



Published in final edited form as:

Gastroenterology. 2025 April ; 168(4): 691–700. doi:10.1053/j.gastro.2025.01.002.

AGA Living Clinical Practice Guideline on Computer-Aided Detection–Assisted Colonoscopy

Shahnaz Sultan^{1,2,*}, Dennis L. Shung^{3,*}, Jennifer M. Kolb^{4,5}, Farid Foroutan^{6,7}, Cesare Hassan^{8,9}, Charles J. Kahi¹⁰, Peter S. Liang^{11,12}, Theodore R. Levin^{13,14}, Shazia Mehmood Siddique^{15,16,17}, Benjamin Lebwohl^{18,19}

¹ Division of Gastroenterology, Hepatology, and Nutrition, University of Minnesota, Minneapolis, Minnesota

² Minneapolis Veterans Affairs Healthcare System, Minneapolis, Minnesota

³ Department of Medicine, Section of Digestive Diseases, Yale School of Medicine, New Haven, Connecticut

⁴ Vatche and Tamar Manoukian Division of Digestive Diseases, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California

⁵ Division of Gastroenterology, Hepatology and Parenteral Nutrition, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California

⁶ MAGIC Evidence Ecosystem Foundation, Oslo, Norway

⁷ Ted Rogers Centre for Heart Research, University Health Network, Toronto, Ontario, Canada

⁸ IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy

⁹ Department of Biomedical Sciences, Humanitas University, Milan, Italy

¹⁰ Department of Gastroenterology, Indiana University Medical Center, Indianapolis, Indiana

¹¹ Department of Medicine, Division of Gastroenterology and Hepatology, NYU Langone Health, New York, New York

¹² Department of Medicine, Veterans Affairs New York Harbor Health Care System, New York, New York

¹³ Division of Research, Kaiser Permanente Northern California, Pleasanton, California

¹⁴ Department of Gastroenterology, Kaiser Permanente Walnut Creek, Walnut Creek, California

¹⁵ Division of Gastroenterology, University of Pennsylvania, Philadelphia, Pennsylvania

Correspondence: Address correspondence to: Chair, Clinical Guidelines Committee, American Gastroenterological Association, National Office, 4930 Del Ray Avenue, Bethesda, Maryland 20814. clinicalpractice@gastro.org.

* Authors share co-first authorship.

Conflicts of interest

A full list of conflicts active at the time of guideline development can be accessed from AGA's National Office in Bethesda, MD.

These authors disclose the following: Cesare Hassan: consultant for Fujifilm and Olympus. Peter S. Liang: research for Freenome and advisory board for Guardant Health and Natera. The remaining authors disclose no conflicts.

¹⁶ Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania

¹⁷ Center for Healthcare Improvement and Patient Safety, University of Pennsylvania, Philadelphia, Pennsylvania

¹⁸ Department of Medicine, Columbia University Irving Medical Center, New York, New York

¹⁹ Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York

Abstract

BACKGROUND & AIMS: This American Gastroenterological Association (AGA) guideline is intended to provide an overview of the evidence and support endoscopists and patients on the use of computer-aided detection (CAdE) systems for the detection of colorectal polyps during colonoscopy.

METHODS: A multidisciplinary panel of content experts and guideline methodologists used the Grading of Recommendations Assessment, Development and Evaluation framework and relied on the following sources of evidence: (1) a systematic review examining the desirable and undesirable effects (ie, benefits and harms) of CAdE-assisted colonoscopy, (2) a microsimulation study estimating the effects of CAdE on longer-term patient-important outcomes, (3) a systematic search of evidence evaluating the values and preferences of patients undergoing colonoscopy, and (4) a systematic review of studies evaluating health care providers' trust in artificial intelligence technology in gastroenterology.

RESULTS: The panel reached the conclusion that no recommendation could be made for or against the use of CAdE-assisted colonoscopy in light of very low certainty of evidence for the critical outcomes, desirable and undesirable (11 fewer colorectal cancers per 10,000 individuals and 2 fewer colorectal cancer deaths per 10,000 individuals), increased burden of more intensive surveillance colonoscopies (635 more per 10,000 individuals), and cost and resource implications. The panel acknowledged the 8% (95% CI, 6%–10%) increase in adenoma detection rate and 2% (95% CI, 0%–4%) increase in advanced adenoma and/or sessile serrated lesion detection rate.

CONCLUSIONS: This guideline highlights the close tradeoff between desirable and undesirable effects and the limitations in the current evidence to support a recommendation. The panel acknowledged the potential for CAdE to continually improve as an iterative artificial intelligence application. Ongoing publications providing evidence for critical outcomes will help inform a future recommendation.

Keywords

Artificial Intelligence; Computer-Aided Detection; Colonoscopy; Colorectal Cancer

Colorectal cancer (CRC) is the third most common cancer worldwide and ranks as the third most common cause of cancer-related mortality in the United States.^{1,2} CRC arises from the slow progression of neoplasia from precursor polyps (eg, adenomas and serrated lesions).^{3,4} The ability to detect and resect polyps during colonoscopy, as part of screening

or surveillance programs, reduces the incidence and mortality of CRC; however, this protection is highly dependent on colonoscopy quality.^{5–9}

More than 15 million colonoscopies are performed annually in the United States, however, studies have demonstrated variable quality of colonoscopies across key quality indicators.¹⁰ Adenoma detection rate (ADR), defined as the percentage of patients undergoing colonoscopy with at least 1 adenoma resected, is currently considered one of the most relevant quality indicators and is inversely correlated with risk of post-colonoscopy CRC (PCCRC, defined as a CRC diagnosed after a colonoscopy in which no cancer was found).^{5,7,11} Studies have found that approximately 25% of adenomas are missed at colonoscopy.¹² Missed polyps may be attributed to a failure of detection (especially of subtle or flat lesions) and/or to inadequate or incomplete mucosal inspection during withdrawal.

In recent years, there has been an exponential rise in studies assessing the impact of artificial intelligence (AI)-based technology with real-time computer-aided detection (CAdE) of polyps.^{13,14} CAdE systems for endoscopy were first described in the early 2000s, but it was only after advancements in machine learning, such as use of deep learning–based algorithms, that the practical applications of CAdE were realized.^{15,16} The most common underlying deep learning architecture used for CAdE is a convolutional neural network, which automatically learns features through filter optimization from colonoscopic imaging data.¹⁷ Each of the CAdE systems learns how to detect polypoid lesions after training on a large data set of expert-labeled endoscopic images and videos and demonstrates remarkable generalizability with robust performance when deployed on previously unseen colonoscopic images and videos.¹⁸ The high performance of CAdE systems and capability to deploy them in real time has led to the development of several CAdE platforms that have sought approval from the US Food and Drug Administration as medical devices for marketing purposes.

Although CAdE-assisted colonoscopy may increase ADR, there is a risk of overdiagnosis, as most of the polyps detected during colonoscopy are diminutive in size (<5 mm) and of low malignant potential. Furthermore, the association between ADR and cancer incidence is nonlinear, with a likely plateau effect over a 45% cutoff.⁹ Overdiagnosis refers to additional ineffective or unnecessary medical consequences of a diagnosis, such as intensive surveillance and follow-up that may not be clinically beneficial.¹⁹ As detection rates increase, the likelihood of both screening effectiveness and overdiagnosis (ie, detection of polyps that would not have progressed to clinical CRC during a lifetime) increases. Thus, improving colonoscopy quality by means of increasing ADR will also increase the proportion of colonoscopy recipients who are put into a more intensive surveillance algorithm, leading to increased costs, risks, resource use, and burden associated with surveillance colonoscopy. There is also an anticipated increase in the resection of hyperplastic polyps. Finally, the impact of CAdE-assisted colonoscopy on long-term patient outcomes, such as risk of CRC or CRC-related mortality, is unclear.

Guideline Objective

The objective of this guideline was to make a recommendation on the role of CAdE-assisted colonoscopy.

Target Audience

The target audience of this guideline includes gastroenterologists, endoscopists, primary care providers, people undergoing colonoscopy, administrators, and policy makers. These guidelines are not intended to impose a standard of care for individual institutions, health care systems, or countries, but to provide guidance as the basis for rational informed decision making.

Methods

Overview

This document represents the official recommendations of the American Gastroenterological Association (AGA) and was commissioned and approved by the AGA Institute Clinical Guidelines Committee and the AGA Governing Board to provide timely, methodologically rigorous guidance on a topic of high clinical importance to the AGA members and the public.

The guideline was developed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework to prioritize clinical questions, identify patient-centered outcomes, and conduct an evidence synthesis. The guideline panel further used the Evidence to Decision framework as a transparent approach to developing recommendations and to provide implementation considerations for clinical practice.²⁰ Therefore, the panel considered the tradeoff between desirable and undesirable effects, patient values and preferences, feasibility, acceptability, resource utilization, health equity, and environmental impact.

Collaboration

This guideline was co-developed with the MAGIC (Making GRADE the Irresistible Choice) Evidence Ecosystem Foundation and the BMJ Rapid Recommendation series.²¹ The collaboration between the BMJ recommendation panel and the AGA guideline panel allowed for collective efforts in conducting²² and updating²³ systematic reviews (SRs), as well as shared evidence profiles.

Guideline Panel Composition and Conflicts of Interest

Members of the guideline panel were selected on the basis of clinical and methodological expertise after review of all conflicts of interest (COIs). The guideline panel included content experts, members with experience in evidence synthesis, and guideline methodologists. In addition, as part of a collaborative effort, the co-chair of the BMJ guideline (F.F.) served as a panel member. The panel contributed to the formulation of clinical questions and outcomes of interest, reviewed the evidence, contributed to the judgments using the Evidence to Decision framework, and participated in the development of guideline recommendations through a series of web-based meetings.

AGA adheres to National Academy of Medicine recommendations for managing COI disclosures in the development of clinical practice guidelines. All members of the guideline development group, including guideline panel chair, guideline panel members,

methodologists, and content experts, completed a disclosure statement before commencing work. Members were expected to update their disclosures in writing as changes occurred throughout the development process. All members of the team were advised not to accept new speaking engagements or consulting arrangements with an honorarium during the guideline development process and until 12 months after publication date. The AGA COI policy is available on request. All COI disclosure forms are maintained at the AGA National Office in Bethesda, Maryland.

Formulation of Clinical Questions and Determining Outcomes of Interest

The guideline panel identified 2 clinically relevant questions on the use of AI technologies. However, to facilitate the collaboration, it was decided that this guideline would focus on CAdE only and a future guideline would address the role of computer-aided diagnosis, which allows for real-time diagnosis of polyp histology.

The following clinical question was formulated using the PICO format, which frames a clinical question by defining a specific patient population (P), intervention (I), comparator (C), and outcomes (O): In patients undergoing colonoscopy (screening, surveillance, and follow-up for positive fecal immunochemical test), should CAdE-assisted colonoscopy vs standard colonoscopy be used? The panel rated the following outcomes as being of critical importance: CRC-related mortality, CRC incidence, and PCCRC incidence. The panel deemed the following outcomes as important (but not critical): ADR, advanced ADR, sessile serrated lesions detection rate, adenomas per colonoscopy, adenoma miss rate, polypectomies of nonadenomatous polyps per colonoscopy, withdrawal time (inspection time), number of colonoscopies per lifetime, perforations, and bleeding events. The panel used predefined thresholds for the minimum clinically important difference to aid the certainty of evidence assessment determined by clinical experts.

Evidence Review and Synthesis

The panel conducted multiple SRs to summarize and synthesize the evidence regarding the use of CAdE, including a commissioned modeling study to examine the long-term impact of CAdE. This guideline is based on the following bodies of evidence: (1) an SR examining the desirable and undesirable effects (ie, benefits and harms) of CAdE-assisted colonoscopy, (2) a microsimulation study estimating the effects of CAdE on longer-term patient-important outcomes (not evaluated by existing randomized controlled trials [RCTs]), (3) a systematic search of evidence evaluating the values and preferences of patients undergoing colonoscopy, and (4) an SR of studies evaluating health care providers' trust in AI technology for polyp detection during gastroenterology.

Certainty of Evidence

The certainty of evidence was judged using the GRADE approach. Briefly, using this approach, evidence from a body of RCTs starts as high-quality evidence and evidence from observational studies starts at low-quality evidence.²⁴ This evidence can be further rated down for risk of bias in the evidence, indirectness, inconsistency, imprecision, and publication bias. In selected cases, evidence from observational studies may be rated up if a large treatment effect is observed, if there is a dose–response relationship, or if all plausible

confounding and bias would reduce a demonstrated effect or suggest a spurious effect if no effect was observed. For modeling studies, we relied on the following GRADE guidance to determine the certainty of evidence model (gradeworkinggroup.org). Evidence profiles and summary of finding tables were developed for each intervention using the MAGIC platform, MAGICapp (<https://magicvidence.org/magicapp>).

Development of Recommendations

The panel used the GRADE approach and the Evidence to Decision framework. The Evidence to Decision framework considers the balance of desirable and undesirable effects of the intervention, certainty of evidence, resource use, cost, equity and health disparities, acceptability, and feasibility. The recommendations are labeled as “strong” or “conditional” according to the GRADE approach. The phrase “the guideline panel recommends” is used for strong recommendations and “the guideline panel suggests” is used for conditional recommendations. Rarely, the panel may make no recommendation when the tradeoff is closely balanced and the certainty of evidence around the critical outcomes is uncertain. The interpretations of the strength of recommendations are summarized in Table 1.

Panel members met online and developed the recommendations on the basis of the evidence summarized in the evidence to decision table (see Supplementary Material). For this recommendation, the panel took a population perspective and reached a consensus on the certainty of evidence; balance of benefits and harms; and assumptions about the values and preferences associated with the decision, health equity, acceptability, and feasibility. The panel did not explicitly incorporate resource considerations (such as direct or indirect costs or cost-effectiveness). The panel agreed on the final recommendation and remarks by means of consensus, with the majority of panel members agreeing with the final recommendation.

Review Process

This guideline was submitted for public comment and external peer review and was approved by the AGA Governing Board.

Results

Systematic Review and Meta-Analysis of Randomized Controlled Trials of Computer-Aided Detection–Assisted Colonoscopy

A recently published SR and meta-analysis of RCTs evaluating the efficacy and safety of CADe was used to inform this guideline.²³ This SR and meta-analysis included 44 RCTs with more than 30,000 participants who underwent CADe-assisted colonoscopy. Based on data from 32,108 participants in 41 RCTs, CADe-assisted colonoscopy was associated with a higher polyp detection rate (56.1% vs 47.9%; relative risk [RR], 1.22; 95% CI, 1.15–1.28) and a higher ADR (44.8% vs 37.4%; RR, 1.22; 95% CI, 1.16–1.29) compared with standard colonoscopy. Across 6 studies with tandem colonoscopy design (total patients, n = 2769), the authors found that adenoma miss rate was lower with CADe (missed adenomas, n = 154 of 957 [16.1%] in CADe vs 374 of 1061 [35.3%] in standard colonoscopy; RR, 0.47; 95% CI, 0.36–0.60). Advanced colorectal neoplasia detection rate, defined as at least 1 histologically proven advanced adenoma, advanced serrated lesion, or both, detected on

colonoscopy, was also reported. Combining data from 22 studies (N = 19,645), CAdE outperformed the standard group in advanced colorectal neoplasia detection rate (12.7% vs 11.5%; RR, 1.16; 95% CI, 1.02–1.32).

With respect to non-neoplastic polyps and withdrawal time, there was an 11% increase in the likelihood of resected polyps being non-neoplastic (34.0% in CAdE group vs 28.8% in the standard group; RR, 1.11; 95% CI, 1.04–1.19). Pooling data from 30,276 procedures in 2 groups showed using CAdE resulted in trivial prolongation of both total withdrawal time (9.17 vs 8.60 minutes; mean difference, 0.57 minutes; 95% CI, 0.31–0.83 minutes) and inspection time (8.34 vs 7.95 minutes; mean difference, 0.31 minutes; 95% CI, 0.14–0.48 minutes).

Microsimulation Modeling Study of Computer-Aided Detection and Colorectal Cancer–Related Outcomes

Data from a modeling microsimulation study were used to inform the effect on critical and important outcomes of CRC incidence, CRC-related mortality, post-colonoscopy cancer, as well as perforation, bleeding, and burden of increased surveillance colonoscopies.²⁵ This modeling study was conducted as part of the development of BMJ Rapid Recommendations and was used to inform both guideline panels. The modeling team developed a microsimulation Markov model to simulate 10-year follow-up for 100,000 individuals aged 60–69 years participating in CRC screening programs (ie, primary screening colonoscopy and colonoscopy after a positive fecal immunochemical test). The model incorporated data from large RCTs and registry studies, providing information on CRC risk, survival rates by cancer stage, and adenoma prevalence. For colonoscopy screening, data from a multinational RCT provided evidence of the effects of screening on cancer incidence and survival.²⁶ Data from an Italian registry study, validated against outcomes from the Dutch national screening program, provided evidence for patients undergoing colonoscopy after a positive fecal immunochemical test.^{27,28}

The model assumed 100% participation in screening and surveillance and incorporated different ADR ranges to simulate the performance of endoscopists with varying skill levels.²⁹ Sensitivity analyses were conducted to test the robustness of these assumptions by evaluating different ADR categories, nonlinear relationships between ADR and CRC risk, and various impacts of CAdE on ADR. The model also incorporated data on surveillance guidelines, polypectomy practices, and adverse event rates from screening trials and pooled data from various programs; however, cost was not explicitly incorporated into the model. RRs, absolute effect estimates, and certainty of the evidence are outlined in the summary of findings table (Table 2).

Systematic Review of Patient Values and Preferences

An SR (based on a protocol) was conducted to assess patients' values and preferences regarding the use of CAdE during colonoscopy. A comprehensive search, performed on August 7, 2023 across 2 major databases (ie, Embase and Medline) identified 7101 records, of which only 2 relevant articles were identified.^{30,31} The authors then broadened the search to include articles related to patient values and preferences of screening tests and identified

1 SR (including 23 studies) on stated preference elicitation for different CRC screening tests.³²

In 1 of these individual studies, the authors evaluated the perspectives of Dutch patients undergoing endoscopic procedures on the potential use of AI in health care, willingness to implement AI, and perceived advantages and disadvantages of its use.³⁰ The authors found that patients expressed concerns about a potential reduction in personal contact with clinicians, with >60% of participants listing this as a key issue. However, a similar proportion of patients believed that AI could enhance the quality of care and reduce diagnostic waiting times. An important limitation is that this study was not specifically focused on AI-assisted colonoscopy. A second study evaluated patients' perception regarding the importance of AI use during colonoscopy and found that >65% would choose an endoscopist who uses AI with CAdE during colonoscopy over one who does not, with 58% reporting the use of AI to be very or somewhat important in colonoscopy.³¹

The SR of 23 studies revealed that patients placed a considerable value on screening test characteristics that could reduce CRC mortality.³² These studies highlighted the importance of mortality reduction and improvements in care quality as critical factors influencing patient preferences for the use of AI tools, such as CAdE. Nonetheless, patients were also concerned about potential burdens, including increased health-related anxiety, overdiagnosis, and the possibility of requiring more frequent surveillance colonoscopies.

Systematic Review of Studies Evaluating Provider Trust in Artificial Intelligence Technology

An SR was conducted of providers' perspectives and trust towards AI for colonoscopy (S. Soleymanjahi, MD, unpublished data, January 2025). A comprehensive search, performed through July 2024 across multiple databases, identified 2044 unique records with 48 more references from manual search and citation chasing. A total of 45 full-text studies were assessed and 13 were included. These studies addressed interest, willingness, and perceived concerns about and barriers to integrating CAdE into practice.

This review identified 12 studies evaluating attitudes of health care providers. The review found that a high portion (85%; 95% CI, 80%–90%) of health care providers and stakeholders find the use of AI technology in colonoscopy interesting, and a large portion (74%; 95% CI, 68%–80%) believed that its use would result in improvements in adenoma detection. A considerable portion (43%; 95% CI, 16%–71%), however, were concerned about increased costs; 40% (95% CI, 16%–67%) were concerned about accountability for incorrect diagnosis; and 23% (95% CI, 12%–37%) were concerned about medicolegal issues. Some health care providers found the use of CAdE to be too distracting in their daily practice.

Rationale

Although the panel initially made a draft recommendation suggesting the use of CAdE (conditional recommendation), after significant feedback during the public comment period, the panel reconvened and ultimately made a NO RECOMMENDATION statement. This was based on a vote, as there was a lack of consensus among the panel members.

The panel agreed that the increase in ADR is an important benefit because it is an established quality indicator. The ADR increase, however, was primarily attributable to detection of diminutive to small polyps with low potential for progression to adenocarcinoma. This results in surveillance colonoscopies at shorter intervals and a greater lifetime number of colonoscopies, most of which are not individually likely to prevent CRC. The panel placed a higher value on long-term outcomes of CRC incidence and mortality and deliberated the close tradeoff between the uncertain and trivial to small benefits on reducing CRC incidence and CRC-related mortality, and the potential increased burden of more frequent, potentially unnecessary, surveillance colonoscopies, increased health-related anxiety, overdiagnosis, and increased resources.

The panel acknowledged the limited data on patient values and preferences, but also recognized that, as an intervention used primarily by endoscopists, the decision to use CAdE may not be conducive to a shared decision-making approach at the level of an individual person.

Summary of the Evidence

Desirable and Undesirable Effects and Certainty of the Evidence

The panel acknowledged the close tradeoff between the desirable and undesirable effects. Across all of the patient-important outcomes, CAdE is possibly associated with an absolute reduction in CRC incidence of 11 fewer cancers per 10,000 individuals (over 10 years), with a number needed to treat of 909 (but we are very unsure). These cancers are presumably post-colonoscopy cancers that are prevented due to use of CAdE. However, the true benefit on these critical outcomes remains very uncertain (because of limitations in the conclusions and inputs from the modeling study and limited events). For CRC-related mortality, the evidence suggests there may be a trivial or no benefit with 2 fewer CRC-related deaths per 10,000 individuals over 10 years, with a number needed to treat of 5000 (low certainty of evidence). For colonoscopy outcomes, CAdE may lead to an absolute increase in ADR of 8%, an absolute increase in the advanced colorectal neoplasia detection rate of 2%, and an absolute increase in sessile serrated lesion detection rate of 1%. The adenoma miss rate may be reduced by 19% (Table 2).

With respect to the undesirable effects, CAdE may have little to no impact on adverse events from colonoscopy (eg, perforation and bleeding) and no clinically meaningful increase in withdrawal time; however, CAdE likely leads to an increase in non-neoplastic polyp detection (and resection) associated with unnecessary costs. Furthermore, CAdE may increase the number of colonoscopies performed per patient over a 10-year period. With CAdE, 635 more individuals per 10,000 will need a follow-up (surveillance) colonoscopy (low certainty of evidence). This is due to the increased detection of adenomas, especially diminutive or small adenomas, which are of low malignant potential, but contribute to a higher number of polyps, resulting in more frequent colonoscopy.

Finally, the panel also reviewed a systematic review of 8 nonrandomized studies including 9782 participants, which demonstrated that CAdE did not lead to an increase in ADR (44% vs 38%; risk ratio, 1.11; 95% CI, 0.97–1.28).³³

Other Considerations in the Evidence to Decision Framework

In addition to considering the desirable and undesirable effects, the panel also reviewed the limited evidence on patient values and preferences. There was no study that specifically evaluated patient values and preferences for AI-assisted colonoscopy. However, based on the SR of patient values for CRC screening tests, the panel concluded that patients may value the potential reduction in CRC incidence with CAdE-assisted colonoscopy and its ability to enhance colonoscopy quality. The panel also acknowledged that some patients voiced concern about AI misuse, decrease in quality care, and privacy issues. The panel concluded that there may be important uncertainty or variability in patient values and preferences and further research is needed to directly assess patient attitudes toward CAdE-assisted colonoscopy, especially in terms of balancing its potential advantages with the perceived drawbacks and potential burdens (increased health-related anxiety and more frequent surveillance colonoscopies).

Although CAdE was judged to be feasible to implement, the panel acknowledged the limited data on acceptability, especially by gastroenterologists. The panel noted that effectiveness of CAdE may be lower in clinical practice compared with its performance in RCTs.³³ In addition, endoscopy unit leadership may need to engage with endoscopists to promote uptake and monitor use.³⁴ The studies evaluating provider attitudes and trust demonstrated that although the use of AI would improve adenoma and polyp detection, concerns about cost of the technology; accountability for incorrect diagnosis; and medicolegal issues were important.

The panel discussed the potential impact of CAdE on resource utilization. The modeling study suggested that the implementation of CAdE may result in a substantial increase in follow-up procedures. These, in turn, may result in substantial increase in costs and a reduction of available resources for the coverage of the population. The panel noted that the impact of CAdE on resource utilization might change, depending on future surveillance interval recommendations that are issued, considering the increased polyp detection (especially of diminutive lesions) resulting from CAdE. Furthermore, the cost of purchasing CAdE (through a flat fee or subscription-based model) for an individual health system or for a national CRC screening program must also be factored in. Cost-effectiveness studies are needed to help understand the long-term impact of CAdE on cost-effectiveness.

The impact of CAdE on health equity was also discussed. A recommendation for the use of CAdE may negatively impact available resources and reduce equitable access to clinicians who perform AI-assisted colonoscopy. Furthermore, CAdE systems are costly and only available to those who have the resources and infrastructure for implementing their use, and the costs of implementing a CAdE system are not reimbursable through insurance. From a global perspective, a recommendation for CAdE may result in inequity. Furthermore, there may be a negative impact on environmental health ([gradeworkinggroup.org](https://www.gradeworkinggroup.org)). The panel also felt that there may be perceived concerns around the development of these AI models, particularly with regard to the quality of the data sets used for their developments. Concerns about model derivation in relation to racial and ethnic bias may introduce mistrust among patients of different backgrounds (see Supplementary Material).

Knowledge Gaps

The guideline panel identified the following knowledge gaps that need to be addressed in future studies:

- The role of AI in health care is rapidly evolving and the impact of widespread adoption of AI tools, such as CAdE, on population health are unknown. Studies focused on long-term patient-important outcomes are needed.
- The role of a combined strategy of CAdE plus computer-aided diagnosis may offset the increase in non-neoplastic polyps that are detected and resected with CAdE alone. Trials assessing these combination strategies must be performed.
- Future guidelines that address recommended post-polypectomy surveillance intervals should consider the changing landscape of polyp detection in the context of this technology.
- There are ongoing ethical and legal implications with concerns about breaches in patient privacy and data security, as well as ownership of health care information used by AI platforms.³⁵
- In addition, there is currently no legal framework for health care errors. For example, if a PCCRC is detected with CAdE-assisted colonoscopy, some authors have suggested that the responsibility should be shared between the device constructors and the physicians who use it.³³
- More transparency is needed around the development and evaluation of commercially available CAdE systems, with more detailed information about the algorithms and image data sets used for training and validation, including their limitations, ideally before commercial use.³⁶
- There are several CAdE devices in the market, and periodic updates are to be expected for most of them. The individual differences among them or among upgrades are unclear and potentially impactful on clinical outcomes.
- The role of CAdE in educational settings is still unclear, with concerns around the potential negative impact of AI in skill development, especially on recognition and detection of polyps.

Plans for Updating This Guideline

This guideline will be updated in approximately 1–2 years when new data become available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This guideline was developed in close collaboration with the MAGIC Evidence Ecosystem and BMJ Rapid Recommendation guideline group with shared methodologic support and contributions to guideline development. The authors gratefully acknowledge Siddharth Singh for providing oversight of the collaboration and Ms Krisie

Gardner for her administrative assistance. The authors acknowledge Prateek Sharma for his participating in the guideline panel and contributions to earlier drafts of this manuscript.

Funding

These guidelines were fully funded by the AGA Institute with no funding from any outside agency or industry. Shahnaz Sultan is supported by US Food and Drug Administration grant U01 FD008147-01. Shazia Mehmood Siddique is supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grant K08-DK120902. Dennis L. Shung is supported by grant DK125718. Benjamin Lebwohl is supported by NIDDK grant U01 DK136523. Theodore R. Levin is supported by the TPMG Delivery Science and Applied Research Program.

Abbreviations used in this paper:

ADR	adenoma detection rate
AGA	American Gastroenterological Association
AI	artificial intelligence
CADe	computer-aided detection
COI	conflict of interest
CRC	colorectal cancer
GRADE	Grading of Recommendations Assessment, Development and Evaluation
PCCRC	post-colonoscopy colorectal cancer
RCT	randomized controlled trial
RR	relative risk
SR	systematic review

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2024;74:229–263. [PubMed: 38572751]
2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin* 2024;74:12–49. [PubMed: 38230766]
3. Nguyen LH, Goel A, Chung DC. Pathways of colorectal carcinogenesis. *Gastroenterology* 2020;158:291–302. [PubMed: 31622622]
4. Sullivan BA, Noujaim M, Roper J. Cause, epidemiology, and histology of polyps and pathways to colorectal cancer. *Gastrointest Endosc Clin N Am* 2022;32:177–194. [PubMed: 35361330]
5. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014;370:1298–1306. [PubMed: 24693890]
6. Kahi CJ, Vemulapalli KC, Johnson CS, et al. Improving measurement of the adenoma detection rate and adenoma per colonoscopy quality metric: the Indiana University experience. *Gastrointest Endosc* 2014;79:448–454. [PubMed: 24246797]
7. Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. *N Engl J Med* 2010;362:1795–1803. [PubMed: 20463339]

8. Rex DK, Anderson JC, Butterly LF, et al. Quality indicators for colonoscopy. *Gastrointest Endosc* 2024;100:352–381. [PubMed: 39177519]
9. Schottinger JE, Jensen CD, Ghai NR, et al. Association of physician adenoma detection rates with postcolonoscopy colorectal cancer. *JAMA* 2022;327:2114–2122. [PubMed: 35670788]
10. Gangwani MK, Aziz A, Dahiya DS, et al. History of colonoscopy and technological advances: a narrative review. *Transl Gastroenterol Hepatol* 2023;8:18. [PubMed: 37197258]
11. Shaukat A, Holub J, Greenwald D, et al. Variation over time and factors associated with detection rates of sessile serrated lesion across the united states: results form a national sample using the GIQuIC registry. *Am J Gastroenterol* 2021;116:95–99. [PubMed: 32833735]
12. Zhao S, Wang S, Pan P, et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. *Gastroenterology* 2019;156:1661–1674.e11. [PubMed: 30738046]
13. Jerebko A, Lakare S, Cathier P, et al. Symmetric curvature patterns for colonic polyp detection. *Med Image Comput Comput Assist Interv* 2006;9:169–176. [PubMed: 17354769]
14. Karkanis SA, Iakovidis DK, Maroulis DE, et al. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Trans Inf Technol Biomed* 2003;7:141–152. [PubMed: 14518727]
15. Hassan C, Wallace MB, Sharma P, et al. New artificial intelligence system: first validation study versus experienced endoscopists for colorectal polyp detection. *Gut* 2020;69:799–800. [PubMed: 31615835]
16. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 2018;155:1069–1078.e8. [PubMed: 29928897]
17. Arbib MA, ed. *Handbook of Brain Theory and Neural Networks*. MIT Press, 1995.
18. Spadaccini M, Massimi D, Mori Y, et al. Artificial intelligence-aided endoscopy and colorectal cancer screening. *Diagnostics (Basel)* 2023;13:1102. [PubMed: 36980409]
19. Kalager M, Wieszczyn P, Lansdorp-Vogelaar I, et al. Overdiagnosis in colorectal cancer screening: time to acknowledge a blind spot. *Gastroenterology* 2018;155:592–595. [PubMed: 30076834]
20. Alonso-Coello P, Schunemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016;353:i2016. [PubMed: 27353417]
21. Foroutan F, Olav Vandvik P, Helsing LM, et al. Computer aided detection and diagnosis of polyps in adult patients undergoing colonoscopy: a living clinical practice guideline. *BMJ*. Forthcoming 2025.
22. Hassan C, Spadaccini M, Mori Y, et al. Real-time computer-aided detection of colorectal neoplasia during colonoscopy: a systematic review and meta-analysis. *Ann Intern Med* 2023;176:1209–1220. [PubMed: 37639719]
23. Soleymanjahi S, Huebner J, Elmansy L, et al. Artificial intelligence-assisted colonoscopy for polyp detection: a systematic review and meta-analysis. *Ann Intern Med* 2024;177:1652–1663. [PubMed: 39531400]
24. Sultan S, Falck-Ytter Y, Inadomi JM. The AGA Institute process for developing clinical practice guidelines part one: grading the evidence. *Clin Gastroenterol Hepatol* 2013;11:329–332. [PubMed: 23517554]
25. Halvorsen N, Hassan C, Correale L, et al. Benefits, burden, and harms of computer-aided polyp detection with artificial intelligence in colorectal cancer screening: a microsimulation modelling study. *BMJ Medicine*. Forthcoming 2025.
26. Bretthauer M, Loberg M, Wieszczyn P, et al. Effect of colonoscopy screening on risks of colorectal cancer and related death. *N Engl J Med* 2022;387:1547–1556. [PubMed: 36214590]
27. Wisse PHA, Erler NS, de Boer SY, et al. Adenoma detection rate and risk for interval postcolonoscopy colorectal cancer in fecal immunochemical test-based screening: a population-based cohort study. *Ann Intern Med* 2022;175:1366–1373. [PubMed: 36162114]
28. Zorzi M, Antonelli G, Barbiellini Amidei C, et al. Adenoma detection rate and colorectal cancer risk in fecal immunochemical test screening programs: an observational cohort study. *Ann Intern Med* 2023;176:303–310. [PubMed: 36802754]

29. Pilonis N, Wieszczy P, Bugajski M, et al. 602: Thresholds of adenoma detection rate and post-colonoscopy colorectal cancer risk. *Gastroenterology* 2022;162:S–S150.
30. van der Zander QEW, van der Ende-van Loon MCM, Janssen JMM, et al. Artificial intelligence in (gastrointestinal) healthcare: patients' and physicians' perspectives. *Sci Rep* 2022;12:16779. [PubMed: 36202957]
31. Burton SJ, Shung D, Chung S, et al. Patient perspective of use of artificial intelligence during colonoscopy. *Gastro Hep Adv* 2025;4:100543. [PubMed: 39790241]
32. Brinkmann M, Fricke LM, Diedrich L, et al. Attributes in stated preference elicitation studies on colorectal cancer screening and their relative importance for decision-making among screenees: a systematic review. *Health Econ Rev* 2022;12:49. [PubMed: 36136248]
33. Patel HK, Mori Y, Hassan C, et al. Lack of effectiveness of computer aided detection for colorectal neoplasia: a systematic review and meta-analysis of nonrandomized studies. *Clin Gastroenterol Hepatol* 2024;22:971–980. e15. [PubMed: 38056803]
34. Shaikat A, Tuskey A, Rao VL, et al. Interventions to improve adenoma detection rates for colonoscopy. *Gastrointest Endosc* 2022;96:171–183. [PubMed: 35680469]
35. Naik N, Hameed BMZ, Shetty DK, et al. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? *Front Surg* 2022;9:862322. [PubMed: 35360424]
36. Gimeno-Garcia AZ, Hernandez-Perez A, Nicolas-Perez D, et al. Artificial intelligence applied to colonoscopy: is it time to take a step forward? *Cancers (Basel)* 2023;15:2193. [PubMed: 37190122]

Recommendation: In adults undergoing colonoscopy, the AGA makes no recommendation on the use of CAdE-assisted colonoscopy. (No recommendation, very low certainty of evidence)

Remarks:

- This recommendation underscores the overall very low certainty of evidence on long-term outcomes deemed critical for decision making (ie, CRC incidence, CRC mortality, and PCCRC). The panel acknowledged the evidence for CAdE in improving ADR as an important surrogate outcome.
- As an iterative tool, the panel highlights the likely potential for CAdE as an AI application to continually improve with further training and validation over time. As this technology improves, and with additional studies evaluating long-term critical outcomes (ie, CRC incidence, CRC mortality, and PCCRC), an update of this recommendation is planned.
- Evidence gaps to be addressed include data from diverse patient populations and settings (eg, community and academic endoscopists), studies evaluating impact on colon polyps of greater clinical significance (eg, advanced adenomas), or more difficult to detect polyps (eg, sessile serrated lesions), information on resource implications, access to colonoscopy, patient values and preferences, cost-effectiveness, and impact on long-term patient-important outcomes.

Table 1.
Interpretation of Strong and Conditional Recommendations Using the Grading of Recommendations Assessment, Development and Evaluation Framework

Implications	Strong recommendation	Conditional recommendation
For patients	Most individuals in this situation would want the recommended course of action and only a small proportion would not.	The majority of individuals in this situation would want the suggested course of action, but many would not.
For clinicians	Most individuals should receive the intervention. Formal decision aids are not likely to be needed to help individuals make decisions consistent with their values and preferences.	Different choices will be appropriate for individual patients consistent with their values and preferences. Use shared decision making. Decision aids may be useful in helping patients make decisions consistent with their individual risks, values, and preferences.
For policy makers	The recommendation can be adapted as policy or performance measure in most situations.	Policy making will involve various stakeholders. Performance measures should assess whether decision making is appropriate.

No Recommendation: A recommendation cannot be made based on limitations in the evidence or the evidence is not adequate to make a recommendation. Further research is needed.

Table 2.

Summary of Findings Used to Inform the Recommendation

Outcome/ timeframe	Study results and measurements	Absolute effect estimates		Certainty of the evidence (quality of evidence)	Narrative summary
		Routine colonoscopy	CADe colonoscopy		
CRC incidence ^{a/10} y	RR, 0.87 (95% CI, 0.57–1.21) Based on data from 10,000 participants in 1 study Follow-up: 10 y	82 per 10,000 Difference: 11 fewer per 10,000 (NNT = 909) (95% CI, 35 fewer to 17 more)	71 per 10,000 (95% CI, 35 fewer to 17 more)	Very low due to serious risk of bias, due to serious imprecision ^b	We are uncertain whether use of CADe-assisted colonoscopy decreases CRC incidence
CRC-related deaths/10 y	RR, 0.84 (95% CI, 0.32–2.19) Based on data from 10,000 participants (modeling study) Follow-up: 10 y	15 per 10,000 Difference: 2 fewer per 10,000 (NNT = 5000) (95% CI, 10 fewer to 18 more)	13 per 10,000 (95% CI, 10 fewer to 18 more)	Low due to serious imprecision ^c	CADe-assisted colonoscopy may have little or no difference on CRC- related deaths
PCCRC incidence ^{d/10} y	RR, 0.67 (95% CI, 0.34–1.34) Based on data from 10,000 participants (modeling study) Follow-up: 10 y	34 per 10,000 Difference: 11 fewer per 10,000 (NNT = 909) (95% CI, 22 fewer to 12 more)	23 per 10,000 (95% CI, 22 fewer to 12 more)	Very low due to serious risk of bias, due to serious imprecision ^e	We are uncertain whether use of CADe-assisted colonoscopy decreases PCCRC incidence
ADR	RR, 1.22 (95% CI, 1.16–1.29) Based on data from 30,674 participants in 40 studies	37.4 per 100 Difference: 8 more per 100 (95% CI, 6 more to 11 more)	44.8 per 100 (95% CI, 6 more to 11 more)	Low due to serious risk of bias, due to serious publication bias ^f	CADe-assisted colonoscopy may increase ADR
AMR	RR, 0.47 (95% CI, 0.36–0.6) Based on data from 2018 participants in 6 studies	35.3 per 100 Difference: 19 fewer per 100 (95% CI, 22 fewer to 14 fewer)	16.1 per 100 (95% CI, 22 fewer to 14 fewer)	Moderate due to serious risk of bias ^g	CADe-assisted colonoscopy probably decreases AMR
Surveillance colonoscopy/10 y	RR, 1.24 (95% CI, 1.22–1.26) Based on data from 10,000 participants (modeling study) Follow-up: 10 y	2645 per 10,000 Difference: 635 more per 10,000 (95% CI, 582 more to 688 more)	3280 per 10,000 (95% CI, 582 more to 688 more)	Low due to serious risk of bias ^h	CADe-assisted colonoscopy may increase number of colonoscopies per lifetime
Perforation ^{i/10} y	RR, 1.0 (95% CI, 1.0–1.0) Based on data from 10,000 participants (modeling study) Follow-up: 10 y	10 per 10,000 Difference: 0 fewer per 10,000 (95% CI, 0 fewer to 0 fewer)	10 per 10,000 (95% CI, 0 fewer to 0 fewer)	Low due to serious risk of bias ⁱ	CADe-assisted colonoscopy may have little or no difference on perforation
Bleeding ^{k/10} y	RR, 1.06 (95% CI, 0.53–2.12) Based on data from 10,000 participants (modeling study) Follow-up: 10 y	19 per 10,000 Difference: 1 more per 10,000 (95% CI, 9 fewer to 11 more)	20 per 10,000 (95% CI, 9 fewer to 11 more)	Low due to serious risk of bias ^j	CADe-assisted colonoscopy may have little or no difference on bleeding
Advanced colorectal neoplasia detection rate ^m	RR, 1.16 (95% CI, 1.02–1.32) Based on data from 19,645 participants in 22 studies	12 per 100 Difference: 2 more per 100 (95% CI, 0 more to 4 more)	14 per 100 (95% CI, 0 more to 4 more)	Low due to serious risk of bias, due to serious imprecision ⁿ	CADe-assisted colonoscopy may increase advanced colorectal neoplasia detection rate

Outcome/ timeframe	Study results and measurements	Absolute effect estimates		Certainty of the evidence (quality of evidence)	Narrative summary
		Routine colonoscopy	CADe colonoscopy		
Serrated sessile polyp detection rate	RR, 1.21 (95% CI, 1.04–1.42) Based on data from 16,511 participants in 22 studies	6.0 per 100 Difference: 1 more per 100 (95% CI, 0 more to 3 more)	7.0 per 100 95% CI, 0 more to 3 more)	Low due to serious risk of bias, due to serious imprecision ^d	CADe-assisted colonoscopy may increase serrated sessile polyp detection rate
Non-neoplastic polyps	RR, 1.11 (95% CI, 1.04–1.19) Based on data from participants in 19 studies	29 per 100 Difference: 3 more per 100 (95% CI, 1 more to 6 more)	32 per 100 95% CI, 1 more to 6 more)	Moderate due to serious risk of bias ^d	CADe-assisted colonoscopy probably increases detection of non- neoplastic polyps
Withdrawal time	Measured in minutes Based on data from 19,235 participants in 25 studies	Mean, 8.60 MD, 0.57 more (95% CI, 0.31 more to 0.83 more)	Mean, 9.17 95% CI, 0.31 more to 0.83 more)	Moderate due to serious risk of bias ^d	CADe-assisted colonoscopy probably has little or no difference on withdrawal time

NOTE. The population is adults undergoing colonoscopy for screening, surveillance, or follow-up of positive fecal immunochemical test. The intervention is CADe colonoscopy. The comparator is routine or standard colonoscopy.

AMR, adenoma miss rate; MD, mean difference; MID, minimum important difference; NNT, number needed to treat.

^a All CRCs detected during and after the first (index) colonoscopy.

^b Risk of bias: serious. A major source of input into the modeling work, the SR on CADe, had low certainty of evidence on the impact of CADe on ADR; imprecision: serious. The 95% CI is wide and includes increased harms.

^c Risk of bias: not serious. Input to modeling has low certainty; imprecision: serious. The 95% CI contains increased harm.

^d This definition of PCCRC includes any CRC detected after the index colonoscopy performed during the screening.

^e Risk of bias: serious. Input to modeling has low certainty; imprecision: serious. The 95% CI includes harms.

^f Risk of bias: serious. Inadequate/lack of blinding of outcome assessors, resulting in potential for detection bias; inconsistency: no serious. Although there is important inconsistency observed, this is not clinically important. Any 1% increase in ADR is potentially important (although extremely dependent on baseline risk). When we use the 1% absolute risk difference threshold, only 1 study is inconsistent (suggesting protective effect). This study, however, only contributes 3% to the overall pooled estimate. Likely no impact on pooled estimate; publication bias: serious. Asymmetrical funnel plot with a significant Egger's test.

^g Risk of bias: serious. Inadequate/lack of blinding of outcome assessors, resulting in potential for detection bias; inconsistency: no serious. MID of 1%. All studies showed >1% reduction in AMR. Therefore, no inconsistency; imprecision: no difference in conclusion based on lower or upper bound of 95% CI.

^h Risk of bias: serious. Input to modeling work has low certainty.

ⁱ Involve transfusion or hospitalization within 30 days.

^j Risk of bias: serious. Input to modeling work has low certainty.

^k Involve transfusion or hospitalization within 30 days.

^l Risk of bias: serious. Input to modeling work had low certainty.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

u Advanced adenoma or advanced serrated lesion.

u Risk of bias: serious. Inadequate/lack of blinding of outcome assessors, resulting in potential for detection bias; inconsistency: no serious. No important inconsistency using MID of 1%; imprecision: serious. The 95% CIs cross MID of 1%. Lower bound suggests no important increase in advanced/serrated polyps.

o Risk of bias: serious. Inadequate/lack of blinding of outcome assessors, resulting in potential for detection bias; inconsistency: no serious. All except 1 study showed overlapping CIs; imprecision: serious. Lower bound of 95% CI suggests that there may be no important increase in detection of serrated sessile polyps. Our inference is that there may be an important increase.

p Risk of bias: serious. Inadequate/lack of blinding of outcome assessors, resulting in potential for detection bias.

q Risk of bias: serious. Inadequate/lack of blinding of outcome assessors, resulting in potential for detection bias, inadequate/lack of blinding of participants and personnel, resulting in potential for performance bias.