Evaluation of a real-time computer-aided polyp detection system during screening colonoscopy: AI-DETECT study

Authors
Ahmir Ahmad[1], Ana Wilson[1], Adam Haycock[1], Adam Humphries[1], Kevin Monahan[1], Noriko Suzuki[1], Siwan Thomas-Gibson[1], Margaret Vance[1], Paul Bassett[2], Kowshika Thiruvilangam[1], Angad Dhillon[3], Brian P. Saunders[1]

Institutions
1 Wolfson Unit for Endoscopy, St Mark's Hospital, London, United Kingdom
2 Statsconsultancy Ltd, Amersham, United Kingdom
3 Queen Elizabeth Hospital, Lewisham and Greenwich NHS Trust, London, United Kingdom

submitted 18.5.2022
accepted after revision 16.9.2022
published online 12.12.2022

Bibliography
Endoscopy 2023; 55: 313–319
DOI 10.1055/a-1966-0661
ISSN 0013-726X

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Corresponding author
Ahmir Ahmad, MRCP, Wolfson Unit for Endoscopy, St Mark's Hospital, Watford Road, Harrow, London, HA1 3UJ, United Kingdom
ahmir.ahmad@nhs.net

Supplementary material
Supplementary material is available under https://doi.org/10.1055/a-1966-0661

Scan this QR-Code for the author commentary.

ABSTRACT

Background Polyp detection and resection during colonoscopy significantly reduce long-term colorectal cancer risk. Computer-aided detection (CADe) may increase polyp identification but has undergone limited clinical evaluation. Our aim was to assess the effectiveness of CADe at colonoscopy within a bowel cancer screening program (BCSP).

Methods This prospective, randomized controlled trial involved all eight screening-accredited colonoscopists at an English National Health Service (NHS) BCSP center (February 2020 to December 2021). Patients were randomized to CADe or standard colonoscopy. Patients meeting NHS criteria for bowel cancer screening were included. The primary outcome of interest was polyp detection rate (PDR).

Results 658 patients were invited and 44 were excluded. A total of 614 patients were randomized to CADe (n = 308) or standard colonoscopy (n = 306); 35 cases were excluded from the per-protocol analysis due to poor bowel preparation (n = 10), an incomplete procedure (n = 24), or a data issue (n = 1). Endocuff Vision was frequently used and evenly distributed (71.7 % CADe and 69.2 % standard). On intention-to-treat (ITT) analysis, there was a borderline significant difference in PDR (85.7 % vs. 79.7 %; $P = 0.05$) but no significant difference in adenoma detection rate (ADR; 71.4 % vs. 65.0 %; $P = 0.09$) for CADe vs. standard groups, respectively. On per-protocol analysis, no significant difference was observed in these rates. There was no significant difference in procedure times.

Conclusions In high-performing colonoscopists in a BCSP who routinely used Endocuff Vision, CADe improved PDR but not ADR. CADe appeared to have limited benefit in a BCSP setting where procedures are performed by experienced colonoscopists.

Introduction

Detection of polyps during colonoscopy is critical in reducing colorectal cancer-related morbidity and mortality. Screening programs are particularly effective at reducing these risks, with high-performing operators undertaking colonoscopy. Colonoscopists with a higher adenoma detection rate (ADR) more effectively protect patients from development of post-colonoscopy colorectal cancer (PCCRC) compared with colonoscopists with low detection rates [1].

Among endoscopists, there is a recognized polyp miss rate of 26 % for adenomas and 27 % for serrated lesions, according to a recent systematic review and meta-analysis of tandem colonoscopy studies [2]. This may, to some extent, explain the significant level of PCCRC occurring within 36 months of a

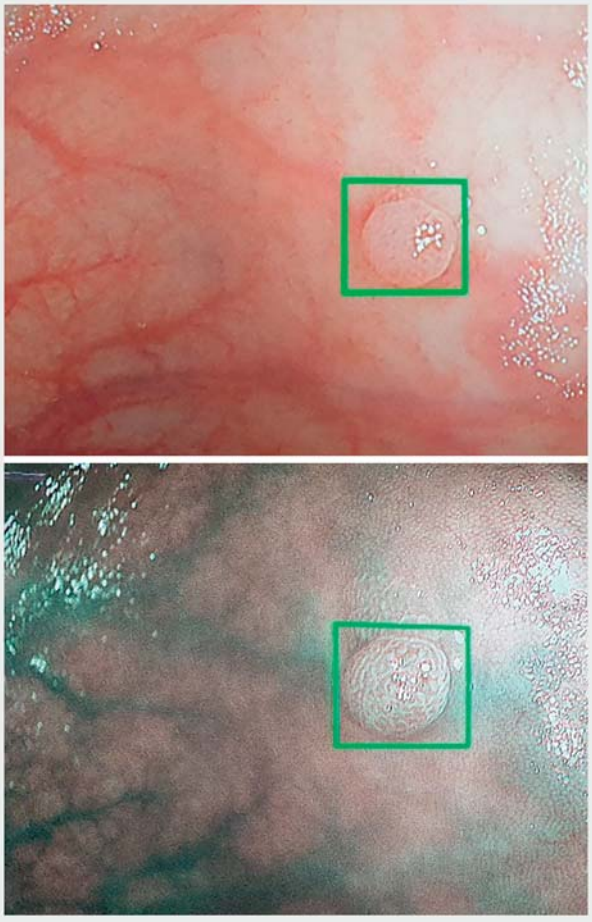
"negative" index colonoscopy [3, 4]. However, within a bowel cancer screening program (BCSP) setting, PCCRC 3-year rates have been shown to be lower, at 3.6 % compared with an overall unadjusted rate of 6.5 % [3]. This may be due to level of experience and technical skill of colonoscopists participating in BCSPs.

Human detection of polyps has limitations due to issues such as the amount of time spent cleaning the mucosa when bowel preparation is inadequate, optical diagnostic skills for recognition of subtle flat polyps, and fatigue or distraction resulting in human error.

In recent decades, there have been significant advances in artificial intelligence (AI)-based systems, which use algorithms to perform tasks that would usually require human intelligence and input [5]. These algorithms can be trained to perform tasks by recognizing patterns in data (machine learning) rather than being programed. Using this technology, computer-aided polyp detection systems (CADe) have been developed to automatically highlight detected polyps in real time during colonoscopy.

AI systems might reduce the degree of variability in polyp detection among endoscopists and could help reduce human error. However, prospective studies in a real-life clinical setting are lacking; therefore, studies are required to assess effectiveness and acceptability of these novel technologies before widespread clinical use can be recommended.

The aim of this study was to evaluate the first clinically available polyp detection system (GI Genius; Medtronic, Minneapolis, Minnesota, USA) in terms of effectiveness in real-life clinical practice among a group of experienced colonoscopists performing bowel cancer screening colonoscopies.



▶ **Fig. 1** Output of the computer-aided detection system (GI Genius; Medtronic, Minneapolis, Minnesota, USA) in the event of polyp detection.

Methods

Study design

This prospective, randomized controlled trial was conducted between February 2020 and December 2021 and involved all eight screening colonoscopists at an English National Health Service (NHS) Bowel Cancer Screening Centre (London North West University Healthcare NHS Trust). In the parallel design, patients were randomized to either the GI Genius system in the CADe arm or standard colonoscopy in the control group, with a 1:1 allocation ratio. Pediatric and adult high definition colonoscopes were used, with colonoscopists free to use Endocuff Vision (Olympus, Tokyo, Japan) or a transparent plastic cap (Olympus) at their discretion. Post-procedure histology results were reviewed within 2 weeks.

In the CADe arm, the first commercially available GI Genius system was used (product code CB1708-EU). In practice this involved attaching a box (module) to the endoscopy stack to integrate the system into the existing set-up; there were no other onsite/offsite requirements for usage. The system input was the real-time video display on the standard colonoscopy video monitor. CADe was switched on immediately before scope insertion until procedure completion, and highlighted possible polyp detections to the colonoscopist in real time with an output of green boxes automatically superimposed on the colonoscopy monitor screen over the area of interest (▶ **Fig.1**). This provided an opportunity for colonoscopists to evaluate the "polyp detections" and to undertake polypectomy where appropriate. The experienced colonoscopists used the CADe outputs as an adjunct to their normal colonoscopy practice. Ultimately, the colonoscopist was responsible for making decisions relating to CADe detections. There were no issues with poor quality or unavailable input data, as the system fully integrated with the existing system set-up.

Inclusion criteria were patients aged 60–74 years with a positive fecal immunochemical test attending for screening colonoscopy within the NHS BCSP or with an established history of adenomas attending for surveillance colonoscopy within the BCSP. In addition, patients aged 55 years were included if they were referred for colonoscopy due to large or multiple adenomas being found during screening flexible sigmoidoscopy.

Exclusion criteria were patients with a risk profile (due to family history or other reasons), whose follow-up was conducted outside the BCSP, and those who did not give consent to the study.

At the level of input data, all cases in which polyp datasets were recorded during the colonoscopy were included. Input data were excluded from the analysis if there was poor bowel preparation to the extent that the clinician felt a repeat colonoscopy was required, cases with incomplete polyp datasets, and where procedures were incomplete (e.g. cecum not reached due to a malignant stricture).

There were no important changes to methods after trial commencement.

Randomization

Patients were block randomized with each clinic list considered a block. The blocks were of size 4 or 6, depending on the size of the clinic list. The randomization was generated using a computer-generated list produced by the study statistician to either CADe or standard colonoscopy (control) in a 1:1 ratio. Patients were enrolled by a dedicated research nurse who assigned the interventions based on the randomization list, which was not accessible to colonoscopists. The randomization sequence was therefore not known to colonoscopists until the intervention had been assigned. There was no colonoscopist or patient blinding.

Outcomes

The primary outcome of interest was the polyp detection rate (PDR), which was defined as the number of patients with at least one polyp identified divided by the total number of colonoscopies performed. As per our usual practice, typical-appearing, small, shiny, hyperplastic, rectosigmoid polyps were left in situ and not included in the assessment.

The secondary outcomes were adenoma detection rate (ADR), sessile serrated lesion (SSL) detection rate, "significant polyp" detection rate (adenoma + SSL), polyps per colonoscopy (total number of polyps divided by the total number of colonoscopies), adenomas per colonoscopy (APC), and serrated polyps per colonoscopy. The impact of CADe on procedure times was assessed, including insertion (intubation to cecum), withdrawal (cecum to extubation), and total (intubation to extubation) times. We also calculated the SP6 score (number of adenomas and SSLs detected per 6-minute withdrawal time at colonoscopy) to assess the impact of CADe on efficiency of polyp detection and management [6].

There were no changes to trial outcomes after the trial commenced. The trial was stopped after the recruitment target had been achieved. We do not report on the technical performance of the AI system, as this was not the aim of the trial.

Statistical analysis

This study was powered to detect a 10 % increase in polyp detection, from a detection rate of 20 % in the control group up to 30 % in the CADe group. With a 5 % significance level and a power of 80 %, 294 patients in each group, 588 patients in total, were required.

All analyses were compared between the two study groups. Demographic characteristics of the two groups were compared descriptively. All continuous outcomes were found to have positively skewed distributions and were compared between groups using the Mann–Whitney test. Categorical outcomes were compared between groups using the chi-squared test. Intention to treat (ITT) and per-protocol analyses were performed.

The study was reported according to CONSORT-AI guidelines (see the online-only Supplementary material) [7].

Results

Patients

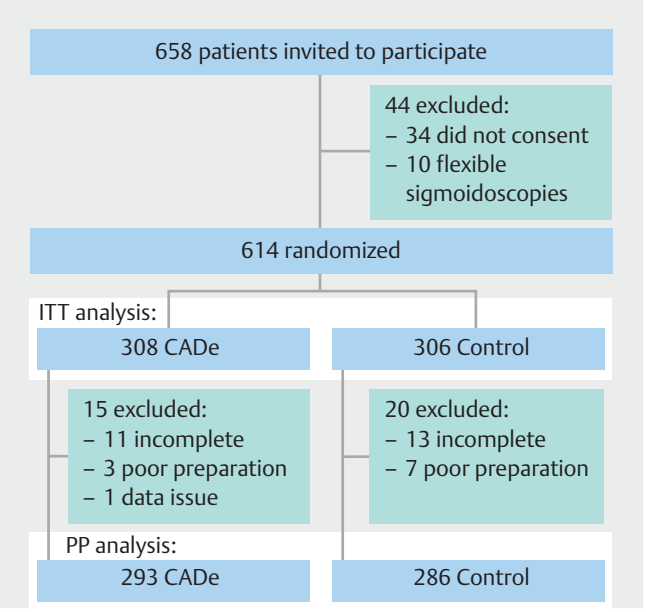
A total of 658 patients were invited, of whom 614 were randomized for the primary analysis (ITT) to CADe (n = 308) or control (n = 306) (▶ **Table 1**, ▶ **Fig. 2**). There were 35 exclusions post-randomization (24 incomplete procedures, 10 poor preparation requiring repeat procedure, and 1 data issue). In the per-protocol analysis, there were 579 patients, with 293 cases in the CADe group and 286 cases in the control group.

In the CADe group, Endocuff Vision was used in 71.7 % (210/293) of cases and a cap was used in 2.0 % (6/293) of cases (no

▶ **Table 1** Baseline characteristics of patients.

	Total	Control	CADe
Patients, n	614	306	308
Sex, n (%)			
▪ Male	208 (33.9)	98 (32.0)	110 (35.7)
▪ Female	406 (66.1)	208 (68.0)	198 (64.3)
▪ Age, mean (SD), years	66.3 (5.4)	66.4 (5.4)	66.2 (5.4)

CADe, computer-aided detection.



▶ **Fig. 2** Study overview. CADe, computer-aided detection; ITT, intention-to-treat; PP, per-protocol.

▶ **Table 2** Patient-level procedure outcomes with and without computer-aided detection (intention-to-treat analysis).

Outcome	All patients	Control	CADe	*P* value
Total procedures, n	614	306	308	
Total polyps, n	2104	1001	1103	
▪ Adenomas	1378	654	724	
▪ Serrated polyps	528	263	265	
– SSLs	242	115	127	
– Hyperplastic polyps	286	148	138	
▪ Inflammatory	24	10	14	
▪ Normal	124	54	70	
▪ Other	33	15	18	
▪ Left in situ	12	2	10	
▪ Not retrieved	5	3	2	
Polyps per colonoscopy, mean (SD)				
▪ Total	3.4 (3.5)	3.3 (3.3)	3.6 (3.7)	0.23
▪ Adenomas	2.2 (2.8)	2.1 (2.6)	2.4 (2.9)	0.25
▪ Serrated polyps	0.9 (1.4)	0.9 (1.4)	0.9 (1.4)	0.93
▪ SSLs	0.4 (0.9)	0.4 (0.9)	0.4 (0.9)	0.40
▪ Hyperplastic polyps	0.5 (0.9)	0.5 (0.9)	0.4 (0.9)	0.45
▪ Other (inflammatory and normal mucosa)	0.3 (0.7)	0.3 (0.6)	0.3 (0.7)	0.06
Detection rates, %				
▪ Polyp	82.7	79.7	85.7	0.05
▪ Significant polyp[1]	75.4	71.6	79.2	0.03
▪ Adenoma	68.2	65.0	71.4	0.09
▪ SSL	23.1	21.6	24.7	0.36
SP6[2], mean (SD)	0.9 (0.8)	0.8 (0.9)	0.9 (0.8)	0.10
Procedure times, median (IQR), minutes				
▪ Total procedure time	24.7 (19.0–32.3)	24.3 (18.5–32.0)	24.9 (19.7–32.5)	0.18
▪ Insertion time	7.3 (5.5–10.0)	7.3 (5.4–9.9)	7.3 (5.7–10.0)	0.43
▪ Withdrawal time	14.5 (9.6–21.1)	13.9 (9.7–20.9)	14.9 (9.5–21.4)	0.34
Cecal intubation rate, n (%)	590 (96.1)	294 (96.1)	296 (96.1)	0.99

CADe, computer-aided detection; SSL, sessile serrated lesions; IQR, interquartile range.
[1] Adenoma + SSL.
[2] Number of adenomas and SSLs detected per 6-minute withdrawal time at colonoscopy.

adjuncts were used in the remaining cases). In the control group, Endocuff Vision was used in 69.2% (198/286) of cases and a cap was used in 3.5% (10/286) of cases. There were no adverse or unintended effects in the groups.

Demographic characteristics of included patients were similar in the two groups (▶ **Table 1**).

Procedure outcomes

On ITT analysis (▶ **Table 2**), a total of 2104 polyps were identified, with a mean of 3.3 (SD 3.3) polyps per colonoscopy in the control group and 3.6 (SD 3.7) in the CADe group ($P=0.23$). There was a borderline statistically significant increase in PDR with CADe (85.7%) compared with the control group at (79.7%; $P=0.05$). There was no statistically significant difference in ADR or APC between the CADe and control groups (ADR 71.4% vs. 65.0%, $P=0.09$; APC 2.4 [SD 2.9] vs. 2.1 [SD 2.6], $P=0.25$) or

▶ **Table 3** Polyp characteristics (intention-to-treat analysis).

Outcome	All patients, n (%)	Control, n (%)	CADe, n (%)
Total polyps, n	2104	1001	1103
Paris classification, n (%)			
▪ Is	1107 (52.6)	528 (52.7)	579 (52.5)
▪ Isp	83 (3.9)	45 (4.5)	38 (3.4)
▪ Ip	139 (6.6)	56 (5.6)	83 (7.5)
▪ IIa	729 (34.7)	350 (35.0)	379 (34.4)
▪ IIb	32 (1.5)	16 (1.6)	16 (1.5)
▪ IIc	5 (0.2)	4 (0.4)	1 (0.1)
▪ III	0 (0.0)	0 (0.0)	0 (0.0)
▪ LST-G	5 (0.2)	1 (0.1)	4 (0.4)
▪ LST-NG	4 (0.2)	1 (0.1)	3 (0.3)
Site of polyps, n (%)			
▪ Cecum	257 (12.2)	108 (10.8)	149 (13.5)
▪ Ascending colon	439 (20.9)	208 (20.8)	231 (20.9)
▪ Hepatic flexure	89 (4.2)	38 (3.8)	51 (4.6)
▪ Transverse colon	465 (22.1)	233 (23.3)	232 (21.0)
▪ Splenic flexure	78 (3.7)	35 (3.5)	43 (3.9)
▪ Descending colon	195 (9.3)	96 (9.6)	99 (9.0)
▪ Sigmoid colon	357 (17.0)	177 (17.7)	180 (16.3)
▪ Rectosigmoid junction	10 (0.5)	6 (0.6)	4 (0.4)
▪ Rectum	214 (10.2)	100 (10.0)	114 (10.3)
Polyp size, n (%)			
▪ 1–5 mm	1611 (76.6)	765 (76.4)	846 (76.7)
▪ 6–9 mm	279 (13.3)	142 (14.2)	137 (12.4)
▪ 10 + mm	214 (10.2)	94 (9.4)	120 (10.9)

CADe, computer-aided detection; LST, laterally spreading tumor; G, granular-type; NG, nongranular type.

SSL detection rate between the groups. However, the significant polyp (adenoma + SSL) detection rate was statistically significantly higher in the CADe group than in the control group (79.2 % vs. 71.6 %; $P = 0.03$) and this implies a small but clinically relevant benefit of CADe. There was also no difference in procedure times (total, insertion, and withdrawal) or SP6 between the groups.

On per-protocol analysis (Table 1 s), 2089 polyps were identified, with a mean of 3.5 (SD 3.3) polyps per colonoscopy in the control group and 3.7 (SD 3.7) in the CADe group ($P = 0.43$). In contrast to the ITT analysis, there was no statistically significant difference in PDR, ADR, or significant polyp detection rate between the groups. As with the ITT analysis, there was no difference in procedure times (total, insertion, and withdrawal) or SP6 between the groups.

Polyp characteristics

In patients with detected polyps, there was no difference in the distribution, Paris classification, or polyp size (▶ **Table 3**, Table 2 s).

Discussion

In this prospective randomized study among high performing colonoscopists within a BCSP setting, polyp detection was marginally increased when using CADe but there was no increase in ADR, or detection of flat and diminutive polyps. CADe did not adversely impact procedure times.

A systematic review including five randomized controlled trials (4354 patients) showed a significantly higher pooled ADR in the CADe group compared with the control group

(36.6 % vs. 25.2 %; $P < 0.01$) [8]. The CADe system used in these studies varied and therefore the results may not be generalizable to all CADe systems (only one of the five systems used was GI Genius). For example, in one study, a "real-time automatic quality control system" was used, which provided feedback on withdrawal stability, bowel preparation, and polyp detection [9]. In combination, these metrics resulted in an increase in ADR but the specific impact of the polyp detection aspect of this system could not be fully assessed. Caution should therefore be taken when making generalizations about CADe from systematic reviews where there is significant heterogeneity in the data and systems used.

The AID-1 study used GI Genius as the CADe system [10]. In this study, 685 patients undergoing colonoscopy performed by six experienced colonoscopists (>2000 screening colonoscopies) at three centers were randomized to procedures with or without CADe. ADR and mean APC were significantly higher in the CADe group (ADR 54.8 % vs. 40.4 %; APC 1.07 [SD 1.54] vs 0.71 [SD 1.20]). Small adenomas up to 9 mm were detected in a higher proportion of patients with CADe than in the control group. These findings contrast with our study, which found no statistically significant difference in ADR or APC between the CADe and control groups (ADR 71.4 % vs. 65.0 %, $P = 0.09$; APC 2.4 [SD 2.9] vs. 2.1 [SD 2.6], $P = 0.25$), and no difference when size or morphology of the polyps was assessed.

The only other published Western, randomized trial using GI Genius was a multicenter, randomized, controlled, noninferiority trial (AID-2) [11]. This differed from AID-1 as it involved 10 nonexpert endoscopists (<2000 colonoscopies) from five centers performing colonoscopy in 660 patients. With CADe, ADR increased by 22 % compared with the control group (53.3 % vs. 44.5 %) and APC increased by 21 % (1.26 [SD 1.82] vs. 1.04 [SD 1.75]). Despite the lower experience level of endoscopists participating in AID-2, the ADR and APC findings were similar between the studies. A post hoc analysis pooling both studies showed ADR improved by 29 % with CADe and endoscopist experience did not have a significant effect on ADR.

Both the AID-1 and AID-2 studies showed no difference in withdrawal time, which concurs with our finding that CADe does not adversely lengthen the procedure.

A more recent multicenter, randomized, back-to-back, tandem colonoscopy study evaluated adenoma miss rate, defined as the number of histologically verified lesions detected at second colonoscopy divided by the total number of lesions detected at first and second colonoscopy [12]. This study showed an adenoma miss rate of 15.5 % where CADe was the first colonoscopy and 32.4 % where standard colonoscopy was performed first. This lower miss rate with CADe was thought to be due to a reduction in missed flat and small lesions. This is contrary to our study, which found no difference in the morphology or size of polyps detected with or without CADe. Although this study presents convincing data for a benefit with CADe, tandem studies are unblinded, open to bias, and do not represent usual clinical practice.

Several studies have assessed the use of alternative CADe systems within a screening setting but have not incorporated the use of Endocuff Vision [13–15]. Shaukat et al. assessed the SKOUT CADe device (Iterative Scopes, Cambridge, Massachusetts, USA) in a randomized study with 1359 patients included in the analysis, and showed a significant improvement in APC when using CADe (0.83 vs. 1.05; $P = 0.002$) for screening and surveillance colonoscopies [13]. However, there was no significant difference in ADR (43.9 % vs. 47.8 %; $P = 0.065$). In a randomized study with 800 patients, the CAD EYE (Fujifilm, Tokyo, Japan) CADe system was assessed within a fecal immunochemical test-based colorectal screening setting and was found to significantly increase ADR (45.3 % vs. 53.6 %) and APC (0.90 vs. 1.13; $P = 0.028$) [14]. In another randomized study evaluating AI-assisted colonoscopy, a significant improvement in PDR and ADR was observed [15]. Although CADe systems may on the surface appear to be similar in terms of their outputs, difference in underlying AI algorithms between devices may contribute to variation in findings between studies.

A number of factors might explain why we did not find stronger evidence of a difference in polyp detection. The high baseline polyp detection performance within this group of experienced colonoscopists may have limited the potential for CADe to impact on outcomes. With an overall PDR of 82.7 % in the ITT analysis and 86.0 % in the per-protocol analysis, there is a "ceiling effect" with little room left for improvement in the intervention group. Previous studies have shown improvements in ADR, when CADe is used even among experienced colonoscopists who had performed >1000 [2], >2000 [10], or >5000 [9] colonoscopies. In our study, unlike others, Endocuff Vision was used in the majority of procedures, which may improve visibility and detection, minimizing any potential advantage from additional use of CADe systems [16].

The key strength of this study is the randomized design. In addition, evaluation of performance in a homogeneous group of accredited BCSP colonoscopists may support generalizability of the results in a screening setting, although further studies are required. As Endocuff Vision was used widely by colonoscopists in this study, as part of their usual practice, the CADe system had a higher polyp detection threshold to exceed to show a statistically significant improvement. If Endocuff Vision had not been used, a larger improvement in polyp detection might have been demonstrated with CADe. However, following the ADENOMA study, which showed significant improvement in ADR with Endocuff Vision, our standard practice has been to use Endocuff Vision within the bowel cancer screening setting [16].

The sample size calculation was originally based on a mixed cohort of patients, in which a relatively low PDR was expected. Thus, the observed PDR in the study was much higher than that assumed in the sample size calculation. The power calculation was based on a PDR of 25 % in the two groups combined, a 10 % group difference, and used an 80 % power. The observed PDR in the study was approximately 80 % in the two groups combined. With the same sample size (n = 588), the study would have a higher power of 86 % to show a 10 % difference (e. g. 75 % vs. 85 % PDR) between groups. Thus, although the assumptions made in the original calculation were not met, this is unlikely to have inversely impacted on the power of the study.

As with other similar studies we were unable to blind colonoscopists to the use of CADe. This risks introducing observer

bias whereby endoscopists are more attentive to mucosal visualization in procedures where CADe is used.

Further studies should assess the impact of CADe systems among endoscopists with a low PDR and with endoscopists undergoing training.

Conclusion

In this comparison of a CADe system with standard colonoscopy performed in a BCSP setting, a borderline statistically significant difference in PDR was observed with CADe in the ITT analysis. However, there was no increase in ADR and no significant detection differences in the per-protocol analysis. CADe therefore appears to have limited benefit in the screening setting and may prove most effective for low polyp detectors and those undergoing training outside screening programs.

Acknowledgments

We would like to acknowledge the support of the entire team at the St Mark's Hospital Bowel Cancer Screening Centre.

Competing interests

A. Ahmad has received research funding and equipment loan from Olympus. N. Suzuki has undertaken consultancy work for Olympus. S. Thomas-Gibson has received speaker fees from Olympus. B.P. Saunders has undertaken consultancy work and received speaker fees, research funding, and equipment loan from Olympus; and speaker fees and equipment loan from Fuji. A. Wilson, A. Haycock, A. Humphries, K. Monahan, M. Vance, P. Bassett, K. Thiruvilangam, and A. Dhillon declare that they have no conflict of interest.

References

[1] 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

[2] 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

[3] Burr NE, Derbyshire E, Taylor J et al. Variation in post-colonoscopy colorectal cancer across colonoscopy providers in English National Health Service: population based cohort study. BMJ 2019; 367: l6090

[4] Rutter MD, Beintaris I, Valori R et al. World Endoscopy Organization consensus statements on post-colonoscopy and post-imaging colorectal cancer. Gastroenterology 2018; 155: 909–925

[5] Pannala R, Krishnan K, Melson J et al. Artificial intelligence in gastrointestinal endoscopy. VideoGIE 2020; 5: 598–613

[6] Rameshshanker R, Saunders BP. Number of significant polyps detected per six-minute withdrawal time at colonoscopy (SP6): a new measure of colonoscopy efficiency and quality. Frontline Gastroenterol 2020; 11: 491–493

[7] Liu X, Cruz Rivera S, Moher D et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020; 26: 1364–1374

[8] 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

[9] 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

[10] Repici A, Badalamenti M, Maselli R et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology 2020; 159: 512–520

[11] Repici A, Spadaccini M, Antonelli G et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. Gut 2022; 71: 757–765

[12] Wallace MB, Sharma P, Bhandari P et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. Gastroenterology 2022; 163: 295–304

[13] 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: 732–741

[14] 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: doi:10.1055/a-1849-6878

[15] Pesatori EV, Milluzzo SM, Cesaro P et al. Incremental yield of artificial intelligence in a fecal occult blood test based organized screening population program. Gastrointest Endosc 2022; 95: AB243–AB244

[16] Ngu WS, Bevan R, Tsiamoulos ZP et al. Improved adenoma detection with Endocuff Vision: the ADENOMA randomised controlled trial. Gut 2019; 68: 280–288