NEJM Evidence

Published April 13, 2022

DOI: 10.1056/EVIDoa2200003

ORIGINAL ARTICLE

Real-Time Artificial Intelligence–Based Optical Diagnosis of Neoplastic Polyps during Colonoscopy

Ishita Barua, M.D.,[1,2] Paulina Wieszczy, Ph.D.,[1,2,3] Shin-ei Kudo, M.D.,[4] Masashi Misawa, M.D.,[4] Øyvind Holme, M.D.,[1,2,5] Shraddha Gulati, M.D.,[6] Sophie Williams, M.D.,[6] Kensaku Mori, Ph.D.,[7] Hayato Itoh, Ph.D.,[7] Kazumi Takishima, M.D.,[4] Kenichi Mochizuki, M.D.,[4] Yuki Miyata, M.D.,[4] Kentaro Mochida, M.D.,[4] Yoshika Akimoto, M.D.,[4] Takanori Kuroki, M.D.,[4] Yuriko Morita, M.D.,[4] Osamu Shiina, M.D.,[4] Shun Kato, M.D.,[4] Tetsuo Nemoto, M.D.,[8] Bu Hayee, M.D.,[6] Mehul Patel, M.D.,[6] Nishmi Gunasingam, M.D.,[6] Alexandra Kent, M.D.,[6] Andrew Emmanuel, M.D.,[6] Carl Munck, M.D.,[9] Jens Aksel Nilsen, M.D.,[9] Stine Astrup Hvattum, M.D.,[9] Svein Oskar Frigstad, M.D.,[9] Petter Tandberg, M.D.,[9] Magnus Løberg, M.D.,[1,2] Mette Kalager, M.D.,[1,2] Amyn Haji, M.D.,[6] Michael Bretthauer, M.D.,[1,2] and Yuichi Mori, M.D.[1,2,4]

Abstract

BACKGROUND Artificial intelligence using computer-aided diagnosis (CADx) in real time with images acquired during colonoscopy may help colonoscopists distinguish between neoplastic polyps requiring removal and nonneoplastic polyps not requiring removal. In this study, we tested whether CADx analyzed images helped in this decision-making process.

METHODS We performed a multicenter clinical study comparing a novel CADx-system that uses real-time ultra-magnifying polyp visualization during colonoscopy with standard visual inspection of small (≤5 mm in diameter) polyps in the sigmoid colon and the rectum for optical diagnosis of neoplastic histology. After committing to a diagnosis (i.e., neoplastic, uncertain, or nonneoplastic), all imaged polyps were removed. The primary end point was sensitivity for neoplastic polyps by CADx and visual inspection, compared with histopathology. Secondary end points were specificity and colonoscopist confidence level in unaided optical diagnosis.

RESULTS We assessed 1289 individuals for eligibility at colonoscopy centers in Norway, the United Kingdom, and Japan. We detected 892 eligible polyps in 518 patients and included them in analyses: 359 were neoplastic and 533 were nonneoplastic. Sensitivity for the diagnosis of neoplastic polyps with standard visual inspection was 88.4% (95% confidence interval [CI], 84.3 to 91.5) compared with 90.4% (95% CI, 86.8 to 93.1) with CADx (P=0.33). Specificity was 83.1% (95% CI, 79.2 to 86.4) with standard visual inspection and 85.9% (95% CI, 82.3 to 88.8) with CADx. The proportion of polyp assessment with high confidence was 74.2% (95% CI, 70.9 to 77.3) with standard visual inspection versus 92.6% (95% CI, 90.6 to 94.3) with CADx.

Drs. Barua, Wieszczy, and Kudo contributed equally as co-first authors, and Drs. Haji, Bretthauer, and Mori contributed equally as co-last authors.

The author affiliations are listed at the end of the article.
Dr. Bretthauer can be contacted at michael.bretthauer@medisin.uio.no or at Clinical Effectiveness Research Group, Institute of Health and Society, University of Oslo, Oslo 0027, Norway.

CONCLUSIONS Real-time polyp assessment with CADx did not significantly increase the diagnostic sensitivity of neoplastic polyps during a colonoscopy compared with optical evaluation without CADx. (Funded by the Research Council of Norway [Norges Forskningsråd], the Norwegian Cancer Society [Kreftforeningen], and the Japan Society for the Promotion of Science; UMIN number, UMIN000035213.)

Introduction

Colorectal cancer is the third most common cancer and the second leading cause of cancer deaths worldwide.[1] Removal of precancerous polyps during colonoscopy is the cornerstone of colorectal cancer screening. Most colorectal polyps are small (≤5 mm in diameter) and located in the sigmoid colon and the rectum. Although most colorectal cancers develop from polyps, many small polyps are not neoplastic and do not have any malignant potential.[2]

With current standard colonoscopy equipment, many endoscopists, especially those with less experience, cannot reliably distinguish between neoplastic and nonneoplastic polyps on visual inspection, a procedure known as "optical diagnosis."[3,4] Therefore, the current standard of care is to remove all polyps and submit them for histopathologic diagnosis. Reliable real-time optical diagnosis of small polyps during colonoscopy could enable targeted removal only of polyps classified as neoplastic, while small, nonneoplastic polyps could be left behind.[5]

In a recent single-center, proof-of-concept study of a novel artificial intelligence (AI) system for computer-aided polyp diagnosis (CADx), we achieved a reliable distinction between small neoplastic and nonneoplastic polyps in the distal colon and the rectum.[6] The CADx system combines colonoscopes with 520× magnification of polyp surfaces during colonoscopy in real time, and it enables AI-derived automated optical diagnosis of neoplastic and nonneoplastic polyps in about 40 seconds. The automated diagnosis is signaled to the colonoscopist by an acoustic and optical alarm during each polyp assessment.[6-8]

The current multicenter clinical study was designed to compare the clinical performance of AI CADx-based optical diagnosis in distinguishing neoplastic from nonneoplastic small polyps in the sigmoid colon and the rectum during colonoscopy with standard visual inspection–based optical diagnosis in routine clinical colonoscopy practice.

Methods

STUDY DESIGN AND OVERSIGHT

We performed a multicenter clinical study of AI CADx polyp classification and visual inspection versus standard visual inspection alone. Study procedures were performed at three participating endoscopy centers: Baerum Hospital (Norway), King's College Hospital London (United Kingdom), and Showa University Northern Yokohama Hospital (Japan).

The institutional review board (IRB) at each of the three participating centers approved the conduct of the study. The study protocol and statistical analysis plan are available with the full text of this article at evidence.nejm.org. Patient consent was implemented at the three study sites according to local IRB practice; In Norway, only participants enrolled in the national screening program pilot were eligible for participation and written informed patient consent was included in the consent of the screening program. In Japan, the IRB approved an opt-out consent approach because of the low risk related to the study intervention (standard treatment was performed for all polyps detected). In London, all patients provided informed consent.

All co-authors agreed on publishing the article and vouch for the completeness and accuracy of the data and the adherence to the protocol.

PATIENTS

Eligible patients were individuals 18 years of age or older who were scheduled for colonoscopy for colorectal cancer screening, polyp surveillance, or evaluation of clinical signs or symptoms at the participating centers between May 2019 and May 2021. Exclusion criteria were inflammatory bowel disease, polyposis syndrome (familial adenomatous polyposis, serrated polyposis), history of or current chemotherapy or radiation for rectosigmoid tumors, inability to undergo polypectomy (e.g., anticoagulants, comorbidities), pregnancy, and referral for removal of polyps with known histology.

All patients with small polyps (≤5 mm in diameter) in the sigmoid colon or the rectum (jointly called rectosigmoid

colon) detected during colonoscopy were included in this study. For patients with more than five eligible polyps, the first five polyps were included and evaluated according to the study interventions (described next).

COLONOSCOPY PROCEDURES

All colonoscopies were performed according to routine standards at the participating centers, including preprocedure assessment, bowel preparation, sedation practices, and postprocedure recovery and care.

The following information was assessed and was registered in the study database immediately during and after each procedure: indication for colonoscopy, quality of bowel preparation assessed by the Boston Bowel Preparation Score (a 9-point assessment scale for cleaning quality during colonoscopy, with higher numbers indicating better preparation)[9]; most proximal segment of the colon reached during colonoscopy; insertion and withdrawal duration; and size, shape, and location of all detected polyps. All detected polyps were removed for histologic assessment for final diagnosis. By study design, study colonoscopists were nonexperts, defined as having between 1 and 5 years of colonoscopy experience or having independently performed between 200 and 1000 procedures before joining the study as an endoscopist. This aspect of the study design was included because we wanted to determine whether CADx improved the performance of reasonably trained, but nonexpert, endoscopists and thus shortened the learning curve in endoscopy training so the study colonoscopists behaved like experts. The study endoscopists were accredited for standard colonoscopy in the participating countries, but they did not have additional training in optical polyp diagnosis before the study. For the purpose of this study, study endoscopists received training on handling the study colonoscopes and devices and image interpretation. Novice endoscopists were not included because they are unlikely to make optical diagnoses independently from supervisors in clinical practice.

EQUIPMENT

The study centers were provided with high-resolution magnification colonoscopes (CF-H290ECI; Olympus Corp., Tokyo, Japan). These appear to be standard instruments by design, feel, and function, including narrow band imaging. In addition, the study colonoscope featured a light microscopy system integrated into the distal tip of the colonoscope. The extra feature provided 520-fold magnification at a focusing depth of 35 μm, and a field of view of 570 × 500 μm, for high-resolution magnified images on demand, which the colonoscopist controlled with a hand-operated lever.[6] This feature enabled real-time, in vivo evaluation of polyp microvascular morphology.

AI SYSTEM

The study centers were also provided with a real-time polyp classification CADx device (EndoBRAIN; Cybernet Systems Corp., Tokyo, Japan), connected to a standard colonoscopy processor unit (EVIS LUCERA ELITE, CV-290; Olympus Corp.). As noted earlier, the CADx system provides an automated diagnosis of rectosigmoid polyps by analyzing images obtained in the magnification mode of the colonoscopes for detected polyps, as previously described.[6-8]

Briefly, the CADx algorithm comprises three steps. The first is feature extraction, which is the analysis of textures characterized by differences in contrast for polyp vessels and lumens, quantified in 312 validated variables. Second is classification, which comprises support-vector machine classification of polyps as nonneoplastic or neoplastic on the basis of the 312 variables through machine learning. For the system training and validation, more than 35,000 polyp images were used which were collected from five Japanese endoscopy centers, as described previously.[10] Finally, in the diagnostic output step, the predicted diagnosis is displayed (Fig. 1) for the colonoscopist as "neoplastic" or "nonneoplastic" with a confidence probability for neoplasia (0 to 100%).

If the CADx diagnosis has a confidence probability of less than 70%, the system flags it as "low confidence," on the basis of a previous preclinical study.[10] If the quality of the captured image is not appropriate for system diagnosis (e.g., artifacts caused by mucus, low image quality), the analysis is flagged as "not a good sample," and no diagnosis is provided.

The nonneoplastic category comprises polyps with no neoplastic features, such as hyperplastic polyps, inflammatory polyps, and juvenile polyps. The neoplastic category comprises polyps with neoplastic features, such as adenomas and cancers.

POLYP HANDLING

For each detected polyp, four consecutive steps were applied. Step 1 comprised the standard endoscopic assessment. First, colonoscopists assessed the size, shape, and appearance of each detected polyp 5 mm or less in diameter in the rectosigmoid colon. Morphology was categorized

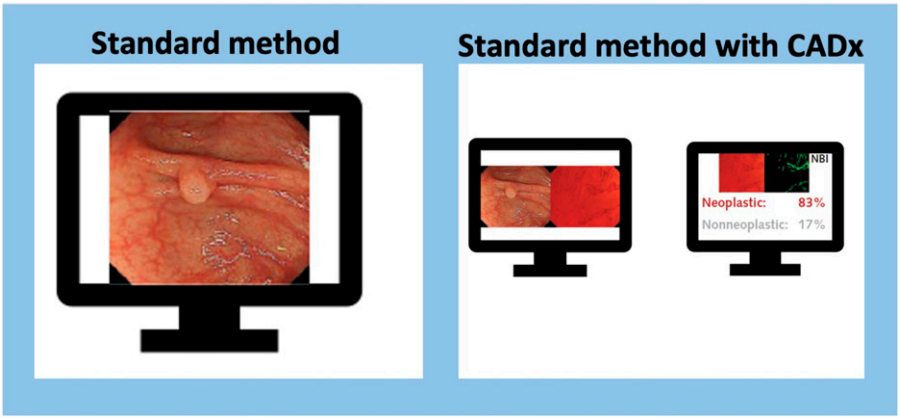

Figure 1. The Standard Method and the Combined Use of the Standard Method and the CADx System.
The Cybernet Systems EndoBRAIN system was used in this study. CADx denotes computer-aided diagnosis.

according to the Paris classification.[11] The endoscopists then classified polyps as either neoplastic (adenoma) or nonneoplastic (nonadenoma) using a binary scale (i.e., low or high confidence level in a nonneoplastic diagnosis, following recommendations in current guidelines[12-14]). Once the endoscopist registered their optical diagnosis, the CADx predicted classification was reported immediately for each polyp and registered in the study database.

Step 2 was the CADx assessment. After the standard assessment as described earlier, colonoscopists captured at least five images from each polyp using narrow band imaging and magnification mode to feed the CADx system. The CADx system then immediately provided the suggested diagnosis of the polyp as either neoplastic or nonneoplastic according to the algorithms described earlier (Fig. 1).

Step 3 was performed after standard clinical assessment and after CADx assessment, respectively. The colonoscopist again scored the confidence level of classification prediction of each polyp as either "high" or "low" and relayed it to the study nurse for immediate capture in the study database.

In step 4, all polyps were removed by snare polypectomy, biopsy forceps, or endoscopic mucosal resection and submitted for histopathologic evaluation. All polyps were evaluated by board-certified (the local board for each country of practice) gastrointestinal pathologists at each center. All pathologists were blinded to colonoscopic diagnoses of the polyps.

All polyps that were diagnosed histopathologically as nonneoplastic but had been considered by the colonoscopist as neoplastic with high confidence after standard assessment were submitted for a second histopathologic review by a different pathologist. The second pathologist was blinded to the first histopathologic diagnosis. See Supplementary Appendix, Section 2 for details.

STUDY END POINTS

The primary endpoint of the study was to compare the sensitivity of identifying small (≤5 mm in diameter) polyps in the rectosigmoid colon as adenomas during colonoscopy with the combination of standard visual inspection and the CADx system, and of standard visual inspection alone, compared with gold-standard histopathology.

Secondary outcome measures included specificity, positive predictive value (PPV), negative predictive value (NPV), rate of high-confidence optical diagnosis, and rate of rectosigmoid polyps of 5 mm or less with adequate images captured for CADx analysis.

Polyps that were not removed, those that were nonepithelial (neuroendocrine polyps, lymphoid aggregates), and those with unsuccessful image capturing were excluded from analyses.

SAMPLE SIZE CALCULATION

On the basis of a pilot study in Japan, we assumed a 6.7-percentage-point increase in sensitivity with the CADx system compared with the standard method, assuming discordance between the two methods of 14.4 percentage points (see the study protocol at evidence.nejm.org). We considered this difference to be clinically meaningful

to uncover. With a statistical power of 90%, the required sample size using a two-sided 5% significance level was 345 neoplastic polyps. We estimated that we needed to enroll 767 patients on the basis of a 25% prevalence of neoplastic eligible polyps, a mean of two eligible polyps per patient, and 90% of polyps with satisfactory prediction by the CADx system. The 90% threshold was motivated by U.S. guidelines recommending an NPV of 90% or greater for optical diagnosis of small neoplastic polyps.[5]

STATISTICAL ANALYSES

Sensitivity, specificity, PPV, and NPV for the standard method and the CADx method compared with histopathology, respectively, were estimated using generalized estimating equation analyses with exchangeable correlation accounting for correlation between multiple polyps within one patient. We did not account for clustering within colonoscopist, site, or country. We calculated 95% confidence intervals (CIs) using sandwich estimates of the variance. Sensitivity and specificity of the two interventions were compared using an exact version of the McNemar test. We did not adjust for multiple comparisons. Polyps that were not removed, from which specimens were lost after removal, or that had nonepithelial histology were excluded from analyses. All tests were performed in relation to the 0.05 significance level and used R version 3.4.1 and Stata version 17 software.

In primary analyses of sensitivity and specificity, sessile serrated lesions were classified as neoplastic (similar to adenomas). In secondary analyses, sessile serrated lesions were classified as nonneoplastic (no adenomas).

No interim analysis was planned at the study start in 2019. Because of slow recruitment during the Covid-19 pandemic, the study team decided to amend the protocol and performed a blinded interim analysis in April 2020. The interim analysis applied an a priori stopping rule for futility (see details in the study protocol on evidence.nejm.org), which was not met. Thus, the study was continued until preplanned recruitment was fulfilled. Because of the blinded nature of the interim analysis, we did not adjust for it in the final analysis.

Results

PATIENTS

The median age of patients included in analyses was 67 years (interquartile range [IQR], 60 to 74), and 63.1% were men (Table 1). Of the 1242 patients who underwent study colonoscopy, 525 had 903 eligible rectosigmoid polyps that received visual inspection.

Of the 903 eligible polyps, 11 were not included in analyses. Of these, 5 were not removed, 3 were lost after removal, and 3 were nonepithelial (two neuroendocrine tumors and one leiomyoma). Consequently, 892 polyps (359 neoplastic polyps and 533 nonneoplastic polyps) from 518 patients were included in the analyses (Fig. 2). The distribution of sex and age of the participants reflects real-world clinical practice (Table S2). We did not register the race and ethnicity of participants.

Twenty-two colonoscopists, including 20 physicians and two nurse endoscopists, performed the study procedures.

COLONOSCOPY PERFORMANCE AND COMPLICATIONS

Baseline characteristics of patients and colonoscopy performance are shown in Table 2. Most colonoscopies were for colorectal cancer screening or polyp surveillance. The median colonoscopy insertion time was 12 minutes (IQR, 8 to 19), and the median withdrawal time with polyp assessments and polypectomies was 28 minutes (IQR, 20 to 40). We did not observe any complications or adverse events related to the colonoscopy or to polyp assessment or removal.

POLYP CHARACTERISTICS

The 518 eligible patients had 892 detected and removed polyps that were 5 mm or less in the rectosigmoid colon. On the basis of the histopathologic examination of the removed polyps, 359 were neoplastic. Of these, 319 were tubular adenomas with low-grade dysplasia, 2 were tubular adenomas with high-grade dysplasia, 9 were tubulovillous adenomas with low-grade dysplasia, and 3 were tubulovillous adenomas with high-grade dysplasia. Of the 26 remaining polyps that were categorized as neoplastic, 7 were traditional serrated adenomas with low-grade dysplasia and 19 were sessile serrated lesions without dysplasia. On the basis of histopathologic examination, 533 polyps were found to be nonneoplastic. Of these, 485 were hyperplastic polyps, 8 were inflammatory polyps, and 40 had other nonneoplastic histology.

PERFORMANCE OF OPTICAL DIAGNOSIS

In primary analyses, the sensitivity for neoplastic polyps was 88.4% (95% CI, 84.3 to 91.5) with the standard

Table 1. Baseline Characteristics of 518 Included Patients and Their Colonoscopies.*

Characteristic	Value
Median age — yr	67 (60 to 74)
Sex	
Men	327 (63.1)
Women	191 (36.9)
Colonoscopy Indication	
Screening colonoscopy (primary screening or fecal testing)	266 (51.4)
Polyp surveillance colonoscopy	161(31.1)
Clinical signs or symptoms	67 (12.9)
Therapy of large polyps	23 (4.4)
Other	1 (0.2)
Median insertion time — min	12 (8 to 19)
Median withdrawal time — min	28 (20 to 40)
Preparation quality good or very good†	481 (92.9)

* Data are presented as the median (interquartile range) or no. (%).

† The Boston Bowel Preparation Scale is a 9-point assessment scale for cleaning quality during colonoscopy. The colon is divided into three segments: proximal, transverse, and distal. Each segment is classified from 0 to 3 depending on the degree of soiling. The sum total of the three segments represents the degree of soiling (≤5 points indicates poor bowel preparation; 6–7 good bowel preparation, and ≥8 very good bowel preparation).[9]

method and 90.4% (95% CI, 86.8 to 93.1) with the CADx method (P=0.33). The percentage of discordant pairs between the standard method and the CADx method was 7.2% (Fig. 3).

The specificity for neoplastic polyps was 83.1% (95% CI, 79.2 to 86.4) with the standard method and 85.9% (95% CI, 82.3 to 88.8) with the CADx method. The discordance between the standard method and the CADx method was 7.9%.

The percentage of polyp assessments with high confidence for categorization into neoplastic or nonneoplastic polyp increased from 74.2% (95% CI, 70.9 to 77.3) with the standard method to 92.6% (95% CI, 90.6 to 94.3) with the CADx method.

In secondary analyses classifying sessile serrated lesions as nonneoplastic, the sensitivity for neoplastic polyps was 91.2% (95% CI, 87.5 to 93.9) with the standard method and 94.1% (95% CI, 91.2 to 96.2) with the CADx method. The specificity for neoplastic polyps was 82.3% (95% CI, 78.4 to 85.6) with the standard method and 85.5% (95% CI, 81.9 to 88.5) with the CADx method. For separate center analyses, see Tables S3 through S8.

Discussion

Implementation of AI in cancer screening and clinical diagnosis requires proof of benefits from high-quality clinical studies. Our international multicenter study assessed the incremental gain of a specific CADx AI system for real-time polyp assessment during colonoscopy. Our study indicates that real-time AI with CADx may not significantly increase the sensitivity for small neoplastic polyps. However, CADx may improve specificity for optical diagnosis of small neoplastic polyps and increase colonoscopist confidence with visual diagnosis of polyps.

AI polyp detection tools (so-called computer-aided polyp detection) during colonoscopy could potentially increase detection of small polyps by up to 50%.[15] While this potentially could increase screening benefit, it also increases health care costs, risk of overtreatment, and patient burden.[16] Most additionally detected polyps are small ones in the distal colon and the rectum, and many of these are nonneoplastic; that is, they do not need to be removed if reliable, real-time classification were possible. One may further argue that removal of small polyps contributes little in terms of cancer prevention.[17]

The "diagnose-and-leave" strategy recently proposed by the American Society for Gastrointestinal Endoscopy (ASGE) suggests not to remove small polyps during colonoscopy if they can be reliably classified (defined as NPV of ≥90%) by optical diagnosis as nonneoplastic.[5] This strategy is not easy to apply because such reliable diagnosis is difficult to achieve with standard colonoscopy systems. Our study provides high-quality data to address this critical issue.

Our main outcome did not reach the prespecified increase of 6.7% in sensitivity with CADx, which was based on preclinical testing, observational studies, and a single-center study. Our study thus emphasizes the importance of rigorous clinical studies to assess AI performance and quantifies the added value and the limitation of CADx in colonoscopy.

According to our results, CADx may not reduce overlooking adenomas during visual inspection of polyps. However, our study showed a potential improvement in specificity for neoplastic polyps, albeit one in which we cannot declare statistical significance because our primary outcome failed to reach that level with the CADx system. There was also a trend toward improved confidence in

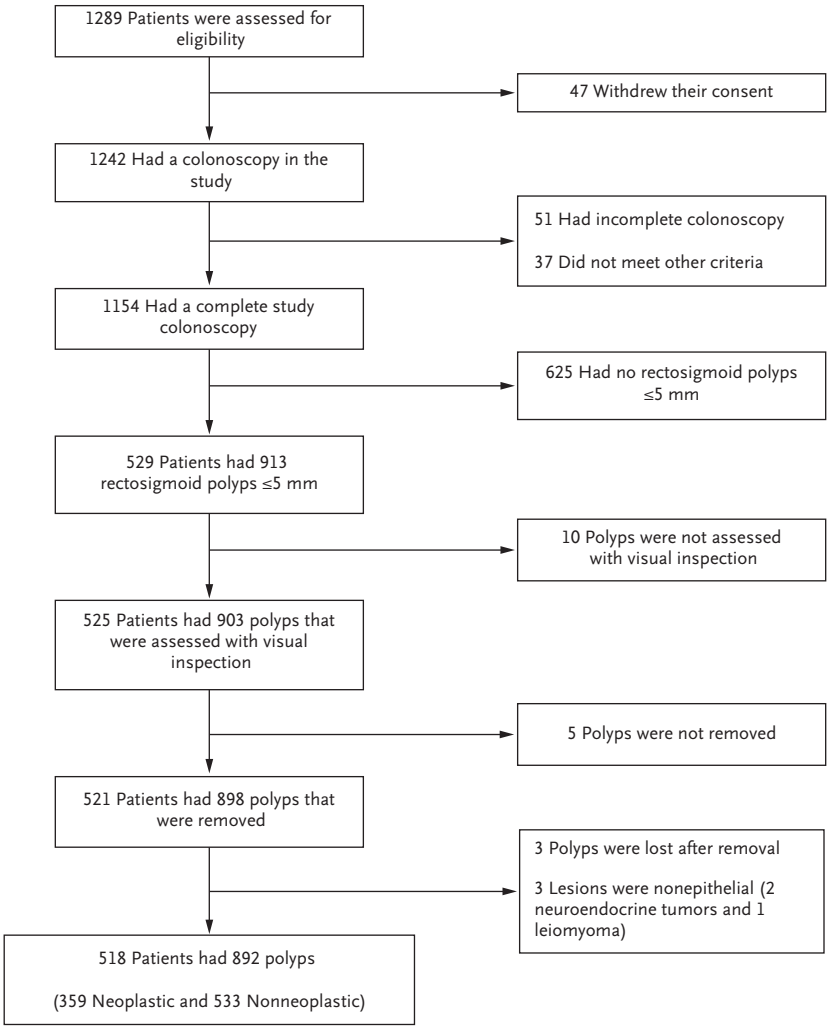

Figure 2. Study Flow Chart.

optical diagnosis of polyps. If this can be established through additional clinical trials, it could potentially contribute to a clinically important reduction in the unnecessary removal of small nonneoplastic polyps by giving the operator the ability to make a high-confidence prediction during a procedure.[5]

PPVs and NPVs are influenced by the prevalence of disease (polyps) and do not adequately assess tools or devices as such. Therefore, our primary outcomes of interest were sensitivity and specificity. However, we also analyzed the predictive values of CADx and observed increments of 1.3% for NPV and 3.1% for PPV with CADx (Table 3). Our results are consistent with the hypothesis that CADx can fulfill the criteria for the diagnose-and-leave strategy with 95% CIs above the NPV threshold of 90%.

The strengths of the current study are the comparison with both non-AI optical diagnosis and gold-standard histopathology for all included polyps; the inclusion of centers from different countries and continents; and the focus on endoscopists with average experience and workload, mimicking real-world colonoscopy practice. A limitation of this study is the inability of the CADx tool to identify sessile serrated polyps, a recently recognized polyp type with likely neoplastic potential. To alleviate this challenge, we conducted two analyses (one classifying sessile serrated polyps as neoplastic and the other classifying them as nonneoplastic) without significant differences in the performance of the CADx tool. Another limitation is the learning curve of the colonoscopists during the study period due to the prospective study design, which may contribute to underestimation of the CADx performance. However, we

Table 2. Characteristics of the 892 Small Polyps (≤5 mm in diameter) in the Distal Colon and the Rectum.*		
Characteristic	**Neoplastic Polyps (n=359)**	**Nonneoplastic Polyps (n=533)**
Median size — mm	4 (3 to 5)	3 (2 to 3)
Location		
Sigmoid colon	274 (76.3)	260 (48.8)
Rectum	85 (23.7)	273 (51.2)
Morphology†		
Polypoid (type Is or Ip)	175 (48.7)	109 (20.5)
Nonpolypoid (type IIa)	184 (51.3)	424 (79.5)
Removal method		
Snare polypectomy	247 (68.8)	265 (49.7)
Forceps	65 (18.1)	258 (48.4)
Endoscopic mucosal resection	46 (12.8)	10 (1.9)

* Data are presented as the median (interquartile range) or no. (%). Sessile serrated lesions were classified as neoplastic polyps in the primary analysis.

† The Paris classification was used. Morphologic classification systems for polyps during colonoscopy classify polyps into polypoid and nonpolypoid, with six different subtypes.[12]

may also have overestimated nonexpert endoscopists' performance because the sensitivity we found to predict adenomas, without the aid of CADx, was 88.4%, which is slightly higher than that reported in previous studies.[18,19] This may be related to the fact that our study was conducted at teaching hospitals with endoscopy training programs.

Finally, the colonoscopes used in the current study are not widely used today, although they are commercially available in Europe, the Middle East, and Asia. Provided that colonoscopes with surface enhancement functions facilitating CADx systems like the one we tested prove to be useful, they would likely become used more widely.

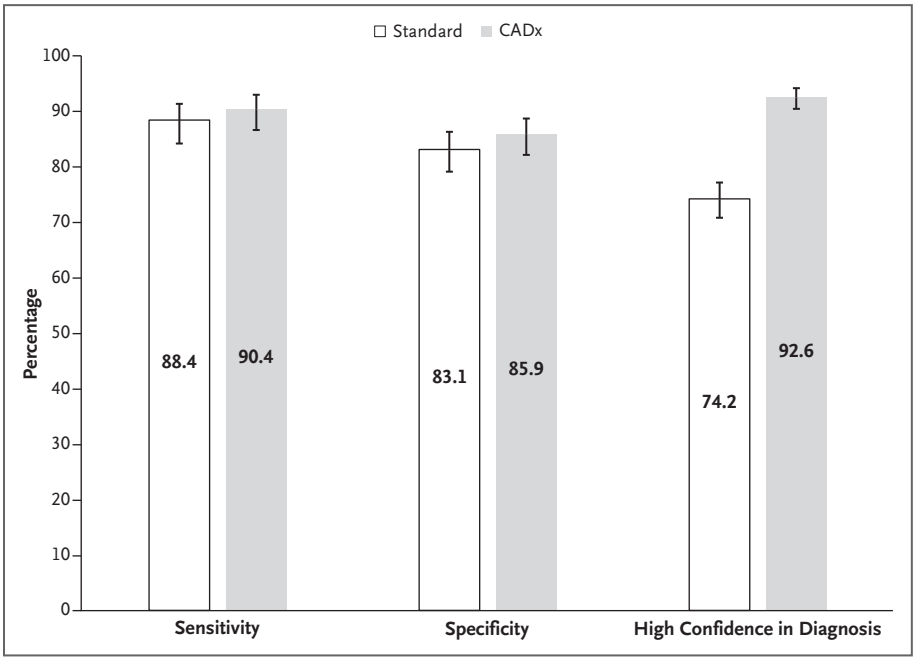

Figure 3. Sensitivity, Specificity, and Confidence of Diagnosis of Standard and AI-Derived CADx Optical Diagnosis of Small Rectosigmoid Polyps during Colonoscopy Compared with Histopathology.

All bars are represented with corresponding 95% confidence intervals. AI denotes artificial intelligence and CADx computer-aided diagnosis.

Table 3. Performance of Standard and AI-Derived CADx Optical Diagnosis of Small Rectosigmoid Polyps during Colonoscopy Compared with Histopathology.*		
Parameter	**Standard Diagnosis**	**CADx Diagnosis**
Sensitivity	88.4 (84.3 to 91.5)	90.4 (86.8 to 93.1)
Specificity	83.1 (79.2 to 86.4)	85.9 (82.3 to 88.8)
Positive predictive value	78.9 (74.3 to 82.9)	82.0 (77.6 to 85.6)
Negative predictive value	91.5 (88.5 to 93.8)	92.8 (90.1 to 94.9)
High confidence in optical diagnosis	74.2 (70.9 to 77.3)	92.6 (90.6 to 94.3)

* Sessile serrated lesions were classified as neoplastic polyps according to the primary analysis plan. Values are presented as percentages (95% confidence intervals). AI denotes artificial intelligence and CADx denotes computer-aided diagnosis.

Our study suggests that the use of CADx helped the provider have higher confidence in optical diagnosis. If this can be replicated, it could contribute to cost reduction because more polyps could be left in situ. Better confidence comes at a cost; CADx assessment prolongs colonoscopy procedure time, which increases health care cost. In previous studies, we demonstrated that the time necessary for CADx assessment of one small polyp, as applied in this study, is about 40 seconds.[6] We consider this additional time well spent with regard to the gain in terms of reduction of unnecessary removal of polyps and histopathologic assessment. Future cost-effectiveness studies may explore whether the prolonged procedure time pays off with the benefit of reduced polypectomies.

In conclusion, real-time assessment with CADx did not significantly increase sensitivity for neoplastic polyps during colonoscopy. There are promising signals for increased specificity and improved confidence of optical diagnosis, but our statistical approach precludes us from making any definitive statements about the identification and removal of small rectosigmoid polyps using the colonoscopy system we employed.

Disclosures

Author disclosures and other supplementary materials are available at evidence.nejm.org.

This study was investigator initiated and funded by public research grants from the Research Council of Norway (Norges Forskningsråd), the Norwegian Cancer Society (Kreftforeningen), and the Japan Society for the Promotion of Science. Cybernet Systems and Olympus loaned the AI software and endoscopes, respectively, to the study at no cost. The companies and funders did not have any influence on and did not play any role in the design, conduct, analysis of the study, or interpretation of the results and were not involved in the writing of the manuscript. Olympus provided a research grant for the study.

Author Affiliations

[1] Clinical Effectiveness Research Group, University of Oslo, Oslo
[2] Clinical Effectiveness Research Group, Department of Transplantation Medicine, Oslo University Hospital, Oslo
[3] Department of Gastroenterology, Hepatology and Clinical Oncology, Centre of Postgraduate Medical Education, Warsaw, Poland
[4] Digestive Disease Center, Showa University Northern Yokohama Hospital, Yokohama, Japan
[5] Department of Medicine, Sørlandet Hospital Kristiansand, Kristiansand, Norway
[6] King's Institute of Therapeutic Endoscopy, King's College Hospital NHS Foundation Trust, London
[7] Graduate School of Informatics, Nagoya University, Nagoya, Japan
[8] Department of Diagnostic Pathology, School of Medicine, Showa University Northern Yokohama Hospital, Kanagawa, Japan
[9] Department of Medicine, Baerum Hospital, Vestre Viken Hospital Trust, Gjettum, Norway

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209-249. DOI: 10.3322/caac.21660.
2. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;366:687-696. DOI: 10.1056/NEJMoa1100370.
3. Abu Dayyeh BK, Thosani N, Konda V, et al. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2015;81:502.e1-502.e16. DOI: 10.1016/j.gie.2014.12.022.
4. Wadhwa V, Alagappan M, Gonzalez A, et al. Physician sentiment toward artificial intelligence (AI) in colonoscopic practice: a survey of US gastroenterologists. Endosc Int Open 2020;8:E1379-E1384. DOI: 10.1055/a-1223-1926.
5. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2011;73:419-422. DOI: 10.1016/j.gie.2011.01.023.

6. Mori Y, Kudo SE, Misawa M, et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. Ann Intern Med 2018;169:357-366. DOI: 10.7326/M18-0249.

7. Misawa M, Kudo SE, Mori Y, et al. Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy. Gastroenterology 2016;150:1531-1532.e3. DOI: 10.1053/j.gastro.2016.04.004.

8. Mori Y, Kudo SE, Chiu PW, et al. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. Endoscopy 2016;48:1110-1118. DOI: 10.1055/s-0042-113609.

9. Lai EJ, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc 2009;69:620-625. DOI: 10.1016/j.gie.2008.05.057.

10. Kudo SE, Misawa M, Mori Y, et al. Artificial intelligence-assisted system improves endoscopic identification of colorectal neoplasms. Clin Gastroenterol Hepatol 2020;18:1874-1881. DOI: 10.1016/j.cgh.2019.09.009.

11. Participants in the Paris Workshop. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon. Gastrointest Endosc 2003;58(Suppl):S3-S4. DOI: 10.1016/S0016-5107(03)02159-X.

12. Hewett DG, Kaltenbach T, Sano Y, et al. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. Gastroenterology 2012;143:599-607.e1. DOI: 10.1053/j.gastro.2012.05.006.

13. IJspeert JEG, Bastiaansen BAJ, van Leerdam ME, et al. Development and validation of the WASP classification system for optical diagnosis of adenomas, hyperplastic polyps and sessile serrated adenomas/polyps. Gut 2016;65:963-970. DOI: 10.1136/gutjnl-2014-308411.

14. Sumimoto K, Tanaka S, Shigita K, et al. Diagnostic performance of Japan NBI Expert Team classification for differentiation among noninvasive, superficially invasive, and deeply invasive colorectal neoplasia. Gastrointest Endosc 2017;86:700-709. DOI: 10.1016/j.gie.2017.02.018.

15. Barua I, Vinsard DG, Jodal HC, et al. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. Endoscopy 2021;53:277-284. DOI: 10.1055/a-1201-7165.

16. Hassan C, Pickhardt PJ, Rex DK. A resect and discard strategy would improve cost-effectiveness of colorectal cancer screening. Clin Gastroenterol Hepatol 2010;8:865-869, 869.e1-869.e3. DOI: 10.1016/j.cgh.2010.05.018.

17. Wieszczy P, Kaminski MF, Løberg M, et al. Estimation of overdiagnosis in colorectal cancer screening with sigmoidoscopy and faecal occult blood testing: comparison of simulation models. BMJ Open 2021;11:e042158. DOI: 10.1136/bmjopen-2020-042158.

18. Rees CJ, Rajasekhar PT, Wilson A, et al. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. Gut 2017;66:887-895. DOI: 10.1136/gutjnl-2015-310584.

19. Ladabaum U, Fioritto A, Mitani A, et al. Real-time optical biopsy of colon polyps with narrow band imaging in community practice does not yet meet key thresholds for clinical decisions. Gastroenterology 2013;144:81-91. DOI: 10.1053/j.gastro.2012.09.054.