Accepted Manuscript

Submission Date: 2024-02-10
Accepted Date: 2024-05-15
Accepted Manuscript online: 2024-05-15

Endoscopy

A computer-aided detection system in the everyday setting of diagnostic, screening and surveillance colonoscopy: an international, randomized trial

Michiel H Maas, Timo Rath, Cristiano Spada, Elsa Soons, Nauzer Forbes, Sergey Kashin, paola cesaro, Axel Eickhoff, Geoffroy Vanbiervliet, Daniele Salvi, Paul J Belletrutti, Peter D Siersema.

Affiliations below.

DOI: 10.1055/a-2328-2844

Please cite this article as: Maas M H, Rath T, Spada C et al. A computer-aided detection system in the everyday setting of diagnostic, screening and surveillance colonoscopy: an international, randomized trial. Endoscopy 2024. doi: 10.1055/a-2328-2844

Conflict of Interest: Maas, M: (none); Rath, T: (received speaker fees from: Olympus Medical, Pentax Medical, Mauna Kea Technologies, Medtronic, Takeda, Galapagos, Falk, Janssen, Abbvie, Repha, Medical Tribune, Lilly); Spada, C: (consultancy fees from Medtronic and AnX Robotics and received speaker fees from Olympus, PENTAX Medical and Norgine); Soons, E: (none); Forbes, N: (Consultant for Boston Scientific corp., PENTAX Medical, received speaker fees from PENTAX Medical, received research funding from PENTAX Medical); Kashin, S: (none); Cesaro, P: (none); Eickhoff, A: (none); Vanbiervliet, G: (consultancy fees from Boston Scientific corp., Ambu A/S, Lecture and conference fees from Fujifilm Medical, Norgine, Tillots, PENTAX Medical); Salvi, D: (none); Belletrutti, P.J.: (received speaker fees from Pendopharm, consultancy fees from Vantage Endoscopy); Siersema, P: (research support from The E-Nose company, Lucid Diagnostics, Micro Tech, Magentiq Eye LTD., Norgine, and Endo Tools Therapeutics, consultancy fees from Magentiq Eye LTD.).

This study was supported by PENTAX Medical

Trial registration: NL9135, Uniform Trial Number (http://www.who.int/ictrp/unambiguous_identification/en/) , Prospective, Randomized, Multicenter study

Abstract:

Background and study aim: Computer-aided detection (CADe) has been developed to improve detection during colonoscopy. After initial reports of high efficacy, there has been an increasing recognition of variability in the effectiveness of CADe systems. The aim of this study was to evaluate a CADe system (PENTAX Medical, Tokyo, Japan) in a varied colonoscopy population.

Patients and methods: A multicenter, randomized trial was conducted at 7 hospitals (both university and non-university) in Europe and Canada. Participants referred for diagnostic, non-iFOBT screening, or surveillance colonoscopy were randomized (1:1) to undergo CADe-assisted or conventional colonoscopy (CC) by experienced endoscopists. Participants with insufficient bowel preparation were excluded from the analysis. Primary outcome was adenoma detection rate (ADR). Secondary outcomes included adenomas per colonoscopy (APC) and sessile serrated lesions per colonoscopy (SSLPC).

Results: In total, 581 participants were enrolled, of which 497 were included in the final analysis: 250 in the CADe-arm and 247 in the CC-arm. Surveillance was the indication in 202/497 (40.6%) colonoscopies, diagnostic in 199/497 (40.0%), and non-iFOBT screening in 96/497 (19.3%). Overall, ADR (38.4% vs. 37.7%; p=0.43) and APC (0.66 vs. 0.66; p=0.97) were similar between CADe and CC. SSLPC was increased (0.30 vs. 0.19; p=0.049) in the CADe-arm vs. CC.

Conclusions: In this study conducted by experienced endoscopists, CADe did not result in a statistically significant increase in ADR. However, the ADR of our control group substantially surpassed our sample size assumptions, increasing the risk of an underpowered trial. (Trialsearch.who.int:NL9135).

Corresponding Author:

Drs. Michiel H Maas, Radboudumc, Gastroenterology and Hepatology, P.O. Box 9101, 6500 HB Nijmegen, Netherlands, Michiel_maas@outlook.com, Michiel.maas@radboudumc.nl

Affiliations:
Michiel H Maas, Radboudumc, Gastroenterology and Hepatology, Nijmegen, Netherlands
Timo Rath, Universitatsklinikum Erlangen, Department of Medicine I, Division of Gastroenterology, Erlangen, Germany
Cristiano Spada, Università Cattolica del Sacro Cuore, Fondazione Policlinico Universitario A. Gemelli IRCCS, Milano, Italy
[...]
Peter D Siersema, Radboudumc, Gastroenterology and Hepatology, Nijmegen, Netherlands

Title: A computer-aided detection system in the everyday setting of diagnostic, screening and surveillance colonoscopy: an international, randomized trial

Authors and affiliations: M.H.J. Maas[1], T. Rath[2], C. Spada[3,9], E. Soons[1], N. Forbes[4], S. Kashin[5], P. Cesaro[3], A. Eickhoff[6], G. Vanbiervliet[7], D. Salvi[3,9], P.J. Belletrutti[4], P.D. Siersema[1,8]; for the Discovery study team.

[1]Radboud University Medical Center, Gastroenterology & Hepatology, Nijmegen, the Netherlands, [2]Universitätsklinikum Erlangen, Erlangen, Germany, [3]Fondazione Poliambulanza Istituto Ospedaliero, Department of Gastroenterology and Endoscopy, Brescia, Italy [4]University of Calgary, Calgary, Canada, [5]Yaroslavl Regional Cancer Hospital, Yaroslavl, Russia, [6]Klinikum Hanau, Hanau, Germany, [7]Hôpital L'Archet, Nice, France, [8]ErasmusMC - University Medical Center, Rotterdam, the Netherlands, [9]Università Cattolica del Sacro Cuore, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy

Grant support: This study was supported by a research grant provided by PENTAX Medical, Tokyo, Japan.

Abbreviations and relevant definitions: ADR, adenoma detection rate; AI, artificial intelligence; APC, adenoma per colonoscopy; ASA, American society of anesthesiologists; BBPS, Boston bowel preparation scale; BMI, body mass index; CADe, computer-aided detection; CI, confidence interval; CC, conventional colonoscopy; CRC, colorectal cancer; iFOBT, immunochemical fecal occult blood test; ITT, intention-to-treat; IQR, interquartile range; mITT, modified intention-to-treat; PCCRC, post-colonoscopy colorectal cancer; PDR, polyp detection rate; PPC, polyp per colonoscopy; RCT, randomized controlled trial; SD, standard deviation; SDR, sessile serrated lesion detection rate; SSL, sessile serrated lesion; SSLPC, sessile serrated lesions per colonoscopy; SPSS, statistical package for the social sciences

Correspondence: Michiel H.J. Maas, MD, Department of Gastroenterology & Hepatology. Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, the Netherlands. Geert Grooteplein-Zuid 10, 6525 GA Nijmegen (route 455). Phone: +31 24 361 3999. Email: Michiel.Maas@Radboudumc.nl

Disclosures: Maas, M: (none); Rath, T: (received speaker fees from: Olympus Medical, Pentax Medical, Mauna Kea Technologies, Medtronic, Takeda, Galapagos, Falk, Janssen, Abbvie, Repha, Medical Tribune, Lilly); Spada, C: (consultancy fees from Medtronic and AnX Robotics and received speaker fees from Olympus, PENTAX Medical and Norgine); Soons, E: (none); Forbes, N: (Consultant for Boston Scientific corp., PENTAX Medical, received speaker fees from PENTAX Medical, received research funding from PENTAX Medical); Kashin, S: (none); Cesaro, P: (none); Eickhoff, A: (none); Vanbiervliet, G: (consultancy fees from Boston Scientific corp., Ambu A/S, Lecture and conference fees from Fujifilm Medical, Norgine, Tillots, PENTAX Medical); Salvi, D: (none); Belletrutti, P.J.: (received speaker fees from Pendopharm, consultancy fees from Vantage Endoscopy); Siersema, P: (research support from The E-Nose company, Lucid Diagnostics, Micro Tech, Magentiq Eye LTD., Norgine, and Endo Tools Therapeutics, consultancy fees from Magentiq Eye LTD.).

Author contributions: Maas, M: (Formal analysis; Project administration; Validation; Visualization; Writing – original draft; Writing – review & editing); Rath, T: (Investigation; Methodology; Resources; Writing – review & editing); Spada, C: (Investigation; Methodology: Resources); Soons, E: (Conceptualization; Methodology; Project administration; Writing – review & editing); Forbes, N: (Investigation; Resources; Writing – review & editing); Kashin, S: (Investigation; Resources); Cesaro, P: (Investigation; Resources); Eickhoff, A: (Investigation; Resources); Vanbiervliet, G: (Investigation; Resources); Salvi, D: (Investigation; Project administration; Writing – review & editing); Belletrutti, P.J.: (Investigation; Resources; Writing – review &editing); Siersema, P.D.: (Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing – review & editing).

Writing assistance: none.

Data transparency statement: Individual de-identified patient data that underlie the results reported in this article will be made available with investigators whose proposed use of the data has been approved by the corresponding author and with a signed data access agreement. Access to the data is granted solely for the purposes outlined in the approved proposal. For inquiries or proposals, please contact the corresponding author. Interested parties seeking access or inquiries regarding to the CADe system and/or its code can contact the manufacturer PENTAX Medical for further information.

Word count:

Abstract: 241 words

Main Text: 3487 words

Abstract

Background and study aim: Computer-aided detection (CADe) has been developed to improve detection during colonoscopy. After initial reports of high efficacy, there has been an increasing recognition of variability in the effectiveness of CADe systems. The aim of this study was to evaluate a CADe system (PENTAX Medical, Tokyo, Japan) in a varied colonoscopy population.

Patients and methods: A multicenter, randomized trial was conducted at 7 hospitals (both university and non-university) in Europe and Canada. Participants referred for diagnostic, non-iFOBT screening, or surveillance colonoscopy were randomized (1:1) to undergo CADe-assisted or conventional colonoscopy (CC) by experienced endoscopists. Participants with insufficient bowel preparation were excluded from the analysis. Primary outcome was adenoma detection rate (ADR). Secondary outcomes included adenomas per colonoscopy (APC) and sessile serrated lesions per colonoscopy (SSLPC).

Results: In total, 581 participants were enrolled, of which 497 were included in the final analysis: 250 in the CADe-arm and 247 in the CC-arm. Surveillance was the indication in 202/497 (40.6%) colonoscopies, diagnostic in 199/497 (40.0%), and non-iFOBT screening in 96/497 (19.3%). Overall, ADR (38.4% vs. 37.7%; p=0.43) and APC (0.66 vs. 0.66; p=0.97) were similar between CADe and CC. SSLPC was increased (0.30 vs. 0.19; p=0.049) in the CADe-arm vs. CC.

Conclusions: In this study conducted by experienced endoscopists, CADe did not result in a statistically significant increase in ADR. However, the ADR of our control group substantially surpassed our sample size assumptions, increasing the risk of an underpowered trial. (Trialsearch.who.int:NL9135).

Introduction

Colonoscopy is considered the gold standard for the detection and removal of premalignant colorectal lesions. Despite its effectiveness, a notable number of lesions are still missed during colonoscopy,[1] increasing the risk of post-colonoscopy colorectal cancer (PCCRC).[2] This risk is inversely correlated with adenoma detection rate (ADR), which is widely considered the main quality parameter in colonoscopy.[3, 4]

Recently, artificial intelligence (AI) systems have emerged to assist in the detection of colorectal polyps during colonoscopy, also known as computer-aided detection (CADe). A recent systematic review of 21 randomized controlled trials (RCTs) assessing CADe versus conventional colonoscopy (CC) in over 18,000 patients demonstrated an approximate 24% relative increase in ADR due to CADe.[5] Despite this significant benefit in overall ADR, there were no significant differences between CADe and CC in the detection of advanced adenomas or sessile serrated lesions (SSLs),[5] raising concerns regarding the efficacy of CADe in these lesions with a higher risk of CRC. Furthermore, despite reports of high efficacy from RCTs, there has been an increasing recognition of the variability in the performance of CADe systems across different colonoscopy indications and pragmatic trials.[6, 7]

The aim of this study was to compare ADR and other quality indicators in CADe-assisted colonoscopy versus CC in patients with diagnostic, non-immunochemical fecal occult blood test (iFOBT) screening, or surveillance indications for colonoscopy.

Patients and Methods

Study design and participants

This multicenter RCT, involving seven hospitals in Canada (n=1), France (n=1), Germany (n=2), Italy (n=1), The Netherlands (n=1), and Russia (n=1), was conducted by 14 endoscopists. Eligible participants, aged 18 years or older, were scheduled for non-iFOBT screening, surveillance, or diagnostic (excluding iFOBT+ referrals) colonoscopy. Exclusion criteria included known colorectal tumors or polyps upon referral, referral for therapeutic procedures, inadequately corrected coagulation disorder, inadequately continued use of anticoagulation medication, ASA score of ≥3, or known or suspected inflammatory bowel disease. Participants with insufficient bowel preparation (BBPS <6), active colitis, polyposis syndrome, colonic stricture, or obstructing CRC impeding complete colonoscopy were excluded from the final analysis. Participants in the Yaroslavl, Russia study site enrolled after February 24, 2022, were excluded from the final analysis following a directive from the Dutch Federation of University Hospitals (NFU), mandating the temporary suspension of all collaborations with Russian study sites. The study was registered at the Netherlands national trial register (https://trialsearch.who.int/) under NL9135, received approval from independent institutional review boards at each site, adhered to the Declaration of Helsinki, and followed applicable Good Clinical Practice guidelines. Data verification and monitoring complied with national and local guidelines where appropriate. The study was reported per CONSORT-AI guidelines for RCTs, and all participants provided written informed consent. All authors had access to the study data and reviewed and approved the final manuscript.

Randomization

Participants were randomized after eligibility was assessed and informed consent was obtained. Participants were randomized in a 1:1 ratio to either CADe-assisted colonoscopy or CC. Randomization employed varying block sizes of 4, 6, and 8. Stratification for randomization was based on whether the subject was undergoing an index colonoscopy, defined as the first lifetime

colonoscopy of a participant. Randomization was performed on-site, within 24 hours before the scheduled colonoscopy, by a central, cloud-based randomization service (CastorEDC, Ciwit B.V., Amsterdam, The Netherlands). Endoscopists, participants and the data analyst were not blinded to the study allocation.

Artificial Intelligence system

The CADe device, DISCOVERY system (PENTAX Medical, Tokyo, Japan), used for the CADe-assisted colonoscopies is a real-time computing device that acquires the video output from the processor during colonoscopy. The CADe device uses a deep neural network to generate a bounding box around a suspected polyp as an output on the monitor screen in real-time (Figure 1). The device is used as an auxiliary device and aims to improve the detection rate by highlighting potential lesions. Final assessment of the highlighted region was the responsibility of the endoscopist. The endoscopist could choose to be acoustically notified of detections. During the study, CADe software versions 1.0.3.1 and 1.0.4 were used.

Study investigators

All endoscopists underwent training in CADe-assisted colonoscopy, completing a minimum of five CADe procedures to confirm their familiarity with the device. The study was conducted at sites with an annual performance exceeding 5000 colonoscopies. Participation was limited to experienced endoscopists to mitigate potential improvements in ADR due to training throughout the study. Endoscopists were eligible if they had independently performed over 500 colonoscopies, reflecting procedural experience rather than a specific minimum ADR. This approach aimed to approach the real-world variability in ADR among endoscopists. Notably, all endoscopists had completed over 2000 independent colonoscopies prior to the start of the study.

Study procedures

In the CADe-arm, the device was switched on at the beginning of the CADe colonoscopy, and the use of CADe was mandatory during the withdrawal phase. Endoscopists were advised to primarily use the

CADe monitor; however, the use of a dual monitor setup, with separate displays for the conventional image and the CADe overlay, was not explicitly prohibited. Each study site used local bowel preparation protocols and sedation administration. All participants randomized to CC underwent colonoscopy as per standard of care. During the study, conventional PENTAX high-definition colonoscopes were used for both arms. To ensure adequate bowel inspection, all participating endoscopists were instructed to aim for a minimum withdrawal time of 6 minutes (excluding time spent on polypectomies or other interventions) following societal guidelines.[3] In addition, an upper withdrawal time of 10 minutes was recommended to reflect everyday procedural scheduling and reduce observation related bias. All lesions were to be collected for histopathological examination in separate containers for each polyp. Diminutive (1-5mm) polyps located in the rectum and considered to be hyperplastic by the performing endoscopist could be left in place according to endoscopists' individual judgement and standard of care. Experienced pathologists blinded to the endoscopic diagnosis determined the histopathological diagnosis according to the Vienna classification.[8]

Study outcomes

The primary outcome was ADR, calculated as the proportion of colonoscopies with at least one histologically confirmed detected adenoma. Additionally, ADR was evaluated across various variables, including colonoscopy indication and a per-endoscopist analysis. Secondary outcomes were mean number of adenomas per colonoscopy (APC; total number of histologically confirmed adenomas divided by the total number of colonoscopies), polyp detection rate (PDR; proportion of colonoscopies with at least one histologically confirmed detected polyp), sessile serrated lesions per colonoscopy (SSLPC; total number of histologically confirmed SSLs divided by the total number of colonoscopies), and SSL detection rate (SDR; proportion of colonoscopies with at least one histologically confirmed detected SSL). Other secondary outcomes included withdrawal time without interventions and the number of false positives during CADe-assisted colonoscopy (defined as a non-neoplastic, non-hyperplastic area highlighted by CADe for >3 consecutive seconds). The reasons for

false positives were reported and were calculated as the number of CADe colonoscopies with at least one subcategory of the reason for false positives.

Sample size calculation

This study was powered to detect a significant difference in ADR. The sample size calculation was performed using G*Power, version 3.1.9.7, Heinrich-Heine-Universität, Düsseldorf, Germany. During the design of the study protocol, only one RCT had been published regarding the effect of CADe on ADR, reporting an increase in ADR from 20.3% to 29.1%.[9] With this limited prior data, we set baseline ADR at 18% for CC, expecting a 50% relative increase with CADe to 27%. Furthermore, we expected no decrease in detection given the working mechanism of CADe, as it is used as an auxiliary device to conventional colonoscopy, enhancing the displayed image output without directly interfering with colonoscope handling. Consequently, we assumed a one-directional effect and used a one-sided test for sample size calculations. Using a one-sided Z-test for independent proportions (5% alpha, 80% power), the sample size was 532 participants. To account for a dropout rate of 5%, the final sample size was set at 560, evenly distributed between the two study arms (280 each).

Statistical analysis

Analyses of the primary and secondary outcomes followed a modified intention-to-treat (mITT) approach, excluding participants with an inadequate BBPS score or inability to perform a quality colonoscopy. Analysis of the primary outcome was performed using the Chi-square test, dividing the two-sided p-value by two to calculate the one-sided p-value. Statistical analyses were performed using IBM SPSS 27 or R Studio 4.1.3

Continuous variables were presented as means (standard deviation, SD) or medians (interquartile range, IQR), and categorical data as numbers/percentages. Differences between study arms for secondary outcomes were assessed using t-tests, Mann-Whitney U tests, or Chi-square tests, as appropriate. Two-sided p-values were reported for the secondary outcomes. Wilson Score Method was used to calculate 95% CIs where applicable. A logistic regression model evaluated ADR.

Predetermined potential confounding factors, including gender, age, BMI, smoking status, reason for colonoscopy, and study site, were excluded from the model following study protocol, as these variables appeared evenly distributed across study arms. Sensitivity analysis compared APC and SSLPC using Poisson regression. Post-hoc analysis explored the effect of CADe among low-, medium-, and high-detectors, categorizing endoscopists based on ADR tertiles. Additionally, to address our relatively high dropout rate, a post-hoc analysis of the primary outcome was conducted on an intention-to-treat (ITT) cohort. Statistical significance was set at $p<0.05$, unless otherwise specified.

Results

The study was performed from March 9, 2021, to February 6, 2023. The relatively long inclusion period was partly related to the COVID-19 pandemic in the initial period of this study. A total of 581 participants were enrolled and randomized (1:1) to either CADe (n=293) or CC (n=288). A total of 84 participants were excluded, leaving 497 participants in the final analysis (mITT); 250 participants in the CADe-arm and 247 in the CC-arm (Figure 2). While the dropout rate of our mITT analysis (14.5%; 497 out of 581 included) exceeded our expected dropout rate, we were unable to replace these participants due to IRB guidelines. Baseline characteristics were similar in both study arms (table 1). No missing values were observed for the calculation of the primary and secondary outcomes. All procedures in the CADe-arm were performed with the CADe modality activated.

Overall findings

ADR was similar in the CADe-arm compared to the CC-arm (38.4% vs. 37.7%, p=0.432; total colonoscopies with at least 1 adenoma, 96 vs. 93). Logistic regression analysis calculated an odds ratio (OR) of 1.032 [95% CI: 0.719 – 1.483] for CADe relative to CC. Similarly, APC was comparable in the CADe-arm compared to the CC-arm (0.66 vs. 0.66, p=0.971; total detected adenomas 165 vs. 163). While PDR was numerically increased in the CADe-arm compared to the CC-arm, the difference was not significant (55.2% vs. 51.4%, p=0.398; total colonoscopies with at least 1 polyp, 138 vs. 127). Furthermore, SSLPC was significantly higher in the CADe-arm compared to the CC-arm (0.30 vs. 0.19, p=0.049; total detected SSLs, 76 vs. 46) and SDR was increased in the CADe-arm compared to the CC-arm (18.4% vs. 12.1%; p=0.053, total colonoscopies with at least 1 SSL, 46 vs. 30, respectively). Median withdrawal time was similar between study arms (withdrawal time without interventions [IQR] of 9.2 [8.0 – 11.0] vs. 9.0 [8.0 – 11.0] minutes, p=0.052; for CADe and CC, respectively).

When stratified by colonoscopy indication the results were similar. For diagnostic colonoscopies (n=199), ADR was increased by 5.5% in the CADe-arm compared to the CC-arm (33.3% vs. 27.8%,

p=0.400; total colonoscopies with at least 1 adenoma, 34 vs. 27) and for SSLPC the increase was 0.09 in the CADe-arm compared to the CC-arm (0.25 vs. 0.16, p=0.150; total detected SSLs, 26 vs. 16). For surveillance colonoscopies (n=202), ADR was equal in the CADe-arm compared to the CC-arm (43.9% vs 43.9%, p=0.931; total colonoscopies with at least 1 adenoma, 43 vs. 45) and for SSLPC the increase was 0.15 in the CADe-arm compared to the CC-arm (0.36 vs. 0.21, p=0.405; total detected SSLs, 15 vs. 8). For non-iFOBT screening colonoscopies (n=96), ADR was decreased by 7.7% in the CADe-arm compared to the CC-arm (38.0 vs. 45.7%, p=0.447; total colonoscopies with at least 1 adenoma, 19 vs. 21) and for SSLPC the increase was 0.13 in the CADe-arm compared to the CC-arm (0.30 vs. 0.17, p=0.227; total detected SSLs, 15 vs. 8). Additional outcomes are reported in table 2, table 3 and supplementary table 1.

During the withdrawal phase of the CADe-assisted colonoscopy, the median number of false positives was 2.0 (IQR: 0.0 – 5.0; mean 4.1 (SD): 6.1). Colonic haustral folds were reported as the most frequent reason for false positives during CADe-colonoscopy (40.8%) (supplementary table 2).

Discussion

In this multicenter RCT involving experienced endoscopists from both university and non-university hospitals, the use of CADe did not significantly increase ADR or APC in diagnostic, non-iFOBT screening or surveillance colonoscopies. However, despite the non-significant increase in adenoma detection, use of the CADe system resulted in an absolute increase of 0.11 (relative increase 58%) of SSLPC compared to CC.

Our study did not find a significant increase in ADR or APC with CADe, which contrasts with previously published western RCTs, as well as a recent meta-analysis of 21 RCTs including over 18,000 patients that reported an absolute ADR difference of 8.1% (44.0% vs. 35.9%) with CADe compared to CC.[5, 10-15] Moreover, one of the earliest RCTs on CADe by Repici et al. reported an absolute increase of 14.4% in ADR among expert endoscopists.[16] However, a recent non-university, single-center study by Karsenti et al., with over 2000 participants, reported an ADR of 37.5% with CADe, which was similar to our study. In their CC-arm, the baseline ADR was 33.7%, resulting in an absolute difference of only 3.8% with the use of CADe. Their study reported a similar proportion of diagnostic and screening colonoscopies compared to our study.[15]

Our non-significant increase in adenoma detection is consistent with recently published controlled and real-world studies. An RCT conducted in a screening and surveillance population at four US community hospitals reported a non-significant increase in APC from 0.67 to 0.73, comparable to our study.[17] A pragmatic implementation trial performed by Ladabaum et al., which employed CADe during all colonoscopies without specific instructions, reported no significant differences in detection rates.[6] Notably, their baseline ADR was comparable to our study. Moreover, their study did not identify differences in the efficacy of CADe between low-detectors and high-detectors. Similarly, Levy et al. integrated CADe into all colonoscopies at their high-volume tertiary referral center. Their study also did not find a significant increase in adenoma detection compared to a retrospective cohort.[7] While our study employed an RCT design, our study period extended over nearly two years, which

may have resulted in a reduced level of constant scrutiny and observation. This intermittent exposure to the CADe device could have reduced the Hawthorne effect while using CADe, which is a potential source of bias in controlled studies.[18] The extended duration of our study, combined with the variable colonoscopy indications, may reflect a more real-world clinical setting compared to previous positive RCTs on CADe that had substantially shorter study durations.[9, 12, 16]

On the other hand, our non-significant results may be attributed to the relatively high baseline ADR in the CC-arm, approaching 38%. As a result, this may have limited the potential beneficial effect of CADe, as endoscopists with a higher ADR might derive less benefit from CADe compared to their peers with a lower ADR.[5, 15, 19] We also observed this trend in our post-hoc analysis comparing low-, medium-, and high-detectors based on their baseline ADR, although no statistically significant difference was observed (supplementary figure 1 and supplementary table 3). Due to the post-hoc nature, this analysis should be interpreted with caution. Moreover, the median withdrawal time in our study significantly exceeded the recommended six minutes and approached the upper-end target of ten minutes.[3, 20] Our relatively long withdrawal times may have contributed to the high baseline ADR of our CC-arm, as each additional minute is shown to be associated with an increase in ADR; however, this effect seems to diminish after 10 minutes.[21, 22] Lastly, while our study is the first RCT evaluating this CADe system, we are cautious to attribute our non-significant results to the potential lack of standalone efficacy of the system. While most CADe RCTs have reported an increase in ADR with CADe use, some RCTs did not find an increase,[23, 24] despite previous positive results using the same CADe system.[10] This suggests that factors beyond the CADe system are important when interpreting these results. Nonetheless, our findings suggest a potential increase in detection, as indicated by the significant increase in SSLPC and a trend towards increased PDR. However, additional studies are required to provide a more comprehensive understanding of the performance of this CADe system.

The clinical relevance of our study is supported by the increased detection of SSLs, as reflected by the significant relative 58% increase of SSLPC and a borderline significant but absolute increase of 7.3% of SDR in the CADe-arm. Although our overall SDR of 12.1% in the CC-arm might appear relatively high, it is comparable to previously published CADe studies with variable indications.[6, 12, 23] This relatively high SDR is not unexpected, given the increasing awareness and recognition of SSLs, as demonstrated by the steady increase of SDR since 2008.[25] Furthermore, in a retrospective analysis of the training and evaluation sets of this CADe system (unpublished results), we found that 11% of the used lesions were diagnosed as SSLs. This relatively large proportion of SSLs in the training set may have contributed to our significant increase in the detection of these notably hard-to-detect lesions. Our study is, to the best of our knowledge, the first to demonstrate a significant increase in the detection of sessile serrated lesions with the use of CADe. These SSLs are nowadays recognized for their clinical importance in the CRC pathway, and SDR is increasingly recognized as a potential quality parameter in colonoscopy.[26] Furthermore, a recent study has shown that endoscopists with an increased SDR have a lower risk of PCCRC, even when corrected for ADR.[27] However, we acknowledge that the detection of sessile serrated lesions was a secondary outcome in our study.

The strengths of this study include the balanced distribution of participants across six countries, in both Europe and Canada, among both university and non-university hospitals. This approach reduced the potential risk of bias associated with endoscopists who conduct a substantial number of procedures and show significantly improved detection with CADe. In addition, the inclusion criteria reflect everyday colonoscopy populations by incorporating varied colonoscopy indications. Furthermore, distal attachments were not used. Finally, the validation of the CADe system was not performed on the included study populations of participating study sites, reducing the risk of potential overfitting, which is a well-known risk of AI systems.

However, our study had some limitations. First, in hindsight, our assumptions for the sample size calculation were rather conservative. Initially, we assumed a baseline ADR of only 18% in the CC-arm,

influenced by limited data, notably a single Chinese CADe RCT.[9] Despite our initial assumptions proving to be underestimated, particularly with the inclusion of experienced endoscopists, our total calculated sample size did not significantly differ from early Western CADe studies.[10, 16, 28] Additionally, detecting a significant result with our modest difference in ADR would require a substantially larger sample size. As exemplified by a recent RCT with over 2000 participants, where use of CADe resulted in a borderline significant absolute increase in ADR of only 3.8% (p=0.051).[15] Nevertheless, we acknowledge the potential increased risk of a type-II statistical error resulting from our sample size calculation. Second, pathology slides were not evaluated by a second, independent, expert pathologist. This could have introduced some bias in the diagnosis of SSLs in our study, considering that even expert pathologists only have a moderate interobserver agreement when diagnosing SSLs.[29] Nonetheless, this risk is likely limited due to the fact that the pathologists in our study demonstrated proficiency in recognizing SSLs, as indicated by our relatively high detection rates of SSLs compared to previous CADe RCTs.[9-11, 16] Third, the mITT analysis included 497 (85.5%) of the 581 colonoscopies, reflecting a higher-than-anticipated exclusion rate due to insufficient or missing BBPS scores. This could be attributed to the variable colonoscopy indications and non-standardized bowel preparation. While our dropout rate was higher than expected, the subsequent potential risk of further underpowering the study appears to be limited, as supported by the similar results of the intention-to-treat analysis (supplementary table 4). To mitigate the risk of exclusion due to insufficient bowel preparation in future studies, randomizing after reaching the caecum could be considered. Fourth, although training recommended primarily using the CADe monitor in a single-monitor setup, a dual-monitor setup displaying both the conventional image and CADe output side-by-side was not prohibited, potentially influencing gaze patterns in the select cases such a setup was used.[30]

In conclusion, use of CADe by experienced endoscopists did not result in an increased ADR and APC in everyday diagnostic, non-iFOBT screening and surveillance colonoscopy in our study. CADe

increased the detection of notoriously hard-to-detect SSLs, which are increasingly recognized for their clinical relevance; however, SSLPC was not a primary outcome in our study.

References

1. 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 e1611. doi:10.1053/j.gastro.2019.01.260
2. 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. doi:10.1056/NEJMoa0907667
3. Kaminski MF, Thomas-Gibson S, Bugajski M et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. United European Gastroenterol J 2017; 5: 309-334. doi:10.1177/2050640617700014
4. 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. doi:10.1056/NEJMoa1309086
5. 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. doi:10.7326/M22-3678
6. Ladabaum U, Shepard J, Weng Y et al. Computer-aided Detection of Polyps Does Not Improve Colonoscopist Performance in a Pragmatic Implementation Trial. Gastroenterology 2023; 164: 481-483 e486. doi:10.1053/j.gastro.2022.12.004
7. Levy I, Bruckmayer L, Klang E et al. Artificial Intelligence-Aided Colonoscopy Does Not Increase Adenoma Detection Rate in Routine Clinical Practice. Am J Gastroenterol 2022; 117: 1871-1873. doi:10.14309/ajg.0000000000001970
8. Schlemper RJ, Riddell RH, Kato Y et al. The Vienna classification of gastrointestinal epithelial neoplasia. Gut 2000; 47: 251-255. doi:10.1136/gut.47.2.251
9. Wang P, Berzin TM, Glissen Brown JR et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019; 68: 1813-1819. doi:10.1136/gutjnl-2018-317500
10. Repici A, Spadaccini M, Antonelli G et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. Gut 2022; 71: 757-765. doi:10.1136/gutjnl-2021-324471
11. Rondonotti E, Di Paolo D, Rizzotto ER et al. Efficacy of a computer-aided detection system in a fecal immunochemical test-based organized colorectal cancer screening program: a randomized controlled trial (AIFIT study). Endoscopy 2022; 54: 1171-1179. doi:10.1055/a-1849-6878
12. Shaukat A, Lichtenstein DR, Somers SC et al. Computer-Aided Detection Improves Adenomas per Colonoscopy for Screening and Surveillance Colonoscopy: A Randomized Trial. Gastroenterology 2022; 163: 732-741. doi:10.1053/j.gastro.2022.05.028
13. Gimeno-Garcia AZ, Hernandez Negrin D, Hernandez A et al. Usefulness of a novel computer-aided detection system for colorectal neoplasia: a randomized controlled trial. Gastrointest Endosc 2023; 97: 528-536 e521. doi:10.1016/j.gie.2022.09.029

14. Kamba S, Tamai N, Saitoh I et al. Reducing adenoma miss rate of colonoscopy assisted by artificial intelligence: a multicenter randomized controlled trial. J Gastroenterol 2021; 56: 746-757. doi:10.1007/s00535-021-01808-w
15. Karsenti D, Tharsis G, Perrot B et al. Effect of real-time computer-aided detection of colorectal adenoma in routine colonoscopy (COLO-GENIUS): a single-centre randomised controlled trial. Lancet Gastroenterol Hepatol 2023; 8: 726-734. doi:10.1016/S2468-1253(23)00104-8
16. Repici A, Badalamenti M, Maselli R et al. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. Gastroenterology 2020; 159: 512-520 e517. doi:10.1053/j.gastro.2020.04.062
17. Wei MT, Shankar U, Parvin R et al. Evaluation of Computer-Aided Detection During Colonoscopy in the Community (AI-SEE): A Multicenter Randomized Clinical Trial. Am J Gastroenterol 2023; 118: 1841-1847. doi:10.14309/ajg.0000000000002239
18. Taghiakbari M, Coman DE, Takla M et al. Measuring the observer (Hawthorne) effect on adenoma detection rates. Endosc Int Open 2023; 11: E908-E919. doi:10.1055/a-2131-4797
19. Xu H, Tang RSY, Lam TYT et al. Artificial Intelligence-Assisted Colonoscopy for Colorectal Cancer Screening: A Multicenter Randomized Controlled Trial. Clin Gastroenterol Hepatol 2023; 21: 337-346 e333. doi:10.1016/j.cgh.2022.07.006
20. Rex DK, Schoenfeld PS, Cohen J et al. Quality indicators for colonoscopy. Gastrointestinal Endoscopy 2015; 81: 31-53. doi:https://doi.org/10.1016/j.gie.2014.07.058
21. Desai M, Rex DK, Bohm ME et al. Impact of withdrawal time on adenoma detection rate: results from a prospective multicenter trial. Gastrointest Endosc 2023; 97: 537-543 e532. doi:10.1016/j.gie.2022.09.031
22. Lee TJ, Blanks RG, Rees CJ et al. Longer mean colonoscopy withdrawal time is associated with increased adenoma detection: evidence from the Bowel Cancer Screening Programme in England. Endoscopy 2013; 45: 20-26. doi:10.1055/s-0032-1325803
23. Ahmad A, Wilson A, Haycock A et al. Evaluation of a real-time computer-aided polyp detection system during screening colonoscopy: AI-DETECT study. Endoscopy 2023; 55: 313-319. doi:10.1055/a-1966-0661
24. Mangas-Sanjuan C, de-Castro L, Cubiella J et al. Role of Artificial Intelligence in Colonoscopy Detection of Advanced Neoplasias : A Randomized Trial. Ann Intern Med 2023; 176: 1145-1152. doi:10.7326/M22-2619
25. Edwardson N, Adsul P, Gonzalez Z et al. Sessile serrated lesion detection rates continue to increase: 2008-2020. Endosc Int Open 2023; 11: E107-E116. doi:10.1055/a-1990-0509
26. Erichsen R, Baron JA, Hamilton-Dutoit SJ et al. Increased Risk of Colorectal Cancer Development Among Patients With Serrated Polyps. Gastroenterology 2016; 150: 895-902 e895. doi:10.1053/j.gastro.2015.11.046
27. Anderson JC, Rex DK, Mackenzie TA et al. Higher Serrated Polyp Detection Rates Are Associated With Lower Risk of Postcolonoscopy Colorectal Cancer: Data From the New Hampshire Colonoscopy Registry. Am J Gastroenterol 2023; 118: 1927-1930. doi:10.14309/ajg.0000000000002403

28. Shaukat A, Colucci D, Erisson L et al. Improvement in adenoma detection using a novel artificial intelligence-aided polyp detection device. Endosc Int Open 2021; 9: E263-E270. doi:10.1055/a-1321-1317
29. Vennelaganti S, Cuatrecasas M, Vennalaganti P et al. Interobserver Agreement Among Pathologists in the Differentiation of Sessile Serrated From Hyperplastic Polyps. Gastroenterology 2021; 160: 452-454 e451. doi:10.1053/j.gastro.2020.09.015
30. Almansa C, Shahid MW, Heckman MG et al. Association between visual gaze patterns and adenoma detection rate during colonoscopy: a preliminary investigation. Am J Gastroenterol 2011; 106: 1070-1074. doi:10.1038/ajg.2011.26

Figure legends

A computer-aided detection system in the everyday setting of diagnostic, screening and surveillance colonoscopy: an international, randomized trial

M.H.J. Maas, T. Rath, C. Spada, E. Soons, N. Forbes, S. Kashin, P. Cesaro, A. Eickhoff, G. Vanbiervliet, D. Salvi, P.J. Belletrutti, P.D. Siersema; for the Discovery study team.

Figure 1. Detection by the CADe system

Legend: Computer-aided detection generated overlay of a blue bounding box highlighting a lesion during real-time colonoscopy.

Figure 2. Study flowchart

Legend: CADe=computer-aided detection; CC=conventional colonoscopy. Other exclusions: Russian site enrollment post February 22, 2022 (n=12), new polyposis diagnosis (n=2), ASA score of 3 (n=1), or new IBD diagnosis (n=1).

Supplementary Tables

A computer-aided detection system in the everyday setting of diagnostic, screening and surveillance colonoscopy: an international, randomized trial

M.H.J. Maas, T. Rath, C. Spada, E. Soons, N. Forbes, S. Kashin, P. Cesaro, A. Eickhoff, G. Vanbiervliet, D. Salvi, P.J. Belletrutti, P.D. Siersema; for the Discovery study team.

Supplementary table 1. Primary and secondary outcomes per reason for colonoscopy in the modified intention-to-treat population

	CC (n=247)	CADe (n=250)	Difference (treatment – control)	P value
Adenoma detection rate (ADR)				
Overall*	93/247 = 37.7%	96/250 = 38.4%	0.7 [-7.8 – 9.3]	.432
Non-iFOBT screening (n=96)	21/46 = 45.7%	19/50 = 38.0%	-7.7 [-27.4 – 12.1]	.447
Surveillance (n=202)	45/104 = 43.3%	43/98 = 43.9%	0.6 [-13.1 – 14.3]	.931
Diagnostic (n=199)	27/97 = 27.8%	34/102 = 33.3%	5.5 [-7.3 – 18.3]	.400
Adenoma per colonoscopy (APC)				
Overall	163/247 = 0.66	165/250 = 0.66	0.00 [-0.19 – 0.19]	.971
Non-iFOBT screening (n=96)	37/46 = 0.80	47/50 = 0.94	0.14 [-0.42 – 0.69]	.748
Surveillance (n=202)	84/104 = 0.81	71/98 = 0.72	-0.09 [-0.40 – 0.23]	.699
Diagnostic (n=199)	42/97 = 0.43	48/102 = 0.47	0.04 [-0.18 – 0.25]	.459
Polyp detection rate (PDR)				
Overall	127/247 = 51.4%	138/250 = 55.2%	3.8 [-5.0 – 12.5]	.398
Non-iFOBT screening (n=96)	27/46 = 58.7%	30/50 = 60.0%	1.3 [-18.4 – 21.0]	.897
Surveillance (n=202)	58/104 = 55.8%	59/98 = 60.2%	4.4 [-9.2 – 18.4]	.523

Diagnostic (n=199)	42/97 = 43.3%	49/102 = 48.0%	4.7 [-9.1 – 18.6]	.502
Sessile serrated lesions per colonoscopy (SSLPC)				
Overall	46/247 = 0.19	76/250 = 0.30	0.11 [0.00 – 0.24]	.049
Non-iFOBT screening (n=96)	8/46 = 0.17	15/50 = 0.30	0.13 [-0.13 – 0.38]	.227
Surveillance (n=202)	22/104 = 0.21	35/98 = 0.36	0.15 [-0.09 – 0.38]	.405
Diagnostic (n=199)	16/97 = 0.16	26/102 = 0.25	0.09 [-0.07 – 0.25]	.150
Sessile serrated lesion detection rate (SDR)				
Overall	30/247 = 12.1%	46/250 = 18.4%	6.3 [-0.04 – 12.5]	.053
Non-iFOBT screening (n=96)	5/46 = 10.9%	10/50 = 20.0%	9.1 [-5.1 – 23.4]	.218
Surveillance (n=202)	14/104 = 13.5%	17/98 = 17.3%	3.8 [-6.1 – 13.8]	.444
Diagnostic (n=199)	11/97 = 11.3%	19/102 = 18.6%	7.3 [-2.6 – 17.1]	.151
Mean Polyps per colonoscopy (PPC)				
Overall	270/247 = 1.09	299/250 = 1.20	0.11 [-0.15 – 0.36]	.515
Non-iFOBT screening (n=96)	59/46 = 1.28	80/50 = 1.60	0.32 [-0.38 – 1.02]	.608
Surveillance (n=202)	135/104 = 1.30	125/98 = 1.28	-0.02 [-0.45 – 0.41]	.886
Diagnostic (n=199)	75/97 = 0.77	94/102 = 0.92	0.15 [-0.16 – 0.46]	.358
Withdrawal time without interventions (min)	9.0 [8.0 – 11.0]	9.2 [8.0 – 11.0]	0.2	.052

Total Procedure time (min)	20.0 [15.0 – 24.7]	20.0 [15.0 – 27.6]	0.0	.430

*Statistical analysis of the primary outcome is performed using a one-sided approach to the Chi-square test. Other p-values represent two-sided analyses. Data is n/N(%), or median(IQR). CADe=computer-aided detection, CC=conventional colonoscopy, [95%CI] calculated using Wilson score interval for proportions.

Supplementary Table 2. False positives in the modified intention-to-treat population

	CADe (n=250)
False positives, median (IQR) [mean, ±SD]	2.0(0.0 – 5.0) [
Reason of false positive per colonoscopy	
Colonic fold	102/250 = 40.8%
Bubble	77/250 = 30.8%
Fecal material	82/250 = 32.8%
Ileocecal valve	27/250 = 10.8%
Suction artefact	16/250 = 6.4%
Other	27/250 = 10.8%

CADe=computer-aided detection. Data are n/N(%). False positives were characterized as an unsuspected area highlighted by CADe for longer than 3 seconds, as assessed by the endoscopist.

Supplementary table 3. Post-hoc analysis of ADR between CC and CADe-assisted colonoscopy according to endoscopist basal ADR

	CC (n=247)	CADe (n=250)	Difference (treatment – control)	P value
Adenoma detection rate (ADR)				
Lower-detector tertile	15/65 = 23.1% [14.5 – 34.6]	27/74 = 36.5% [26.4 – 47.9]	13.4 [-1.6 – 28.4]	.086
Medium-detector tertile	28/75 = 37.3% [27.3 – 48.6]	29/72 = 40.3% [29.7 – 51.8]	3.0 [-12.8 – 18.7]	.714
High-detector tertile	47/95 = 49.5% [39.6 – 59.4]	36/91 = 39.6% [30.1 – 49.8]	-9.9 [-24.1 – 4.3]	.174

Endoscopists were categorized in tertiles based on their ADR in the CC study arm. Endoscopists with <5 colonoscopies performed in the CC study arm were excluded from the initial calculation of tertiles. Subsequently, they were added to their corresponding tertile based on their ADR. Low-detectors were the endoscopists in the bottom tertile, medium-detectors in the middle tertile, and high-detectors were the top tertile. ADR=adenoma detection rate, CADe=computer-aided detection, CC=conventional colonoscopy.

Supplementary table 4. Primary and secondary outcomes of the intention-to-treat population

	CC (n=290)	CADe (n=287)	Difference (treatment – control)	P value
Adenoma detection rate (ADR)				
Overall*	105/290 = 36.2%	106/287 = 36.9%	0.7 [-7.1 – 8.6]	.428
Non-iFOBT screening (n=107)	23/52 = 44.2%	21/55 = 38.2%	-6.0 [-24.7 – 12.6]	.525
Surveillance (n=234)	51/124 = 41.1%	47/110 = 42.7%	1.6 [-11.1 – 14.2]	.805
Diagnostic (n=236)	31/114 = 27.2%	38/122 = 31.1%	3.4 [-7.6 – 15.5]	.504
Adenoma per colonoscopy (APC)				
Overall	177/290 = 0.61	184/287 = 0.64	0.03 [-0.14 – 0.20]	.911
Non-iFOBT screening (n=107)	39/52 = 0.75	50/55 = 0.91	0.16 [-0.35 – 0.67]	.537
Surveillance (n=234)	91/124 = 0.73	81/110 = 0.74	0.01 [-0.29 – 0.29]	.980
Diagnostic (n=236)	47/114 = 0.41	53/122 = 0.43	0.02 [-0.17 – 0.22]	.557
Polyp detection rate (PDR)				
Overall	147/290 = 50.7%	158/287 = 55.1%	4.4 [-3.8 – 12.5]	.294
Non-iFOBT screening (n=107)	31/52 = 59.6%	33/55 = 60.0%	0.4 [-18.2 – 19.0]	.968
Surveillance (n=234)	67/124 = 54.0%	67/110 = 60.9%	6.9 [-5.8 – 19.5]	.289

Diagnostic (n=224)	49/114 = 43.0%	58/122 = 47.5%	4.3 [-8.1 – 17.3]	.482
Sessile serrated lesions per colonoscopy (SSLPC)				
Overall	56/290 = 0.19	85/287 = 0.30	0.11 [0.00 – 0.21]	.045
Non-iFOBT screening (n=107)	9/52 = 0.17	16/55 = 0.29	0.12 [-0.11 – 0.35]	.235
Surveillance (n=234)	23/124 = 0.19	36/110 = 0.33	0.14 [-0.05 – 0.35]	.320
Diagnostic (n=236)	24/114 = 0.21	33/122 = 0.27	0.06 [-0.1 – 0.22]	.202
Sessile serrated lesion detection rate (SDR)				
Overall	36/290 = 12.4%	53/287 = 18.5%	6.1 [0.2 – 11.9]	.044
Non-iFOBT screening (n=107)	6/52 = 11.5%	11/55 = 20.0%	9.5 [-5.2 – 22.1]	.231
Surveillance (n=234)	15/124 = 12.1%	18/110 = 16.4%	4.3 [-4.7 – 13.3]	.349
Diagnostic (n=236)	15/114 = 13.2%	24/122 = 19.7%	6.5 [-2.9 – 15.9]	.178
Mean Polyps per colonoscopy (PPC)				
Overall	304/290 = 1.05	335/287 = 1.17	0.12 [-0.12 – 0.35]	.388
Non-iFOBT screening (n=107)	65/52 = 1.25	86/55 = 1.56	0.31 [-0.33 – 0.95]	.569
Surveillance (n=234)	148/124 = 1.19	139/110 = 1.26	0.07 [-0.32 – 0.46]	.708
Diagnostic (n=236)	91/114 = 0.80	110/122= 0.90	0.10 [-0.19 – 0.40]	.444
Withdrawal time without interventions (min)	9.0 (8.0 – 11.0)	9.0 (8.0 – 11.0)	0.0	.026

Total Procedure time (min)	19.0 (15.0 – 24.0)	20.0 (15.0 – 27.0)	1.0	.225

ITT population (n=577) consists of all randomized patients (n=581) after exclusion of new polyposis diagnosis (n=2), ASA score of 3 (n=1), or new IBD diagnosis (n=1). *Statistical analysis of the primary outcome is performed using a one-sided approach to the Chi-square test. Other p-values represent two-sided analyses. CADe=computer-aided detection, CC=conventional colonoscopy, CI=confidence interval. Data are n/N(%) or median (IQR). [95% CI] are calculated using the Wilson score interval for proportions.

Supplementary Table 5. Adenoma detection rate characteristics in the modified intention-to-treat population

	CC (n=247)	CADe (n=250)	Difference (treatment – control)	P value
Localization				
Caecum	15/247 = 6.1%	13/250 = 5.2%	-0.9 [-4.9 – 3.2]	.673
Ascending colon	33/247 = 13.4%	48/250 = 19.2%	5.8 [-0.6 – 12.3]	.078
Transverse colon	32/247 = 13.0%	32/250 = 12.8%	-0.2 [-6.0 – 5.7]	.959
Descending colon	21/247 = 8.5%	12/250 = 4.8%	-3.7 [-8.1 – 0.7]	.097
Sigmoid colon	25/247 = 10.1%	18/250 = 7.2%	-2.9 [-7.9 – 2.0]	.247
Rectum	8/247 = 3.2%	14/250 = 5.6%	2.4 [-1.2 – 6.0]	.201
Proximal colon	63/247 = 25.5%	72/250 = 28.8%	3.3 [-4.5 – 11.1]	.409
Distal colon	49/247 = 19.8%	41/250 = 16.4%	-3.4 [-10.2 – 3.3]	.320
Size				
≤5mm	67/247 = 27.1%	77/250 = 30.8%	3.7 [-4.3 - 11.6]	.367
6-9mm	38/247 = 15.4%	28/250 = 11.2%	-4.2 [-10.1 – 1.8]	.169
≥10mm	14/247 = 5.7%	14/250 = 5.6%	-0.1 [-0.4 – 4.0]	.974
Morphology (Paris classification)				
Pedunculated	11/247 = 4.5%	9/250 = 3.6%	-0.9 [-4.3 – 2.6]	.628
Sessile	72/247 = 29.2%	73/250 = 29.2%	0.0 [-7.9 – 8.0]	.990
Flat elevated	17/247 = 6.9%	29/250 = 11.6%	4.7 [-0.4 – 9.8]	.070
Flat lesion	5/247 = 2.0%	4/250 = 1.6%	-0.4	.723

			[-2.8 – 1.9]	
Slightly depressed	0/247 = 0.0%	0/250 = 0.0%	-	-
Excavated	0/247 = 0.0%	0/250=0.0%	-	-

CC=conventional colonoscopy, CI=confidence interval, CADe=computer-aided detection. Data are n/N(%). [95% CI] are calculated using the Wilson score interval for proportions.

Supplementary table 6. ADR per study site in the modified intention-to-treat population

Study site	CC (n=247)	CADe (n=250)	Difference (treatment – control)	P value
01 (n=112)	22/57 = 38.6%	25/55 = 45.5%	6.9 [-11.4 – 25.1]	.462
02 (n=85)	17/39 = 43.6%	15/46 = 32.6%	-11.0 [-31.6 – 9.7]	.298
03 (n=108)	18/54 = 33.3%	24/54 = 44.4%	11.1 [-7.2 – 29.4]	.236
04 (n=109)	21/54 = 38.9%	20/55 = 36.4%	-2,5 [-20.7 – 15.7]	.786
05 (n=23)	4/12 = 33.3%	4/11 = 36.4%	3.1 [-36.0 – 42.0]	.879
06 (n=38)	8/20 = 40.0%	4/18 = 22.2%	-17.8 [-46.6 – 11.0]	.239
07 (n=22)	3/11 = 27.3%	4/11 = 36.4%	9.1 [-29.6 – 47.8]	.647

CADe=computer-aided detection, CC=conventional colonoscopy, CI=confidence interval. Data are n/N(%). [95% CI] are calculated using the Wilson score interval for proportions.

Supplementary table 7. ADR per endoscopist in the modified intention-to-treat population

Endoscopist	CC (n=247)	CADe (n=250)	Difference (treatment – control)	P value
001	17/30 = 56.7%	11/23 = 47.8%	-8.9 [-35.8 – 18.2]	.523
002	3/15 = 20.0%	8/17 = 47.1%	27.1 [-4.1 – 58.2]	.108
003	0/5 = 0.0%	5/10 = 50%	50.0 [19.0 – 81.0]	.053
004	2/7 = 28.6%	1/5 = 20.0%	-8.6 [-57.0 – 39.9]	.753
005	16/43 = 37.2%	21/43 = 48.8%	11.6 [-9.2 – 32.4]	.276
006	2/7 = 28.6%	2/6 = 33.3%	4.7 [-45.7 – 55.2]	.853
007	0/3 = 0.0%	1/3 = 33.3%	33.3 [-20.0 – 86.7]	.273
008	0/1 = 0.0%	0/2 = 0.0%	-	-
009	17/39 = 43.6%	15/46 = 32.6%	-11.0 [-31.6 – 9.7]	.298
010	8/28 = 28.6%	10/33 = 30.3%	1.7 [30.3 – 28.6]	.883
011	13/26 = 50.0%	10/22 = 45.5%	-4.5 [-32.9 – 23.8]	.753
012	4/12 = 33.3%	4/11 = 36.4%	3.1 [-36.0 – 42.0]	.879
013	3/11 = 27.3%	4/11 = 36.4%	9.1 [-29.6 – 47.8]	.647
014	8/20 = 40.0%	4/18 = 22.2%	-17.8 [-46.6 – 11.0]	.239

CADe=computer-aided detection, CC=conventional colonoscopy, CI=confidence interval,. Data are n/N(%). [95% CI] are calculated using the Wilson score interval for proportions.

Table 1

Table 1. Baseline characteristics in the modified intention-to-treat population

	CADe (n=250)	CC (n=247)
Age, years	61.0 (52 – 69)	61.0 (52 – 69)
Gender		
Female	141/250 (56.4%)	136/247 (55.1%)
Male	109/250 (43.6%)	111/247 (44.9%)
Colonoscopy indication		
Screening (non-iFOBT)	50/250 (20.0%)	46/247 (18.7%)
Surveillance	98/250 (39.2%)	104/247 (42.1%)
Diagnostic*	102/250 (40.8%)	97/247 (39.3%)
Index colonoscopy, yes	112/250 (44.8%)	108/247 (43.7%)
Smoking, yes	26/250 (10.4%)	32/247 (13.0%)
Family history of CRC	60/250 (24.0%)	45/247 (18.2%)
BMI, kg/m2†	25.5 (23.1 – 28.3)	25.0 (22.5 – 28.8)
BBPS score 6	61/250 (24.4%)	65/247 (26.3%)
BBPS score 7	27/250 (10.8%)	27/247 (10.9%)
BBPS score 8	29/250 (11.6%)	34/247 (13.8%)
BBPS score 9	133/250 (53.2%)	121/247 (49.0%)

*Diagnostic indications did not include iFOBT+ referrals, †1 patient missed weight. Data is n/N(%), or median(IQR). BBPS=Boston bowel preparation score, BMI=body mass index, CADe=computer-aided detection, CC=conventional colonoscopy, iFOBT=immunochemical fecal occult blood test

Table 2

Table 2. Primary and secondary outcomes in the modified intention-to-treat population

	CC (n=247)	CADe (n=250)	Difference (treatment – control)	P value
Adenoma detection rate (ADR)*	93/247 = 37.7%	96/250 = 38.4%	0.7 [-7.8 – 9.3]	.432
Adenoma per colonoscopy (APC)	163/247 = 0.66	165/250 = 0.66	0.00 [-0.19 – 0.19]	.971
Polyp detection rate (PDR)	127/247 = 51.4%	138/250 = 55.2%	3.8 [-5.0 – 12.5]	.398
Sessile serrated lesions per colonoscopy (SSLPC)	46/247 = 0.19	76/250 = 0.30	0.11 [0.00 – 0.24]	.049
Sessile serrated lesion detection rate (SDR)	30/247 = 12.1%	46/250 = 18.4%	6.3 [-0.04 – 12.5]	.053
Mean polyps per colonoscopy (PPC)	270/247 = 1.09	299/250 = 1.20	0.11 [-0.15 – 0.36]	.515
Withdrawal time without interventions (min)	9.0 (8.0 – 11.0)	9.2 (8.0 – 11.0)	0.2	.052
Total procedure time (min)	20.0 (15.0 – 24.7)	20.0 (15.0 – 27.6)	0.0	.430

*Statistical analysis of the primary outcome is performed using a one-sided approach to the Chi-square test. Other p-values represent two-sided analyses. Data is n/N(%), or median(IQR). CADe=computer-aided detection, CC=conventional colonoscopy, [95%CI] calculated using Wilson score interval for proportions.

Table 3

Table 3. Additional analysis of primary and secondary outcomes in the modified intention-to-treat population

	CC (n=247)	CADe (n=250)	Odds or effect ratio CADe to CC	P value
Adenoma detection rate (ADR)*	93/247 = 37.7%	96/250 = 38.4%	1.032 [0.719 – 1.483]	.864
Mean number of adenomas per colonoscopy (APC)†	163/247 = 0.66	165/250 = 0.66	1.006 [0.811 – 1.249]	.955
Sessile serrated lesions per colonoscopy (SSLPC)‡	46/247 = 0.19	76/250 = 0.30	1.632 [1.132 – 2.354]	.009
Mean number of polyps per colonoscopy (PPC)‡	270/247 = 1.09	299/250 = 1.20	1.098 [0.863 – 1.397]	.446

Data are n/N(%) [95% CI]. CC=conventional colonoscopy, CI=confidence interval, CADe=computer-aided detection *Odds ratio from logistic binary regression with treatment arm and ADR. †Effect ratio from Poisson regression with log-link function. ‡Effect ratio from Poisson regression with negative binomial function.