recent studies on the efficacy of other CAD systems in improving ADR in screening and surveillance colonoscopy populations. A study by Shaukat and colleagues[8] in a screening and surveillance population in the USA found an increase of 0·22 (27% relative increase) in APC in CAD-assisted colonoscopy. The APC in the conventional colonoscopy group was higher in their study compared with our results, which might be explained by the higher proportion of patients with a screening colonoscopy indication compared with our study. Although this study[8] did not find a significant increase in ADR, the absolute increase in APC was similar to our findings. A single-centre randomised study in France, with more than 2000 participants, reported a comparable absolute increase in APC of 0·17 (0·89 for CAD *vs* 0·71 for conventional colonoscopy). Furthermore, ADR in the CAD group was similar to our study at 37·5%.[15] The overall higher APC of this study could potentially be attributed to its single-centre design in a high-quality endoscopy unit, and longer withdrawal times. A meta-analysis including five other RCTs showed a pooled improvement in APC of 0·22 (0·58 for CAD *vs* 0·36 for conventional colonoscopy) and in ADR of 11·4% (36·6% for CAD *vs* 25·2% for conventional colonoscopy).[5] However, most of these studies were done in China and primarily included patients with a diagnostic colonoscopy indication and had a lower baseline ADR, which is less comparable with the non-iFOBT screening and surveillance of White populations included in our study and the study by Shaukat and colleagues. A pooled analysis of two randomised studies by Repici and colleagues,[16] including iFOBT referrals and non-expert

endoscopists, similarly found a higher APC in the CAD group (1·26) compared with the conventional colonoscopy group (1·04), and a higher ADR with CAD (53·3%) compared with conventional colonoscopy (44·5%). The overall higher detection rates in their study are likely to be attributable to the inclusion of iFOBT colonoscopy indications.[17]

Our study is in contrast with real-world studies from Ladabaum and colleagues[18] and Levy and colleagues,[19] in which the implementation of CAD did not result in a significant increase in adenoma detection in an uncontrolled setting, even among low-detectors. It remains unclear why these results differ from our study and previous RCTs. Notably, the study by Ladabaum and colleagues reported similar rates of bowel cleansing and longer withdrawal times in comparison to our study. The authors suggest that the Hawthorne effect could be of a larger influence than expected.

In this study, withdrawal time without interventions was similar between the study groups, both in the standard parallel study and the second examination of the tandem study groups. Despite the similar withdrawal times, the withdrawal time remained relatively short, at 6 min for all examinations, meeting the minimum recommended duration.[20] However, this is shorter compared with other studies.[7,8,15,16,21] This relatively shorter withdrawal time might have potentially reduced our baseline ADR in the conventional colonoscopy group, as an increased withdrawal time is associated with an increase in ADR.[22] Nevertheless, the overall quality of the colonoscopies in our study was robust, as partially reflected by the exclusion of only a small number of participants due to insufficient BBPS, and the majority of colonoscopies having the maximum BBPS score of 9.

The clinical relevance of our study is supported by the increased detection of proximal lesions, particularly of adenomas located in the caecum and ascending colon. This reflects the benefit of this CAD device in these right-sided locations with a higher risk of colorectal cancer-related mortality after a negative colonoscopy.[23,24] Among a non-significantly increased detection of non-polypoid lesions in the proximal colon, the number of detected SSLs was still relatively low in both study groups, which might have underpowered the study to show an effect in SSL detection rate. However, this is not unexpected, as none of the randomised trials investigating a CAD device have reported a significant increase in detection of SSLs with CAD compared with conventional colonoscopy up to now.[7,8,13–17,21]

Apart from an increased detection of diminutive adenomas, an increased detection of small adenomas sized 6–9 mm was also seen with the use of CAD in our study. The increased detection of small adenomas is of interest and might be clinically relevant, as these patients are at a higher risk of developing metachronous advanced adenomas compared with patients with only diminutive adenomas.[25]

	Conventional colonoscopy (n=467)	CAD (n=449)	p value
Location			
Distal	..	..	0·85
Rectum	23 (5%)	28 (6%)	..
Sigmoid	37 (8%)	40 (9%)	..
Descending colon	32 (7%)	33 (7%)	..
Splenic flexure	1 (<1%)	4 (1%)	..
Proximal	..	..	0·006
Transversal colon	59 (13%)	57 (13%)	..
Hepatic flexure	11 (2%)*	19 (4%)	..
Ascending colon	56 (12%)*	78 (17%)†	..
Cecum	19 (4%)*	55 (12%)†	..
Size‡			
≤5 mm	153 (33%)	203 (45%)	0·035
>5 to <10 mm	47 (10%)	70 (16%)	0·036
≥10 mm	38 (8%)	40 (9%)	0·83
Morphology§			
Pedunculated	13 (3%)	14 (3%)	0·76
Subpedunculated	29 (6%)	39 (9%)	0·55
Sessile	161 (35%)	214 (48%)	0·015
Flat elevated lesions	23 (5%)	35 (8%)	0·36
Broad-based nodule	2 (<1%)	6 (1%)	Not calculated¶
Barely perceptible elevation	6 (1%)	2 (<1%)	Not calculated¶
Depressed	0	0	0
Central depression	0	0	0
Excavated	0	0	0

Data are n/N (%). CAD=computer-aided detection. *Of 466 colonoscopies performed in these locations. †Out of a total of 448 coloscopies performed in these locations. ‡The size information for one polyp in the CAD treatment group was not reported. §The Paris classification information was not reported for four polyps in the CAD arm and four polyps in the conventional colonoscopy group. ¶Was not calculated due to low count.

Table 3: **Per-patient distribution of adenomas according to location, size, and morphology**

Although our study found a significant reduction in AMR, 19% of adenomas in the CAD-first group were still found to have been missed. Post-hoc video analysis of a tandem CAD study[26] found that the majority of missed adenomas during the CAD first procedure were not present in the visual field and, thus, could not have been detected by CAD. This might suggest that there could be a synergistic effect of combining both a mucosal exposure technique and CAD to further improve detection rates. Moreover, AMR in the non-iFOBT screening population was not statistically significant, whereas both the overall AMR and AMR in the surveillance population were. This difference is likely to be attributed to the relatively smaller sample size and the lower adenoma detection overall in the non-iFOBT screening population. Nevertheless, in the non-iFOBT screening group, APC and ADR were both significantly increased, suggesting that CAD might still have a

beneficial impact on the AMR in this population. In addition, although our study showed a relative 47% decrease in missed adenomas with the use of CAD, the AMR in the conventional colonoscopy first group was 36%, which is higher than the pooled miss rate of regular colonoscopy reported in a previous meta-analysis.[1] However, the AMR of both groups in our study was similar to the AMR in tandem studies using CAD in a similar population,[6,9] with a decrease in AMR from 32·4% to 15·5% in the first study and from 31·25% to 20·12% in the second study. Withdrawal times in these studies were longer or similar compared with our study, at 9·9 min (mean) and 6·5 min (median), respectively.

An advantage of our study is the inclusion of two additional tandem study groups alongside a traditional randomised parallel study design. As the parallel study groups and tandem study groups were powered to detect a significant effect in APC and APE, and AMR, respectively, the significant reduction in AMR was supported by the increased adenoma detection found in our parallel study groups. An additional strength of our study is the international, multicentre (university and non-university) setting and inclusion of experienced endoscopists with a baseline ADR of at least 25%. In addition, our study probably has a small risk of overfitting of the CAD device. It is known that systems using artificial intelligence are much more likely to operate accurately when used on similar data as their training and validation data sets. However, the CAD device in this study was largely developed with training and validation sets from non-participating hospitals and only two participating study sites provided training data before study start. This reduces the risk of potential overfitting of the CAD device in the population of this study.

This study has some limitations. First, this study had an unbalanced distribution across study sites, possibly partly attributed to delayed study initiation and halting of inclusions due to the COVID-19 pandemic. Consequently, this uneven distribution could potentially reduce generalisability. Notably, APC in the conventional colonoscopy group of two study sites was substantially higher compared with the APC in the CAD group of other study sites. Nonetheless, these two study sites also showed a similar increase in APC with CAD compared with other sites. It might be that the higher baseline APC at these sites might be explained by the higher BMI (appendix p 4), which is a known risk factor for colorectal adenoma formation.[27] Second, in this study, a subgroup of patients had tandem colonoscopy done by the same endoscopist to calculate AMR. Because the same endoscopist did the repeat procedure, the effect of the CAD device on miss rates could have been overestimated, as tandem studies are more likely to result in positive results than regular parallel trials.[28] However, our study also included sufficiently powered parallel study groups with a significant increase in adenoma detection, thereby probably decreasing the risk of overinterpretation of our reported AMR. Third, pathology slides were not reviewed by a second, independent, expert pathologist, which might explain the relatively low SSL detection rate among an increased detection of proximal lesions. However, diagnosing colonic SSLs remains challenging, as even experienced pathologists using expert panel recommendations have a moderate interobserver agreement in diagnosing SSLs.[29] Adequately powered studies will be required to properly estimate the effect of CAD on SSL detection. Fourth, our study lacks results on cost-effectiveness and long-term outcome results such as the risk of post-colonoscopy colorectal cancer or colorectal cancer-related and overall mortality. Previous studies have shown that an increased ADR is associated with a reduced risk of post-colonoscopy colorectal cancer.[30–32] However, it remains unclear whether the increased adenoma detection with CAD would have a similar effect on reducing the risk of post-colonoscopy colorectal cancer or colorectal cancer-related mortality, and subsequently, associated cost-effectiveness. Nevertheless, Markov Model simulations in two recent studies suggested potential cost-effectiveness of CAD.[33,34]

In conclusion, use of CAD increased adenoma detection in non-iFOBT screening and surveillance colonoscopies and reduced AMR compared with conventional colonoscopy. Notably, this novel CAD system increased the detection of both 6–9 mm and proximal adenomas, which might be of significant clinical relevance.

See Online for appendix

Contributors

MHJM contributed to the investigation, visualisation, data interpretation, writing of the original draft, and the review and editing of the manuscript. HN, HS, LHK, RH, and VS contributed to the resources and investigation. AAB, AK, MJL, BL, SKL, SN, and HJ contributed to the resources, investigation, and the review and editing of the manuscript. ES contributed to the conceptualisation and methodology. PDS contributed to the conceptualisation, methodology, resources, data interpretation, supervision, and the review and editing of the manuscript. MHJM and PDS accessed and verified the reported underlying data. All authors read and approved the final manuscript. MHJM and PDS had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

HN reports equipment from Magentiq Eye. AAB is a consultant for Boston Scientific Corp. SN is a consultant for Boston Scientific Corp, Olympus, and Neptune Medical. HJ reports consultancy fees from Magentiq Eye, and stock options from Magentiq Eye from activities before the clinical trial. PDS reports research support from Pentax, The eNose company, Lucid Diagnostics, Micro Tech, Motus GI, Magentiq Eye, Norgine, and Endo Tools Therapeutics, and consultancy fees from Motus GI and Magentiq Eye. All other authors declare no competing interests.

Data sharing

Individual de-identified patient data that underlie the results reported in this Article will be made available to investigators whose proposed use of the data has been approved by the corresponding author and with a signed data access agreement. Access to the data is granted solely for the purposes outlined in the approved proposal. For inquiries or proposals, please contact the corresponding author. Interested parties seeking access to or inquiring after the computer-aided detection system or its code can contact the manufacturer Magentiq Eye for further information.

Acknowledgments
We thank all the participating physicians, nurses, and auxiliary team members of each study site of this trial. We thank Oren Bar-Ilan from DataSights, Haifa, Israel, who did the formal data analysis. We thank all the patients who participated in this trial. Support for this study was provided by Magentiq Eye, Haifa, Israel.

References

1 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:** 1661–74.
2 Anderson R, Burr NE, Valori R. Causes of post-colonoscopy colorectal cancers based on World Endoscopy Organization system of analysis. *Gastroenterology* 2020; **158:** 1287–99.
3 Robertson DJ, Lieberman DA, Winawer SJ, et al. Colorectal cancers soon after colonoscopy: a pooled multicohort analysis. *Gut* 2014; **63:** 949–56.
4 Pohl H, Robertson DJ. Colorectal cancers detected after colonoscopy frequently result from missed lesions. *Clin Gastroenterol Hepatol* 2010; **8:** 858–64.
5 Hassan C, Spadaccini M, Iannone A, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021; **93:** 77–85.
6 Wallace MB, Sharma P, Bhandari P, et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. *Gastroenterology* 2022; **163:** 295–304.
7 Rondonotti E, Di Paolo D, Rosa Rizzotto E, et al. Efficacy of a computer aided detection (CADe) system in a FIT-based organized colorectal cancer screening program: a randomized controlled trial (AIFIT study). *Endoscopy* 2022; **54:** 1171–79.
8 Shaukat A, Lichtenstein DR, Somers SC, et al. Computer-aided detection improves adenomas per colonoscopy for screening and surveillance colonoscopy: a randomized trial. *Gastroenterology* 2022; **163:** 732–41.
9 Glissen Brown JR, Mansour NM, Wang P, et al. Deep learning computer-aided polyp detection reduces adenoma miss rate: a United States multi-center randomized tandem colonoscopy study (CADeT-CS Trial). *Clin Gastroenterol Hepatol* 2022; **20:** 1499–507.
10 Lai EJ, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. *Gastrointest Endosc* 2009; **69:** 620–25.
11 Schlemper RJ, Riddell RH, Kato Y, et al. The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* 2000; **47:** 251–55.
12 Endoscopic Classification Review Group. Update on the Paris classification of superficial neoplastic lesions in the digestive tract. *Endoscopy* 2005; **37:** 570–78.
13 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:** 1813–19.
14 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:** 343–51.
15 Karsenti D, Tharsis G, Perrot B, et al. Effect of real-time computer-aided detection of colorectal adenoma in routine colonoscopy (COLO-GENIUS): a single-centre randomised controlled trial. *Lancet Gastroenterol Hepatol* 2023; **8:** 726–34.
16 Repici A, Spadaccini M, Antonelli G, et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. *Gut* 2022; **71:** 757–65.
17 Xu H, Tang RSY, Lam TYT, et al. Artificial intelligence-assisted colonoscopy for colorectal cancer screening: a multicenter randomized controlled trial. *Clin Gastroenterol Hepatol* 2023; **21:** 33–46.
18 Ladabaum U, Shepard J, Weng Y, Desai M, Singer SJ, Mannalithara A. Computer-aided detection of polyps does not improve colonoscopist performance in a pragmatic implementation trial. *Gastroenterology* 2023; **164:** 481–83.
19 Levy I, Bruckmayer L, Klang E, Ben-Horin S, Kopylov U. Artificial intelligence-aided colonoscopy does not increase adenoma detection rate in routine clinical practice. *Am J Gastroenterol* 2022; **117:** 1871–73.
20 Kaminski MF, Thomas-Gibson S, Bugajski M, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. *United European Gastroenterol J* 2017; **5:** 309–34.
21 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:** 512–20.
22 Desai M, Rex DK, Bohm ME, et al. Impact of withdrawal time on adenoma detection rate: results from a prospective multicenter trial. *Gastrointest Endosc* 2023; **97:** 537–43.
23 Singh H, Nugent Z, Demers AA, Kliewer EV, Mahmud SM, Bernstein CN. The reduction in colorectal cancer mortality after colonoscopy varies by site of the cancer. *Gastroenterology* 2010; **139:** 1128–37.
24 Lakoff J, Paszat LF, Saskin R, Rabeneck L. Risk of developing proximal versus distal colorectal cancer after a negative colonoscopy: a population-based study. *Clin Gastroenterol Hepatol* 2008; **6:** 1117–21.
25 Anderson JC, Rex DK, Robinson C, Butterly LF. Association of small versus diminutive adenomas and the risk for metachronous advanced adenomas: data from the New Hampshire Colonoscopy Registry. *Gastrointest Endosc* 2019; **90:** 495–501.
26 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:** 1252–61.
27 Ben Q, An W, Jiang Y, et al. Body mass index increases risk for colorectal adenomas based on meta-analysis. *Gastroenterology* 2012; **142:** 762–72.
28 Zimmermann-Fraedrich K, Pohl H, Rösch T, et al. Designs of colonoscopic adenoma detection trials: more positive results with tandem than with parallel studies—an analysis of studies on imaging techniques and mechanical devices. *Gut* 2021; **70:** 268–75.
29 Vennelaganti S, Cuatrecasas M, Vennalaganti P, et al. Interobserver agreement among pathologists in the differentiation of sessile serrated from hyperplastic polyps. *Gastroenterology* 2021; **160:** 452–54.
30 Kaminski MF, Wieszczy P, Rupinski M, et al. Increased rate of adenoma detection associates with reduced risk of colorectal cancer and death. *Gastroenterology* 2017; **153:** 98–105.
31 Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. *N Engl J Med* 2010; **362:** 1795–803.
32 Schottinger JE, Jensen CD, Ghai NR, et al. Association of physician adenoma detection rates with postcolonoscopy colorectal cancer. *JAMA* 2022; **327:** 2114–22.
33 Areia M, Mori Y, Correale L, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. *Lancet Digit Health* 2022; **4:** e436–44.
34 Hassan C, Povero M, Pradelli L, Spadaccini M, Repici A. Cost-utility analysis of real-time artificial intelligent-assisted colonoscopy in Italy. *Endosc Int Open* 2023; **11:** E1046–55.