Review

ANNALS OF MEDICINE & SURGERY

OPEN

The effectiveness of real-time computer-aided and quality control systems in colorectal adenoma and polyp detection during colonoscopies: a meta-analysis

Muhammad Fawad Aslam, MBBS, MRCP(UK)[a], Shehar Bano, MBBS[d], Mariam Khalid, MBBS[b], Zouina Sarfraz, MBBS[d,*], Azza Sarfraz, MBBS[c], Muzna Sarfraz, MBBS[e], Karla Robles-Velasco, MD[f], Miguel Felix, MD[f,g], Kitson Deane, MD[h], Ivan Cherrez-Ojeda, MD, MSc[e]

Aims: This meta-analysis aims to quantify the effectiveness of artificial intelligence (AI)-supported colonoscopy compared to standard colonoscopy in adenoma detection rate (ADR) differences with the use of computer-aided detection and quality control systems. Moreover, the polyp detection rate (PDR) intergroup differences and withdrawal times will be analyzed.

Methods: This study was conducted adhering to PRISMA guidelines. Studies were searched across PubMed, CINAHL, EMBASE, Scopus, Cochrane, and Web of Science. Keywords including the following 'Artificial Intelligence, Polyp, Adenoma, Detection, Rate, Colonoscopy, Colorectal, Colon, Rectal' were used. Odds ratio (OR) applying 95% CI for PDR and ADR were computed. SMD with 95% CI for withdrawal times were computed using RevMan 5.4.1 (Cochrane). The risk of bias was assessed using the RoB 2 tool.

Results: Of 2562 studies identified, 11 trials were included comprising 6856 participants. Of these, 57.4% participants were in the AI group and 42.6% individuals were in in the standard group. ADR was higher in the AI group compared to the standard of care group (OR = 1.51, $P = 0.003$). PDR favored the intervened group compared to the standard group (OR = 1.89, $P < 0.0001$). A medium measure of effect was found for withdrawal times (SMD = 0.25, $P < 0.0001$), therefore with limited practical applications.

Conclusion: AI-supported colonoscopies improve PDR and ADR; however, no noticeable worsening of withdrawal times is noted. Colorectal cancers are highly preventable if diagnosed early-on. With AI-assisted tools in clinical practice, there is a strong potential to reduce the incidence rates of cancers in the near future.

Keywords: adenoma, colorectal, meta-analysis, polyps, trials, withdrawal time

Departments of [a]Medicine, [b]Department of Medical Education, Rai Medical College, Sargodha, [c]Department of Pediatrics, Aga Khan University, Karachi, [d]Department of Research and Publications, Fatima Jinnah Medical University, [e]Department of Medicine, King Edward Medical University, Lahore, Pakistan, [f]Universidad Espíritu Santo, Samborondón, Ecuador, [g]Department of Medicine, New York City Health + Hospitals, Lincoln, Bronx and [h]Woodhull Medical and Mental Health Center, New York City, New York, USA

M.F.A. and S.B. are co-first authors.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding author. Address: Department of Research and Publications, Fatima Jinnah Medical University, Lahore 54000, Pakistan. Tel: +92 42 99200572; fax: +92 (42) 99203716. E-mail address: zouinasarfraz@gmail.com (Z. Sarfraz).*

Annals of Medicine & Surgery (2023) 85:80–91

Received 21 September 2022; Accepted 8 December 2022

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.annalsjournal.com

Published online 1 February 2023

http://dx.doi.org/10.1097/MS9.0000000000000079

HIGHLIGHTS

1. Artificial intelligence systems applied to standard colonoscopy procedures improve adenoma and polyp detection rates.
2. This meta-analysis adds to a limited literature set.
3. It promotes the use of novel diagnostic technologies for use in clinical practice.

Introduction

Colorectal cancer (CRC) is the third most common malignancy and the second most cause of cancer-related mortality worldwide[1]. An estimated 1.93 million cases of CRC were diagnosed in 2020 with 0.94 million deaths. The global burden of CRC is predicted to reach 3.2 million by 2040[1]. While the incidence of CRC is rising in high-income countries, low-income and middle-income countries are showing a high rise in reported cases[1]. The highest estimates of new CRC cases are expected to arise in China and the United States over the next 20 years[1]. Given the rising burden of CRC, colorectal polyp and adenoma screening strategies using artificial intelligence (AI)-supported routine testing may improve the timely detection of CRC[2,3]. Colorectal polyps and adenomas, which are diagnosed on

colonoscopy, are benign glandular tumors of the colon and rectum. They are precursor lesions of colorectal adenocarcinoma (CRC)[4]. The American Society for Gastrointestinal Endoscopy (ASGE) Technology Committee in the Preservation and Incorporation of Valuable endoscope Innovations (PIVI) statement endorses the optical evaluation of colorectal adenomas that present as polyps as acceptable if the endoscopist reaches a threshold of greater than or equal to 90% agreement with histopathology results and greater than or equal to 90% negative predictive value for the diagnosis of adenomatous histology[5,6]. However, despite the high accuracy of optical testing strategies for colorectal polyps, endoscopists have been reluctant due to concerns such as inappropriate surveillance intervals, incorrect diagnoses, and related medicolegal concerns[7].

To overcome the shortcomings in colonoscopies, this study is directed at providing the efficacy of AI-integration into coloscopies in clinical use. AI-assisted colonoscopy diagnostic systems for detection (CADe) or colorectal adenomas are a major area of research and implementation of AI in clinical practice. It must be ascertained whether AI can provide real-time support to physicians as well suggesting probable histology and confidence levels. This meta-analysis aims to quantify primarily, adenoma detection rates (ADR) differences with the use of CADe and quality control systems in adult patients compared to standard methods in colonoscopy practice. Secondarily, polyp detection rate (PDR) intergroup differences and withdrawal times will be analyzed to assess the effectiveness of AI systems paired with coloscopy. We predict that AI-assisted colonoscopy improves the ADR, PDR, and withdrawal times and allows for operator-independent pathology predictions.

Methods

This meta-analysis was conducted adhering to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[8]. The PRISMA-P checklist is appended in Supplementary Materials, Supplemental Digital Content 1, http://links.lww.com/MS9/A10. In accordance with the preidentified objectives, this paper sought to assess the effectiveness of AI systems applied to coloscopies in ADR, PDR, and withdrawal times. PROSPERO 2022 CRD42022333731[9].

Search strategy

Studies were searched across the following databases: (i) PubMed, (ii) CINAHL, (iii) EMBASE, (iv) Scopus, (v) Cochrane, and (vi) Web of Science until April 21, 2022. An additional search of clinicaltrials.gov was conducted to enlist any ongoing trials in this area. The search strategies were created by combining keywords and applying the Boolean logic. The following keywords were utilized: Artificial Intelligence, Polyp, Adenoma, Detection, Rate, Colonoscopy, Colorectal, Colon, Rectal. The full electronic search strategy is given under Supplementary Materials, Supplemental Digital Content 2, http://links.lww.com/MS9/A11. No date or language restrictions were applied; any non-English study was translated into English using Google Translate. The results were exported into EndNote X9 (Clarivate Analytics). An umbrella methodology was applied where the reference lists of examined studies were also examined.

Study selection (inclusion and exclusion)

In this meta-analysis, we included controlled randomized controlled trials that included an intervened group (IG) (with the AI system), and a standard group (SG) (regular colonoscopy procedure). The former group was termed IG, the intervention group, and the former was termed SG. If the trial did not have a control group (i.e. a single arm), it was removed regardless of the outcomes. there were no specific prerequisites for patient age, sex, study sample, follow-up, or geographics. Any study that utilized AI in an IG and SG with enlisting of any (ADR, PDR, withdrawal time) outcome was included. However, studies that were non-clinical in nature (theoretical modeling), had no control groups or did not apply an AI-based model were removed. In case more than one study was published by the same author group, all of them were included to ensure adequate computation of current data. As required, the authors of the trials were contacted to clarify missing or incomplete data.

Data extraction and synthesis

Two authors (Z.S. and A.S.) extracted study-related outcomes into a shared spreadsheet together with the final author present for any disagreements (I.C.-O.). Any disagreements were resolved by active discussion. The data were extracted as author, year, country, title, domain, methodology, technology specifications, ADR/PDR/withdrawal time in intervened and SGs, overall findings, and conclusions. The summary estimates were either provided as odds ratio (OR) with 95% CI or as SMD with 95% CI. The heterogeneity between the included studies was assessed using the I^2 index. The findings of the meta-analysis were illustrated as forest plots. Publication bias was assessed only if the total number of studies was higher than 10. The findings were illustrated in the form of a funnel plot. The data analysis was conducted using Review Manager 5.4.1 (Cochrane).

Risk of bias assessment

Version 2 of the Cochrane risk of bias tool for randomized trials (RoB 2) tool was used to assess the risk of bias. The tool is structured to assess biases focused on trial design, reporting, and conduct. The judgment is made based on low, high, or unclear risks of bias in each domain. The study-by-study findings are depicted in Figure 2, with an overview of every risk of bias item expressed as percentages across all included studies in Figure 3. Based on AMSTAR 2 assessment, this study is a high-quality review (Supplementary Materials)[10], Supplemental Digital Content 3, http://links.lww.com/MS9/A12

Results

Search process

The PRISMA flow diagram is depicted in Figure 1. During the identification phase, 2562 studies were identified from all databases. Of the identified studies, 241 were duplicates. In the screening phase, 2285 studies were screened and 2249 studies were removed as neither the titles nor the abstracts met the inclusion criteria. During the full-text screening, 36 studies were assessed for eligibility. Of these, 22 were removed as they did not provide ADR data or there were no comparators, or they were cohorts. Three studies met the exclusion criteria. In the inclusion phase, 11 studies were added to the meta-analysis (Fig. 1).

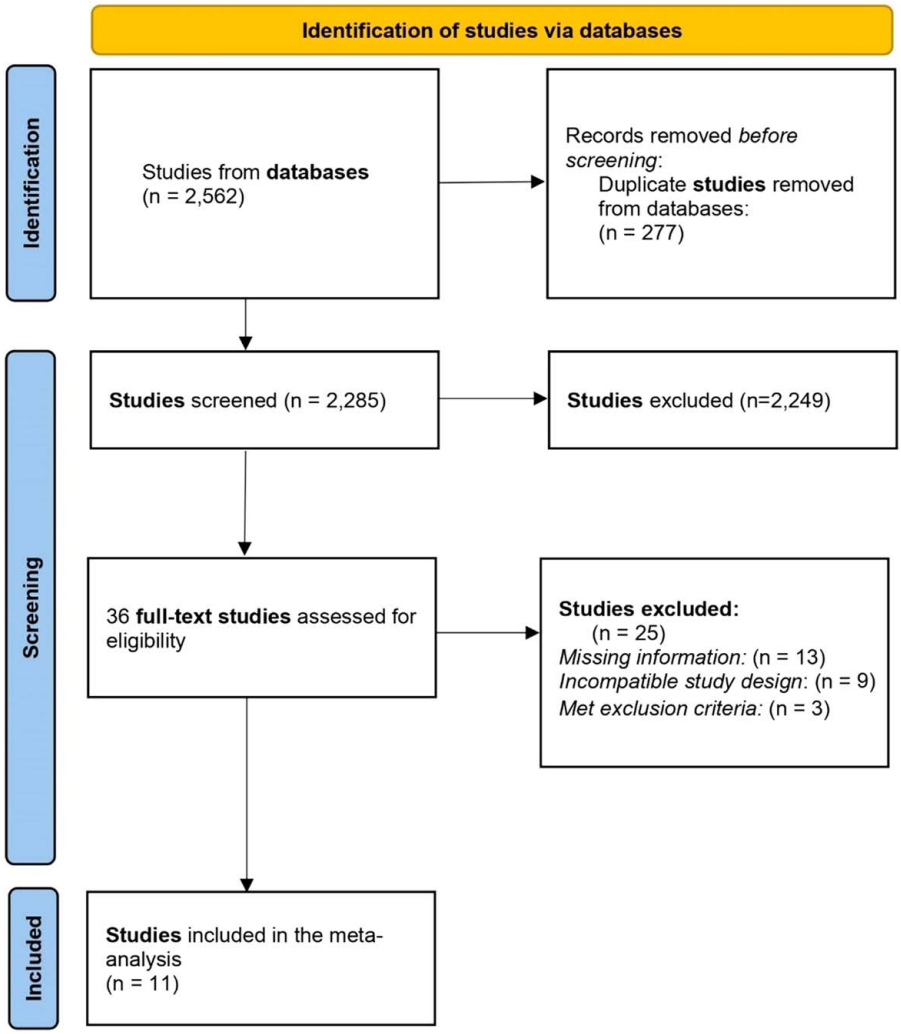

Figure 1. PRISMA flowchart

Additionally, 51 clinical trial registrations were identified through online registries. All the trial records identified from ClinicalTrials.Gov ($n = 51$) were assessed, of which 18 records were excluded. The 33 records of trials are enlisted in Supplementary Materials, Supplemental Digital Content 2, http://links.lww.com/MS9/A11.

Overview of the included studies

The analysis pooled 6856 participants, all of whom were enrolled in clinical trials across China (seven studies, 63.6%), Japan (two studies, 18.2%), Germany (one study, 9.1%), and Italy (one study, 9.1%). A total of 3938 samples of individuals were studied with real-time CAD systems and quality control systems, whereas 2918 others underwent regular colonoscopies with no additional detection and quality control systems. The majority of the detection systems originated from China such as (i) CADe systems, (ii) such real-time automatic poly detection systems, (iii) ENDOANGEL (real-time quality improvement system), and (iv) EndoScreeners. Other CADe systems included one that originated from Japan under the tutelage of the Jikei University School of Medicine and another by Medtronic. These were originally coded AI-assisted systems that the researchers used to determine the outcomes of adenoma and PDRs with colonoscopies. The methodologies, technology specifications, overall findings, conclusion, and appraisal of application to current practice are enlisted in Table 1.

Meta-analytical findings

The meta-analysis pooled in 6856 participants wherein 3938 (57.4%) belonged to the interventional group (facilitated with CADe and quality control systems); 2918 (42.6%) belonged to the SG that underwent regular colonoscopies. Individuals who intervened with the systems had a relatively higher ADR (OR = 1.51, 95% CI = 1.15–1.99, $P = 0.003$). Since the true ADR was higher with the intervention, the systems were beneficial in detecting adenomas that would otherwise not be visible to the naked eye. The forest plot and risks of bias for all studies are displayed in Figure 2. High heterogeneity was noted in the ADR outcome ($I^2 = 84\%$) suggesting differences in structural methodologies of the included studies. the studies with the highest weight were Wang *et al.*[17], Wang *et al.*[12], and Repici *et al.*[15].

Table 1

Characteristics of included studies

No.	References, country	Title	Domain	Methodology	Technology specifications	Adenomas detected in IG	Adenomas detected in SG	Polyp detection rate in the IG vs. SG	Withdrawal time in minutes (SD) IG vs. SG	Overall findings	Conclusions
1	Misawa *et al.*[11], Japan	Artificial intelligence-assisted polyp detection for colonoscopy: initial experience	Original artificial intelligence-assisted CADe system to analyze colonoscopy videos	Colonoscopy videos using the CADe system were analyzed among 73 patients retrospectively in a clinical trial setting	Original-coded artificial intelligence-assisted CADe system	73 adenomas detected ($N=59$ divided into 155 samples)	40 adenomas detected ($N=14$ divided into 135 samples)	30/155 (19.4%) vs. 8/135 (5.9%)	NA	Flat lesions were considered difficult for CADe (64.5% detection rate). The sensitivity (90%), specificity (63.3%), and accuracy (76.5%) were overall acceptable with a 94% overall detection rate of polyps, and 60% false-positive detection	The proposed CADe system found AI to have viable potential to provide automated detection of colorectal polyps
2	Wang *et al.*[12], China	Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomized controlled study	The effect of automatic polyp detection system based on deep learning on polyp detection rate and adenoma detection rate	An open, nonblinded trial comprising consecutive patients was prospectively randomized under coloscopy with (IG, $n=522$) and without (SG, $n=536$) the real-time automatic poly detection system that provides visual notice and sound alarm on polyp detection	Real-time automatic polyp detection system (Shanghai Wision AI Co. Ltd.)	262 adenomas detected ($N=498$)	160 adenomas detected ($N=269$)	235/522 (45%) vs. 156/536 (29.1%)	6.89 (1.79) vs. 6.39 (1.21)	The AI system increased adenoma detection rate (29.1% in IG vs. 20.3% in SG, $P<0.001$) and increased the mean number of adenomas per patient (0.53 in IG vs. 0.31 in SG, $P<0.001$). higher number of diminutive adenomas were found in the IG (185) compared to 102 in the SG (102) ($P<0.001$), while larger adenomas had no advantage (IG: 77 vs. SG:58, $P=0.075$). Hyperplastic polyps significantly increased (IG: 114 vs. SG:52, $P<0.001$)	In populations with low prevalence of polyps, an automatic polyp detection system can significantly increase the rate of diminutive adenomas and hyperplastic polyps detected
3											

Table 1

(*Continued*)

No.	References, country	Title	Domain	Methodology	Technology specifications	Adenomas detected in IG	Adenomas detected in SG	Polyp detection rate in the IG vs. SG	Withdrawal time in minutes (SD) IG vs. SG	Overall findings	Conclusions
	Gong *et al.*[13], China	Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomized controlled study	Real-time quality improvement system (ENDOANGEL) to monitor real-time withdrawal speed and colonoscopy withdrawal time in everyday use	A randomized controlled study with consecutive patients recruited aged 18–75 y where the real-time quality improvement system (ENDOANGEL) was validated on coloscopy videos and images	Real-time quality improvement system based on deep learning – ENDOANGEL	58 adenomas detected ($N=355$)	27 adenomas detected ($N=349$)	166/355 (47%) vs. 118/349 (34%)	6.38 (2.48) vs. 4.76 (2.54)	Among the intention-to-treat population, adenoma detection rates were significantly increased in the ENDOANGEL group ($n=58$ of 355, 16%) compared to the controls ($n=27$ of 349, 8%). With OR = 2.3 ($P=0.001$). Among the per-protocol analysis, ENDOANGEL group (54 of 324, 17%) compared to controls (26 of 318, 8%) yielded an OR of 2.18 ($P=0.026$). The accuracy was reportedly 97.9%, with a sensitivity of 95.8%, and a specificity of 99.3%	With no adverse events reported during the RCT, the ENDOANGEL system significantly improved the adenoma detection rate during coloscopy, thereby seeming both safe and effective for use
4	Liu *et al.*[14], China	Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy	Computer-aided detection of polyps and adenomas in colonoscopy	1026 patients, prospectively randomized for colonoscopy with and without computer-aided detection with visual notification and voice alarm to compare detection rates	The CADe system of polyps; Henan Xuanweitang Medical Information Technology Co., LTD., Zhengzhou City, Henan Province, China	250 adenomas detected ($N=486$)	144 adenomas detected ($N=248$)	221/508 (43.7) vs. 144/518 (27.8%)	6.82 (1.78) vs. 6.74 (1.62)	IG had higher detection rates of adenomas, small adenomas and proliferative polyps ($P<0.001$) compared to SG; larger adenomas showed no significant difference ($P>0.05$)	The computer-aided detection system (AI-powered) was considered feasible in increasing the detection of both adenomas and polyps in colonoscopy
5	Repici *et al.*[15], Italy	Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial	Deep learning systems to allow for real-time computer-aided detection (CADe) of polyps with high accuracy	685 individuals underwent screening colonoscopies for colorectal cancer who underwent high-definition	Artificial intelligence-based medical device (GI-Genius, Medtronic)	187 adenomas detected ($N=341$)	139 adenomas detected ($N=344$)	279/341 (82%) vs. 214/344 (62%)	6.95 (1.68) vs. 7.25 (2.48)	Adenoma detection rate was higher in the CADe group (54.8%) compared to the control (40.4%), with	The multicenter, randomized trial found that CADe in real-time coloscopy increased the adenoma detection rate and

				coloscopies with CADe system compared to those without						adenomas 5 mm or smaller at a higher proportion in the CADe group (33.7%) compared to the control (26.5%)	overall adenomas detected per coloscopy without altering withdrawal times
6	Su *et al.*[16], China	Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos)	Automatic quality control system (AQRS) for improvement of polyp and adenoma detection in clinical practice	To use an automatic quality control system (AQSC) to assess whether adenoma and polyp detection can improve in clinical practice	Images used from standard coloscopies (EC-3490TFi, EC-3490FK, EC-3890MZi, EC-3890Fi, or EC-3890FZi, Pentax Medical, China)	113 adenomas detected ($N=308$)	56 adenomas detected ($N=315$)	118/308 (38.3%) vs. 80/315 (25.4%)	7.03 (1.01) vs. 5.68 (1.26)	659 patients were randomized with AQCS significantly increasing adenoma detection rates (0.289 compared to 0.165). There was a superior withdrawal time in the AQCS group with 7.03 min compared to 5.68 min in the control group	AQCS could improve colonoscopists' performance during the withdrawal phase and improve the detection of adenomas
7	Wang *et al.*[17], China	Effect of a deep learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomized study	Colonoscopy with computer-aided detection (CADe) to improve colon polyps and adenomas detection by providing visual alarms during the procedure	Employing a double-blind randomized trial 1046 patients aged 18–75 were enrolled for diagnostic and screening coloscopy with CADe system compared to a sham system	CADe system (EndoScreener; Wision AI, Shanghai, China)	165 adenomas detected ($N=484$)	132 adenomas detected ($N=478$)	252/484 (52%) vs. 177/478 (37%)	7.46 (2.02) vs. 6.99 (1.57)	The adenoma detection rate was higher in the CADe group (0.34) compared to the sham group (0.28). Polyps initially missed by endoscopists by identified by the CADe system were typically small in size, flat in shape, isochromatic, had unclear boundaries, were on the edge of the visual field, were also partly behind colon folds	Polyps that are difficult for skilled endoscopists to recognize can be detected using high-performance CADe systems during colonoscopies
8	Wang *et al.*[18], China	Lower adenoma miss rate of computer-aided detection-assisted colonoscopy versus routine white-light colonoscopy in a prospective tandem study	Computer-aided detection (CADe) systems based on deep learning to reduce rates of missed adenomas by providing visual alerts identifying precancerous polyps (endoscopy monitor) in real-time	Prospective randomization of 369 patients was done where the rates of missed adenomas by playing visual alters and identifying precancerous polyps in real-time was conducted in CADe colonoscopy and	CADe system (EndoScreener, Shanghai Wision AI Co. Ltd, Shanghai, China)	144 adenomas detected ($N=184$)	120 adenomas detected ($N=185$)	118/184 (64%) vs. 102/185 (55%)	6.74 (1.35) vs. 7.29 (2.01)	The adenoma miss rate was lower in the CADe colonoscopy group (13.9%) compared to the routine group (40%)	CADe colonoscopy lowered the overall miss rate of adenomas when endoscopists used CADe systems; therefore, reducing the incidence of interval colon cancers

Table 1

(*Continued*)

No.	References, country	Title	Domain	Methodology	Technology specifications	Adenomas detected in IG	Adenomas detected in SG	Polyp detection rate in the IG vs. SG	Withdrawal time in minutes (SD) IG vs. SG	Overall findings	Conclusions
				routine white-light colonoscopy							
9	Kamba *et al.*[19], Japan	Reducing adenoma miss rate of colonoscopy assisted by artificial intelligence: a multicenter randomized controlled trial	Developing computer-aided detection (CADe) systems using an original deep learning algorithm based on convolutional neural networks to assist endoscopists in detecting colorectal adenomas during colonoscopies	Multicenter randomized controlled trial with 358 patients aged 40–80 y assigned to the CADe and standard colonoscopy group	CADe system (LPIXEL Inc. and The Jikei University School of Medicine)	111 adenomas detected ($N=172$)	93 adenomas detected ($N=174$)	120/172 (69.8%) vs. 106/174 (60.9%)	7.23 (1.35) vs. 7.48 (1.48)	The adenoma miss rate of the CADe group (13.8%) was lower compared to the standard group (36.7%). The adenoma detection rate for CADe-assisted coloscopies was 64.5% compared to 53.6% of the standard colonoscopy group	The miss rate of adenomatous lesions was lowered with coloscopies assisted with the CADe system
10	Yao *et al.*[20], China	Effect of an artificial intelligence-based quality improvement system on efficacy of a computer-aided detection system in colonoscopy: a four-group parallel study	Real-time monitoring of adenoma detection rate with both computer-aided polyp detection (CADe) system and computer-aided quality improvement (CAQ) system	Single-center placebo-controlled trial with 1076 patients randomized into (CADe, CAQ and CADe + CAQ) groups compared to control, to compute adenoma detection rate	Colonoscopy with the assistance of ENDOANGEL's polyp detection and quality monitoring functions	205 adenomas detected ($N=805$)	40 adenomas detected ($N=271$)	NA	NA	The adenoma detection rate was 14.76% in the control group whereas in the CADe, CAQ, and CADe + CAQ, it was 21.27, 24.54, and 30.6% respectively	CAQ was a useful add-on to the efficacy of CADe in the parallel controlled study. The interaction effect on improving adenoma detection can be stated with significance
11	Zippelius *et al.*[21], Germany	Diagnostic accuracy of a novel artificial intelligence system for adenoma detection in daily practice: a prospective nonrandomized comparative study	Artificial intelligence system in real-time colonoscopy to improve endoscopic quality and reduce the rate of interval cancer	The prospective clinical trial compared and analyzed 150 patients undergoing diagnostic endoscopy with and without the artificial intelligence system	Artificial intelligence system (no further specifications)	76 adenomas detected ($N=150$)	78 adenomas detected ($N=150$)	297/311 (95.5%) vs. 294/311 (94.5%)	NA	The adenoma detection rate in routine colonoscopy was 52% compared to 50.7% in the AI system, yielding no significant differences. Routine colonoscopy detected adenomas in two patients that were otherwise missed by the AI system	The AI system was considered to be comparable to that of experienced endoscopists during real-time colonoscopy with equally high adenoma detection rates (> 50%)

AI indicates artificial intelligence; IG, intervened group; SG, standard group.

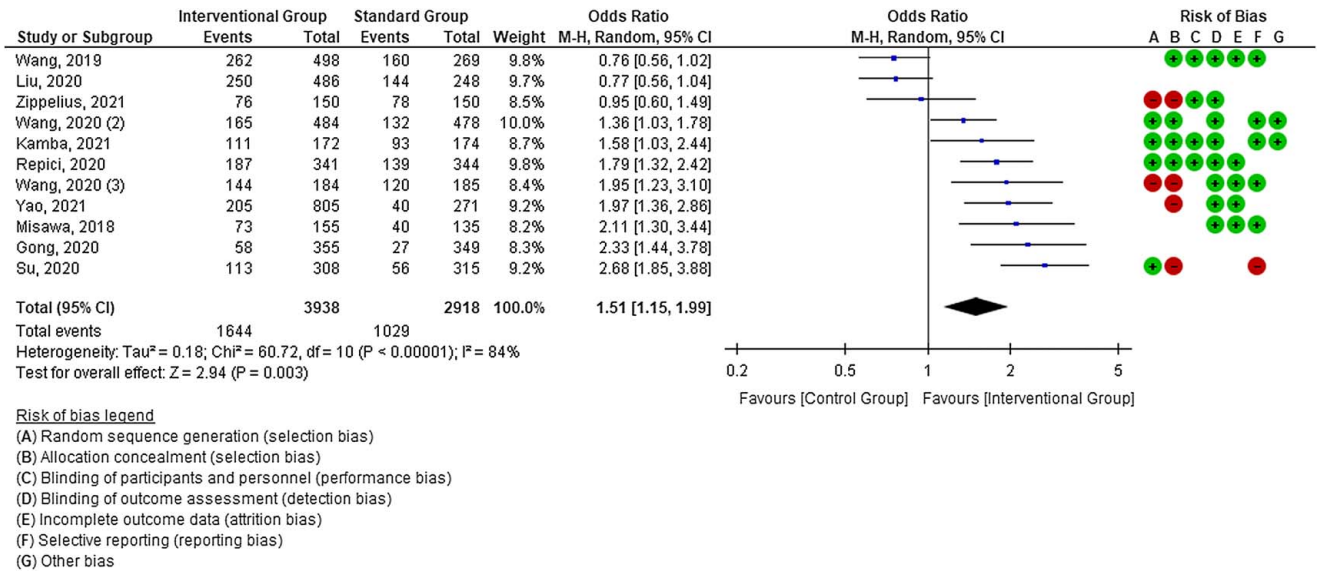

Figure 2. Adenoma detection rate forest plot and risk of bias. Odds ratio = 1.51, 95% CI = 1.15–1.99. Heterogeneity: $Tau^2 = 0.18$; $\chi^2 = 60.72$, $df = 10$ ($P < 0.00001$); $I^2 = 84\%$. Test for overall effect: $Z = 2.94$ ($P = 0.003$).

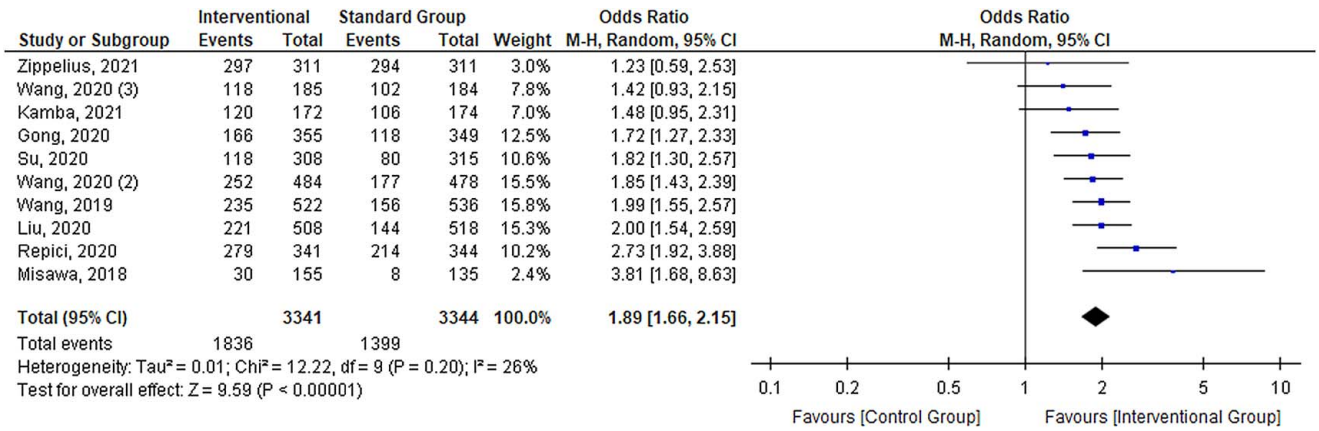

Figure 3. Polyp detection rate forest plot. Odds ratio = 1.89, 95% CI = 1.66–2.15. Heterogeneity: $Tau^2 = 0.01$; $\chi^2 = 12.22$, $df = 9$ ($P = 0.20$); $I^2 = 26\%$. Test for overall effect: $Z = 9.59$ ($P < 0.00001$).

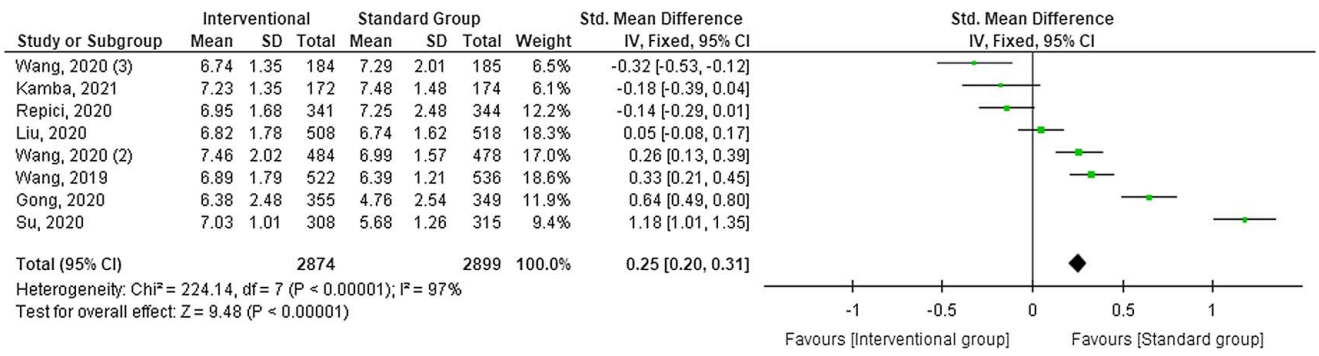

Figure 4. Standardized mean difference for withdrawal time, forest plot. SMD = 0.25, 95% CI = 0.2–0.31. Heterogeneity: $\chi^2 = 224.14$, $df = 7$ ($P < 0.00001$); $I^2 = 97\%$. Test for overall effect: $Z = 9.48$ ($P < 0.00001$).

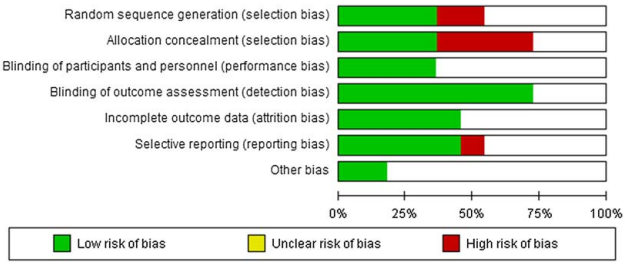

Figure 5. Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies.

Ten of the 11 studies provided information on PDRs; 3341 individuals were intervened with AI-colonoscopy and 3344 underwent standard colonoscopy procedures. Individuals who intervened with AI systems had a higher PDR (OR = 1.89, 95% CI = 1.66–2.15. $P < 0.00001$). There was less heterogeneity among the included studies ($I^2 = 26\%$) (Fig. 3).

Eight of the 11 studies enlisted withdrawal timings (in minutes). The computed findings of withdrawal timings were presented as SMD to denote the effect size. A medium effect size of 0.25 was determined (SMD = 0.25, 95% CI = 0.2–0.31). Given the relatively small effect size for withdrawal, there are limited practical applications of withdrawal time as a measure for RCTs assessing AI-assisted detection of adenomas and polyps (Fig. 4). Large heterogeneity was observed in the SMD assessment for withdrawal time, indicating that there were discrepancies in methodologies, aside from clinical differences; both Su *et al.*[16] and Gong *et al.*[13] deviated most from the weighted average, suggesting sources of heterogeneity.

Assessment of ongoing clinical trials

A total of 33 ongoing clinical trials were identified. These are enlisted in Supplementary Materials, Supplemental Digital Content 2, http://links.lww.com/MS9/A11, with key information pertaining to the following: NCT number, title, status, targeted conditions, interventions, outcome measure, collaborators, sex and age, enrollment, study type, study design, primary completion date, and locations.

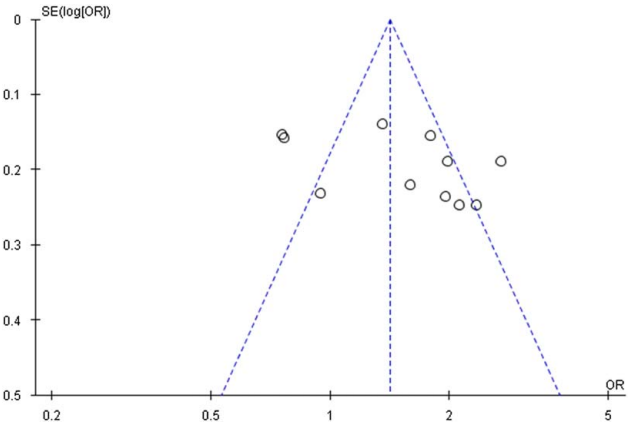

Figure 6. Funnel plot depicting publication bias. OR, odds ratio.

Methodological assessment of the included studies

The methodological quality of included studies is presented individually in Figure 2 and summarized in Figure 5. Concerning the random sequence generation (selection bias), four studies had a low risk of bias, two had a high risk of bias, and five had an unclear risk of bias. This meant that the majority of the studies either had low risk or unclear risks about the nonrandom component in the sequence generation process. On reviewing allocation concealment (selection bias) which is the technique of concealing the allocation sequence from those who assign participants to the interventional groups, four studies had a low risk of bias, whereas three had a high risk of bias; four studies had an unclear risk of bias.

Blinding of participants and healthcare personnel (performance bias), was deemed low risk in four studies, whereas it was unclear in the other seven studies. Blinding of outcome assessment (attrition bias) was considered low risk in eight studies, whereas it was unclear in three studies. Incomplete outcome data (attrition bias), had low risks in five studies, whereas it was unclear in six studies. Selecting reporting (reporting bias) was low risk in five studies, unclear risk in five studies, and high risk in one study.

Publication bias

Assessment for publication bias was done because more than 10 studies were included. Wang *et al.*[12], Liu *et al.*[14], and Su *et al.*[16] deviated from an inverted funnel shape. All the other eight studies were well within the remit of an inverted funnel shape (Fig. 6). Publication bias was less likely to have been present, however, our interpretation should be used with caution since we could not quantify findings of the Egger's test due to software limitations.

Discussion

To our knowledge, ours is the largest meta-analysis exploring ADR, PDR, and withdrawal time with the assistance of AI technological systems in coloscopies comprising 11 studies. We analyzed 6856 participants of which 57.4% were intervened with CADe and quality control systems where the ADR was computed. PDR and withdrawal times were assessed in 10 and 8 studies, respectively. While ADR was various across the trials, Wang *et al.*[12] (OR = 0.76, 95% CI = 0.56–1.02), Liu *et al.*[14] (OR = 0.77, 95% CI = 0.56–1.04), and Zippelius *et al.*[21] (OR = 0.95, 95% CI = 0.6–1.49), reported less favorable odds of adenoma detection with their computer-aided AI systems. However, all other studies favored the technology add-on to colonoscopies.

Deliwala *et al.*[22] performed an extensive search through January 2021 to locate clinical trials reporting ADR/PDR and they included a total of six trials with 4996 participants. The OR for adenoma was 1.77 (95% CI = 1.57–2.08) whereas, polyps were detected with an OR of 1.91 (95% CI = 1.68–2.16)[22]. Their findings suggest that AI-assisted colonoscopy can be a useful proxy to address critical gaps in colorectal adenoma detection[22]. Zhang and colleagues included seven studies with a total of 5427 individuals. AI-aided colonoscopy improved ADR (OR = 1.72, 95% CI = 1.52–1.95) and PDR (OR = 1.95, 95% CI: 1.75–2.19) significantly[23]. Ashat *et al.*[24] performed a

systematic review and meta-analysis to assess the impact of AI-assisted colonoscopy in clinical practice. The authors identified six studies with 5058 participants. The ADR was higher with AI systems (33.7%) compared to controls (22.9%) (OR = 1.76, 95% CI = 1.55–2); similarly PDR favored AI-intervention (OR = 1.9, 95% CI = 1.68–2.15)[24]. Barua *et al.*[25] compared colonoscopy with and without AI by calculating relative risks for detecting adenomas; they included a total of five studies where the ADR with AI was 29.6 and 19.3% with standard procedures. PDR was 45.4% with AI compared to 30.6% without AI. Aziz *et al.*[26] meta-analyzed three studies with a total of 2815 patients and found improved ADR (IG = 32.9% vs. SG = 20.8%) and PDR (IG = 32.9% vs. SG = 20.8%). Mohan *et al.*[27] searched multiple databases until May 2020 and included six RCTs in the analysis that used convolutional neural network-based CADe-assisted colonoscopy. The relative risks of AI-powered groups being detected with an adenoma compared to the SG were RR = 1.5 (95% CI = 1.3–1.72); similarly, the relative risks for PDR were 1.42 (95% CI = 1.33–1.51; $P < 0.00001$)[27]. Overall, our findings are comparative with literature that supports the idea that AI-assisted colonoscopies are superior to standard clinical colonoscopies, characterized by improvement in ADR and PDR.

Li and colleagues published a meta-analysis in 2021 that employed a total of five studies with 4311 patients that fulfilled the selection criteria; their results showed that the ADR and PDR with the assistance of AI had a pooled OR of 1.91 (95% CI = 1.68–2.16) and 1.75 (95% CI = 1.52–2.01), respectively. They found that the morphology of adenomas and polyps in colonoscopy impacted detection rates and the characteristics likely influenced the results[28]. Hassan and colleagues included five RCTs with 4356 patients. The pooled ADR and PDR were significantly higher in the CADe group compared to the control group (36.6 vs. 25.2%) and (50.3 vs. 34.6%), respectively[29]. The authors concluded that AI aids the detection of any colorectal neoplasia and is independent of main adenoma characteristics[29]. Zhang *et al.*[23] stated that polyps and adenomas of smaller sizes were better detected. AI-based systems used during colonoscopy increased the detection of small nonadvanced adenomas, but not advanced adenomas[25]. In light of current gaps in the literature, the shape and pathology recognition of polyps and adenomas may be improved with the AI technique which may be addressed in ongoing and future clinical trials. Taghiakbari *et al.*[30] found AI-assisted colonoscopies to have promising results in the detection of colorectal polyps and adenomas, however, the real-time application in clinical practice is not yet determined because the design, validation, and testing of AI models are underway. The largest ongoing AI-assisted clinical trial, PREEMPT CRC currently ongoing across the United States and the United Arab Emirates, will enroll 25,000 patients to screen for CRC among participants who are undergoing routine screening colonoscopies[31]. Our study appraises some of the aforementioned limitations of the AI technology while also providing support for future applications in colonoscopy procedures.

Lui and Leung[32] in their review express concerns about missed lesions during colonoscopy, which is a central reason for postcolonoscopy CRC. Challenges in the timely detection of CRC have promoted the development of AI-enabled polyp/adenoma detection[32]. However, certain challenges must be considered with the incorporation of AI-assisted screening tools for CRC. Ameen *et al.*[33], state that while AI systems are often promoted as solutions to improve the accuracy and quality of clinical decisions, they often rely on computational determinism and deductive reasoning. Clinical experience, autonomy, and judgment are reduced to inputs and outputs termed binary or multiclass classification problems[33]. In their argumentative review of AI and CRC, the authors write that to optimize the benefits of AI systems and avoid negative consequences for clinical decision making and patient care, there is a need for more balanced and nuanced approaches to AI system deployment and use in CRC[33]. Adenoma miss rates could be as high as 26% with standard techniques in current practice. Our study identified that the odds of detecting adenomas increase by around 50%, and of polyps by around 90%, with the use of AI systems with coloscopies ($P = 0.003$). Nonetheless, there is heterogeneity across the AI models and study designs as well as a lack of any long-term outcomes[32].

Eight of the 11 studies enlisted withdrawal timings (in minutes). The computed findings of withdrawal timings were presented as SMD to denote the effect size. A medium effect size of 0.25 was determined (SMD = 0.25, 95% CI = 0.2–0.31). given the relatively small effect size for withdrawal, there are limited practical applications of withdrawal time as a measure for RCTs assessing AI-assisted detection of adenomas and polyps (Fig. 4).

In the group supplemented with an AI system, the cumulative mean withdrawal time was 6.94 minutes. In the standard colonoscopy group, the overall mean withdrawal time was 6.57 minutes. The effect size computed in the meta-analysis was 0.25, meaning that no setbacks were faced with AI systems. The additional process did not increase withdrawal times worth noting. Current literature and expert opinion suggest that a minimum adequate mean withdrawal time of 6 minutes is required for screening colonoscopy to achieve the target PDR and ADR[34,35]. It is useful for future trials to assess withdrawal times and compare with the current industry standard, and the comparative findings of our study (6.94 min).

Notably, the development of novel techniques and devices aimed at improving ADR comprise of full-spectrum endoscopy, the Third Eye panoramic cap and balloon-colonoscope system[36]. These technologies are not widely applied due to limited expertise and high costs. On the other hand, with the use of simple add-on devices to the tips of endoscopes, direct views behind colonic folds can be facilitated[36]. As per current guidelines, there is a growing interest in said devices to improve ADR, despite mixed results, there is still an assumption of their superiority as compared to standard colonoscopy[36]. There are certain risk factors for the recurrence of advanced colorectal adenoma after polypectomy, however, the potential impact of adenoma detection on screening surveillance must be discussed. In a retrospective study enrolling 3360 patients that underwent polypectomy and colonoscopy, 746 patients were detected with 1017 advanced adenomas[37]. Based on the retrospective analysis by Facciorusso *et al.*[37], the recurrence rate was highest among those with high-grade dysplasia greater than or equal to 15 mm. It was suggested that patients with high-grade dysplasia undergo more intensive surveillance, while those without high-grade dysplasia could consider longer follow-up periods given a lower risk of advanced colorectal adenoma[37].

Appraisal of application to current practice: setbacks and notes for future trials

The proposed CADe system by Misawa *et al.*[11] has potential benefits in automating the detection of colorectal polyps, the nonopen-source nature of the AI platform provides limited benefit to healthcare. The high reproducibility, uniformity, and fidelity of Wang's *et al.*[12] CADe system offer potential large-scale advantages compared with human assistance which requires prospective evaluation. Polyps/adenomas were not successfully detected in the ascending colon and cecum due to the instability of the colonoscope, meaning diagnostic equipment may lead to biased outcomes[12]. Repici *et al.*[15] find that a higher PDR/ADR with CADe was consistent across all centers of the multicenter trial, reducing the possibility of operator bias. However, limitations exist in terms of operators' psychological biases[15]. Inexperienced endoscopists may contribute to larger false-positive results, leading to prolonged withdrawal times[15]. It is pertinent to consider a combined approach to detecting polyps and adenomas in unison with endoscopists and AI systems to overcome human and technical errors.

The ENDOANGEL system by Gong *et al.*[13] has been validated at only one center and the PDR/ADR was improved for different polyp/adenoma sizes with the system; further need for multicenter clinical trials arises to investigate and adopt the effectiveness of the system in other areas of the world. There were 36 false alarms with an average of 0.071 false alarms per colonoscopy by Liu *et al.*[14] and none were missed. However, the findings by Liu *et al.*[14] may not be extrapolated to more burdened areas as the enrolled population had a low overall PDR/ADR due to geocultural factors (i.e. genetics, diet, lifestyle, habits).

Quality control of current and future AI-enabled systems for polyp/adenoma detection requires further consideration. Su *et al.*'s[16] prospective RCT tested an automatic quality control system (AQCS), with results indicating that the quality control system could improve the detection of colorectal polyps and adenomas. A technical challenge is that the system could only review images obtained from the Pentax imaging system, meaning that AQCS requires adaptation to other vendors for endoscope images[16]. PDR/ADR can be increased with a senior endoscopist operating the tool, which may lead to underestimation of the true effectiveness of CADe systems[17]. CADe systems are effective and safe to increase adenoma detection during coloscopies, however, long-term observation is required worldwide[17]. False prompts do exist with AQCS, meaning that further training and larger datasets could correct current errors[16]. Data quality may also be improved with a multicenter site design to adapt AI-enabled systems in clinical practice[20]. Overall, small increases in the quality of colonoscopy programs can lead to net gains in large-scale CRC screening which may be possible with AI quality controlled systems for polyp and adenoma detection[17].

Conclusion

In summary, AI increases the polyp and ADR in colonoscopy, with no noteworthy delay in withdrawal time. AI assistance can be improved with more large-scale prospective trials which account for interoperator differences, technical errors, and tuning of AI quality control systems. CRC is highly preventable and the incorporation of AI-assisted tools into routine practice has the strong potential to reduce incidence rates of CRC in the near future.

Ethical approval

None required.

Consent

None required.

Sources of funding

No funding was obtained.

Author contribution

All authors contributed to the study conception, data collection, analysis, writing (original draft and revision).

Conflict of interest disclosure

All authors declare no conflict of interest.

Research registration unique identifying number (UIN)

1. Name of the registry: PROSPERO.
2. Unique identifying number or registration ID: PROSPERO 2022 CRD42022333731.
3. Hyperlink to your specific registration (must be publicly accessible and will be checked): https://www.crd.york.ac.uk/prospero/display_record.php?RecordID = 333731

Guarantor

Zouina Sarfraz and Ivan Cherrez-Ojeda are co-guarantors of this study.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Data availability statement

All data used for the purpose of this study is available online.

References

[1] Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol 2021;14:101174.
[2] Alberti LR, Garcia DPC, Coelho DL, *et al.* How to improve colon cancer screening rates. World J Gastrointest Oncol 2015;7:484.
[3] Mitsala A, Tsalikidis C, Pitiakoudis M, *et al.* Artificial intelligence in colorectal cancer screening, diagnosis and treatment. A new era. Curr Oncol 2021;28:1581–607.
[4] Fleming M, Ravula S, Tatishchev SF, *et al.* Colorectal carcinoma: pathologic aspects. J Gastrointest Oncol 2012;3:153.

[5] Rex DK. Can we do resect and discard with artificial intelligence-assisted colon polyp "optical biopsy?,". Tech Innov Gastrointest Endosc 2020;22:52–5.
[6] Dayyeh BKA, Thosani N, Konda V, *et al*. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2015;81:502–e1.
[7] Willems P, Djinbachian R, Ditisheim S, *et al*. Uptake and barriers for implementation of the resect and discard strategy: an international survey. Endosc Int Open 2020;8:E684–92.
[8] Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg 2021;88: 105906.
[9] The effectiveness of real-time computer-aided and quality control systems in colorectal adenoma and polyp detection during colonoscopies: a meta analysis, (n.d.). Accessed 21 September 2022. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=333731.
[10] Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. Bmj 2017;358:1–7.
[11] Misawa M, Kudo S, Mori Y, *et al*. Artificial intelligence-assisted polyp detection for colonoscopy: initial experience. Gastroenterology 2018;154:2027–9.
[12] Wang P, Berzin TM, Brown JRG, *et al*. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019;68:1813–9.
[13] 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;5:352–61.
[14] Liu W-N, Zhang Y-Y, Bian X-Q, *et al*. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. Saudi J Gastroenterol Off J Saudi Gastroenterol Assoc 2020;26:13.
[15] 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.
[16] 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: 415–24.
[17] 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.
[18] Wang P, Liu P, Brown JRG, *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.
[19] Kamba S, Tamai N, Saitoh I, *et al*. Reducing adenoma miss rate of colonoscopy assisted by artificial intelligence: a multicenter randomized controlled trial. J Gastroenterol 2021;56:746–57.
[20] Yao L, Zhang L, Liu J. Effect of an artificial intelligence-based quality improvement system on efficacy of a computer-aided detection system in colonoscopy: a four-group parallel study. Endoscopy 2022;54:757–68.
[21] Zippelius C, Alqahtani SA, Schedel J, *et al*. Diagnostic accuracy of a novel artificial intelligence system for adenoma detection in daily practice: a prospective nonrandomized comparative study. Endoscopy 2021;54: 465–72.
[22] Deliwala SS, Hamid K, Barbarawi M, *et al*. Artificial intelligence (AI) real-time detection vs. routine colonoscopy for colorectal neoplasia: a meta-analysis and trial sequential analysis. Int J Colorectal Dis 2021;36: 2291–303.
[23] Zhang Y, Zhang X, Wu Q, *et al*. Artificial intelligence-aided colonoscopy for polyp detection: a systematic review and meta-analysis of randomized clinical trials. J Laparoendosc Adv Surg Tech 2021;31:1143–9.
[24] Ashat M, Klair JS, Singh D, *et al*. Impact of real-time use of artificial intelligence in improving adenoma detection during colonoscopy: a systematic review and meta-analysis. Endosc Int Open 2021;9:E513–21.
[25] Barua I, Vinsard DG, Jodal HC, *et al*. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. Endoscopy 2021;53:277–84.
[26] Aziz M, Fatima R, Dong C, *et al*. The impact of deep convolutional neural network-based artificial intelligence on colonoscopy outcomes: a systematic review with meta-analysis. J Gastroenterol Hepatol 2020;35: 1676–83.
[27] Mohan BP, Facciorusso A, Khan SR, *et al*. Real-time computer aided colonoscopy versus standard colonoscopy for improving adenoma detection rate: a meta-analysis of randomized-controlled trials. EClinicalMedicine 2020;29:100622.
[28] Li J, Lu J, Yan J, *et al*. Artificial intelligence can increase the detection rate of colorectal polyps and adenomas: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol 2021;33:1041–8.
[29] 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.
[30] Taghiakbari M, Mori Y, von Renteln D. Artificial intelligence-assisted colonoscopy: a review of current state of practice and research. World J Gastroenterol 2021;27:8103.
[31] Prevention of Colorectal Cancer Through Multiomics Blood Testing. 2022. Accessed 24 April 2022. https://clinicaltrials.gov/ct2/show/NCT04369053.
[32] Lui TKL, Leung WK. Is artificial intelligence the final answer to missed polyps in colonoscopy? World J Gastroenterol 2020;26:5248.
[33] Ameen S, Wong M-C, Yee K-C, *et al*. AI and clinical decision making: the limitations and risks of computational reductionism in bowel cancer screening. Appl Sci 2022;12:3341.
[34] Barclay RL, Vicari JJ, Doughty AS, *et al*. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med 2006;355:2533–41.
[35] Wong WJ, Arafat Y, Wang S, *et al*. Colonoscopy withdrawal time and polyp/adenoma detection rate: a single-site retrospective study in regional Queensland. ANZ J Surg 2020;90:314–6.
[36] Facciorusso A, Del Prete V, Buccino RV, *et al*. Comparative efficacy of colonoscope distal attachment devices in increasing rates of adenoma detection: a network meta-analysis. Clin Gastroenterol Hepatol 2018;16: 1209–9.
[37] Facciorusso A, Di Maso M, Serviddio G, *et al*. Factors associated with recurrence of advanced colorectal adenoma after endoscopic resection. Clin Gastroenterol Hepatol 2016;14:1148–54.