



Single Versus Second Observer vs Artificial Intelligence to Increase the ADENOMA Detection Rate of Colonoscopy—A Network Analysis

Manesh Kumar Gangwani¹ · Hossein Haghbin² · Rizwan Ishtiaq³ · Fariha Hasan⁴ · Julia Dillard¹ · Fouad Jaber⁵ · Dushyant Singh Dahiya⁶ · Hassam Ali⁷ · Shaharyar Salim⁸ · Wade Lee-Smith⁹ · Amir Humza Sohail¹⁰ · Sumant Inamdar¹¹ · Muhammad Aziz¹² · Benjamin Hart¹³

Received: 23 January 2024 / Accepted: 7 February 2024 / Published online: 4 March 2024
© The Author(s) 2024, corrected publication 2024

Abstract

Background and Aims Screening colonoscopy has significantly contributed to the reduction of the incidence of colorectal cancer (CRC) and its associated mortality, with adenoma detection rate (ADR) as the quality marker. To increase the ADR, various solutions have been proposed including the utilization of Artificial Intelligence (AI) and employing second observers during colonoscopies. In the interest of AI improving ADR independently, without a second observer, and the operational similarity between AI and second observer, this network meta-analysis aims at evaluating the effectiveness of AI, second observer, and a single observer in improving ADR.

Methods We searched the Medline, Embase, Cochrane, Web of Science Core Collection, Korean Citation Index, SciELO, Global Index Medicus, and Cochrane. A direct head-to-head comparator analysis and network meta-analysis were performed using the random-effects model. The odds ratio (OR) was calculated with a 95% confidence interval (CI) and p -value < 0.05 was considered statistically significant.

Results We analyzed 26 studies, involving 22,560 subjects. In the direct comparative analysis, AI demonstrated higher ADR (OR: 0.668, 95% CI 0.595–0.749, $p < 0.001$) than single observer. Dual observer demonstrated a higher ADR (OR: 0.771, 95% CI 0.688–0.865, $p < 0.001$) than single operator. In network meta-analysis, results were consistent on the network meta-analysis, maintaining consistency. No statistical difference was noted when comparing AI to second observer (RR 1.1 (0.9–1.2, $p = 0.3$). Results were consistent when evaluating only RCTs. Net ranking provided higher score to AI followed by second observer followed by single observer.

Conclusion Artificial Intelligence and second-observer colonoscopy showed superior success in Adenoma Detection Rate when compared to single-observer colonoscopy. Although not statistically significant, net ranking model favors the superiority of AI to the second observer.

Keywords Artificial intelligence · Adenoma detection rate · Colorectal cancer · Single observer · Second observer

Introduction

Colonoscopy is regarded as a tier-1 screening modality, which has significantly contributed to the reduction in the incidence of colorectal cancer (CRC) and its associated mortality [1]. Adenoma detection rate (ADR), defined as the percentage of screening colonoscopies that reveal at least one adenoma, is a quality marker for endoscopists. It is worth noting that every 1% improvement in ADR corresponds to a 3% reduction in CRC cases [2]. To further emphasize the importance of adenoma detection, the US Multi-Society

Task Force (MSTF) recommends striving for a minimum ADR of 30% for male and 20% for female patients [3].

In order to improve this important quality marker, numerous solutions have been proposed including the utilization of Artificial Intelligence (AI), employing second observers, employing distal attachment devices, among other strategies [4, 5].

The idea of a second observer, whether in the form of a trainee, nurse, or technician is a time-proven method that has consistently shown effectiveness in improving both ADR and Polyp Detection Rate (PDR) during colonoscopies [6]. However, as we enter the era of AI, there is growing interest in the potential of this technology to improve ADR

Extended author information available on the last page of the article

independently, eliminating the need for a second observer. In a comprehensive network meta-analysis of 94 Randomized Controlled Trials (RCTs), our study has previously demonstrated the superiority of AI in improving both ADR and PDR when compared to various modalities, including different methods of virtual chromoendoscopy and add-on devices [5]. Notably, our study excluded the second observer from its analysis in the recently published findings [5].

Due to the similarity in the operational mechanics of AI and a second observer, there has been a recent interest in comparing these two modalities [7]. Consequently, we decided to perform a network meta-analysis aimed at evaluating the effectiveness of AI, second observer, and a single observer in improving ADR.

Methods and Materials

Search Strategy

A comprehensive search of the following databases was conducted from inception until April 24, 2023, using multiple databases including MEDLINE (PubMed platform, NCBI), Embase (Embase.com, Elsevier), Web of Science Core Collection, Korean Citation Index, and SciELO (Clarivate), Global Index Medicus (World Health Organization), and Cochrane Central Register of Controlled Trials (Cochrane Library, Wiley). Screening of key reference / bibliographic lists for more studies was additionally performed. The keywords/ subject terms used to search were ‘colonoscopy,’ ‘adenoma,’ ‘artificial intelligence,’ ‘adenoma detection rate,’ ‘single blind’ ‘double blind,’ and ‘dual observer’ along with their corresponding medical subject heading terms, in various combinations. There were no language restrictions or filters applied. The search strategy was created by an experienced librarian (WLS) and reviewed by another investigator (MKG). The detailed search strategy can be reviewed in Supplementary Table 1.

Selection and Data Collection Process

Upon completion of the search process, all results were exported to EndNote 20 citation management software (Clarivate, Philadelphia, Penn, USA) where duplicates were removed via EndNote’s duplicate detection algorithms. Subsequently, manual screening was undertaken to identify and eliminate any duplicates. The title and abstract screening were conducted independently by two investigators (J.D and D.S.D), which was resolved by a third senior author (M. A) in cases of disputes. The full-text screening was conducted in the same manner by the same investigators. In cases where full-text articles were not readily available or further data were required, the corresponding authors of the included

studies were contacted and missing or additional data were requested. Throughout the process, strict adherence to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines was practiced [8].

Inclusion and Exclusion Criteria

We included all studies that involved adult patients aged 18 and above undergoing colonoscopy with a single observer, double observer, or AI. Our inclusion criteria comprised relevant randomized studies and prospective and retrospective studies. Conversely, we excluded case reports, case series with fewer than 10 patients, editorials, guidelines, and review articles.

Data Extraction and Study Outcomes

The results were tabulated using Microsoft Excel (Microsoft, Redmond, Wash, USA). The extracted data items included author names, date of publication, type of study design, age, sex, the total number of patients and adenoma detection rate in single, double (operator plus observer), and AI operators.

Data Synthesis and Statistical Analysis

Categorical data were summarized as counts and percentages. We conducted a direct head-to-head comparator analysis and network meta-analysis of all available groups. The direct meta-analysis was performed using the DerSimonian-Laird method and random-effects model on Open Meta Analyst (CEBM, University of Oxford, Oxford, United Kingdom) [9].

Comparison of interventions and visual displaying of the findings was conducted using network meta-analysis using the random-effects model on the ‘R’ package ‘Netmeta’ (Bell Labs, Murray Hill, USA) [10]. The odds ratio (OR) for each outcome was calculated with a 95% confidence interval (CI). Microsoft Excel (Microsoft, Redmond, Wash, USA) was used to tabulate and create tables for this study. A p -value < 0.05 was considered statistically significant. The “frequentist method” was used to rank the intervention and a P-score was generated. Study heterogeneity was assessed using the I² statistic defined by the Cochrane Handbook for systematic reviews and a value $> 50\%$ was considered as substantial heterogeneity [11]. Disagreement between direct and indirect evidence was assessed using the node-splitting technique [10].

Bias Assessment

The risk of bias assessment for the included studies was conducted using the Newcastle–Ottawa Scale (NOS) for observational studies [12], grading of recommendations

assessment, development, and evaluation (GRADE) for RCTs [13]. GRADE approach to evaluate the strength of evidence for results from the NMA. Five domains that affect the level of confidence in the NMA results are considered: (i) risk of bias, (ii) inconsistency (iii) indirectness, (iv) imprecision, and (v) heterogeneity. Due to the nature of the trials evaluating the effects of a second observer and AI, participants and personnel could not be blinded. Therefore, this domain was not used to calculate the overall risk of bias for included studies. GRADEPro GDT was utilized to construct the summary of findings (SoF) table illustrated in Supplementary Table 5. Publication bias was assessed visually using a funnel plot and Egger's regression analysis. A p -value < 0.05 was considered statistically significant publication bias.

Results

The search strategy identified 2952 articles, of which 260 unique studies were subjected to full-text review. After applying strict inclusion and exclusion criteria (Supplementary Fig. 1), a final selection of 26 studies was made. Our study included 20 randomized control trials and 3 retrospective and 3 prospective studies involving 22,560 subjects [7, 14–38]. These studies encompassed the years 2008 to 2023 and comprised of 12,596 'single' observer, 4187 'double' observer and 5777 'AI' observer cases. The mean age across the studies ranged from 46.0 to 66.4 years (Table 1). Table 2 provides details on adenoma detection rate categorized by single, double, and AI operators.

Direct Meta-Analysis

In the direct comparative analysis, single-observer ADR was compared with AI ADR. The results indicated a statistical difference in adenoma detection rate between single operator and AI, favoring AI (OR: 0.668, 95% CI 0.595–0.749, $p < 0.001$), as depicted in Fig. 1. Similarly, a significant difference was observed in ADR between single observer vs second observer, with second observer demonstrating a higher ADR. (OR: 0.771, 95% CI 0.688–0.865, $p < 0.001$), as depicted in Fig. 2.

Network Meta-Analysis

Adenoma detection rate in all modalities is summarized in Table 3. AI (RR: 1.26, 95% CI 1.17–1.35) and second observer (RR: 1.19, 95% CI 1.09–1.30) exhibited statistically significant higher rates of adenoma detection when compared to single operator. Network forest plot is demonstrated in Fig. 3A. (net forest plot). No statistical difference

was noted when comparing AI to second observer. (RR 1.1 (0.9–1.2, $p = 0.3$).

Random-effect models of second observer to single observer had low heterogeneity at 18% exhibiting high confidence in the level of evidence and low risk of bias, while single operator versus AI had high heterogeneity at 63% suggesting high variability among included studies. Net split models were also consistent in direct and indirect evidence when evaluating adenoma detection rates. (Supplementary Table 2) This consistency reinforces the confidence in the overall findings of the network meta-analysis and strengthens the conclusions drawn regarding the comparative effectiveness of the interventions. Corresponding net split plots (Supplementary Fig. 2A) and net graphs (Supplementary Fig. 2B) are provided in supplementary materials.

Randomized Controlled Trials

In RCTs, adenoma detection rate in all modalities is summarized in Table 4. Similarly, AI (RR: 1.27, 95% CI 1.18–1.37) and second observer (RR: 1.21, 95% CI 1.07–1.37) exhibited statistically significant higher rates of adenoma detection when compared to single operator. No statistical difference was noted when comparing AI to second observer. (RR: 1.0 (0.9–1.2), $p = 0.49$). Network forest plot is demonstrated in Fig. 3B. (net forest plot).

Random-effect models of second observer versus single observer had low heterogeneity at 0% exhibiting high confidence in the level of evidence and low risk of bias, while AI versus single operator had high heterogeneity at 65% suggesting high variability among included studies. Net split models were also consistent in direct and indirect evidence when evaluating adenoma detection rates. (Supplementary Table 3). Corresponding net split plots (Supplementary Fig. 3A) and net graphs (Supplementary Fig. 3B) are provided in supplementary materials.

Ranking of Interventions

P -scores were used to assess the effectiveness of AI, second and single observers in detecting adenomas. AI achieved the highest score of 0.92, followed by second observer with a score of 0.58. Single operator scored 0.1. These findings were similar to that of RCTs. A higher P -score indicates a higher probability of detecting adenomas. Net ranking in Supplementary Fig. 5 illustrates these findings.

Publication Bias

The Newcastle–Ottawa scale was utilized to rank cohort studies, with scores ranging from 5 to 6, indicating moderate quality as shown in Supplementary Table 4. GRADE assessment was used for RCTs, which showed high to low

Table 1 Demographic characteristics of included studies

Authors (Country)	Study design	Intervention	Total sample size	Single	Second	AI	Age (years)		Sex ratio (F:M)	
							Single	Double	Single	Double
Ahmad et al. [23] (UK)	RCT	Single vs. AI	614	306		308	66.4		208:98	
Aslanian et al. [24] (USA)	RCT	Single vs. Double	502	249	253		57.8	58.1	116:133	133:120
Buchner et al. [38] (USA)	Retrospective	Single vs. Double	2430	2112	318		63.0	62.0	981:1131	143:175
Eckardt et al. [18] (USA)	Prospective	Single vs. Double	368	187	181		56.0	56.0	92:95	107:74
Gong et al. [16] (China)	RCT	Single vs. AI	704	349		355	49.0		191:158	168:187
Hüneburg et al. [33] (Germany)	RCT	Single vs. AI	96	46		50	46.3		25:21	30:20
Kamba et al. [34] (Japan)	RCT	Single vs. AI	355	177		178	61.4		41:136	42:136
Kim et al. [22] (Korea)	RCT	Single vs. Double	383	191	192		56.4	57.3	76:115	72:120
Lee et al. [31] (Korea)	RCT	Single vs. Double	791	384	407		58.1	58.6	187:197	185:222
Liu et al. [35] (China)	RCT	Single vs. AI	1026	518		508	50.1		231:287	244:264
Liu et al. [37] (China)	RCT	Single vs. AI	790	397		393	48.8		203:194	213:180
Peters et al. [29] (USA)	Retrospective	Single vs. Double	3594	2895	699		58.3	59.6	1587:1308	382:317
Quan et al. [14] (USA)	Prospective	Single vs. AI	600	300		300	63.5		128:172	134:166
Repici et al. [21] (Italy)	RCT	Single vs. AI	685	344		341	61.1		179:165	169:172
Ren et al. [27] (China)	Prospective	Single vs. Double	1192	579	613		52.0	52.2	228:351	245:368
Rogart et al. [24] (USA)	Retrospective	Single vs. Double	309	126	183		55.6	56.9	50:76	72:111
Rondonotti et al. [20] (Italy)	RCT	Single vs. AI	800	395		405	61.0		199:196	192:213
Shaukat et al. [15] (USA)	RCT	Single vs. AI	1359	677		682	59.9		322:355	314:368
Su et al. [25] (China)	RCT	Single vs. AI	623	315		308	51.6		167:148	149:159
Vilkoite et al. [36] (Latvia)	RCT	Single vs. AI	400	206		194	51.2		104:102	103:91
Wang et al. [17] (China)	RCT	Single vs. Double	587	291	296		53.4	53.2	161:130	152:144
Wang et al. [32] (China)	RCT	Single vs. AI	1058	536		522	49.9		287:249	259:263
Wang et al. [19] (China)	RCT	Single vs. AI	962	478		484	49.0		224:254	243:241
Wang et al. [7] (China)	RCT	Double vs. AI	1261		625	636		47.0		299:326
Yamaguchi et al. [26] (Japan)	RCT	Single vs. AI	231	118		113	63.3		52:66	272:364
Yang et al. [28] (China)	RCT	Single vs. Double	840	420	420		50.8	51.1	200:220	55:58

RCT randomized controlled trial, *Single* single operator, *Second* operator plus an observer, *AI* artificial intelligence

Table 2 Intervention outcomes

Author	ADR (%)		
	Single	Double	AI
Ahmad et al. [23]	199/306		220/308
Aslanian et al. [24]	101/249	119/253	
Buchner et al. [38]	549/2112	96/318	
Eckardt et al. [18]	36/187	27/181	
Gong et al. [16]	27/349		58/355
Hüneburg et al. [33]	12/46		18/50
Kamba et al. [34]	93/174		111/172
Kim et al. [22]	57/191	74/192	
Lee et al. [31]	166/384	196/407	
Liu et al. [35]	124/518		199/508
Liu et al. [37]	83/397		114/393
Peters et al. [29]	802/2895	240/699	
Quan et al. [14]	113/300		131/300
Repici et al. [21]	139/344		187/341
Ren et al. [27]	119/579	128/613	
Rogart et al. [24]	29/126	68/183	
Rondonotti et al. [20]	179/395		217/405
Shaukat et al. [15]	297/677		326/682
Su et al. [25]	52/315		89/308
Vilkoite et al. [36]	43/206		59/194
Wang et al. [17]	67/291	89/296	
Wang et al. [32]	112/536		138/522
Wang et al. [19]	132/478		165/484
Wang et al. [7]		150/625	164/636
Yamaguchi et al. [26]	72/118		66/113
Yang et al. [28]	143/420	185/420	

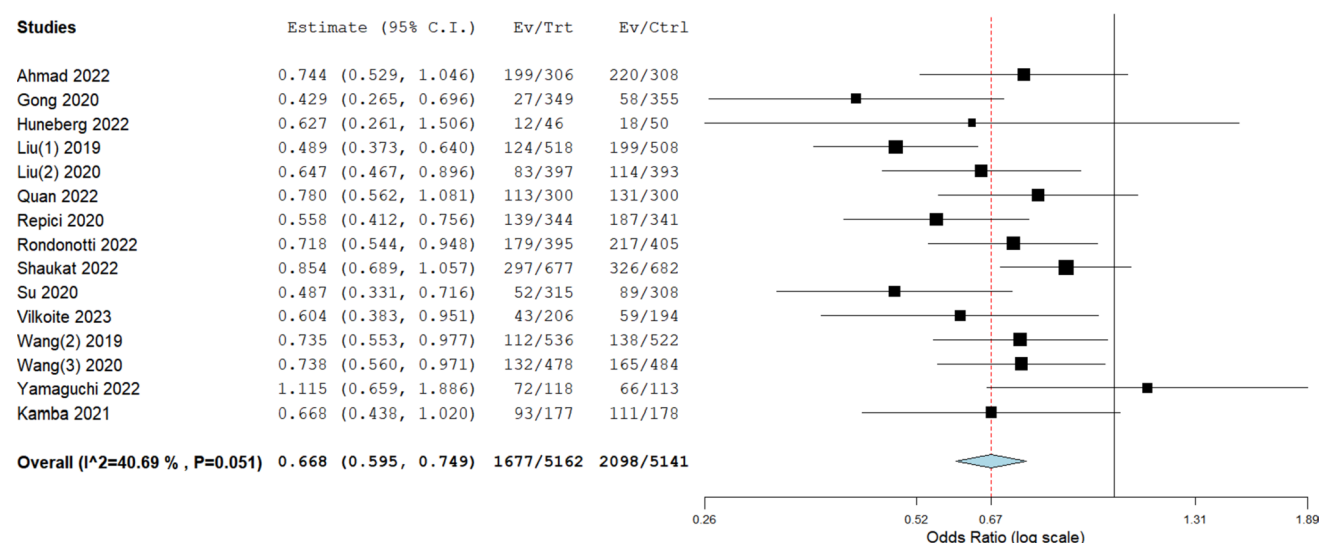
ADR adenoma detection rate

quality of evidence, as shown in Supplementary Table 5. The results of the Egger's publication score ($p=0.0404$) and Thompson-Sharp ($p=0.0475$) revealed significant evidence of publication bias, as shown in Supplementary Fig. 4A. Similarly, for RCTs, Egger's publication score ($p=0.0035$) and Thompson-Sharp ($p=0.0046$) showed evidence of publication bias (Supplementary Fig. 4B).

Discussion

This systematic review and network analysis have demonstrated that both AI and second observer are superior to the single observer in improving ADR. However, when comparing AI to a second observer, we did not find any statistical difference; however, meta-ranking suggested potential preference toward AI although statistical significance could not be achieved.

Our meta-analysis confirmed the commonsense expression that “two pairs of eyes are better than one,” whether those pairs belong to humans or machines. As Wallace aptly suggested, three mechanisms contribute to missing an adenoma: it may not be within the visual field, it might go unrecognized, or it may be unrecognizable [39]. The second observer aids in recognizing “not recognized” polyps and potentially even those in the first scenario if they encourage better technique. In contrast, AI has the potential of assisting in identifying “not recognizable” polyps, in addition to the aforementioned scenarios. Hence, the effectiveness of a second human observer depends on their experience and training. Prior studies have demonstrated that when a trainee serves as the second observer, ADR increases with each year of the trainee's fellowship training [29]. Furthermore,

**Fig. 1** Forest plot showing direct comparison between single operator vs. AI

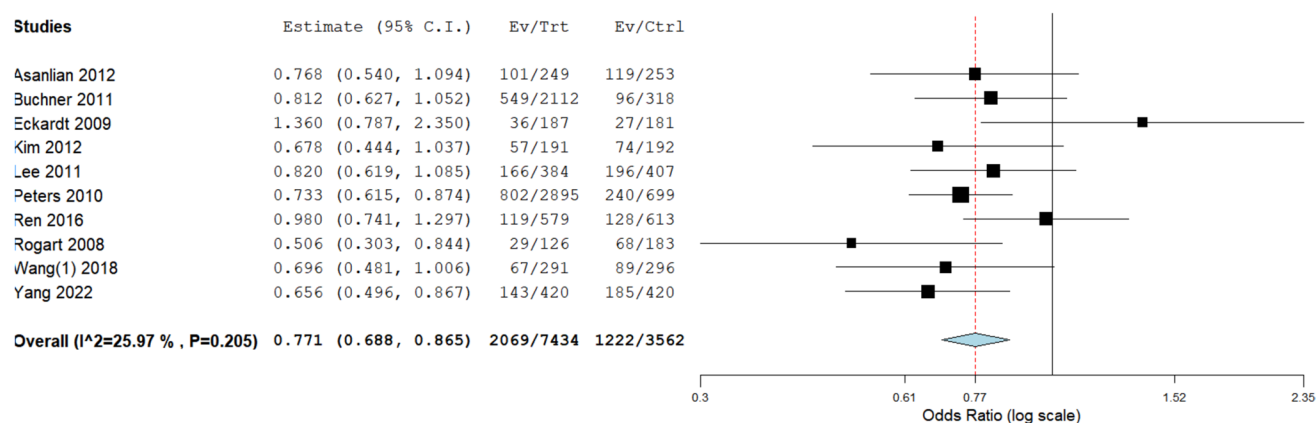


Fig. 2 Forest plot showing direct comparison between single operator vs. double operator

Table 3 Network meta-analysis outcomes evaluating adenoma detection rate in cohort studies

	Single	Second	AI
Single	–	0.8 (0.8–0.9), $p < 0.01$	0.9 (0.8–1.1), $p = 0.30$
Dual	1.2 (1.1–1.3), $p < 0.01$	–	0.8 (0.7–0.9), $p < 0.01$
AI	1.3 (1.2–1.4), $p < 0.01$	1.1 (0.9–1.2), $p = 0.30$	–

Relative risk for interventions, $I^2 = 51.1\%$

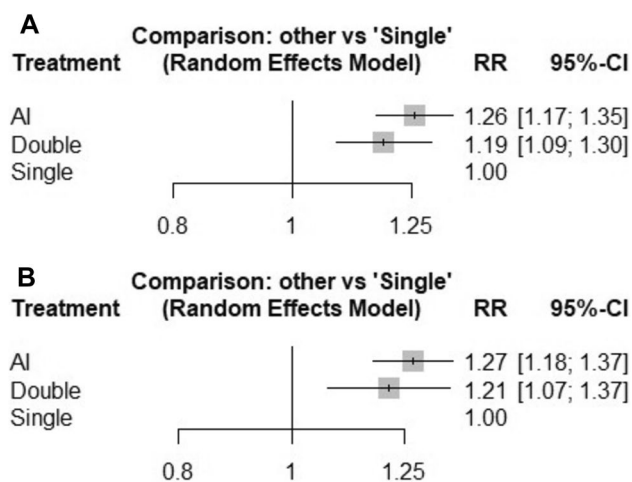


Fig. 3 **A** Network Forest plot of adenoma detection rate in cohort studies. **B** Network Forest plot of adenoma detection rate in RCTs

studies have shown that experienced nurses, when acting as second observers, increase ADR compared to their less experienced counterparts [40]. The studies included in our meta-analysis involved different settings and observers with varying levels of experience. Therefore, standardizing training for second observers, reducing heterogeneity, and evaluating these effects become imperative.

As for AI, with the ever-increasing quality of endoscopic imaging, visual field, and the computational power of processors, AI should theoretically achieve superior potential over the human-eye in diagnosing adenomas. AI can provide real-time pixel-level analysis of every frame, overcoming the human-eye's limitations, such as its propensity to miss briefly visible or partially blocked adenomas. In addition, the human-eye is susceptible to inherent defects such as “inattention blindness” when distractions lead to missed polyps and “change blindness” when alterations are missed during eye movement [41–43]. These intrinsic limitations cannot be fully addressed by another human observer. In spite of these limitations, the presence of another human observer may have unaccounted benefits as Rex et al. showed an increase

Table 4 Network meta-analysis outcomes evaluating adenoma detection rate in RCTs

	Single	Second	AI
Single	–	0.8 (0.8 – 0.9), $p < 0.01$	0.8 (0.7–0.9), $p < 0.01$
Second	1.2 (1.1–1.4), $p < 0.01$	–	1.0 (0.8–1.1), $p = 0.49$
AI	1.3 (1.2–1.4), $p < 0.01$	1.0 (0.9–1.2), $p = 0.49$	–

Relative risk for interventions, $I^2 = 55.4\%$

in withdrawal time when the endoscopist was being observed [44]. This may possibly be due to competition and increased attentiveness, leading to improved inspection techniques, such as inspecting behind folds and spending more time in withdrawal. Meanwhile, endoscopists may become over reliant on AI, inadvertently reducing the quality of inspection, highlighting the need for a second observer.

AI offers solutions to mitigate reduced exam quality due to endoscopists' overconfidence in it. By reminding endoscopists to inspect behind folds and improve withdrawal time, Computer-Aided Quality (CAQ) AI has shown efficacy in improving exam quality [45]. While our studies primarily included the Computer-Aided polyp Detection (CADE) AI, combinations of CADE and CAQ AI have shown superior results compared to CADE [45]. Future research is required to study these effects, especially in head-to-head comparisons with second observers.

A notable issue with AI in diagnostics is its high rate of false positives, mistakenly identifying benign cases as adenomas, which are disproven upon further examination [46]. Nevertheless, evidence suggests that AI outperforms the human counterpart in this aspect; Wang et al.'s recent study found that second observers were significantly more prone to false alarms than AI [7]. We speculate that with advancements in polyp characterization (CADx) AI, the rate of false positives in the AI group will further decrease. This may potentially make AI more efficient than second observers- an area requiring further investigation.

Another aspect to explore is how these modalities benefit endoscopists based on their expertise. Some studies suggest that second observers exclusively benefit inexperienced endoscopists and offer no additional advantage to expert endoscopists [30, 31]. This may be due to the marginal benefit of second observers for experts. Another potential cause might be the reluctance of trainees or nurses to point out missed lesions to expert endoscopists. This is important as the included studies are from different countries with cultural heterogeneity. Another potential source of heterogeneity is the level of experience of the second observer; some studies included trainees, while others included nurses. Despite this variation, all studies showed effectiveness regardless of the type of second observer. With regard to AI, while some studies show greater benefits for inexperienced endoscopists, others do not find significant differences in AI's benefit between high and low detectors [45, 47]. Although subgroup analysis was not possible due to insufficient data, future studies may help identify the subgroups that benefit most from these modalities.

Our study had some limitations. First, the study combines both RCTs and observational studies, with the latter carrying inherent bias. However, most of the data were derived from RCTs that demonstrate lower relative bias rates. Second, the non-blinded methodology in the studies

could affect endoscopist behavior due to the presence of another observer or overconfidence in the assisting modality [48]. The involvement of multiple endoscopists should help reduce this bias. Third, heterogeneity exists in terms of patient populations and timing. In our systematic review, the second-observer studies were published predominantly prior to 2018, while all AI studies emerged after 2019; however, all colonoscopies utilized HD endoscopes. Fourth, due to the limited follow-up duration, long-term outcome data on interval CRC and mortality are scarce and could not be analyzed. Lastly, we were unable to conduct subgroup analysis based on polyp morphology and location, which is essential given that serrated lesions in the right colon are an important cause of interval CRCs.

In conclusion, both AI and second observer led to improvement in ADR compared to single-observer colonoscopy. More standardized RCTs are required to compare AI with second observers, as current data suggest AI's superiority, even though statistical significance was not achieved. As the technology evolves, we recommend utility of AI, if feasible, to improve ADR.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10620-024-08341-9>.

Author's contributions All authors have significantly contributed toward the design of the study, writing of the manuscript, and the review of the manuscript.

Funding No funding was received for the preparation of this manuscript.

Data availability Data, analytic methods, and study materials will not be made available to other researchers.

Declarations

Competing interests The authors declare no conflict of interest.

Research involving in human and animal participants No human subjects/animals were involved in this systematic review and meta-analysis.

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

References

- Shaukat A, Kahi CJ, Burke CA, Rabeneck L, Sauer BG, Rex DK. ACG clinical guidelines: colorectal cancer screening 2021. *Am J Gastroenterol*. 2021;116:458–479. <https://doi.org/10.14309/ajg.0000000000001122>.
- Corley DA, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med*. 2014;26:2541.
- Gupta S, Lieberman D, Anderson JC et al. Recommendations for follow-up after colonoscopy and polypectomy: a consensus update by the US multi-society task force on colorectal cancer. *Gastrointest Endosc*. 2020;91:463–485.e5. <https://doi.org/10.1016/j.gie.2020.01.014>.
- Aziz M, Thoguluva CV. Improving the ADR: narrow band, broad benefits. *Dig Dis Sci*. 2020;6:1586–1587.
- Aziz M, Hagbini H, Sayeh W et al. Comparison of artificial intelligence with other interventions to improve adenoma detection rate for colonoscopy: a network meta-analysis. *J Clin Gastroenterol*. 2022. <https://doi.org/10.1097/mcg.0000000000001813>.
- Aziz M, Weissman S, Khan Z et al. Use of 2 observers increases adenoma detection rate during colonoscopy: systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2019. <https://doi.org/10.1016/j.cgh.2019.07.033>.
- Wang P, Liu XG, Kang M et al. Artificial intelligence empowers the second-observer strategy for colonoscopy: a randomized clinical trial. *Gastroenterol Rep (Oxf)*. 2023. <https://doi.org/10.1093/gastro/goac081>.
- Liberati A, Altman DG, Tetzlaff J et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj*. 2009;339:b2700. <https://doi.org/10.1136/bmj.b2700>.
- Wallace BC, Schmid CH, Lau J, Trikalinos TA. Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. *BMC Med Res Methodol*. 2009;9(1):80. <https://doi.org/10.1186/1471-2288-9-80>.
- Shim SR, Kim SJ, Lee J, Rücker G. Network meta-analysis: application and practice using R software. *Epidemiol Health*. 2019;41:e2019013. <https://doi.org/10.4178/epih.e2019013>.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. 2003;327:557–560. <https://doi.org/10.1136/bmj.327.7414.557>.
- Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25:603–605. <https://doi.org/10.1007/s10654-010-9491-z>.
- Guyatt GH, Oxman AD, Vist GE et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*. 2008;336:924–926. <https://doi.org/10.1136/bmj.39489.470347.AD>.
- Quan SY, Wei MT, Lee J et al. Clinical evaluation of a real-time artificial intelligence-based polyp detection system: a US multicenter pilot study. *Sci Rep*. 2022;12:6598. <https://doi.org/10.1038/s41598-022-10597-y>.
- 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. <https://doi.org/10.1053/j.gastro.2022.05.028>.
- Gong D, Wu L, Zhang J et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol*. 2020;5:352–361. [https://doi.org/10.1016/s2468-1253\(19\)30413-3](https://doi.org/10.1016/s2468-1253(19)30413-3).
- Wang W, Xu L, Bao Z et al. Differences with experienced nurse assistance during colonoscopy in detecting polyp and adenoma: a randomized clinical trial. *Int J Colorectal Dis*. 2018;33:561–566. <https://doi.org/10.1007/s00384-018-3003-0>.
- Eckardt AJ, Swales C, Bhattacharya K, Wassef WY, Leung K, Levey JM. Does trainee participation during colonoscopy affect adenoma detection rates? *Dis Colon Rectum*. 2009;52:1337–1344. <https://doi.org/10.1007/DCR.0b013e3181a80d8f>.
- Wang P, Liu X, Berzin TM et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADE-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol*. 2020;5:343–351. [https://doi.org/10.1016/s2468-1253\(19\)30411-x](https://doi.org/10.1016/s2468-1253(19)30411-x).
- Rondonotti E, Di Paolo D, Rizzotto ER et al. Efficacy of a computer-aided detection system in a fecal immunochemical test-based organized colorectal cancer screening program: a randomized controlled trial (AIFIT study). *Endoscopy*. 2022;54:1171–1179. <https://doi.org/10.1055/a-1849-6878>.
- 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.e7. <https://doi.org/10.1053/j.gastro.2020.04.062>.
- Kim TS, Park DI, Lee DY et al. Endoscopy nurse participation may increase the polyp detection rate by second-year fellows during screening colonoscopies. *Gut Liver*. 2012;6:344–348. <https://doi.org/10.5009/gnl.2012.6.3.344>.
- Ahmad A, Wilson A, Haycock A et al. Evaluation of a real-time computer-aided polyp detection system during screening colonoscopy: AI-DETECT study. *Endoscopy*. 2023;55:313–319. <https://doi.org/10.1055/a-1966-0661>.
- Rogart JN, Siddiqui UD, Jamidar PA, Aslanian HR. Fellow involvement may increase adenoma detection rates during colonoscopy. *Am J Gastroenterol*. 2008;103:2841–2846. <https://doi.org/10.1111/j.1572-0241.2008.02085.x>.
- Su JR, Li Z, Shao XJ et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc*. 2020;91:415–424.e4. <https://doi.org/10.1016/j.gie.2019.08.026>.
- Yamaguchi D, Shimoda R, Miyahara K et al. Impact of an artificial intelligence-aided endoscopic diagnosis system on improving endoscopy quality for trainees in colonoscopy: prospective, randomized, multicenter study. *Dig Endosc*. 2023. <https://doi.org/10.1111/den.14573>.
- Ren KY. Impact of endoscopy nurse involvement on polyp detection rates during colonoscopy. *World Chin J Digestol*. 2016;24:1277. <https://doi.org/10.11569/wcjd.v24.i8.1277>.
- Yang Q, Zhu X, Wu Z, Leng F, Shu X, Yang L. Impact of the second examination of the proximal colon on the adenoma detection rate: a prospective randomized controlled trial. *Clin Transl Gastroenterol*. 2023;14:e00557. <https://doi.org/10.14309/ctg.0000000000000557>.
- Peters SL, Hasan AG, Jacobson NB, Austin GL. Level of fellowship training increases adenoma detection rates. *Clin Gastroenterol Hepatol*. 2010;8:439–442. <https://doi.org/10.1016/j.cgh.2010.01.013>.
- Aslanian HR, Shieh FK, Chan FW et al. Nurse observation during colonoscopy increases polyp detection: a randomized prospective study. *Am J Gastroenterol*. 2013;108:166–172. <https://doi.org/10.1038/ajg.2012.237>.
- Lee CK, Park DI, Lee SH et al. Participation by experienced endoscopy nurses increases the detection rate of colon polyps during a screening colonoscopy: a multicenter, prospective, randomized study. *Gastrointest Endosc*. 2011;74:1094–1102. <https://doi.org/10.1016/j.gie.2011.06.033>.
- 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;68:1813–1819. <https://doi.org/10.1136/gutjnl-2018-317500>.

33. Hüneburg R, Bucksch K, Schmeißer F et al. Real-time use of artificial intelligence (CADEYE) in colorectal cancer surveillance of patients with Lynch syndrome—a randomized controlled pilot trial (CADLY). *United Eur Gastroenterol J*. 2023;11:60–68. <https://doi.org/10.1002/ueg2.12354>.
34. 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;56:746–757. <https://doi.org/10.1007/s00535-021-01808-w>.
35. Liu WN, Zhang YY, Bian XQ et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol*. 2020;26:13–19. https://doi.org/10.4103/sjg.SJG_377_19.
36. Vilkoite I, Tolmanis I, Meri HA et al. The role of an artificial intelligence method of improving the diagnosis of neoplasms by colonoscopy. *Diagnostics (Basel)*. 2023. <https://doi.org/10.3390/diagnostics13040701>.
37. Liu P, Wang P, Glissen Brown JR et al. The single-monitor trial: an embedded CADe system increased adenoma detection during colonoscopy: a prospective randomized study. *Therap Adv Gastroenterol*. 2020;13:1756284820979165. <https://doi.org/10.1177/1756284820979165>.
38. Buchner AM, Shahid MW, Heckman MG et al. Trainee participation is associated with increased small adenoma detection. *Gastrointest Endosc*. 2011;73:1223–1231. <https://doi.org/10.1016/j.gie.2011.01.060>.
39. Wallace MB. Improving colorectal adenoma detection: technology or technique? *Gastroenterology*. 2007;132:1221–1223. <https://doi.org/10.1053/j.gastro.2007.03.017>.
40. Dellon ES, Lippmann QK, Sandler RS, Shaheen NJ. Gastrointestinal endoscopy nurse experience and polyp detection during screening colonoscopy. *Clin Gastroenterol Hepatol*. 2008;6:1342–1347. <https://doi.org/10.1016/j.cgh.2008.06.014>.
41. Memmert D, Unkelbach C, Ganns S. The impact of regulatory fit on performance in an inattentive blindness paradigm. *J Gen Psychol*. Apr-Jun 2010;137:129–139. <https://doi.org/10.1080/00221301003645061>.
42. Wolfe JM, Reinecke A, Brawn P. Why don't we see changes?: The role of attentional bottlenecks and limited visual memory. *Vis Cogn*. 2006;14:749–780. <https://doi.org/10.1080/1350628050195292>.
43. Simons DJ, Rensink RA. Change blindness: past, present, and future. *Trends Cognitive Sci*. 2005;9:16–20. <https://doi.org/10.1016/j.tics.2004.11.006>.
44. Rex DK, Hewett DG, Raghavendra M, Chalasani N. The impact of videorecording on the quality of colonoscopy performance: a pilot study. *Am J Gastroenterol*. 2010;105:2312–2317. <https://doi.org/10.1038/ajg.2010.245>.
45. Yao L, Zhang L, Liu J et al. 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;54:757–768. <https://doi.org/10.1055/a-1706-6174>.
46. Holzwanger EA, Bilal M, Glissen Brown JR et al. Benchmarking definitions of false-positive alerts during computer-aided polyp detection in colonoscopy. *Endoscopy*. 2021;53:937–940. <https://doi.org/10.1055/a-1302-2942>.
47. Repici A, Spadaccini M, Antonelli G et al. Artificial intelligence and colonoscopy experience: lessons from two randomized trials. *Gut*. 2022;71:757–765. <https://doi.org/10.1136/gutjnl-2021-324471>.
48. Wang HS, Pisegna J, Modi R et al. Adenoma detection rate is necessary but insufficient for distinguishing high versus low endoscopist performance. *Gastrointest Endosc*. 2013;77:71–78. <https://doi.org/10.1016/j.gie.2012.08.038>.

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

Authors and Affiliations

Manesh Kumar Gangwani¹ · Hossein Haghbini² · Rizwan Ishtiaq³ · Fariha Hasan⁴ · Julia Dillard¹ · Fouad Jaber⁵ · Dushyant Singh Dahiya⁶ · Hassam Ali⁷ · Shahryar Salim⁸ · Wade Lee-Smith⁹ · Amir Humza Sohail¹⁰ · Sumant Inamdar¹¹ · Muhammad Aziz¹² · Benjamin Hart¹³

✉ Manesh Kumar Gangwani
gangwani.manesh@gmail.com;
maneshkumar.gangwani@utoledo.edu

¹ Department of Medicine, University of Toledo Medical Center, Toledo, OH, USA

² Department of Gastroenterology and Hepatology, Ascension Providence Hospital, Southfield, MI, USA

³ Department of Medicine, St Francis Hospital and Medical Center, Hartford, CT, USA

⁴ Department of Internal Medicine, Cooper University Hospital, Camden, NJ, USA

⁵ Department of Internal Medicine, University of Missouri-Kansas City, Kansas City, MO, USA

⁶ Department of Medicine, Central Michigan University College of Medicine, Mount Pleasant, MI, USA

⁷ Department of Gastroenterology and Hepatology, East Carolina University Health, Greenville, NC, USA

⁸ Department of Internal Medicine, Aga Khan University Hospital, Karachi, Pakistan

⁹ University of Toledo Libraries, University of Toledo, Toledo, OH, USA

¹⁰ Department of General Surgery, New York University Langone Health, Long Island, NY, USA

¹¹ Department of Gastroenterology and Hepatology, University of Arkansas for Medical Sciences, Little Rock, AR, USA

¹² Department of Gastroenterology and Hepatology, University of Toledo Medical Center, Toledo, OH, USA

¹³ Department of Hepatology and Gastroenterology, University of Michigan, Ann Arbor, MI, USA