ELMER PRESS

Original Article

Gastroenterol Res. 2026;19(5):254-272

Artificial Intelligence-Assisted Colonoscopy Improves Adenoma Detection by Trainee and Less-Experienced Endoscopists: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Hamza Altal[a, f], Rami Jadallah[b], Hasan Daher[a], Connor Pollard[a], Kinney Xu[a], Cody Marik[a], Saba Daher[a], Mohammad Darweesh[c], Andrew Storm[d], Mark Young[e]

Abstract

Background: Multiple systematic reviews and meta-analyses have established the benefits of computer-aided detection (CADe) among experienced endoscopists. The role of CADe among endoscopists who are still acquiring competence trainees in formal endoscopic training and independently practicing but less-experienced operators particularly its effect on right-colon adenoma detection, remains incompletely defined. We performed an updated systematic review and meta-analysis of randomized controlled trials (RCTs) to evaluate the impact of CADe on adenoma detection by trainee and less-experienced endoscopists.

Methods: We searched PubMed/MEDLINE, Embase, the Cochrane Central Register of Controlled Trials, and ClinicalTrials.gov from inception through May 2026 for RCTs comparing CADe-assisted with standard colonoscopy performed by trainee or less-experienced endoscopists. The primary outcome was adenoma detection rate (ADR). Secondary outcomes were right-colon ADR (RADR), polyp detection rate (PDR), and advanced adenoma detection rate (AADR). Risk ratios (RRs) were pooled using DerSimonian–Laird random-effects models; heterogeneity was quantified with I^2. Ninety-five percent prediction intervals (PIs), pooled risk differences (RDs), Hartung–Knapp confidence intervals (CIs), leave-one-out sensitivity analyses, a prespecified sensitivity analysis restricted to operators in formal training, Doi plots with the Luis Furuya-Kanamori (LFK) index, and Grading of Recommendations, Assessment, Development and Evaluation (GRADE) certainty ratings were also generated.

Results: Five RCTs enrolling 3,158 patients met the inclusion criteria. CADe significantly increased overall ADR (RR 1.17, 95% CI 1.07–1.29; P = 0.0012; I^2 = 28.8%; 95% PI 0.95–1.46; RD +6.9%, 95% CI +2.7% to +11.0%; number needed to treat (NNT) 15). The effect was larger for RADR (RR 1.31, 95% CI 1.14–1.50; P = 0.0002; I^2 = 26.6%; 95% PI 0.86–1.97; RD +8.2%, 95% CI +4.6% to +11.7%). PDR was also higher with CADe (RR 1.16, 95% CI 1.08–1.25; P = 0.0001; I^2 = 27.8%; 95% PI 0.97–1.39). AADR did not differ between groups (RR 0.97, 95% CI 0.73–1.29; P = 0.82; I^2 = 0.0%; 95% PI 0.52–1.81). Results were robust to leave-one-out analysis and to exclusion of the single trial of independently practicing non-expert operators (ADR RR 1.16, 95% CI 1.01–1.33). Certainty of evidence was moderate for ADR, low for RADR and PDR, and very low for AADR.

Conclusions: CADe-assisted colonoscopy increases adenoma detection by trainee and less-experienced endoscopists, with the largest relative effect for right-colon adenomas. Because the PIs cross the null and the certainty of evidence is moderate at best, these findings support further prospective evaluation of CADe as an adjunct to colonoscopy training rather than establishing that routine integration into fellowship curricula produces more competent independent endoscopists.

Keywords: Computer-aided detection; Deep learning; Proximal colon; Gastroenterology fellowship; Colorectal cancer screening; Endoscopy education; Quality indicators of colonoscopy; Post-colonoscopy colorectal cancer

Manuscript submitted August 4, 2026, accepted September 4, 2026
Published online October 26, 2026

[a]Department of Internal Medicine, East Tennessee State University Quillen College of Medicine, Johnson City, TN, USA
[b]Department of Internal Medicine, University of Tennessee Knoxville, Knoxville, TN, USA
[c]Department of Gastroenterology, NYU Langone Health, New York, NY, USA
[d]Department of Gastroenterology, Wake Forest University School of Medicine, Winston-Salem, NC, USA
[e]Department of Internal Medicine, Division of Gastroenterology, East Tennessee State University Quillen College of Medicine, Johnson City, TN, USA
[f]Corresponding Author: Hamza Altal, Department of Internal Medicine, East Tennessee State University Quillen College of Medicine, Johnson City, TN 37614, USA. Email: Altalh@etsu.edu

doi: https://doi.org/10.14740/gr2183
Gastroenterology Research
1918-2805 (print), 1918-2813 (online)

Introduction

Colorectal cancer ranks as the third most frequently diag-

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com
This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, including commercial use, provided the original work is properly cited

nosed malignancy and the second leading cause of cancer-related mortality worldwide, accounting for an estimated 1.92 million new cases and nearly 900,000 deaths globally in 2022 [1]. Screening colonoscopy with detection and removal of premalignant adenomas remains a cornerstone of colorectal cancer prevention, and adenoma detection rate (ADR) stands as one of the most extensively validated quality indicators of colonoscopy performance. Higher ADR is inversely associated with the risk of post-colonoscopy interval colorectal cancer and colorectal cancer death, with each 1% absolute increase in ADR conferring an approximately 3% relative decrease in interval cancer risk [2, 3]. Recognizing this relationship, the American Society for Gastrointestinal Endoscopy (ASGE) and the European Society of Gastrointestinal Endoscopy (ESGE) have established minimum ADR benchmarks of ≥ 35% and ≥ 25%, respectively, for high-quality screening colonoscopy [4, 5].

Despite these benchmarks, ADR varies markedly between endoscopists, and missed adenomas remain a leading mechanism of colonoscopy failure. A systematic review and meta-analysis of tandem colonoscopy studies (43 publications, > 15,000 procedures) estimated a pooled adenoma miss rate of approximately 26% (95% confidence interval (CI) 23–30), with disproportionately high miss rates for flat adenomas (34%; 95% CI 24–45) and serrated polyps (27%; 95% CI 16–40), as well as for proximal lesions [6]. Post-colonoscopy colorectal cancers are more often located in the proximal colon than cancers detected at first colonoscopy, and most are attributable to procedural factors such as missed lesions, inadequate examination, and incomplete resection, the majority of which are considered potentially preventable. Right-colon ADR (RADR) has therefore been proposed as a location-specific quality indicator that may better capture the risk of proximal neoplasia that is missed at index colonoscopy [7–9].

Endoscopists who have not yet accumulated substantial procedural experience face a steeper version of this challenge, and this group is not homogeneous. It spans complete novices, gastroenterology fellows in formal training programs, and independently practicing endoscopists whose lifetime case volume remains modest. A tandem colonoscopy study demonstrated that trainee adenoma miss rates are substantially elevated in the early phase of fellowship, declining only gradually to below 25% after approximately 450 procedures [10]. Moreover, trainee ADR varies widely among gastroenterology fellows and independently predicts post-training performance, with limited improvement observed after fellowship completion [11]. Strategies capable of accelerating competency acquisition and standardizing colonoscopy quality during this formative period are therefore a recognized priority. One such strategy is real-time computer-aided detection (CADe), which employs deep learning algorithms to identify and highlight candidate lesions on the endoscopy display during withdrawal. Foundational randomized trials of CADe in experienced endoscopist populations have consistently demonstrated significant improvements in ADR across both low- and high-baseline-ADR settings [12, 13]. Pooled analyses further confirm a 24% to 44% relative ADR increase across diverse clinical settings, with consistent benefit across CADe platforms [14–17].

Whether the magnitude of this benefit translates to endoscopists who are still building experience, however, has only recently begun to be examined. A preliminary meta-analysis presented in abstract form at Digestive Disease Week (DDW) 2024 pooled four randomized trials in trainees and reported a non-significant trend toward improved ADR (risk ratio (RR) 1.15, 95% CI 0.99–1.34) but did not analyze right-colon detection and remains available only in abstract form [18]. We therefore performed an updated systematic review and meta-analysis of randomized controlled trials (RCTs) to evaluate the effect of CADe-assisted colonoscopy on both overall and right-colon adenoma detection by trainee and less-experienced endoscopists.

Materials and Methods

Study design and registration

This systematic review and meta-analysis was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [19]. The completed PRISMA 2020 checklist is provided as Supplementary Material 1 (gr.elmerpub.com). The review protocol was not prospectively registered in a public systematic review registry, and no protocol was published. To make the review process transparent, we state explicitly which elements were fixed before data extraction, and which were not. The review question, the eligibility criteria, the primary outcome (ADR), and the secondary outcomes (RADR, polyp detection rate (PDR), and AADR) were specified in an internal written protocol agreed by the review team before screening began and before any outcome data were extracted; these were not altered after the results were seen. The following elements were added after the initial analysis, in response to peer review, and should be regarded as post hoc: the sensitivity analysis restricted to operators in formal endoscopic training; the reporting of prediction intervals (PIs), pooled risk differences (RDs), and Hartung–Knapp CIs; leave-one-out analyses for the secondary outcomes; the Doi plot and Luis Furuya-Kanamori (LFK) index; and the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) certainty assessment. No outcome was added, removed, or redefined after the pooled estimates were known, and no analysis reported here was selected based on its result.

Search strategy and data sources

A comprehensive literature search was performed across PubMed/MEDLINE, Embase, the Cochrane Central Register of Controlled Trials (CENTRAL), and ClinicalTrials.gov from inception through 31 May 2026. Search strings combined controlled vocabulary and free-text terms encompassing colonoscopy, artificial intelligence (AI), CADe, deep learning, adenoma detection, and operator-experience terms (training, fellowship, trainee, novice, non-expert). Reference lists of

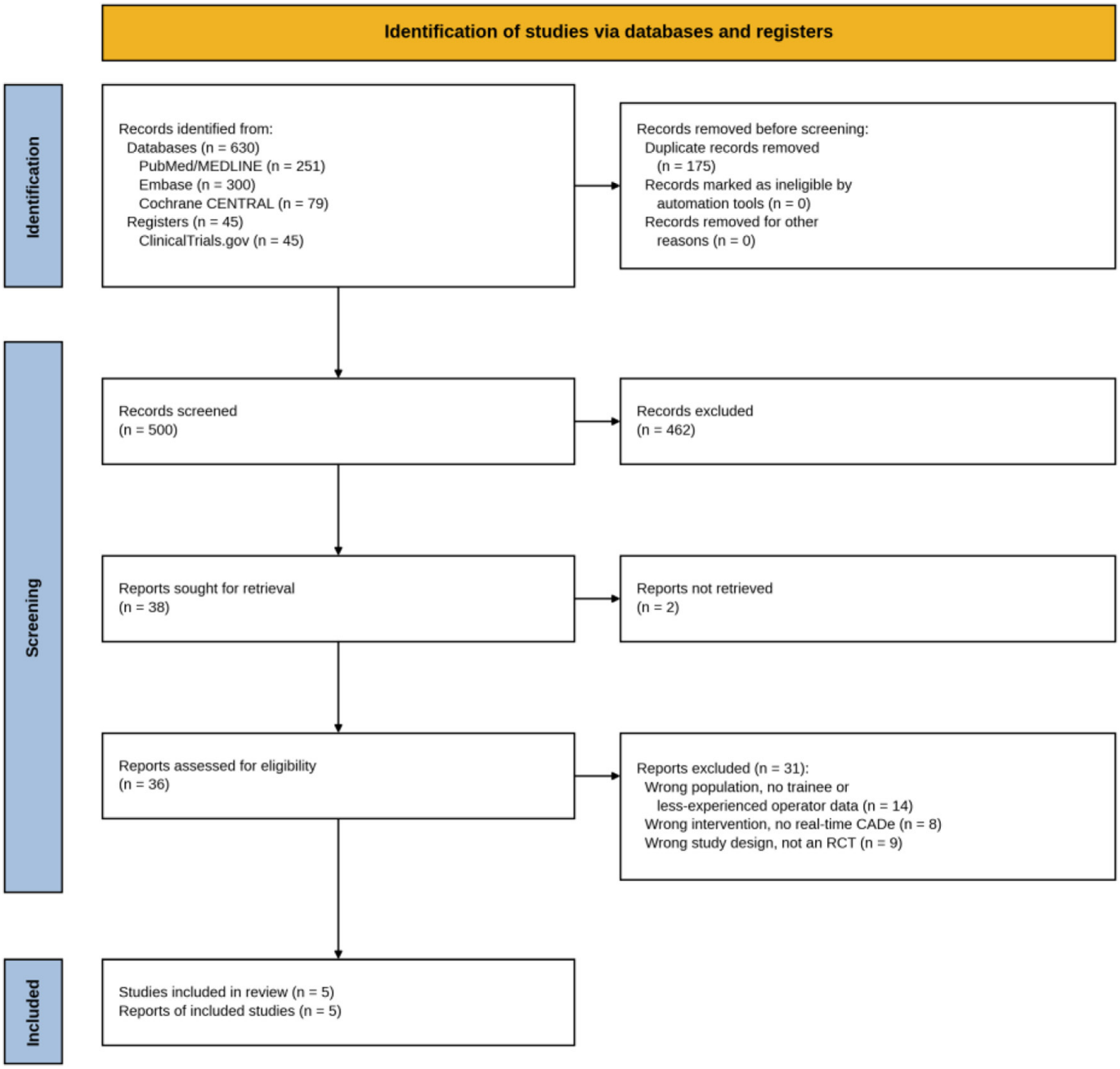


Figure 1. PRISMA 2020 flow diagram of study selection, showing the number of records retrieved from each database and register searched.

identified articles and prior systematic reviews were manually screened to capture additional eligible studies. No language restrictions were applied. Conference abstracts were screened but excluded from the primary pooled analysis when no corresponding peer-reviewed full-text publication of the same dataset could be obtained. The complete search string for each database and register, the date each search was last run, and the number of records retrieved from each source are reported in Supplementary Material 2 (gr.elmerpub.com); database-specific and register-specific yields are also displayed in the PRISMA flow diagram (Fig. 1).

Eligibility criteria

Studies were eligible if they satisfied all of the following criteria: (1) RCT design; (2) adult patients (≥ 18 years) undergoing colonoscopy for any indication; (3) an intervention arm employing a real-time CADe system during withdrawal; (4) a control arm using standard white-light colonoscopy with-

 Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

out CADe; (5) procedures performed entirely or substantially by endoscopists who had not yet attained expert procedural volume, comprising both (a) operators enrolled in formal endoscopic training (novices, gastroenterology fellows, and endoscopists-in-training) and (b) independently practicing endoscopists explicitly designated by the trial investigators as non-expert on the basis of a prespecified lifetime procedure threshold; and (6) reporting of at least one prespecified outcome of interest. Studies were excluded if they involved exclusively experienced or expert endoscopists, evaluated computer-aided diagnosis without a detection component, or were available only as conference abstracts when the underlying full-text dataset could not be obtained.

We adopted this broader population definition deliberately, and we state the rationale explicitly because it determines how the pooled estimates should be read. The operators in Repici 2022 (AID-2) [20] were qualified to perform screening colonoscopy and non-complex therapeutic procedures independently and had each completed fewer than 2,000 lifetime colonoscopies; they were therefore independently practicing, relatively less-experienced endoscopists rather than trainees undergoing supervised training. Restricting the review to operators in formal training would produce a more homogeneous but substantially smaller evidence base and would exclude the largest single contributor of patient-level right-colon data. We therefore defined the review population as trainee and less-experienced endoscopists, and, to keep the narrower clinical question answerable, we prespecified a sensitivity analysis restricted to the four trials whose operators were unambiguously in formal endoscopic training (Section "Statistical analysis"). Throughout the manuscript, we use "trainee" only for operators in formal training and "less-experienced" for independently practicing operators below an investigator-defined volume threshold; the composite population is referred to as "trainee and less-experienced endoscopists."

Outcomes

The prespecified primary outcome was overall ADR, defined as the proportion of colonoscopies in which at least one histologically confirmed adenoma was identified. Secondary outcomes included RADR, defined according to each trial's anatomic boundary (cecum through the transverse colon in Repici 2022 [20] and Lau 2024 [21], and cecum through the ascending colon in Chang 2026 [22]); PDR, defined as the proportion of colonoscopies in which at least one polyp of any histology was identified; and advanced adenoma detection rate (AADR), defined as the proportion of colonoscopies in which at least one adenoma ≥ 10 mm, with villous component, or with high-grade dysplasia was identified. Sessile serrated lesion (SSL) detection rate was added as an outcome of interest during revision but could not be pooled (Section "SSL detection").

Two of the five trials used a back-to-back tandem design (Yamaguchi 2024 [23], Yao 2024 [24]) rather than a parallel-group design, and the appropriateness of combining them requires justification. In both tandem trials, patients were randomized before the first colonoscopy, and the first pass was performed by the trainee under the assigned condition (CADe or standard white light); the second pass served only to estimate the miss rate and did not influence the first-pass result. The first pass of a tandem trial is therefore itself a randomized comparison of trainee detection with versus without CADe, measured on the same patient-level outcome scale as a parallel-group trial. Accordingly, we extracted first-pass, trainee-performed detection rates for both tandem trials for every outcome, so that all pooled estimates reflect lesions detected by the operator under the randomized condition rather than cumulative detection across both passes. We did not pool adenoma miss rate, the primary endpoint of Yamaguchi 2024, with ADR, because these are different quantities. This distinction is important when interpreting Yamaguchi 2024, which was powered for miss rate rather than for a between-arm ADR difference (Section "Primary outcome: overall ADR").

Study selection and data extraction

Two reviewers (HA and RJ) independently screened titles and abstracts, performed full-text review of potentially eligible reports, and extracted data using a standardized form. Extracted variables included study identifier, year, country, trial design, randomization unit, CADe system, operator characteristics (number of operators, level of training, formal training status, and prior procedural experience), patient characteristics (age, sex, colonoscopy indication, and bowel preparation quality), intervention and comparator platform characteristics, primary and secondary outcome events and denominators, withdrawal time, and adverse events. Discrepancies were resolved by discussion and, when necessary, settlement by a third reviewer (HD). The full set of clinical, operator and technical characteristics is presented in Tables 1 and 2.

Risk of bias assessment

Two reviewers independently assessed risk of bias using the Cochrane Risk of Bias 2 (RoB 2) tool, rendering judgments across five domains: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result [25]. Disagreements were resolved by consensus. Per-domain and overall judgments for each included trial are reported in Section "Risk of bias".

Statistical analysis

For dichotomous outcomes, study-level RRs with 95% CIs were calculated and pooled using DerSimonian–Laird random-effects models [26]. Statistical heterogeneity was quantified using the I^2 statistic, with values of approximately 25%, 50%, and 75% interpreted as low, moderate, and substantial heterogeneity, respectively, supplemented by Cochran's Q test [27]. A two-sided P value < 0.05 was considered statistically significant. All analyses were performed in R (version 4.3.2).

Table 1. Characteristics of Included Studies

Study (year)	Country	Design	CADe system	Operator category	Operator definition	CADe n	Control n	Total n
Repici 2022 (AID-2) [20]	Italy; Switzerland	Parallel RCT	GI Genius (Medtronic)	Less experienced	Independently practicing non-experts (< 2,000 lifetime colonoscopies)	330	330	660
Lau 2024 (ENDOAID-TRAIN) [21]	Hong Kong	Parallel RCT	ENDO-AID (Olympus)	Trainee	Endoscopists-in-training (< 500 procedures, < 3 years of training)	386	380	766
Yamaguchi 2024 [23]	Japan	Back-to-back tandem RCT	CAD EYE (Fujifilm)	Trainee	Trainees (0–20 prior colonoscopies)	113	118	231
Yao 2024 [24]	China	Multicenter back-to-back tandem RCT	Custom CNN (EndoAngel)	Trainee	Novices (no prior colonoscopy experience)	227	229	456
Chang 2026 [22]	USA	Pragmatic parallel RCT (room- and fellow-level randomization)	GI Genius (Medtronic)	Trainee	Gastroenterology fellows (PGY 4–6)	401	644	1,045
Total						1,457	1,701	3,158

"Trainee" denotes operators enrolled in formal endoscopic training; "less-experienced" denotes independently practicing operators below an investigator-defined lifetime procedure threshold. Counts are the numbers of patients analyzed in each arm. For the two tandem trials (Yamaguchi 2024, Yao 2024), counts and outcomes refer to the trainee-performed first pass. AID-2: Artificial Intelligence Detection 2 trial; CADe: computer-aided detection; CNN: convolutional neural network; ENDOAID-TRAIN: ENDO-AID in Endoscopists-in-Training trial; PGY: postgraduate year; RCT: randomized controlled trial; USA: United States of America.

For every outcome we additionally calculated a 95% PI, using a t distribution and the estimated between-study variance, to describe the range within which the true effect of a future comparable trial would be expected to lie [28]. Because DerSimonian–Laird CIs can be anticonservative when the number of trials is small, we repeated each pooled analysis using Hartung–Knapp–Sidik–Jonkman inference as a more conservative sensitivity analysis [29]. To convey the practical magnitude of any benefit, absolute RDs were pooled with the same random-effects approach and, where the pooled RD excluded zero, converted to a number needed to treat (NNT).

Robustness was examined in three ways. First, leave-one-out sensitivity analyses were performed for the primary outcome and for every secondary outcome, sequentially omitting each contributing trial. Second, a sensitivity analysis restricted to the four trials of operators in formal endoscopic training was performed for every outcome, excluding Repici 2022 (AID-2), whose operators practiced independently. Third, because a formal subgroup analysis by level of operator experience is not interpretable with five trials, differences across experience strata were examined through a structured narrative comparison of trial-level effect estimates ordered by operator experience (Section "Heterogeneity and sensitivity analyses").

Small-study effects were explored using a Doi plot with the LFK index rather than funnel-plot asymmetry tests, because the latter have very low power with fewer than 10 trials [30]. An LFK index within ± 1 indicates no asymmetry, ± 1 to ± 2 minor asymmetry, and beyond ± 2 major asymmetry. These analyses are exploratory rather than confirmatory and, with three to five contributing trials per outcome, are not a reliable basis for inferring publication bias. Certainty of evidence for each outcome was rated independently by two reviewers using the GRADE framework across the domains of risk of bias, inconsistency, indirectness, imprecision, and publication bias, with disagreements resolved by consensus [31]; the resulting summary-of-findings table is reported in Section "Certainty of evidence".

Institutional review board approval

Not applicable. This study is a synthesis of previously published aggregate data and did not involve the collection of new human or animal subject data; institutional review board approval was therefore not required.

Ethical compliance with human/animal study

Not applicable. This study is a meta-analysis of previously published randomized controlled trials and did not involve direct human or animal subjects research. Each included trial obtained its own institutional review board approval and informed consent as described in their respective publications.

Results

Study selection

The PRISMA 2020 flow diagram is presented in Figure 1. Database searches identified 630 records (PubMed/MEDLINE

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

Table 2. Detailed Participant, Operator, Intervention, and Comparator Characteristics of the Included Trials

Characteristic	Repici 2022 (AID-2) [20]	Lau 2024 [21]	Yamaguchi 2024 [23]	Yao 2024 [24]	Chang 2026 [22]
Participant characteristics					
Patients analyzed (CADe/control)	330/330	386/380	113/118	227/229	401/644
Age, years	62.3 ± 10 (whole cohort)	66.00 ± 10.05/65.36 ± 11.33	63.1 ± 10.8/63.3 ± 11.8	50.63 ± 12.25/49.85 ± 11.68	54.6 ± 11.6/54.2 ± 12.2
Female sex	330/660 (50.0%) (whole cohort)	181/386 (46.9%)/169/380 (44.5%)	55/113 (48.7%)/52/118 (44.1%)	110/227 (48.5%)/106/229 (46.3%)	168/401 (41.9%)/287/644 (44.6%)
Colonoscopy indication	Screening, surveillance, or diagnostic	Screening 7.3%/6.1%; surveillance 32.6%/31.8%; symptomatic 60.1%/62.1%	FIT-positive 46.9%/54.2%; detailed examination 19.5%/22.0%; polyp surveillance 13.3%/11.0%; abdominal symptoms 11.5%/8.5%; other 8.9%/4.2%	Screening 63.9%/63.8%; diagnostic 26.4%/28.0%; surveillance 9.7%/8.3%	GI bleeding 52%/57%; symptoms 17%/17%; surveillance for neoplasia 13%/12%; screening 14%/11%; imaging follow-up 3%/2%
History of previous adenoma	Not separately reported; surveillance included as an eligible indication	Not separately reported; surveillance 32.6%/31.8%. Family history of colorectal cancer 19.5%/15.6%	Polyp surveillance 15/113 (13.3%)/13/118 (11.0%); family history of colorectal cancer 11.5%/11.0%	Not separately reported; surveillance 9.7%/8.3%. Family history 8.8%/7.0%	Not separately reported; surveillance for neoplasia 13%/12%. Family history of colon cancer 8%/7%
Bowel preparation quality	NR	BBPS total 7.85 ± 1.17/7.84 ± 1.21; patients with BBPS 0-1 in any segment excluded	Modified Aronchick excellent or good 86.7%/83.9%; BBPS whole colon 8.4 ± 1.2/8.2 ± 1.5	BBPS adequate (sum ≥6.0 and each segment ≥2.0) 86.3%/88.2%	NR by arm
Baseline control-arm ADR	44.5%	44.5%	61.0%	16.2%	34.9%
Operator characteristics					
Number of operators	10	22 junior endoscopists	6 trainees	8 novices (plus 10 experts for the second pass)	16 fellows
Specialty	Gastroenterology/ endoscopy	Gastroenterologists (78.8%/76.3% of procedures) and surgeons-in-training (21.2%/23.7%)	Gastroenterology	Gastroenterology	Gastroenterology
Formal training status	No, independently practicing	Yes, endoscopists-in-training	Yes, trainees who had started fellowship training	Yes, fellows in training	Yes, GI fellows, PGY 4–6
Years of endoscopy experience	NR; qualified for independent screening colonoscopy and non-complex therapy	< 3 years of training	First year of fellowship training	0 years of independent colonoscopy; > 1 year of gastroenterology fellowship	1st to 3rd year of fellowship

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

Table 2. Detailed Participant, Operator, Intervention, and Comparator Characteristics of the Included Trials - *(continued)*

Characteristic	Repici 2022 (AID-2) [20]	Lau 2024 [21]	Yamaguchi 2024 [23]	Yao 2024 [24]	Chang 2026 [22]
Lifetime colonoscopies performed	< 2,000 per operator	< 500 per operator	0–20 cases per operator	0 prior independent colonoscopies; 20 supervised colonoscopies with expert hands-on training before study entry	Median 84 (IQR 54–127) for PGY 4; 327 (IQR 272–373) for PGY 5; 468 for PGY 6
Experience strata used by investigators	Single stratum (non-expert, < 2,000 procedures)	Beginner (< 200 procedures, n = 12) and intermediate (200–500 procedures, n = 10)	Single stratum	Novice versus expert (expert defined as ≥5,000 colonoscopies)	Fellow year 1 (PGY 4), year 2 (PGY 5), year 3 (PGY 6)
Prior experience with CADe	NR	Yes, all junior endoscopists received training on the CADe system before study initiation	NR	None reported	NR
Intervention characteristics					
CADe platform	GI Genius	ENDO-AID	CAD EYE	EndoAngel (custom CNN)	GI Genius
Manufacturer	Medtronic, Dublin, Ireland	Olympus, Tokyo, Japan	Fujifilm, Tokyo, Japan	Renmin Hospital of Wuhan University, Wuhan, China	Medtronic, Dublin, Ireland
Software/version	NR	Preinstalled on the EVIS X1 CV-1500 processor	NR	Custom deep-learning system on a dedicated host computer	US FDA-approved system; version NR
Visual alerts	Yes, real-time on-screen marking of candidate lesions	Yes, alert flag with picture-in-picture display in normal mode; a target mode is also available	Yes, real-time on-screen marking of candidate lesions	Yes, real-time on-screen marking of candidate lesions	Yes, real-time on-screen marking of candidate lesions
Auditory alerts	NR	NR	NR	NR	NR
CADe active during insertion	No	No	No, colonoscope inserted without CAD EYE in both groups	No	No
CADe active during withdrawal	Yes	Yes	Yes	Yes	Yes
Additional computer-aided diagnosis (CADx)	No	No	Yes, CADx used to characterize lesions when measuring polyps	No	No
Additional computer-aided quality (CAQ) or withdrawal-speed feedback	No	No	No	Yes, computer-aided quality system providing withdrawal-speed and mucosal-exposure feedback	No

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

Table 2. Detailed Participant, Operator, Intervention, and Comparator Characteristics of the Included Trials - *(continued)*

Characteristic	Repici 2022 (AID-2) [20]	Lau 2024 [21]	Yamaguchi 2024 [23]	Yao 2024 [24]	Chang 2026 [22]
Other co-interventions	None	On-site or next-door supervisor available for safety	Expert second pass without CAD EYE (tandem design)	Expert second pass with AI assistance (tandem design)	Usual attending supervision, not standardized
Comparator characteristics					
Endoscopy platform	High-definition colonoscopy	Olympus EVIS X1 CV-1500 processor with CF-HQ290L/I (97.4%/98.4%) or CF-EZ1500DL/I colonoscopes	Fujifilm EC-L600ZP7 colonoscope	Fujinon EC-590WM or EC-600WM; Olympus EVIS LUCERA CV260 or CV290	Standard colonoscopy in a non-AI room
High definition versus standard definition	High definition	High definition	High definition	High definition	High definition
White-light imaging	Yes	Yes	Yes	Yes	Yes
Permitted image-enhancement modalities	NR	Narrow-band imaging or texture and color enhancement imaging permitted for polyp characterization only; magnification and chromoendoscopy not allowed	Narrow-band-type imaging available through the CAD EYE platform for characterization	NR	Not mandated or tracked
Distal attachment devices	NR	Prohibited (transparent cap, Endocuff Vision)	NR	NR	Mucosal exposure devices such as distal caps were neither mandated nor tracked
Other detection-enhancing techniques	None reported	None reported	None reported	None reported	Right-colon examination techniques such as retroflexion and second view were neither mandated nor tracked

Values are n, n (%), or mean ± standard deviation unless otherwise stated, and are presented as CADe arm/control arm where both are available. Data are as reported in the primary publications; where a trial reported a characteristic only for the whole cohort rather than by arm, this is indicated, and NR denotes a characteristic not reported in the source publication. ADR: adenoma detection rate; AID-2: Artificial Intelligence Detection 2 trial; BBPS: Boston Bowel Preparation Scale; CADe: computer-aided detection; CADx: computer-aided diagnosis; CAQ: computer-aided quality; CNN: convolutional neural network; FDA: United States Food and Drug Administration; FIT: fecal immunochemical test; GI: gastrointestinal; IQR: interquartile range; NR: not reported in the source publication; PGY: postgraduate year.

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

Table 3. Pooled Outcome Summary

Outcome	Trials (n)	Patients (n)	Pooled RR (95% CI)	P value	I^2 (%)	95% PI	Hartung–Knapp 95% CI	Pooled RD (95% CI); NNT
Overall ADR (primary)	5	3,158	1.17 (1.07–1.29)	0.0012	28.8	0.95–1.46	1.02–1.35	+6.9% (+2.7 to +11.0); NNT 15
Right-colon ADR	3	2,471	1.31 (1.14–1.50)	0.0002	26.6	0.86–1.97	0.96–1.77	+8.2% (+4.6 to +11.7); NNT 12
Polyp detection rate	4	2,498	1.16 (1.08–1.25)	0.0001	27.8	0.97–1.39	1.02–1.32	+8.9% (+3.7 to +14.2); NNT 11
Advanced adenoma detection rate	3	1,882	0.97 (0.73–1.29)	0.8240	0.0	0.52–1.81	0.56–1.68	+0.2% (−2.2 to +2.6); NNT NA

Risk ratios and risk differences were pooled with DerSimonian–Laird random-effects models. Event counts and denominators for every trial are shown in Figures 2 to 5 and in Supplementary Material 3 (gr.elmerpub.com). Prediction intervals describe the range within which the true effect of a future comparable trial would be expected to lie. Hartung–Knapp confidence intervals are presented as a more conservative alternative to DerSimonian–Laird inference given the small number of trials. NNT is derived from the pooled risk difference and is reported only where that difference excluded zero (overall ADR 95% CI 9 to 37; right-colon ADR 95% CI 9 to 22; polyp detection rate 95% CI 7 to 27). Advanced adenoma detection rate values were revised during peer review to use the trainee-performed first pass counts reported by Yao 2024, for consistency with the extraction rule stated in Section "Outcomes". ADR: adenoma detection rate; CI: confidence interval; NA: not applicable; NNT: number needed to treat; PI: prediction interval; RD: risk difference; RR: risk ratio.

251, Embase 300, and Cochrane CENTRAL 79) and register searches a further 45; the full search strategy for each source is reported in Supplementary Material 2 (gr.elmerpub.com). After removal of 175 duplicates, 500 records underwent title and abstract screening, of which 462 were excluded. Thirty-eight reports were sought for retrieval and 36 were assessed for full-text eligibility; 31 were excluded (14 with no trainee- or less-experienced-operator data, eight without a real-time CADe intervention, and nine with a non-randomized design). Five RCTs enrolling a total of 3,158 patients met all inclusion criteria and were included in the quantitative synthesis [20–24].

Study characteristics

Characteristics of the included trials are summarized in Table 1, and a detailed comparison of participant, operator, intervention, and comparator characteristics is provided in Table 2. The five trials were conducted between 2022 and 2026 across Europe, Asia, and the United States, evaluating four distinct CADe platforms: GI Genius (Medtronic, Dublin, Ireland) in Repici 2022 [20] and Chang 2026 [22]; ENDO-AID (Olympus, Tokyo, Japan) in Lau 2024 [21]; CAD EYE (Fujifilm, Tokyo, Japan) in Yamaguchi 2024 [23]; and a custom convolutional neural network (EndoAngel) in Yao 2024 [24]. Two trials employed a parallel-group design (Lau 2024 [21], Chang 2026 [22]), two used a back-to-back tandem design (Yamaguchi 2024 [23], Yao 2024 [24]), and Repici 2022 (AID-2) was a parallel-group trial conducted exclusively in independently practicing non-expert endoscopists [20]. Operator experience varied substantially across studies, ranging from endoscopists with no prior colonoscopy experience (Yao 2024 [24]) to gastroenterology fellows in postgraduate years 4–6 (Chang 2026 [22]) and to independently practicing endoscopists with up to 2,000 prior procedures (Repici 2022 [20]). Four trials (Lau 2024, Yamaguchi 2024, Yao 2024, Chang 2026) enrolled operators in formal endoscopic training; only Repici 2022 did not.

Primary outcome: overall ADR

CADe-assisted colonoscopy was associated with a significantly higher overall ADR compared with standard colonoscopy. Across the five included trials (n = 3,158 patients; 673 of 1,457 (46.2%) with CADe versus 650 of 1,701 (38.2%) with standard colonoscopy), the pooled RR was 1.17 (95% CI 1.07–1.29; P = 0.0012; I^2 = 28.8%) (Table 3; Fig. 2). Study-level event counts and denominators for every trial and every outcome are provided in Supplementary Material 3 (gr.elmerpub.com). The corresponding pooled absolute RD was +6.9% (95% CI +2.7% to +11.0%; P = 0.0012), equivalent to one additional patient with at least one adenoma detected for every 15 colonoscopies performed with CADe (95% CI 9 to 37). The 95% PI was 0.95 to 1.46, and the more conservative Hartung–Knapp CI was 1.02 to 1.35 (P = 0.033). The direction of effect favored CADe in four of five trials [20, 21, 23, 24]. The sole trial in which the point estimate did not favor CADe was Yamaguchi 2024, which employed a tandem design with adenoma miss rate as its primary endpoint and was not powered for between-arm ADR comparison (the trial did, however, demonstrate a significant reduction in adenoma miss rate, 25.6% versus 38.6%, P = 0.033) [23].

Secondary outcomes

RADR

Three trials (Repici 2022 [20], Lau 2024 [21], Chang 2026 [22]) reported patient-level RADR (n = 2,471; 396 of 1,117 (35.5%) versus 342 of 1,354 (25.3%)). The pooled effect fa-

 Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

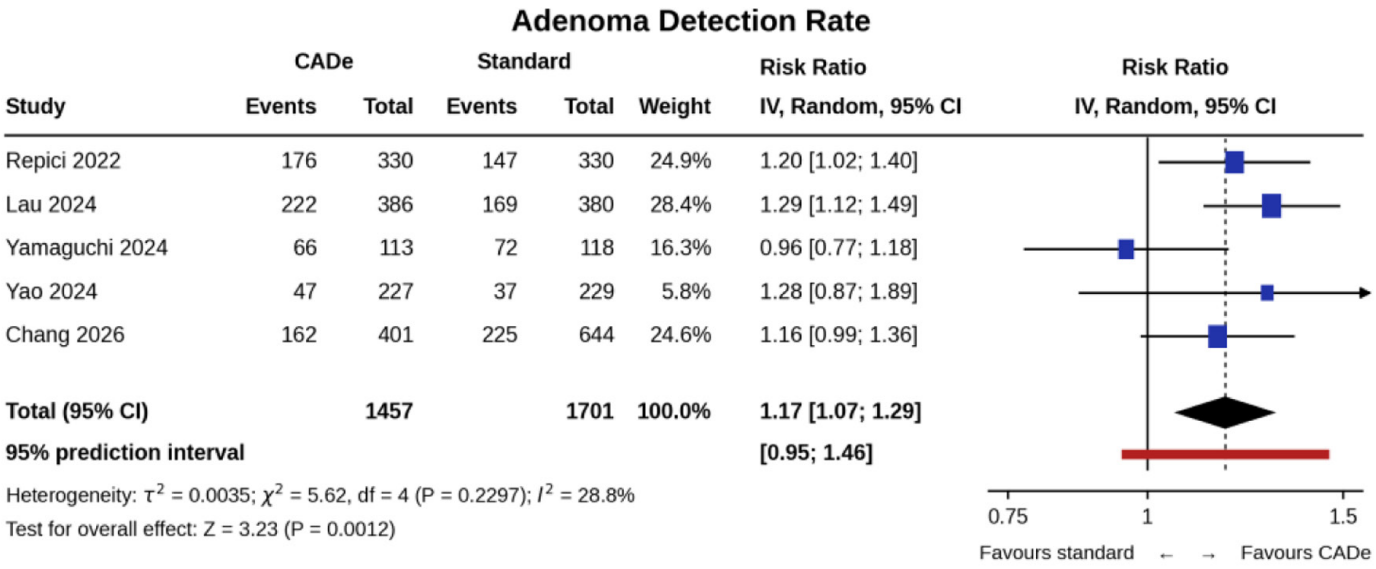


Figure 2. Forest plot for the primary outcome (overall adenoma detection rate), with the 95% prediction interval shown beneath the pooled estimate. CADe: computer-aided detection; CI: confidence interval; IV: inverse variance.

vored CADe with a larger magnitude than observed for overall ADR, with low heterogeneity: RR 1.31 (95% CI 1.14–1.50; P = 0.0002; I^2 = 26.6%) (Table 3; Fig. 3). The pooled absolute RD was +8.2% (95% CI +4.6% to +11.7%; P < 0.0001; NNT 12, 95% CI 9 to 22). The 95% PI was 0.86 to 1.97, and the Hartung–Knapp CI was 0.96 to 1.77 (P = 0.064).

PDR

Four trials reported PDR (n = 2,498; 749 of 1,127 (66.5%) versus 774 of 1,371 (56.5%)). CADe was associated with a significantly higher PDR: RR 1.16 (95% CI 1.08–1.25; P = 0.0001; I^2 = 27.8%) (Table 3; Fig. 4). The pooled absolute RD was +8.9% (95% CI +3.7% to +14.2%; P = 0.0008; NNT 11, 95% CI 7 to 27). The 95% PI was 0.97 to 1.39, and the Hartung–Knapp CI was 1.02 to 1.32 (P = 0.035).

AADR

Across the three trials reporting AADR (n = 1,882; 85 of 943 (9.0%) versus 87 of 939 (9.3%)), the pooled effect did not differ significantly between groups: RR 0.97 (95% CI 0.73–1.29; P = 0.82; I^2 = 0.0%) (Table 3; Fig. 5). The pooled absolute RD was +0.2% (95% CI −2.2% to +2.6%; P = 0.86), the 95% PI was 0.52 to 1.81, and the Hartung–Knapp CI was 0.56 to 1.68 (P = 0.83). Because only three trials contrib-

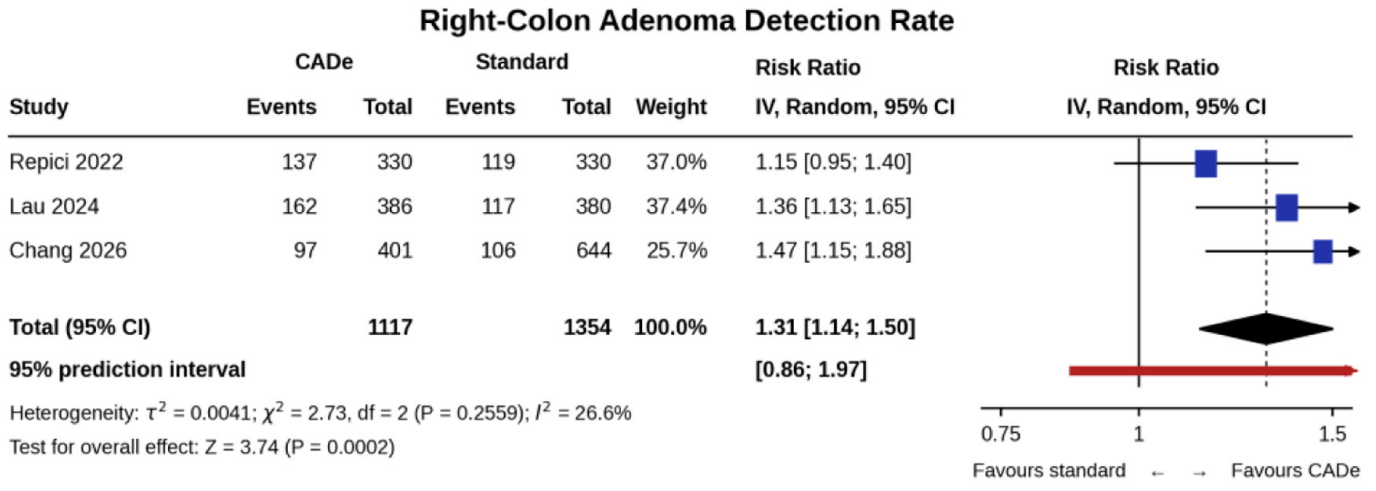


Figure 3. Forest plot for right-colon adenoma detection rate, with the 95% prediction interval shown beneath the pooled estimate. CADe: computer-aided detection; CI: confidence interval; IV: inverse variance.

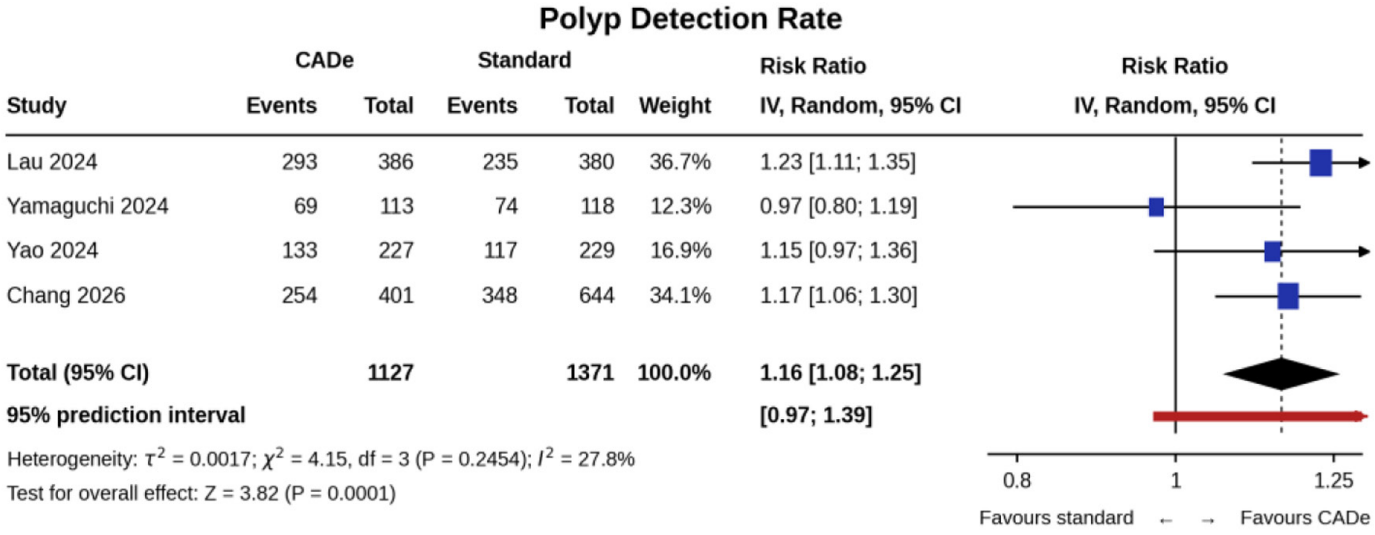


Figure 4. Forest plot for polyp detection rate, with the 95% prediction interval shown beneath the pooled estimate. CADe: computer-aided detection; CI: confidence interval; IV: inverse variance.

uted 172 events in total, this analysis is underpowered, and the CI remains compatible with both a clinically meaningful reduction and a clinically meaningful increase in advanced adenoma detection.

SSL detection

Serrated lesion detection was reported by only one of the five included trials and could not be pooled. Lau 2024 reported an SSL detection rate of 2.1% (8 of 386) with CADe versus 1.8% (7 of 380) with standard colonoscopy (adjusted RR 1.14, 95% CI 0.42–3.11; P = 0.801) [21]. Yamaguchi 2024 reported per-lesion rather than per-patient histology, with one SSL and no traditional serrated adenomas detected in the CADe arm compared with four SSLs and two traditional serrated adenomas in the control arm [23]. Repici 2022, Yao 2024, and Chang 2026 did not report a patient-level serrated lesion detection rate [20, 22, 24]. Serrated lesion detection is therefore an important missing outcome in this evidence base, and RADR should not be assumed to capture proximal serrated neoplasia (Section "Strengths and limitations").

Safety and adverse events

No quantitative synthesis of safety outcomes was performed, because adverse events were reported too sparsely and too heterogeneously across the five trials to support meta-analysis; study-level safety and adverse-event data for all included tri-

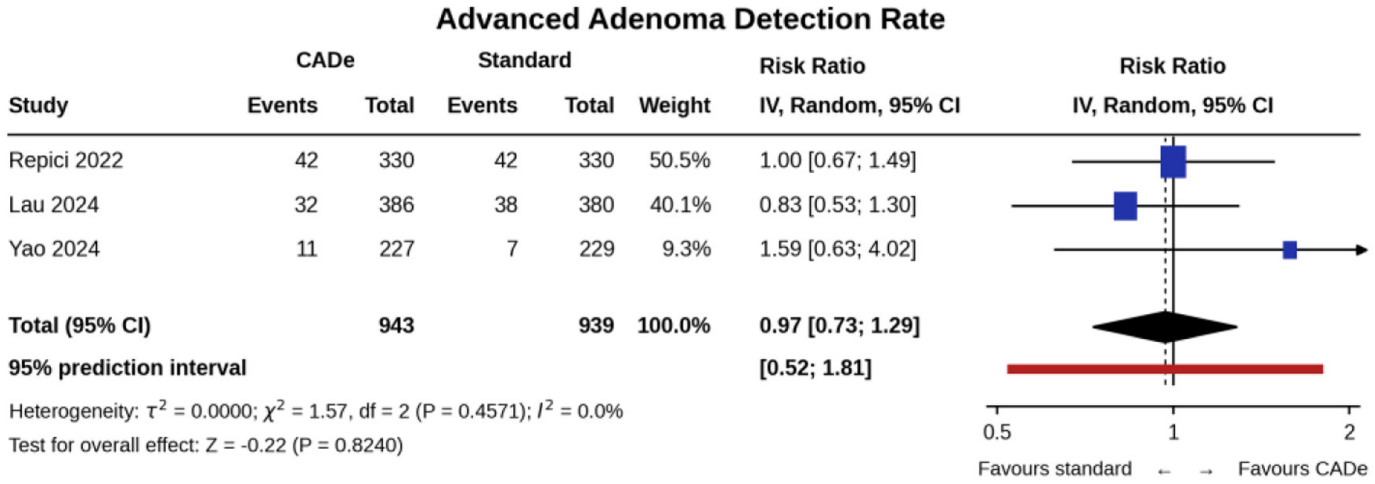


Figure 5. Forest plot for advanced adenoma detection rate, with the 95% prediction interval shown beneath the pooled estimate. Advanced adenoma counts for Yao 2024 are the trainee-performed first-pass values, consistent with the extraction rule stated in Section "Outcomes". CADe: computer-aided detection; CI: confidence interval; IV: inverse variance.

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

Table 4. Risk of Bias Assessment (Cochrane RoB 2)

Study	D1: Randomization process	D2: Deviations from intended interventions	D3: Missing outcome data	D4: Measurement of the outcome	D5: Selection of the reported result	Overall
Repici 2022 [20]	Low	Some concerns	Low	Low	Some concerns	Some concerns
Lau 2024 [21]	Low	Low	Low	Low	Low	Low
Yamaguchi 2024 [23]	Low	Some concerns	Low	Low	Low	Some concerns
Yao 2024 [24]	Low	Some concerns	Low	Low	Low	Some concerns
Chang 2026 [22]	Some concerns	Some concerns	Low	Low	Low	Some concerns

Judgments in domain 2 reflect the fact that operator exposure to the CADe display cannot be blinded in any trial of this intervention. The domain 1 judgment for Chang 2026 reflects randomization at the level of the procedure room and fellow day rather than the individual patient. D: domain; RoB 2: Cochrane Risk of Bias 2 tool.

als are summarized in Supplementary Material 4 (gr.elmerpub.com). Lau 2024 was the only trial to report adverse-event counts, and it reported no significant excess in the CADe arm: three procedure-related serious adverse events occurred across both arms, one post-polypectomy coagulation syndrome in the CADe arm and two delayed post-polypectomy bleeds in the control arm [21]. Yamaguchi 2024 listed perioperative severe adverse events among the variables recorded but did not report them, Yao 2024 stated safety of the tandem procedure as a study aim without reporting adverse-event counts, and Repici 2022 and Chang 2026 reported no adverse-event data at all [20, 22–24]. A between-arm safety comparison was therefore possible in only one of the five trials, and the safety of CADe-assisted colonoscopy in this population remains essentially untested (Section "Strengths and limitations"). Lau 2024 additionally reported a substantially higher non-neoplastic resection rate with CADe (52.1% versus 35.0%; adjusted RR 1.70, 95% CI 1.37–2.11; P < 0.001), which is relevant to the downstream burden of CADe-driven detection [21].

Heterogeneity and sensitivity analyses

Statistical heterogeneity was low for the primary outcome (I^2 = 28.8%), RADR (I^2 = 26.6%), and PDR (I^2 = 27.8%), and absent for AADR (I^2 = 0.0%). Leave-one-out sensitivity analyses were performed for all four outcomes and are displayed in Supplementary Material 5 (panels A to D, gr.elmerpub.com). For the primary outcome, the pooled RR ranged from 1.13 (95% CI 1.02–1.26) when Lau 2024 was omitted to 1.22 (95% CI 1.12–1.33) when Yamaguchi 2024 was omitted and remained statistically significant in every iteration. For RADR, the pooled RR ranged from 1.25 (95% CI 1.06–1.48, omitting Chang 2026) to 1.40 (95% CI 1.21–1.63, omitting Repici 2022), remaining significant throughout. For PDR, the pooled RR ranged from 1.12 (95% CI 1.02–1.24, omitting Lau 2024) to 1.19 (95% CI 1.12–1.27, omitting Yamaguchi 2024), again significant throughout, although omission of Chang 2026 brought the lower confidence bound to the null (RR 1.14, 95% CI 1.00–1.29). For AADR, the pooled RR ranged from 0.92 (95% CI 0.68–1.24) to 1.07 (95% CI 0.74–1.55) and was non-significant in every iteration.

In the prespecified sensitivity analysis restricted to the four trials of operators in formal endoscopic training (Repici 2022 excluded; n = 2,498), CADe remained associated with a significantly higher ADR, although with a slightly attenuated point estimate, a wider CI and greater heterogeneity: RR 1.16 (95% CI 1.01–1.33; P = 0.030; I^2 = 46.3%; 95% PI 0.81–1.67; RD +6.2%, 95% CI +0.9% to +11.5%) (Supplementary Material 6, gr.elmerpub.com). The right-colon effect was, if anything, larger in this restricted population (two trials, n = 1,811): RR 1.40 (95% CI 1.21–1.63; P < 0.0001; I^2 = 0.0%; RD +9.0%, 95% CI +4.9% to +13.0%). PDR was unchanged, because Repici 2022 did not contribute patient-level PDR data (RR 1.16, 95% CI 1.08–1.25). AADR remained null (two trials, n = 1,222; RR 1.00, 95% CI 0.56–1.79). The direction and statistical significance of the primary and right-colon findings therefore do not depend on inclusion of the single trial of independently practicing non-expert operators.

A formal subgroup analysis stratified by level of operator experience is not interpretable with five trials, so trial-level estimates were compared narratively in order of increasing operator experience. Yao 2024 enrolled complete novices with no prior colonoscopy experience (ADR RR 1.28, 95% CI 0.87–1.89); Yamaguchi 2024 enrolled trainees with 0–20 prior cases (RR 0.96, 95% CI 0.77–1.18); Lau 2024 enrolled endoscopists-in-training with fewer than 500 procedures and under 3 years of training (RR 1.29, 95% CI 1.12–1.49); Chang 2026 enrolled gastroenterology fellows whose median prior colonoscopy volume ranged from 84 in the first year to 468 in the third year (RR 1.16, 95% CI 0.99–1.36); and Repici 2022 enrolled independently practicing endoscopists with fewer than 2,000 procedures (RR 1.20, 95% CI 1.02–1.40). No monotonic relationship between operator experience and the magnitude of CADe benefit is apparent, and the two trials at the extremes of the experience spectrum produced similar point estimates. The single trial that did not favor CADe (Yamaguchi 2024) is distinguished by its endpoint and power rather than by its position on the experience spectrum. These

Table 5. GRADE Summary of Findings: CADe-Assisted Versus Standard Colonoscopy Performed by Trainee and Less-Experienced Endoscopists

Outcome (trials; patients)	Anticipated absolute effects	Relative effect (95% CI)	Certainty of evidence	Reasons for downgrading
Overall adenoma detection rate (5 trials; 3,158 patients)	38.2% with standard colonoscopy versus 46.2% with CADe; risk difference +6.9% (+2.7 to +11.0); NNT 15	RR 1.17 (1.07–1.29)	Moderate	Risk of bias (−1): operator exposure to the CADe display cannot be blinded and 4 of 5 trials were rated at some concerns using RoB 2. Inconsistency, indirectness, and imprecision not serious; publication bias undetected.
Right-colon adenoma detection rate (3 trials; 2,471 patients)	25.3% versus 35.5%; risk difference +8.2% (+4.6 to +11.7); NNT 12	RR 1.31 (1.14–1.50)	Low	Risk of bias (−1): as above. Imprecision (−1): only 3 contributing trials, and the Hartung-Knapp confidence interval (0.96–1.77) crosses the null. Inconsistency and indirectness not serious; publication bias undetected.
Polyp detection rate (4 trials; 2,498 patients)	56.5% versus 66.5%; risk difference +8.9% (+3.7 to +14.2); NNT 11	RR 1.16 (1.08–1.25)	Low	Risk of bias (−1): as above. Indirectness (−1): polyp detection rate counts lesions of any histology rather than neoplasia, and CADe substantially increased the non-neoplastic resection rate in the single trial that measured it. Inconsistency and imprecision not serious; publication bias undetected.
Advanced adenoma detection rate (3 trials; 1,882 patients)	9.3% versus 9.0%; risk difference +0.2% (−2.2 to +2.6)	RR 0.97 (0.73–1.29)	Very low	Risk of bias (−1): as above. Imprecision (−2): 3 trials contributing 172 events, with a confidence interval compatible with both appreciable benefit and appreciable harm. Inconsistency and indirectness not serious; publication bias undetected.

All outcomes began as high-certainty evidence because all included studies were randomized controlled trials; (−1) and (−2) denote the number of levels by which certainty was downgraded in that domain. Doi plots and LFK indices were generated for all outcomes but are exploratory with three to five trials and were not used as a basis for downgrading. Certainty definitions: Moderate, the true effect is likely close to the estimate but may be substantially different; Low, the true effect may be substantially different from the estimate; Very low, the estimate is very uncertain. CADe: computer-aided detection; CI: confidence interval; GRADE: Grading of Recommendations, Assessment, Development and Evaluation; NNT: number needed to treat; RoB 2: Cochrane Risk of Bias 2 tool; RR: risk ratio.

trial-level comparisons are confounded by CADe platform, trial design, and baseline ADR, and should be regarded as hypothesis-generating.

Risk of bias

RoB 2 judgments are summarized in Table 4. Overall, one trial was judged at low risk of bias (Lau 2024 [21]) and four at some concerns, most commonly attributable to the inherently unblindable nature of operator exposure to the CADe display and to design-specific considerations, including tandem second-pass outcome ascertainment and room-level rather than patient-level randomization.

Assessment of small-study effects

Doi plots with the LFK index were generated for all four outcomes (Supplementary Material 7, panels A to D, gr.elmerpub.com). For the primary outcome, the LFK index was −0.79, within the range conventionally interpreted as no asymmetry. The LFK index was −0.26 for RADR, +1.91 for AADR (minor asymmetry), and −2.39 for PDR (major asymmetry). We emphasize that these indices are computed from three to five trials per outcome, that the LFK index has been shown to have limited reliability in this setting, and that the values reported here are exploratory and should not be interpreted as evidence for or against publication bias.

Certainty of evidence

GRADE ratings for all four outcomes are presented in Table 5. Certainty was rated moderate for overall ADR, downgraded once for risk of bias because operator awareness of the CADe display cannot be blinded and four of five trials were rated at some concerns. Certainty was rated low for RADR (downgraded for risk of bias and for imprecision, given three contributing trials and a Hartung–Knapp CI that crosses the null) and low for PDR (downgraded for risk of bias and for indirectness, because PDR counts lesions of any histology and CADe increased non-neoplastic resection in the one trial that measured it). Certainty was rated very low for AADR, downgraded once for risk of bias and twice for very serious imprecision, the CI being compatible with both appreciable benefit and appreciable harm.

Discussion

Principal findings

In this updated meta-analysis of five RCTs encompassing 3,158

Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

patients, CADe-assisted colonoscopy was associated with a significant improvement in overall ADR among trainee and less-experienced endoscopists (RR 1.17, 95% CI 1.07–1.29; absolute RD +6.9%, NNT 15). The magnitude of benefit was larger for right-colon adenoma detection (RR 1.31, 95% CI 1.14–1.50; absolute RD +8.2%, NNT 12), with low heterogeneity across the three contributing trials. PDR was likewise significantly higher with CADe, whereas detection of advanced adenomas did not differ between arms. The findings were robust to leave-one-out analysis and to restriction of the analysis to operators in formal endoscopic training. However, the 95% PIs for every outcome crossed the null, and the certainty of the evidence was moderate at best, considerations that temper the strength of the inferences that can be drawn (Sections "Interpretation of the PIs" and "Strengths and limitations").

Comparison with prior literature

Prior pooled analyses of CADe in experienced or mixed-experience endoscopist populations have consistently demonstrated ADR improvements of comparable or somewhat larger magnitude than the operator-restricted estimate reported here. Hassan et al pooled five randomized trials (4,354 patients) and reported a pooled ADR RR of 1.44 (95% CI 1.27–1.62), while a subsequent and larger meta-analysis encompassing 21 randomized trials in 18,232 patients estimated a pooled RR of 1.24 (95% CI 1.16–1.33) [14, 15]. A network meta-analysis comparing CADe with advanced imaging modalities and a separate systematic review focused on polyp detection yielded broadly concordant estimates [16, 17]. The most directly comparable prior synthesis restricted to trainees was a 2024 conference abstract that pooled four randomized trials and reported a non-significant trend toward improved ADR (RR 1.15, 95% CI 0.99–1.34); to our knowledge, this work is available only in abstract form and has not been published as a peer-reviewed full manuscript as of the date of this analysis [18]. Our synthesis additionally incorporates the recently published Chang 2026 trial, a US fellowship-based pragmatic, randomized trial in which the overall ADR difference did not itself reach statistical significance (40.5% vs. 35.0%; mean difference 5.5%, 95% CI −4.3% to 15.3%) but which reported a significant right-colon benefit (24.1% vs. 16.5%; mean difference 7.6%, 95% CI 1.7% to 13.5%). With this trial included, our random-effects synthesis yields a pooled overall ADR estimate that crosses the threshold of statistical significance, accompanied by a particularly strong effect in the right colon [22].

The right-colon finding carries clinical weight, but its interpretation requires care. Post-colonoscopy interval colorectal cancers arise disproportionately in the proximal colon, and root-cause analyses have attributed most of these failures to procedural factors, including missed lesions [6, 7]. The observation that CADe improves right-colon detection across three methodologically distinct trials enrolling operators at different career stages suggests that CADe addresses one of the most clinically consequential gaps in the early learning curve. It does not, however, demonstrate that interval colorectal cancer is reduced. None of the included trials measured post-colonoscopy colorectal cancer, colorectal cancer mortality, or surveillance intervals as outcomes; all measured detection. The inference from higher ADR to fewer interval cancers rests on observational cohort data [2, 3] rather than on the trials synthesized here, and whether a CADe-associated increase in ADR translates into a reduction in interval colorectal cancer remains unproven. A further caveat is that SSLs contribute substantially to post-colonoscopy colorectal cancer in the proximal colon and were not analyzed here (Section "SSL detection"); RADR should therefore not be treated as a complete measure of proximal detection quality.

Notably, several real-world implementation studies in general endoscopist populations have reported attenuated or absent CADe benefits, including a single-center pragmatic implementation trial, a pre/post observational study in routine practice, and a recent systematic review of real-world studies [32–34]. Whether the benefit reported here is preserved upon broad implementation across training programs remains an important open question. Non-randomized studies in trainees, including a retrospective fellow cohort, a comparative cohort of AI-trained versus conventionally trained endoscopists, a comparison of trainee versus expert ADR with CADe, and a tandem trainee–expert study generally report directionally consistent benefits but cannot be pooled with the randomized evidence and remain subject to confounding [35–38].

Interpretation of the PIs

CIs describe the precision of the average effect across the trials analyzed; PIs describe where the true effect of a future comparable trial would be expected to fall and therefore incorporate between-study heterogeneity. For every outcome in this review, the 95% PI crossed the null: 0.95 to 1.46 for overall ADR, 0.86 to 1.97 for RADR, 0.97 to 1.39 for PDR, and 0.52 to 1.81 for AADR. Read literally, this means that although the average effect of CADe on adenoma detection in this population is a relative increase of approximately 17%, a future trial in a comparable setting could plausibly observe no benefit, and for RADR could plausibly observe either a substantially larger benefit or a small decrement. This tempers the confidence with which the pooled estimates can be extrapolated to a specific training program or CADe platform.

Two caveats apply. First, PIs estimated from three to five trials are themselves imprecise, because the between-study variance is estimated with very few degrees of freedom; the RADR PI (0.86 to 1.97) is wide chiefly for this reason. Second, the Hartung–Knapp CIs, which are more conservative than DerSimonian–Laird intervals when the number of trials is small, continued to exclude the null for ADR (1.02–1.35) and PDR (1.02–1.32) but not for RADR (0.96–1.77). Taken together, these analyses indicate that the overall ADR benefit is reasonably robust to the method of inference, whereas the right-colon estimate is the least statistically stable and should be regarded as the most in need of confirmation.

CADe as a performance aid versus an educational intervention

The trials synthesized here establish that trainees and less-ex-

perienced endoscopists detect more adenomas while a CADe system is switched on. They do not establish that CADe trains them to become better unaided endoscopists, and these are different claims that warrant separate treatment. Every included trial measured detection during the assisted procedure; none tested operators during a subsequent AI-free period. The evidence therefore supports CADe as a performance-assistance tool and is silent on CADe as an educational intervention.

There is a plausible counterargument that prolonged reliance on visual prompts could work against skill acquisition. A trainee who learns to withdraw with an algorithm flagging candidate lesions may attend less to subtle mucosal changes, to systematic visual scanning, to fold and flexure examination, and to the cognitive decision of whether a fleeting abnormality warrants a second look. Automation bias which is the tendency to defer to an automated recommendation and to under-weight contradictory evidence is well described in other domains of clinical decision support, and CADe could in principle improve immediate procedural quality while partially substituting for development of the very pattern-recognition skills that trainees ultimately need to exercise independently. This is not an argument against using CADe during fellowship; it is an argument for evaluating CADe with the same educational standard applied to other training technologies.

That standard is instructive. Simulation-based endoscopy training has generally been judged not by whether performance improves while the simulator is in use, but by whether skills transfer to real procedures. A systematic review of virtual-reality tools for gastrointestinal endoscopy training found that virtual-reality training improved technical skills and reduced procedure time both in the simulated environment and in real patients and concluded that it functions best as a supplement to, rather than a replacement for, conventional bedside instruction [39]. CADe training should probably be evaluated against the same principle of durable skill transfer. Future randomized trials could allocate CADe exposure during training and then assess trainees during a prespecified AI-free period, with outcomes including unaided ADR, unaided proximal ADR, serrated lesion detection, withdrawal technique and mucosal exposure, false-positive reactions, and, where feasible, visual-attention measures. Until such trials are performed, the effect of CADe exposure during training on subsequent independent competency remains the single most important unanswered question raised by this literature, and it is not answered by the data synthesized here.

CADe within the wider landscape of technology-assisted colonoscopy

Detection is only one component of an increasingly automated colonoscopy workflow, and training curricula will need to address the others. Artificial intelligence now extends beyond CADe into real-time prediction of colorectal polyp histology, where a systematic review and meta-analysis of diagnostic accuracy during the procedure itself suggests that computer-aided diagnosis offers particularly useful support to less experienced operators [40]. This raises a question specific to training program design: should an algorithm merely alert a trainee to the presence of a lesion, or should trainees also be required to characterize a lesion independently before viewing the algorithm's prediction? A staged interaction of this kind, in which the trainee commits to an assessment before the machine's output is revealed, may help preserve independent judgment and reduce cognitive dependence.

Technological augmentation of colorectal screening is also broadening beyond the colonoscope. Colon capsule endoscopy, for example, has achieved broadly comparable PDRs to conventional colonoscopy in pooled analyses, with both modalities yielding detection rates of approximately 61%, although it cannot perform therapeutic intervention and a positive study still requires subsequent colonoscopy [41]. The relevant implication for endoscopy education is not that any single technology will displace conventional colonoscopy, but that trainees will increasingly need to be taught how to work with a range of automated and semi-automated systems and critically appraising their outputs rather than simply deferring to them.

Clinical and training implications

These findings carry implications for both training programs and healthcare systems considering routine CADe deployment. CADe may function as a "virtual second observer" during the early learning curve, reducing inter-operator variability in colonoscopy quality and particularly improving right-colon detection. Modeling studies of CADe in general screening populations suggest favorable cost-effectiveness profiles, and current World Endoscopy Organization guidance endorses cautious clinical integration [42, 43]. The lack of a detectable effect on advanced adenoma detection in this analysis should be interpreted with caution: advanced lesions are uncommon, and the available data are underpowered for this outcome. Nonetheless, this null result deserves emphasis rather than a footnote. It remains possible that a meaningful part of the observed ADR improvement is driven by smaller or lower-risk lesions, a possibility reinforced by the substantially higher non-neoplastic resection rate observed with CADe in Lau 2024 [21]. Any implementation decision should therefore weigh the additional detection of small lesions, the associated polypectomy and pathology burden, and the potential for shortened surveillance intervals against a benefit for advanced neoplasia that has not been demonstrated.

Strengths and limitations

This meta-analysis has several strengths. Inclusion was restricted to RCTs conducted specifically in trainee or less-experienced endoscopist populations, minimizing confounding by operator experience relative to mixed-experience syntheses. The most recent available randomized evidence was incorporated, including a 2026 US fellowship-based pragmatic trial not included in prior trainee-focused syntheses [22]. Right-colon adenoma detection was evaluated as a discrete outcome,

 Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

revealing a particularly strong effect for this quality indicator, an analysis that prior trainee-focused syntheses have not separately reported. In addition, the robustness of every outcome was examined through leave-one-out analysis, a restricted-population sensitivity analysis, conservative Hartung–Knapp inference, and PIs, and the certainty of the evidence was formally rated using GRADE.

Several important limitations must be acknowledged, and the most consequential concern the clinical and methodological heterogeneity of the included trials. First, the operator populations were heterogeneous by design: they ranged from endoscopists with no prior colonoscopy experience (Yao 2024) through trainees with 0–20 prior cases (Yamaguchi 2024), endoscopists-in-training with fewer than 500 procedures (Lau 2024), and fellows with a median of 84 to 468 prior procedures (Chang 2026), to independently practicing operators with up to 2,000 procedures (Repici 2022). An endoscopist approaching 2,000 colonoscopies differs educationally from one performing a first 20 cases, and the pooled estimate is an average across this spectrum rather than an estimate for any single training stage. We have addressed this by broadening and stating the population definition explicitly, by reserving the term "trainee" for operators in formal training, and by reporting the restricted-population sensitivity analysis; it nevertheless remains a limitation.

Second, the interventions themselves were not equivalent. The trials evaluated four distinct CADe platforms, and the accompanying software differed in kind and not only in brand: Yamaguchi 2024 used CAD EYE, which incorporates computer-aided diagnosis alongside detection, and Yao 2024 combined a detection algorithm with a computer-aided quality system providing withdrawal-speed and mucosal-exposure feedback, whereas Repici 2022, Lau 2024, and Chang 2026 evaluated detection alone. Pooled estimates should therefore be read as an average effect across a family of related but non-identical interventions rather than as a system-specific effect. Third, trial designs differed: three were parallel-group trials and two were back-to-back tandem trials, and among the parallel-group trials Chang 2026 randomized at the level of the procedure room and fellow-day rather than the individual patient. Our rationale for pooling first-pass tandem data with parallel-group data is stated in Section "Outcomes", but the designs are not interchangeable, and Yamaguchi 2024 was powered for adenoma miss rate rather than for ADR.

Fourth, colonoscopy indications varied substantially, from predominantly fecal immunochemical test-positive and surveillance populations (Yamaguchi 2024) to a mixed diagnostic population in which only 12% of procedures were performed for screening (Chang 2026); baseline control-arm ADR consequently ranged from 16.2% (Yao 2024) to 61.0% (Yamaguchi 2024), an almost fourfold difference that plausibly influences the achievable magnitude of any CADe effect. Fifth, the degree of expert supervision and second-pass verification differed: the tandem trials incorporated an expert second pass, Lau 2024 employed supervisor verification, and Chang 2026 excluded procedures in which the attending physician intervened. Sixth, the anatomic definition of the right colon was not uniform across the contributing trials: Repici 2022 and Lau 2024 both defined the right colon as the cecum through the transverse colon, whereas Chang 2026 restricted right-colon detection to the cecum and ascending colon. Collectively, these differences in operator experience, intervention, design, indication, supervision, baseline ADR, and anatomic definition constitute the principal sources of heterogeneity in this review, and they should be borne in mind when interpreting the pooled estimates and their generalizability [20–24].

Seventh, only five trials met the inclusion criteria, limiting statistical power for the secondary outcomes and rendering any assessment of small-study effects exploratory; the PIs for all four outcomes crossed the null, and only three trials contributed patient-level RADR data. Eighth, the unblindable nature of operator exposure to the CADe display is intrinsic to all CADe randomized trials, and a Hawthorne-effect contribution to observed differences cannot be excluded; this consideration is reflected in the RoB 2 judgments under the "deviations from intended interventions" domain [25]. Ninth, serrated lesion detection which a class of lesion that contributes disproportionately to proximal post-colonoscopy colorectal cancer was reported by only one trial and could not be synthesized. Finally, downstream clinical consequences of CADe-driven detection including interval cancer incidence, surveillance burden, and the proportion of non-neoplastic resections were not consistently reported across trials and could not be pooled. Safety reporting was a particular weakness: only Lau 2024 reported adverse-event counts, so no between-arm safety comparison was possible in four of the five trials and the safety of CADe-assisted colonoscopy in this population remains essentially untested (Section "Safety and adverse events" and Supplementary Material 4, gr.elmerpub.com). As a limitation of the review process itself, the protocol was not prospectively registered in a public systematic review registry, and the analyses listed in Section "Study design and registration" were added after the initial results were seen in response to peer review.

Future research

Future randomized trials should examine the durability of CADe-associated benefits as operators accumulate procedural experience, the effect of CADe on long-term outcomes such as interval colorectal cancer incidence, and whether the right-colon benefit observed here extends to CADe platforms beyond ENDO-AID and GI Genius. Trials should report serrated lesion detection as a distinct outcome, given its relevance to proximal post-colonoscopy colorectal cancer, and should report absolute as well as relative effects. Trials should also investigate whether CADe-augmented training during fellowship influences post-fellowship performance in unaided practice, ideally by testing trainees during a prespecified AI-free period (Section "CADe as a performance aid versus an educational intervention"), given the well-described persistence of trainee-stage ADR into independent practice [11]. Where the number of trials permits, future syntheses should examine CADe effect by operator experience stratum, by platform, and by trial design, analyses that the present evidence base cannot support. Cost-effectiveness analyses tailored to training contexts, alongside pragmatic implementation studies across

geographically and economically diverse training programs, are also warranted.

Conclusions

In five RCTs enrolling 3,158 patients, CADe-assisted colonoscopy increased adenoma detection by trainee and less-experienced endoscopists, with the largest relative effect observed in the right colon. The evidence base is nevertheless small and clinically heterogeneous: five trials, four CADe platforms, two trial designs, and operators ranging from complete novices to independently practicing endoscopists with up to 2,000 procedures. The PIs for every outcome crossed the null, certainty of evidence was moderate for overall ADR and low or very low for the secondary outcomes, no effect on advanced adenoma detection was demonstrated, and no included trial measured post-colonoscopy colorectal cancer or any other patient-important long-term outcome. These findings therefore support further prospective evaluation of CADe as an adjunct to colonoscopy training for improving adenoma detection. They do not yet establish that routine integration of CADe into gastroenterology fellowship curricula produces more competent independent endoscopists, and the effect of CADe exposure during training on unaided post-training performance should be regarded as the priority question for future research.

Supplementary Material

Suppl 1. Completed PRISMA 2020 checklist.

Suppl 2. Full search strategy and number of records retrieved for each database and register.

Suppl 3. Study-level outcome data for all trials and all outcomes.

Suppl 4. Study-level safety and adverse-event data for all included trials.

Suppl 5. Leave-one-out sensitivity analyses for (A) overall adenoma detection rate, (B) right-colon adenoma detection rate, (C) polyp detection rate, and (D) advanced adenoma detection rate.

Suppl 6. Forest plot for overall adenoma detection rate restricted to the four trials of operators in formal endoscopic training.

Suppl 7. Doi plots with the LFK index for (A) overall adenoma detection rate, (B) right-colon adenoma detection rate, (C) polyp detection rate, and (D) advanced adenoma detection rate.

Acknowledgments

None to declare.

Financial Disclosure

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.

Informed Consent

Not applicable. This meta-analysis used data exclusively from previously published randomized controlled trials.

Author Contributions

Hamza Altal: Conceptualization, methodology, screening, data extraction, formal analysis, writing the original draft, and corresponding author. Rami Jadallah: Screening, data extraction, risk of bias assessment, and writing - review and editing. Hasan Daher: Literature search, screening, and writing the review and editing. Connor Pollard: Data extraction, validation, and writing the review and editing. Kinney Xu: Data extraction, validation, and writing the review and editing. Cody Marik: Literature search, screening, and writing the review and editing. Saba Daher: Data extraction and writing the review and editing. Mohammad Darweesh: Methodology, critical review, and writing the review and editing. Mark Young: Supervision, critical review, and writing the review and editing. Andrew Storm: Supervision, critical review, and writing the review and editing. All authors reviewed the final version of the manuscript and approved it for publication.

Data Availability

The authors declare that data supporting the findings of this study are available within the article and its supplementary materials. All data analyzed were extracted from previously published RCTs [20–24]; study-level event counts and denominators for every outcome are provided in Supplementary Material 3 (gr.elmerpub.com).

Abbreviations

AADR: advanced adenoma detection rate; ADR: adenoma detection rate; AI: artificial intelligence; ASGE: American Society for Gastrointestinal Endoscopy; CADx: computer-aided diagnosis; CAQ: computer-aided quality; CADe: computer-aided detection; CENTRAL: Cochrane Central Register of Controlled Trials; CI: confidence interval; CNN: convolutional neural network; DDW: Digestive Disease Week; ESGE: European Society of Gastrointestinal Endoscopy; FIT: fecal immunochemical test; GRADE: Grading of Recommendations, Assessment, Development and Evaluation; HD: high definition; LFK: Luis Furuya-Kanamori; NNT: number needed to treat; NR: not reported; PDR: polyp detection rate; PGY: postgraduate year; PI: prediction interval; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses;

 Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com

RADR: right-colon adenoma detection rate; RCT: randomized controlled trial; RD: risk difference; RoB 2: Cochrane Risk of Bias 2 tool; RR: risk ratio; SSL: sessile serrated lesion

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-263. doi pubmed
2. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med. 2014;370(14):1298-1306. doi pubmed
3. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, Zwierko M, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362(19):1795-1803. doi pubmed
4. Rex DK, Anderson JC, Butterly LF, Day LW, Dominitz JA, Kaltenbach T, Ladabaum U, et al. Quality indicators for colonoscopy. Gastrointest Endosc. 2024;100(3):352-381. doi pubmed
5. Kaminski MF, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, Hoff G, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy. 2017;49(4):378-397. doi pubmed
6. Zhao S, Wang S, Pan P, Xia T, Chang X, Yang X, Guo L, et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. Gastroenterology. 2019;156(6):1661-1674.e11. doi pubmed
7. Anderson R, Burr NE, Valori R. Causes of post-colonoscopy colorectal cancers based on world endoscopy organization system of analysis. Gastroenterology. 2020;158(5):1287-1299.e2. doi pubmed
8. le Clercq CM, Bouwens MW, Rondagh EJ, Bakker CM, Keulen ET, de Ridder RJ, Winkens B, et al. Postcolonoscopy colorectal cancers are preventable: a population-based study. Gut. 2014;63(6):957-963. doi pubmed
9. Stoffel EM, Erichsen R, Froslev T, Pedersen L, Vyberg M, Koeppe E, Crockett SD, et al. Clinical and molecular characteristics of post-colonoscopy colorectal cancer: a population-based study. Gastroenterology. 2016;151(5):870-878.e3. doi pubmed
10. Munroe CA, Lee P, Copland A, Wu KK, Kaltenbach T, Soetikno RM, Friedland S. A tandem colonoscopy study of adenoma miss rates during endoscopic training: a venture into uncharted territory. Gastrointest Endosc. 2012;75(3):561-567. doi pubmed
11. van Doorn SC, Klanderman RB, Hazewinkel Y, Fockens P, Dekker E. Adenoma detection rate varies greatly during colonoscopy training. Gastrointest Endosc. 2015;82(1):122-129. doi pubmed
12. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology. 2020;159(2):512-520.e7. doi pubmed
13. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut. 2019;68(10):1813-1819. doi pubmed
14. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, Antonelli G, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. Gastrointest Endosc. 2021;93(1):77-85.e6. doi pubmed
15. Hassan C, Spadaccini M, Mori Y, Foroutan F, Facciorusso A, Gkolfakis P, Tziatzios G, et al. Real-time computer-aided detection of colorectal neoplasia during colonoscopy: a systematic review and meta-analysis. Ann Intern Med. 2023;176(9):1209-1220. doi pubmed
16. Spadaccini M, Iannone A, Maselli R, Badalamenti M, Desai M, Chandrasekar VT, Patel HK, et al. Computer-aided detection versus advanced imaging for detection of colorectal neoplasia: a systematic review and network meta-analysis. Lancet Gastroenterol Hepatol. 2021;6(10):793-802. doi pubmed
17. Barua I, Vinsard DG, Jodal HC, Loberg M, Kalager M, Holme O, Misawa M, et al. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. Endoscopy. 2021;53(3):277-284. doi pubmed
18. Patel H. Benefits and harms of incorporating AI during colonoscopy for trainees: a systematic review and meta-analysis of published literature. Presented at: Digestive Disease Week (DDW); 2024. Washington, DC. Abstract.
19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi pubmed
20. Repici A, Spadaccini M, Antonelli G, Correale L, Maselli R, Galtieri PA, Pellegatta G, et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. Gut. 2022;71(4):757-765. doi pubmed
21. Lau LHS, Ho JCL, Lai JCT, Ho AHY, Wu CWK, Lo VWH, Lai CMS, et al. Effect of Real-Time Computer-Aided Polyp Detection System (ENDO-AID) on adenoma detection in endoscopists-in-training: a randomized trial. Clin Gastroenterol Hepatol. 2024;22(3):630-641.e634. doi pubmed
22. Chang PW, Nguyen DD, Kong N, Wang D, Wang S, Ong J, Amini MM, et al. Impact of artificial intelligence-assisted colonoscopy on gastroenterology fellow performance: a pragmatic randomized controlled trial. Gastrointest Endosc. 2026;103(5):1043-1051.e1043. doi pubmed
23. Yamaguchi D, Shimoda R, Miyahara K, Yukimoto T, Sakata Y, Takamori A, Mizuta Y, 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. 2024;36(1):40-48. doi pubmed
24. Yao L, Li X, Wu Z, Wang J, Luo C, Chen B, Luo R, et

al. Effect of artificial intelligence on novice-performed colonoscopy: a multicenter randomized controlled tandem study. Gastrointest Endosc. 2024;99(1):91-99.e99. doi pubmed
25. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. doi pubmed
26. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188. doi pubmed
27. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-560. doi pubmed
28. IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247. doi pubmed
29. IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014;14:25. doi pubmed
30. Furuya-Kanamori L, Barendregt JJ, Doi SAR. A new improved graphical and quantitative method for detecting bias in meta-analysis. Int J Evid Based Healthc. 2018;16(4):195-203. doi pubmed
31. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-926. doi pubmed
32. Ladabaum U, Shepard J, Weng Y, Desai M, Singer SJ, Mannalithara A. Computer-aided detection of polyps does not improve colonoscopist performance in a pragmatic implementation trial. Gastroenterology. 2023;164(3):481-483.e486. doi pubmed
33. Levy I, Bruckmayer L, Klang E, Ben-Horin S, Kopylov U. Artificial intelligence-aided colonoscopy does not increase adenoma detection rate in routine clinical practice. Am J Gastroenterol. 2022;117(11):1871-1873. doi pubmed
34. Wei MT, Fay S, Yung D, Ladabaum U, Kopylov U. Artificial intelligence-assisted colonoscopy in real-world clinical practice: a systematic review and meta-analysis. Clin Transl Gastroenterol. 2024;15(3):e00671. doi pubmed
35. Khouri A, Dickson C, Green A, Hanjar A, Sonnier W. Effect of computer aided detection device on the adenoma detection rate and serrated detection rate among trainee fellows. JGH Open. 2024;8(9):e70018. doi pubmed
36. Orzeszko Z, Gach T, Necka S, Ochwat K, Major P, Szura M. The implementation of computer-aided detection in an initial endoscopy training improves the quality measures of trainees' future colonoscopies: a retrospective cohort study. Surg Endosc. 2025;39(8):5276-5286. doi pubmed
37. Bernhofer S, Prosenz J, Venturi D, Maieron A. The impact of artificial intelligence on the adenoma detection rate: comparison between experienced, intermediate and trainee endoscopists' adenoma detection rate. Wien Klin Wochenschr. 2026;138(1-2):1-8. doi pubmed
38. Biscaglia G, Cocomazzi F, Gentile M, Loconte I, Mileti A, Paolillo R, Marra A, et al. Real-time, computer-aided, detection-assisted colonoscopy eliminates differences in adenoma detection rate between trainee and experienced endoscopists. Endosc Int Open. 2022;10(5):E616-E621. doi pubmed
39. Dương QT, Soldera J. Virtual reality tools for training in gastrointestinal endoscopy: a systematic review. Artif Intell Gastrointest Endosc. 2024;5(2):92090. doi
40. Curlej P, Soldera J. Artificial intelligence in predicting colorectal polyp histology: systematic review and meta-analysis of diagnostic accuracy in real-time procedures. World J Gastrointest Endosc. 2026;18(3):116381. doi
41. Woods M, Soldera J. Colon capsule endoscopy polyp detection rate vs colonoscopy polyp detection rate: systematic review and meta-analysis. World J Meta-Anal. 2024;12(4):100726. doi
42. Areia M, Mori Y, Correale L, Repici A, Bretthauer M, Sharma P, Taveira F, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. Lancet Digit Health. 2022;4(6):e436-e444. doi pubmed
43. Mori Y, East JE, Hassan C, Halvorsen N, Berzin TM, Byrne M, von Renteln D, et al. Benefits and challenges in implementation of artificial intelligence in colonoscopy: World Endoscopy Organization position statement. Dig Endosc. 2023;35(4):422-429. doi pubmed

 Articles © The authors | Journal compilation © Gastroenterol Res and Elmer Press Inc™ | https://gr.elmerpub.com