Development of an artificial intelligence tool for detecting colorectal lesions in inflammatory bowel disease

Daniela Guerrero Vinsard, MD,[1] Jeffrey R. Fetzer, MS,[1,2] Upasana Agrawal, MBBS,[1,3] Jassimran Singh, MBBS,[1,4] Devanshi N. Damani, MBBS,[5] Priyadharshini Sivasubramaniam, MBBS,[6] Shivaram Poigai Arunachalam, PhD, DBA,[1,2,7] Cadman L. Leggett, MD,[1] Laura E. Raffals, MD,[1] Nayantara Coelho-Prabhu, MBBS[1]

Rochester, Minnesota; Shreveport, Louisiana; Worcester, Massachusetts; El Paso, Texas, USA

Background and Aims: Artificial intelligence has the potential to enhance the endoscopist's ability to detect polypoid and nonpolypoid dysplasia, which could result in decreased rates of colorectal cancer in patients with inflammatory bowel disease (IBD). We aimed to develop the first known model for computer-aided detection (CADe) of colorectal lesions in patients with IBD.

Methods: A CADe model developed at our institution with colorectal lesions from patients without IBD was tested for baseline performance in a dataset of high-definition white-light endoscopy (HDWLE) images of IBD-associated colorectal lesions. Subsequently, we retrained the original CADe to build an IBD-CADe model using 1266 HDWLE still images and 426 dye-based chromoendoscopy still images depicting histologically proven IBD-associated colorectal lesions. The lesions were annotated by histopathology, size, morphology, and inflammation score surrounding the lesion. We evaluated the model's performance metrics before and after retraining.

Results: The sensitivity of the original CADe was 50%, positive predictive value (PPV) was 97%, and accuracy was 64%. With the retrained IBD-CADe model, performance metrics for detecting lesions on HDWLE were as follows: sensitivity, 95.1%; specificity, 98.8%; PPV, 98.9%; negative predictive value, 94.7%; accuracy, 96.8%; and area under the curve, .85. IBD-CADe for chromoendoscopy images showed a sensitivity of 67.4%, specificity of 88.0%, PPV of 83.3%, negative predictive value of 74.3%, accuracy of 77.8%, and area under the curve of .65. Subgroup analysis showed a 93% sensitivity for detecting lesions 5 mm or smaller, 91% for lesions 6 to 10 mm, and 85% for those larger than 10 mm. IBD-CADe performed best for Paris classification types Ip, Is, and IIa. Of 9 lesions missed by IBD-CADe, most had a Mayo endoscopic subscore of 0 or 1.

Conclusions: This model is the first step toward developing other artificial intelligence–based endoscopic tools to enhance dysplasia detection for patients with IBD. (iGIE 2023;2:91-101.)

Long-standing diagnoses of ulcerative colitis or Crohn's disease carry increased risk of colorectal cancer (CRC). CRC represents 10% to 15% of all-cause mortality in persons with inflammatory bowel disease (IBD).[1] CRC risk increases with disease duration, with a 1% risk 10 years after IBD diagnosis, 4% at 20 years, and 14% at 30 years.[2-5] The American Gastroenterological Association in its clinical practice update on endoscopic surveillance and management of colorectal dysplasia in IBD recommended colonoscopy to screen for dysplasia 8 to 10 years after the initial diagnosis and whenever concomitant primary sclerosing cholangitis is diagnosed.[6] Multiple studies showed that surveillance colonoscopy enables early detection of premalignant and malignant lesions, thereby improving cancer-related survival.[7-9] Colonoscopy, however, can be limited by human error, and instances of missed colorectal lesions are well documented.[10] Colorectal lesions are more likely to be missed in patients with IBD because background mucosal inflammation can make lesions, especially flat ones, harder to identify.[11,12]

IBD-associated lesions can have various morphologic and histopathologic findings. *Polypoid* lesions are defined as lesions protruding more than 2.5 mm from the mucosa into the lumen. They can be characterized as "pedunculated" when attached to the mucosa by a stalk or as "sessile" when the entire base is contiguous with the mucosa. *Nonpolypoid* lesions are classified as minimally elevated, flat, or depressed.[13] Pseudopolyps in the setting of inflammation may resemble polypoid dysplastic lesions, prompting biopsy sampling or resection. Furthermore, adenomatous lesions are difficult to recognize when they are located within a cluster of pseudopolyps. Serrated

RESEARCH INTO CONTEXT

Evidence before this study

Artificial intelligence (AI) has been widely developed and validated for adenoma detection during screening colonoscopy, but colorectal lesions from patients with inflammatory bowel disease (IBD) have been excluded from training and testing datasets.

Added value of this study

We developed the first known AI system capable of detecting polypoid and nonpolypoid colorectal lesions on images from patients with IBD using high-definition white-light endoscopy (HDWLE) and dye-based chromoendoscopy. Our IBD-specific computer-aided detection (IBD-CADe) system had an area under the curve of .85 for HDWLE images, detecting lesions even localized in a background of moderately inflamed mucosa. The performance of IBD-CADe was lower for chromoendoscopy images, with an area under the curve of .65.

Implications of all available evidence

IBD-CADe is expected to serve as a tool to augment dysplasia detection through targeted biopsy sampling and reduce the need for random biopsy sampling during surveillance colonoscopy of patients with Crohn's disease and ulcerative colitis.

epithelial changes (SECs), recognized endoscopically and histologically, are unique to IBD. SECs are mostly flat lesions that histologically resemble hyperplastic-like mucosal changes and are being discovered more frequently with advanced imaging technologies. Their clinical significance is not clear, with some study authors reporting a higher risk of dysplasia in patients with SECs and other study findings showing no clinical significance.[14,15]

Artificial intelligence (AI), acting as a second observer for adenoma detection during screening and surveillance colonoscopy, has been validated in patients without IBD.[16-23] These AI algorithms have yet to be trained or validated in patients with IBD. To our knowledge, successful use of AI to identify dysplasia in patients with IBD is limited to a single case report of a system trained with conventional non-IBD colorectal lesions.[24] Developing AI algorithms to detect and classify IBD-associated dysplasia is challenging because of the variety of lesions encountered in these patients and because of background mucosal inflammation. These image characteristics must be accounted for in study design and methodology to achieve an acceptable performance for the AI tool. The primary aim of our study was to train and test a computer-aided detection (CADe) model specifically for automated detection of IBD-associated polypoid and nonpolypoid colorectal lesions with all degrees of background mucosal inflammation.

METHODS

This study was approved by the Mayo Clinic Institutional Review Board (#21-000968). All patient images were deidentified.

Patient population

Patients were included in the study if they were 18 years or older, had IBD (Crohn's disease, ulcerative colitis, or indeterminate colitis), underwent surveillance colonoscopy at our center from January 2016 through June 2021, and had a macroscopically visible lesion during surveillance colonoscopy with corresponding pathology. Data collection started in 2016, when our center implemented high-definition white-light endoscopy (HDWLE).

In total, 3437 surveillance colonoscopies were performed for patients with IBD at our tertiary care center during the study period. One or more colorectal lesions with corresponding pathology were identified from 728 patients during 977 of the 3437 procedures (28.4%), totaling 1692 endoscopic images (1266 HDWLE and 426 chromoendoscopy images).

Endoscopy and exclusion criteria for images

Colonoscopy was performed using high-definition colonoscopes (Olympus Medical Systems, Tokyo, Japan). Chromoendoscopy was performed after the colon was sprayed with .2% FD&C (Food, Drug, and Cosmetic Act) blue 2 solution (Professional Compounding Centers of America), which is chemically similar to indigo carmine.

Endoscopic images were excluded from the image datasets if they were blurry or dark; had an obscured surface with stool, blood, mucus, glare, or endoscopic instruments in the region of interest (ROI); or had markers over the ROI that could not be removed (eg, arrows, circles, or words). Additionally, images of lesions found in the terminal ileum were excluded. Also, images containing more than 1 polyp were excluded unless the pathology report clearly described each lesion separately or described pathology that was the same for all visible lesions. Finally, images were excluded if they contained 2 or more lesions that were biopsy sampled or resected and the specimens were placed in the same bottle with other lesions without clear differentiation in pathology.

Data extraction and annotation

Eight physicians collected characteristics about patients, procedures, polyps, and endoscopic images (Table 1).

TABLE 1. Baseline characteristics of patients, procedures, and polyps

Characteristics	Values
Sex (n = 721)	
Men	436 (60.5)
Women	285 (39.5)
Median age, y (interquartile range)	
At colonoscopy	57 (43-66)
At IBD diagnosis	37 (26-52)
Type of IBD (n = 714)	
Ulcerative colitis	471 (66.0)
Crohn's disease	237 (33.2)
Indeterminate disease	6 (.8)
Bowel preparation quality (n = 964)	
Good	655 (67.9)
Fair	259 (26.9)
Poor	50 (5.2)
Chromoendoscopy (n = 970)	
Yes	326 (33.6)
No	644 (66.4)
Polyp size, mm (n = 1574)	
1-5	1080 (68.6)
6-9	306 (19.4)
10-19	138 (8.8)
≥20	50 (3.2)
Polyp location (n = 1589)	
Cecum/ascending colon	399 (25.1)
Hepatic flexure	74 (4.7)
Transverse colon	282 (17.7)
Splenic flexure	29 (1.8)
Descending colon	234 (14.7)
Sigmoid and rectosigmoid colon	325 (20.5)
Rectum	204 (12.8)
Multiple sites	42 (2.6)
Polyp histopathology (n = 1572)	
Tubular adenoma	356 (22.6)
Inflammatory pseudopolyp	345 (21.9)
Hyperplastic polyp	366 (23.3)
Pseudopolyp without inflammation	182 (11.6)
Sessile serrated adenoma	172 (10.9)
Serrated epithelial change	122 (7.8)
Tubulovillous adenoma	12 (.8)
Adenocarcinoma	11 (.7)
Other (lipoma, perineurioma)	6 (.4)

Values are n (%) unless otherwise defined. Data from some patients were missing or unknown for some characteristics.
IBD, Inflammatory bowel disease.

Images were annotated and classified into 5 groups by using the pathology as ground truth: dysplastic lesions (tubular adenomas, tubulovillous adenomas, and adenocarcinomas), nondysplastic lesions (hyperplastic polyps), SECs, sessile serrated adenomas (SSAs), and pseudopolyps.

HDWLE and chromoendoscopy images were separated. All 1692 images were resized from their native 1920 × 1080 pixels to 1350 × 1080 pixels and hand labeled with bounding boxes by 6 physicians. When inputted to the model, these images were resized to 256 × 256 pixels. The labeling accuracy was confirmed by the study's first author (D.G.V.) and senior author (N.C.-P.). Visual Geometry Group Image Annotator (University of Oxford, Oxford, UK), an open-source image annotator software based on HTML (hypertext markup language), JavaScript, and CSS (cascading style sheets), was used to annotate the ROI—the visible lesion—with bounding boxes. Each image was also annotated by the endoscopist-reported polyp size (≥5 mm, 6-10 mm, and >10 mm), polyp morphology according to the Paris classification, and, for HDWLE images only, the degree of background mucosal inflammation by using the Mayo endoscopic subscore (MES). The MES ranges from 0, indicating normal mucosa or inactive disease, to 3, indicating severe activity.

Algorithm development

Pretraining phase for original CADe. We first built a CADe model for patients without IBD. This original CADe model was pretrained on 8000 endoscopic images using CSPDarkNet53 (DarkNet with cross-stage partial networks) architecture from a Mayo Clinic database containing HDWLE images. Subsequently, the model was trained with 2550 annotated HDWLE images of colon polyps from patients without IBD and tested on 284 non-IBD polyp images (Fig. 1).

Testing the original CADe tool in the IBD lesion dataset. The original CADe system (trained with non-IBD polyps only) was tested in our entire dataset of 2016 unlabeled IBD lesions from 1266 HDWLE images. The lesions were classified into 5 categories by their pathology: dysplastic lesions, nondysplastic lesions, pseudopolyps, SECs, and SSAs. We measured the sensitivity (recall), positive predictive value (PPV; precision), false-positive rate, and accuracy (F1 score) of the original CADe in our IBD image dataset.

IBD-CADe training, validation, and testing. A training dataset comprising 80% of the pool of hand-labeled IBD lesions was used to retrain the CADe model, another 10% of the dataset was used for validation (iterations), and the remaining 10% was used for testing. The model used to build the algorithm for detecting IBD lesions was the Scaled-YOLOv4: Scaling Cross Stage Partial Network with the pretrained weights from CSPDarkNet53. This model is redesigned from YOLOv4 to achieve a more optimal processing speed versus accuracy trade-off. A 5-fold cross-validation was performed to minimize overfitting. The IBD-CADe algorithm (Fig. 2) was trained and tested separately in both HDWLE and chromoendoscopy images (Supplementary Table 1, available online at www.igiejournal.org).

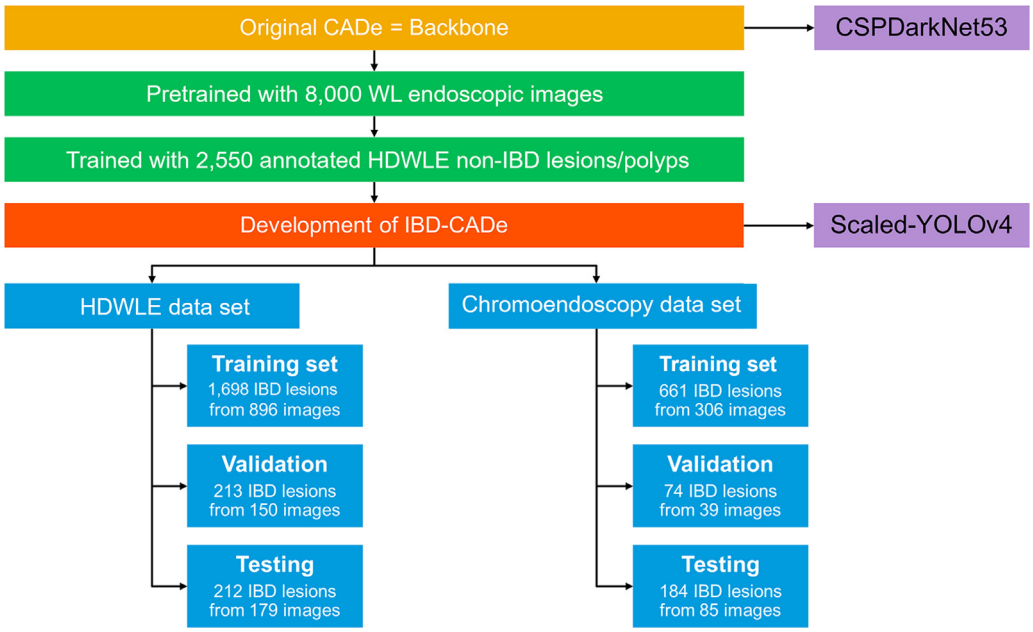


Figure 1. Study flowchart for development of IBD-CADe. *CADe*, Computer-aided detection; *HDWLE*, high-definition white-light endoscopy; *IBD*, inflammatory bowel disease; *WL*, white light.

Image datasets. The HDWLE training image dataset comprised 896 labeled HDWLE images: 204 dysplastic lesions, 154 nondysplastic lesions, 388 pseudopolyps, 47 SECs, and 103 SSAs. The validation dataset was composed of 191 images: 29 dysplastic lesions, 22 nondysplastic lesions, 120 pseudopolyps, 6 SECs, and 14 SSAs. The HDWLE testing dataset included 179 images: 30 dysplastic lesions, 21 nondysplastic lesions, 108 pseudopolyps, 7 SECs, and 13 SSAs. Also included in the testing dataset were 185 nonpolyp IBD-associated HDWLE images for analysis of true negatives and IBD-CADe specificity and area under the curve.

The dye-based chromoendoscopy dataset comprised 426 chromoendoscopy-labeled images: 106 dysplastic lesions, 114 nondysplastic lesions, 111 pseudopolyps, 42 SECs, and 53 SSAs. Of these images, approximately 80% were used for training and 20% for validation and testing. A set of 95 nonpolyp chromoendoscopy images was included for analysis of true negatives for IBD-CADe performance.

HDWLE video pilot dataset. Ten short HDWLE video clips from patients with IBD, showing histologically proven polypoid lesions, were used to test and validate the IBD-CADe system in videos. The videos were randomly selected from our video library. The IBD-CADe output was analyzed at 30 frames per second. The degree of mucosal inflammation immediately adjacent to all polypoid lesions in the endoscopic videos was graded by 2 authors (D.G.V. and N.C.-P.) using the MES. Polyp size, morphology, and histopathology also were abstracted.

Performance metrics and statistical analysis

We measured sensitivity, PPV, false-positive rate, accuracy, F1 score, and processing speed before and after retraining the original CADe model with visible IBD lesions for HDWLE images. We hypothesized that the performance of the original CADe system for detecting IBD lesions would increase after retraining the algorithm.

The CADe model performance was measured using intersection over union, a popular evaluation metric used in the object detection benchmarks.[25] An intersection-over-union threshold of .25 was used in this study. The performance of IBD-CADe with HDWLE and chromoendoscopy images was reported with standard performance metrics (Supplementary Table 2, available online at www.igiejournal.org) and receiver operating curves for each endoscopic modality. The trained IBD-CADe highlighted the ROI—the IBD polypoid lesion—with rectangular bounding boxes (convolutional neural network boxes), with a confidence threshold specified at .25. (The system outputted a bounding box only when confidence of the presence of a polyp in the ROI exceeded 25%.) The confidence of the ROI containing a polyp was also depicted above the bounding box.

We conducted 3 subgroup analyses. First, inflammation as a variable affecting the model's performance was addressed as a secondary endpoint with the use of HDWLE images. We analyzed whether images with false-negative or false-positive results had a higher degree of inflammation that could affect IBD-CADe's performance. Second, flat morphology

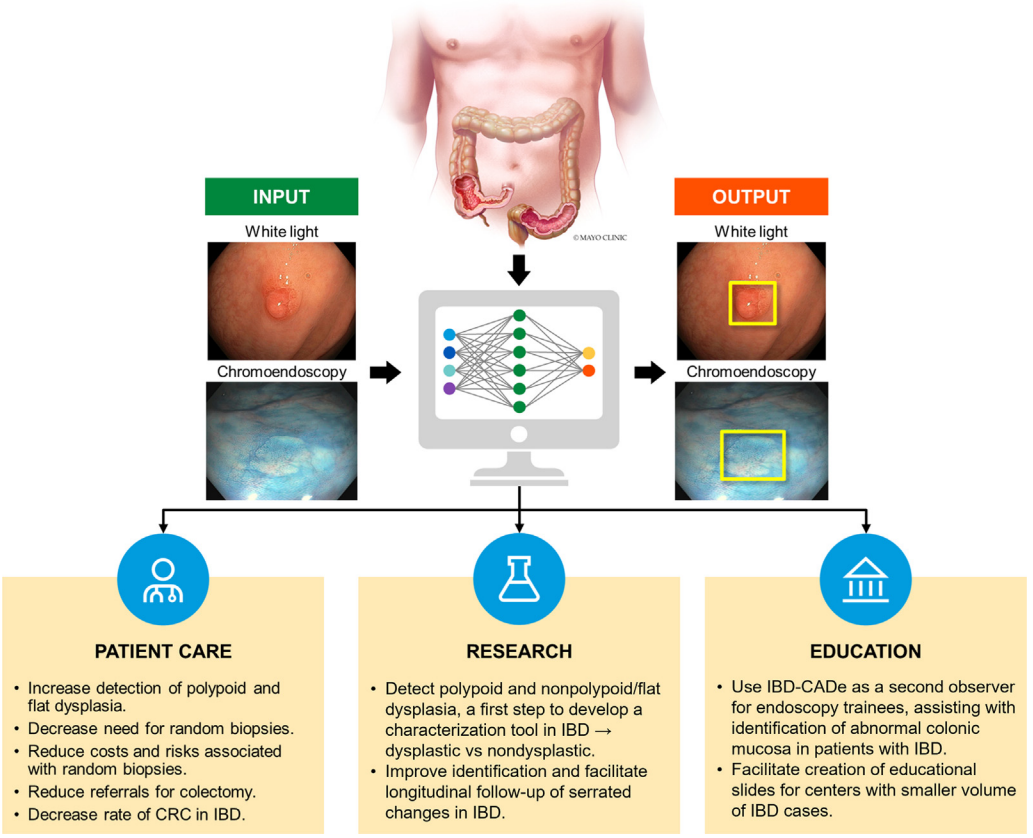


Figure 2. Artificial intelligence for detecting colorectal polypoid and nonpolypoid lesions in IBD. *CADe*, Computer-aided detection; *CRC*, colorectal cancer; *IBD*, inflammatory bowel disease.

affecting performance was addressed as a secondary endpoint with the use of HDWLE and chromoendoscopy images. We analyzed whether flat lesions were more likely to result in false negatives. Third, polyp size as a variable affecting IBD-CADe's ability for detection was addressed in a subgroup analysis.

RESULTS

Performance of the original CADe

The original CADe algorithm was tested for detecting unlabeled IBD-associated colorectal lesions. The algorithm performed best for detecting dysplastic polyps and worst for pseudopolyps and SECs, likely because pseudopolyps and SECs are unique to patients with IBD and because the original CADe system was initially trained only with conventional non-IBD polyps. For detecting IBD lesions, the original CADe had an overall sensitivity of .50, PPV of .97, false-positive rate of 1.7%, and F1 score of .64. Performance metrics of the original CADe model are shown in Table 2 according to histopathology of the IBD lesions.

Performance of IBD-CADe after retraining

After the original CADe system was retrained with labeled IBD lesions, the performance of IBD-CADe was tested in both image datasets and a video dataset.

HDWLE images. Performance metrics for detecting lesions in the HDWLE testing set (n = 179) were as follows: sensitivity, 95.1%; specificity, 98.8%; PPV, 98.9%; negative predictive value, 94.7%; F1 score, 96.4%; and accuracy, 96.8% (Supplementary Table 3, available online at www.igiejournal.org). IBD-CADe had an area under the curve of .85 for HDWLE lesions (Fig. 3). Table 2 shows performance metrics of IBD-CADe versus the original CADe. Results of the subgroup analysis of the performance of IBD-CADe in the HDWLE image dataset appear in Table 3.

Histopathology. Nondysplastic and dysplastic polyps were detected at comparable rates, with 91% true positives for nondysplastic polyps and 90% true positives for dysplastic polyps. Detection of SECs and SSAs was slightly lower at 85.7% true positives in both groups. Of 107 images of pseudopolyps, 21 images had a single pseudopolyp, and all single pseudopolyps were accurately detected by IBD-CADe.

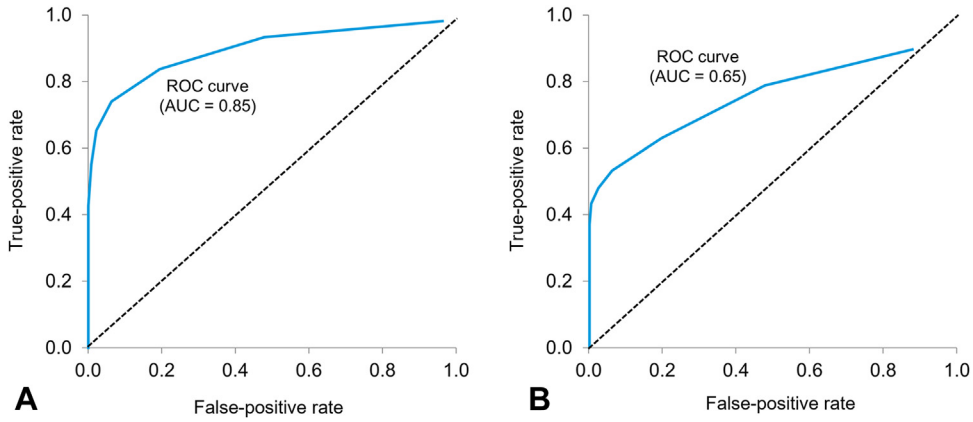


Figure 3. ROC curves for IBD-CADe performance. **A,** High-definition white-light endoscopy. **B,** Dye-based chromoendoscopy. *AUC*, Area under the curve; *IBD-CADe*, inflammatory bowel disease computer-aided detection; *ROC*, receiver-operating characteristic.

TABLE 2. Performance of original CADe vs IBD-CADe for detecting polypoid lesions in patients with IBD

	Original CADe model								IBD-CADe model							
Histopathology of IBD lesion*	Total no. of IBD images†	TP	FN	FP	Sensitivity	PPV	FPR (%)	F1 score	Total no. of test set images†	TP	FN	FP	Sensitivity	PPV	FPR (%)	F1 score
Dysplastic	254	208	46	8	.82	.96	3.1	.89	30	27	3	0	.90	1.0	0	.95
Nondysplastic	197	150	47	8	.76	.95	4.1	.85	21	19	2	1	.90	.95	4.7	.93
Pseudopolyps	1380	468	912	10	.34	.98	.7	.50	108	107	1	8	.99	.93	7.4	.96
Serrated epithelial change	57	35	22	4	.61	.90	7.0	.73	7	6	1	0	.85	1.0	0	.92
Sessile serrated adenoma	128	101	27	5	.79	.95	3.9	.86	13	11	2	0	.84	1.0	0	.92
Total	2016	962	1054	35	.50	.97	1.7	.64	179	170	9	9	.95	.95	5.0	.95

CADe, Computer-aided detection; *FN*, false negative; *FP*, false positive; *FPR*, false-positive rate; *IBD*, inflammatory bowel disease; *PPV*, positive predictive value; *TP*, true positive.
*Dysplastic includes tubular adenomas, polypoid mucosa with dysplasia of any grade, tubulovillous adenomas, and adenocarcinomas. Nondysplastic includes hyperplastic polyps and polypoid lymphoid aggregates. Pseudopolyps include inflammatory and noninflammatory pseudopolyps; in this group, the performance of IBD-CADe is reported per image because some pseudopolyp images contained >30 pseudopolyps.
†Some images had >1 lesion.

Eighty-six images had multiple pseudopolyps within the same image, and 43 of 86 had all pseudopolyps in the same image detected with bounding boxes. A false negative was considered when no bounding boxes appeared on any pseudopolyp within an image. Polyp detection was estimated in a per-polyp fashion for all types of polyps except for pseudopolyps, which were estimated in a per-image fashion, given that several images contained too many pseudopolyps to count manually (Fig. 4A).

Lesion size. IBD-CADe had comparable performance for lesions ≤5 mm and those measuring 6 to 10 mm (93% vs 91%, respectively) but lower performance for lesions larger than 10 mm (85%).

Morphology. Flat lesions (Paris IIb) and lesions with mixed morphology were most frequently missed by IBD-CADe. The algorithm missed 3 of 9 lesions classified as Paris IIb and 1 of 8 lesions with mixed Paris classification. IBD-CADe performed best with Paris Ip, Is, and IIa lesions (Fig. 4B).

Inflammation. The HDWLE test dataset included 41 lesions with an MES of 0, 133 MES 1 lesions, 23 MES 2 lesions, and 23 MES 3 lesions. For the 9 lesions missed by IBD-CADe in the HDWLE test set, 3 of 41 (7.3%) had a corresponding MES of 0, 2 of 133 (1.5%) had a score of 1, 2 of 23 (8.7%) had a score of 2, and 2 of 23 (8.7%) had a score of 3, with more lesions missed at higher scores of inflammation (Table 4 and Fig. 5). Individual characteristics of dysplastic polyps included in the HDWLE dataset are described in Supplementary Table 4 (available online at www.igiejournal.org).

Dye-based chromoendoscopy images. The algorithm did not perform as well for detecting lesions with chromoendoscopy images as with HDWLE images, likely because of the robust pretraining backbone of IBD-CADe

TABLE 3. Subgroup analysis for true positives and false negatives for polyp histopathology, morphology, size, and Mayo endoscopic subscore

Polyp characteristics	Total no. of images	True positive	False negative
High-definition white-light endoscopy (n = 179)			
Size of single polyps,* mm			
≤5	41	38 (92.7)	3 (7.3)
6-10	33	30 (90.9)	3 (9.1)
>10	20	17 (85.0)	3 (15.0)
Morphologic (Paris) classification			
Ip	71	71 (100.0)	0 (0)
Is	52	49 (94.2)	3 (5.8)
IIa	35	32 (91.4)	3 (8.6)
IIb	12	9 (75.0)	3 (25.0)
IIc	0	0 (0)	0 (0)
III	0	0 (0)	0 (0)
Mixed	9	8 (88.9)	1 (11.1)
Mayo endoscopic subscore			
0	19	14 (73.7)	5 (26.3)
1	117	117 (100.0)	0 (0)
2	37	35 (94.6)	2 (5.4)
3	6	4 (66.7)	2 (33.3)
Histopathology			
Dysplastic	30	27 (90.0)	3 (10.0)
Nondysplastic	21	19 (90.5)	2 (9.5)
Pseudopolyp†	107	106 (99.1)	1 (.9)
Serrated epithelial change	7	6 (85.7)	1 (14.3)
Sessile serrated adenoma	14	12 (85.7)	2 (14.3)
Chromoendoscopy (n = 89)			
Size,* mm			
<5	39	26 (66.7)	13 (33.3)
6-10	22	13 (59.1)	9 (40.9)
>10	11	8 (72.7)	3 (27.3)
Morphologic (Paris) classification			
Ip	14	10 (71.4)	4 (28.6)
Is	26	22 (84.6)	4 (15.4)
IIa	23	17 (73.9)	6 (26.1)
IIb	22	9 (40.9)	13 (59.1)
IIc	1	0 (0)	1 (100.0)
III	0	0 (0)	0 (0)
Mixed	3	2 (66.7)	1 (33.3)
Histopathology			
Dysplastic	22	15 (68.2)	7 (31.8)
Nondysplastic	20	14 (70.0)	6 (30.0)
Pseudopolyp‡	23	16 (69.6)	7 (30.4)
Serrated epithelial change	13	11 (84.6)	2 (15.4)
Sessile serrated adenoma	11	4 (36.4)	7 (63.6)

Values are n (%).

*Excludes images with multiple pseudopolyps.

†Pseudopolyps on high-definition white-light endoscopy: Of 107 images, 21 contained a single pseudopolyp and 86 contained many pseudopolyps; 43 of 86 had all the multiple pseudopolyps detected within 1 image.

‡Pseudopolyps on chromoendoscopy: 6 of 23 images had a single pseudopolyp, and 3 of 17 had all pseudopolyps detected when >1 pseudopolyp was present in the image.

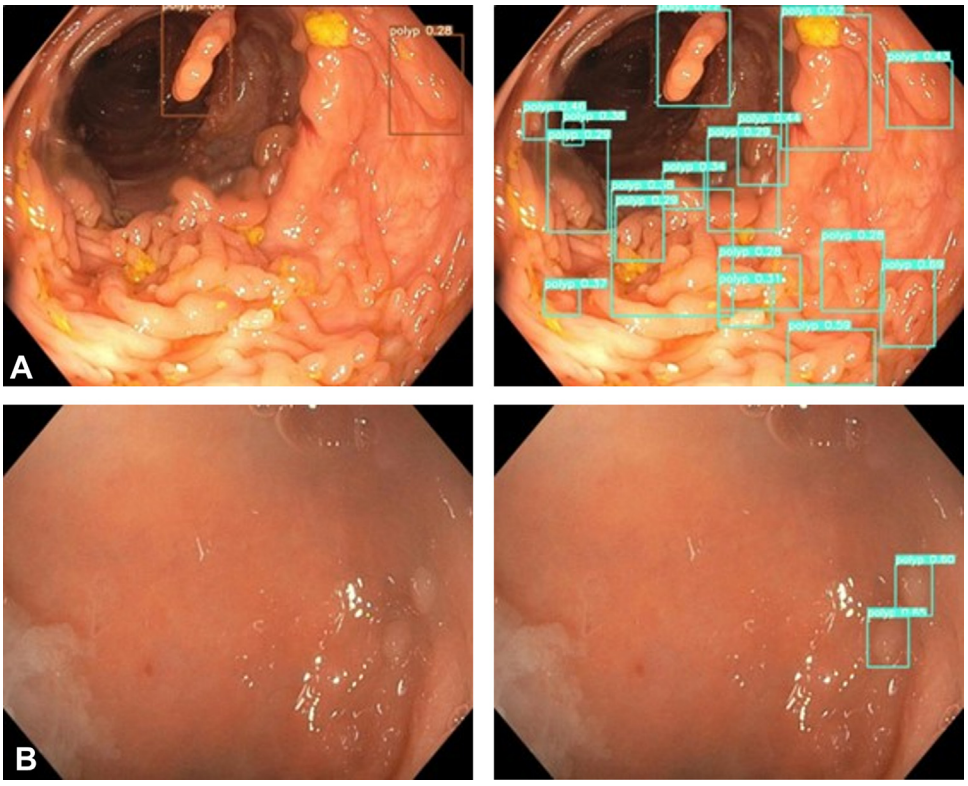


Figure 4. Examples of colorectal lesions. **A,** Pseudopolyps (*top left* and *right*). **B,** Serrated epithelial changes (*bottom left* and *right*) that were missed before and detected after retraining the model with inflammatory bowel disease lesions. The value above each bounding box indicates confidence of the region of interest containing a polyp.

TABLE 4. Mayo endoscopic subscore for the testing image set of IBD polypoid lesions and their corresponding FNs

Pathology of IBD lesion	No. of FNs/Mayo 0 images	No. of FNs/Mayo 1 images	No. of FNs/Mayo 2 images	No. of FNs/Mayo 3 images	Total FNs per polyp type*
Dysplastic	0/1	0/15	1/9	2/5	3/29 (10.3)
Nondysplastic	2/9	0/12	0/0	0/0	2/21 (9.5)
Pseudopolyps	0/27	0/49	1/14	0/18	1/108 (1)
Serrated changes	0/0	1/7	0/0	0/0	1/7 (14.3)
Sessile serrated adenomas	1/4	1/9	0/0	0/0	2/13 (15.4)
Total FNs per Mayo endoscopy score*	3/41 (7.3)	2/133 (1.5)	2/23 (8.7)	2/23 (8.7)	9 missed lesions

IBD, Inflammatory bowel disease; *FN*, false negative.
*Values are n/N (%).

with HDWLE colonic images (Supplementary Fig. 1, available online at www.igiejournal.org). The overall performance of IBD-CADe for chromoendoscopy images was consistent with a sensitivity of 67.4%, specificity of 88.0%, PPV of 83.3%, negative predictive value of 74.3%, F1 score of 73.3%, and accuracy of 77.8%, with an area under the curve of .65 (Fig. 3). Table 3 shows results of the subgroup analysis of the performance of IBD-CADe in the dye-based chromoendoscopy image dataset.

Histopathology. The highest sensitivity was found in SECs (85%) and the lowest in SSAs (36%). Detection was comparable for nondysplastic lesions, dysplastic lesions, and pseudopolyps (70%, 65%, and 70%, respectively). Pseudopolyps were graded in a per-image fashion and not a per-polyp fashion. Six of 23 images contained a single pseudopolyp, all detected with bounding boxes. Only 3 of 17 images with multiple pseudopolyps had all pseudopolyps detected with bounding boxes.

Lesion size. IBD-CADe in chromoendoscopy images detected 73% of lesions larger than 10 mm, 67% of lesions ≤5 mm, and 59% of lesions between 6 and 10 mm.

Morphology. Similar to HDWLE images, the most challenging morphologic classification for detecting lesions

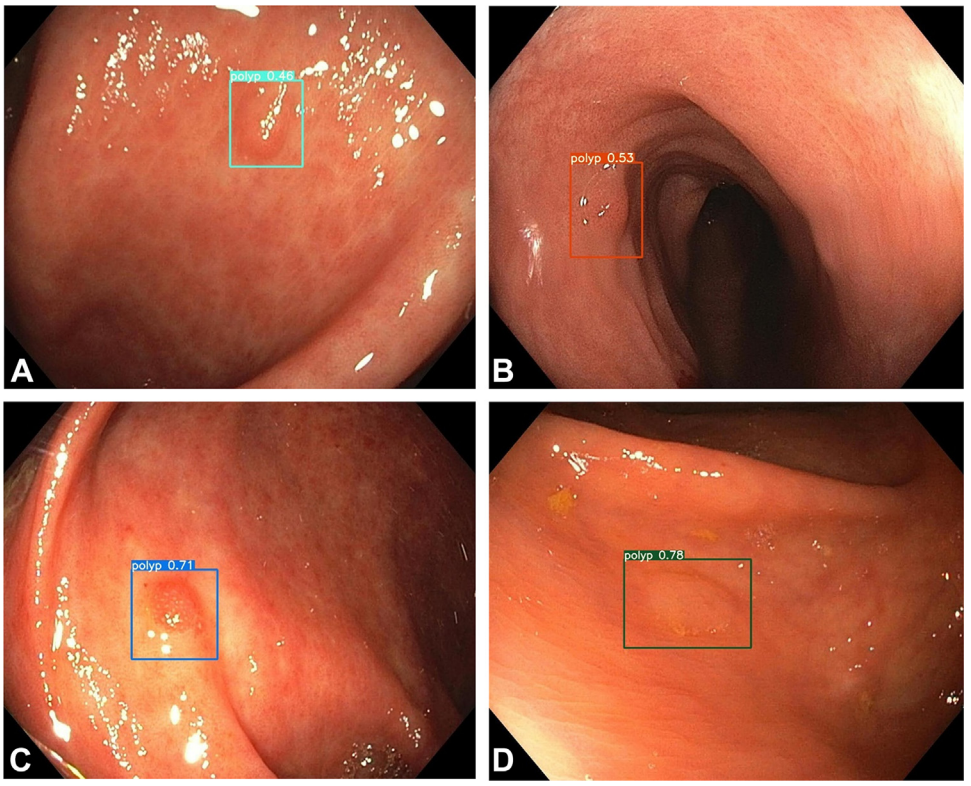


Figure 5. Examples of lesions with a background of inflammation detected by the inflammatory bowel disease computer-aided detection system. **A,** pseudopolyp; **B,** tubular adenoma; **C,** sessile serrated adenoma; **D,** serrated epithelial change.

with chromoendoscopy was Paris IIb (flat; Supplementary Fig. 2, available online at www.igiejournal.org), with 9 of 22 lesions detected and 13 missed. IBD-CADe detected 85% of Paris Is sessile lesions and 74% of IIa minimally elevated lesions, comparable with Ip lesions and lesions with mixed morphology. Individual characteristics of dysplastic polyps included in the dye-based chromoendoscopy test set are described in Supplementary Table 4.

HDWLE videos. The length of the 10 video clips showing IBD lesions varied between 5 and 13 seconds with 30 frames per second. IBD-CADe was tested in every video frame. True positives, false negatives, and false positives were reported, and each case was graded according to MES, Paris morphologic classification, and polyp size.

Case 9 containing a 4-mm, Paris Is tubular adenoma achieved the highest performance, with a bounding box correctly placed in 100% of the video frames and no false negatives or false positives. Case 8 containing a 2-mm, Paris IIa tubular adenoma had the lowest performance, with 42% true-positive frames and 58% false-negative frames, as well as false positives for biopsy forceps and colonic mucosa (Supplementary Fig. 3, available online at www.igiejournal.org). For this case, the colonoscope's camera lenses had intermittent glare, which limited the algorithm's performance, likely affecting the model's performance. The performance of IBD-CADe for each additional video is summarized in Supplementary Table 5 (available online at www.igiejournal.org).

DISCUSSION

CADe of colorectal lesions has been widely explored for patients without IBD. Study findings have shown that AI has a large effect on decreasing adenoma miss rates and increasing rates of adenoma detection per colonoscopy.[16,20,26] Also, AI has performed better than standard colonoscopy for detecting diminutive and nonpolypoid colorectal lesions.[23] If implemented in screening colonoscopy in the United States, AI annually could help prevent an additional 7194 CRC cases and 2089 related deaths and save $290 million.[27]

Detection of IBD-associated dysplasia enables real-time therapeutic decisions, accurate documentation, and appropriate surveillance recommendations. Improved detection of previously "invisible," flat, dysplastic lesions using HDWLE and chromoendoscopy—both virtual and dye-based—has resulted in better patient outcomes, particularly lower colectomy rates when resectable dysplasia is found.[4,6,28,29] To our knowledge, this is the first cohort study showing the efficacy of an AI algorithm (IBD-CADe) for detecting visible colorectal lesions in patients with IBD. We trained, tested, and validated this model

using both HDWLE and dye-based chromoendoscopy. Our proposed IBD-CADe system had high accuracy in detecting colorectal lesions of various morphologic classifications, sizes, and histopathology from patients with IBD. This technology could enhance the endoscopist's ability to detect polypoid and nonpolypoid dysplastic mucosa in patients with IBD and could increase the number of targeted biopsy samplings performed during surveillance colonoscopy.

Our study has several strengths. We used a robust variety of histologically confirmed polypoid and nonpolypoid lesions from patients with IBD to train the model. We graded the level of inflammation in the background of the lesion and assessed the effect of inflammation on the algorithm's performance, showing that IBD-CADe can still allow lesions with more than 90% true positives to be detected in a background of mild to moderate inflammation. Additionally, we examined the effect of lesion size and morphologic classification and found that IBD-CADe had a true-positive rate above 90% for lesions 5 mm or smaller and for Paris IIa (slightly elevated) lesions with use of HDWLE. A potential explanation for why larger lesions (>10 mm) had a lower sensitivity is that these lesions in IBD often are pseudopolyps with long stalks or have a mixed morphology and overlying mucus. As the algorithm is further defined, we recognize the importance of improving this outcome. Another strength of this study was that both SSAs and SECs were included in our training and testing datasets. To our knowledge, this is the first CADe algorithm capable of detecting SECs. Serrated lesions often are subtle and flat and can be covered by a mucous cap, making them easier to miss by the human eye.[30] IBD-CADe helped detect SECs and SSAs with a lower true-positive rate than for other dysplastic and nondysplastic lesions (85.7% vs >90%, respectively). As the model is further trained, especially on videos, we expect this accuracy to continue to improve. Use of HDWLE video clips to test IBD-CADe is another strength of this study.

IBD-CADe had significantly superior performance when we used HDWLE images versus chromoendoscopy images, likely because the backbone of IBD-CADe was pretrained with many white-light images and fewer chromoendoscopy images. As we obtain more data using dye-based and virtual chromoendoscopy, this model can be further trained to detect lesions with these modalities. Also, HDWLE has been shown to be noninferior to chromoendoscopy for lesion detection. However, advanced imaging techniques can help differentiate dysplastic versus nondysplastic lesions. Hence, we propose the use of these modalities for future work on characterizing colorectal lesions in patients with IBD.

We acknowledge several study limitations. First, this retrospective study used still images from a single large academic center for algorithm development. Therefore, prospective internal validation and external validation of the performance of this algorithm are warranted. Second, lesions located in severely inflamed mucosa and lesions with Paris classification IIb were more frequently missed by IBD-CADe, possibly because the training dataset had a limited number of images with these characteristics. Larger numbers of images and video frames containing these important characteristics are needed to refine this algorithm. Third, pseudopolyps were a challenging group of lesions when the algorithm's performance was assessed, particularly when many pseudopolyps were located in the same colonic segment. However, the clinical utility of pseudopolyp detection by an AI system is trivial because they are easily detectable by endoscopists. We believe the true benefit of AI assistance is characterizing and therefore differentiating an adenoma located in a cluster of pseudopolyps. This will require larger datasets. Last, several false-positive frames were detected when we tested IBD-CADe in videos, mainly related to glare, snares, biopsy forceps, bubbles, and prominent folds. Therefore, the algorithm should be refined with annotated endoscopic images containing these characteristics. Our next step is to refine this model with more video data. Also, we plan to develop a lesion characterization model that will differentiate dysplastic from nondysplastic lesions.

In conclusion, we developed and evaluated the first known IBD-CADe system capable of detecting polypoid and nonpolypoid colorectal lesions in patients with IBD, particularly on HDWLE images. IBD-CADe may enhance detection of visible dysplasia in patients with IBD.

ACKNOWLEDGMENTS

Kathleen Louden, ELS, senior scientific/medical editor, Mayo Clinic, substantively edited the manuscript. The Scientific Publications staff, Mayo Clinic, provided proofreading, administrative, and clerical support. The data that support the findings of this study are available from the corresponding author on reasonable request (coelho prabhu.nayantara@mayo.edu).

DISCLOSURE

The following authors disclosed financial relationships: C. L. Leggett: Consultant for Verily Life Sciences. L. E. Raffals: Advisory board member for Janssen Pharmaceuticals. All other authors disclosed no financial relationships.

Abbreviations: AI, artificial intelligence; CADe, computer-aided detection; CRC, colorectal cancer; HDWLE, high-definition white-light endoscopy; IBD, inflammatory bowel disease; MES, Mayo endoscopic sub score; PPV, positive predictive value; ROI, region of interest; SEC, serrated epithelial change; SSA, sessile serrated adenoma.

REFERENCES

1. Stidham RW, Liu W, Bishu S, et al. Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis. JAMA Netw Open 2019;2:e193963.

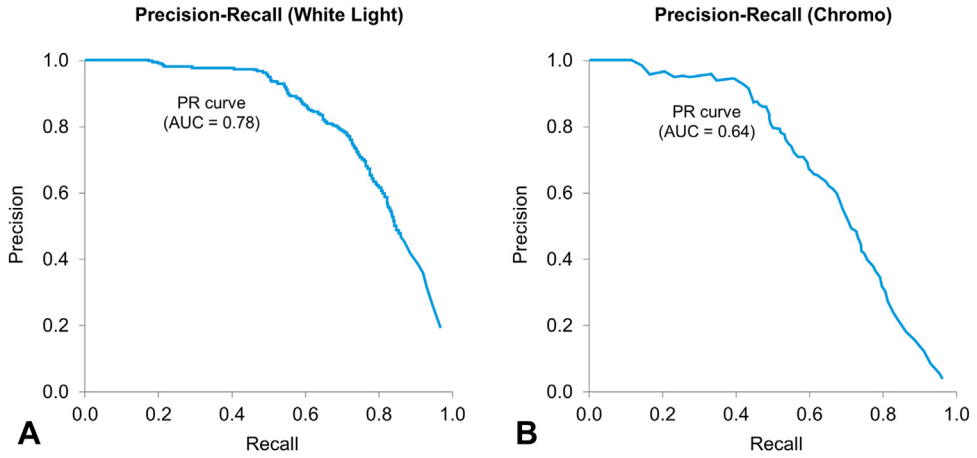
2. Beaugerie L, Itzkowitz SH. Cancers complicating inflammatory bowel disease. N Engl J Med 2015;372:1441-52.
3. Herrinton LJ, Liu L, Levin TR, et al. Incidence and mortality of colorectal adenocarcinoma in persons with inflammatory bowel disease from 1998 to 2010. Gastroenterology 2012;143:382-9.
4. Iacucci M, Cannatelli R, Tontini GE, et al. Improving the quality of surveillance colonoscopy in inflammatory bowel disease. Lancet Gastroenterol Hepatol 2019;4:971-83.
5. Rutter MD, Saunders BP, Wilkinson KH, et al. Thirty-year analysis of a colonoscopic surveillance program for neoplasia in ulcerative colitis. Gastroenterology 2006;130:1030-8.
6. Murthy SK, Feuerstein JD, Nguyen GC, et al. AGA clinical practice update on endoscopic surveillance and management of colorectal dysplasia in inflammatory bowel diseases: expert review. Gastroenterology 2021;161:1043-51.
7. Wijnands AM, de Jong ME, Lutgens M, et al. Prognostic factors for advanced colorectal neoplasia in inflammatory bowel disease: systematic review and meta-analysis. Gastroenterology 2021;160:1584-98.
8. Bye WA, Nguyen TM, Parker CE, et al. Strategies for detecting colon cancer in patients with inflammatory bowel disease. Cochrane Database Syst Rev 2017;9:CD000279.
9. Ananthakrishnan AN, Cagan A, Cai T, et al. Colonoscopy is associated with a reduced risk for colon cancer and mortality in patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol 2015;13:322-9.
10. 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;156:1661-74.
11. Deepak P, Hanson GJ, Fletcher JG, et al. Incremental diagnostic yield of chromoendoscopy and outcomes in inflammatory bowel disease patients with a history of colorectal dysplasia on white-light endoscopy. Gastrointest Endosc 2016;83:1005-12.
12. Leong RW, Ooi M, Corte C, et al. Full-spectrum endoscopy improves surveillance for dysplasia in patients with inflammatory bowel diseases. Gastroenterology 2017;152:1337-44.
13. Laine L, Kaltenbach T, Barkun A, et al. SCENIC international consensus statement on surveillance and management of dysplasia in inflammatory bowel disease. Gastrointest Endosc 2015;81:489-501.
14. Parian AM, Lazarev MG. Serrated colorectal lesions in patients with inflammatory bowel disease. Gastroenterol Hepatol 2018;14:19-25.
15. Johnson DH, Khanna S, Smyrk TC, et al. Detection rate and outcome of colonic serrated epithelial changes in patients with ulcerative colitis or Crohn's colitis. Aliment Pharmacol Ther 2014;39:1408-17.
16. Glissen Brown JR, Mansour NM, Wang P, et al. Deep learning computer-aided polyp detection reduces adenoma miss rate: a United States multi-center randomized tandem colonoscopy study (CADeT-CS Trial). Clin Gastroenterol Hepatol 2022;20:1499-507.
17. Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol 2020;5:352-61.
18. Hassan C, Wallace MB, Sharma P, et al. New artificial intelligence system: first validation study versus experienced endoscopists for colorectal polyp detection. Gut 2020;69:799-800.
19. Lui TKL, Guo CG, Leung WK. Accuracy of artificial intelligence on histology prediction and detection of colorectal polyps: a systematic review and meta-analysis. Gastrointest Endosc 2020;92:11-22.
20. Repici A, Badalamenti M, Maselli R, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology 2020;159:512-20.
21. Su JR, Li Z, Shao XJ, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). Gastrointest Endosc 2020;91:415-24.
22. Wang P, Liu X, Berzin TM, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. Lancet Gastroenterol Hepatol 2020;5:343-51.
23. Wallace MB, Sharma P, Bhandari P, et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. Gastroenterology 2022;163:295-304.
24. Maeda Y, Kudo SE, Ogata N, et al. Can artificial intelligence help to detect dysplasia in patients with ulcerative colitis? Endoscopy 2021;53:E273-4.
25. Jiang P, Ergu D, Liu F, et al. A review of Yolo algorithm developments. Proc Comput Sci 2022;199:1066-73.
26. Repici A, Spadaccini M, Antonelli G, et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. Gut 2022;71:757-65.
27. Areia M, Mori Y, Correale L, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. Lancet Digit Health 2022;4:e436-44.
28. Lutgens MW, van Oijen MG, van der Heijden GJ, et al. Declining risk of colorectal cancer in inflammatory bowel disease: an updated meta-analysis of population-based cohort studies. Inflamm Bowel Dis 2013;19:789-99.
29. Shah SC, Itzkowitz SH. Management of inflammatory bowel disease-associated dysplasia in the modern era. Gastrointest Endosc Clin North Am 2019;29:531-48.
30. Zhou G, Xiao X, Tu M, et al. Computer aided detection for laterally spreading tumors and sessile serrated adenomas during colonoscopy. PLoS One 2020;15:e0231880.

https://doi.org/10.1016/j.igie.2023.03.004
Received February 1, 2023. Accepted March 21, 2023.

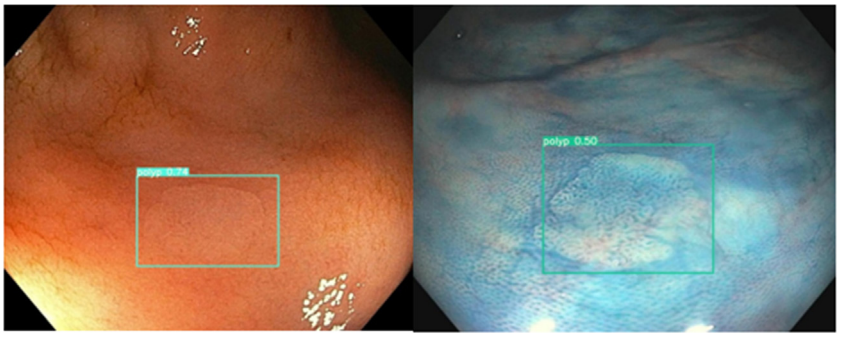
Current affiliations: Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota, USA (1), Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA (2), Department of Laboratory Medicine and Pathology (6), Department of Internal Medicine (7), Mayo Clinic, Rochester, Minnesota, USA; Department of Internal Medicine, Louisiana State University Health Sciences Center, Shreveport, Louisiana, USA (3), Department of Internal Medicine, Saint Vincent Hospital Worcester (4), Department of Internal Medicine Texas Tech University Health Sciences Center, El Paso, Texas, USA (5).

Reprint requests: Nayantara Coelho-Prabhu, MBBS, Division of Gastroenterology and Hepatology, Mayo Clinic, 200 First St SW, Rochester, MN 55905.

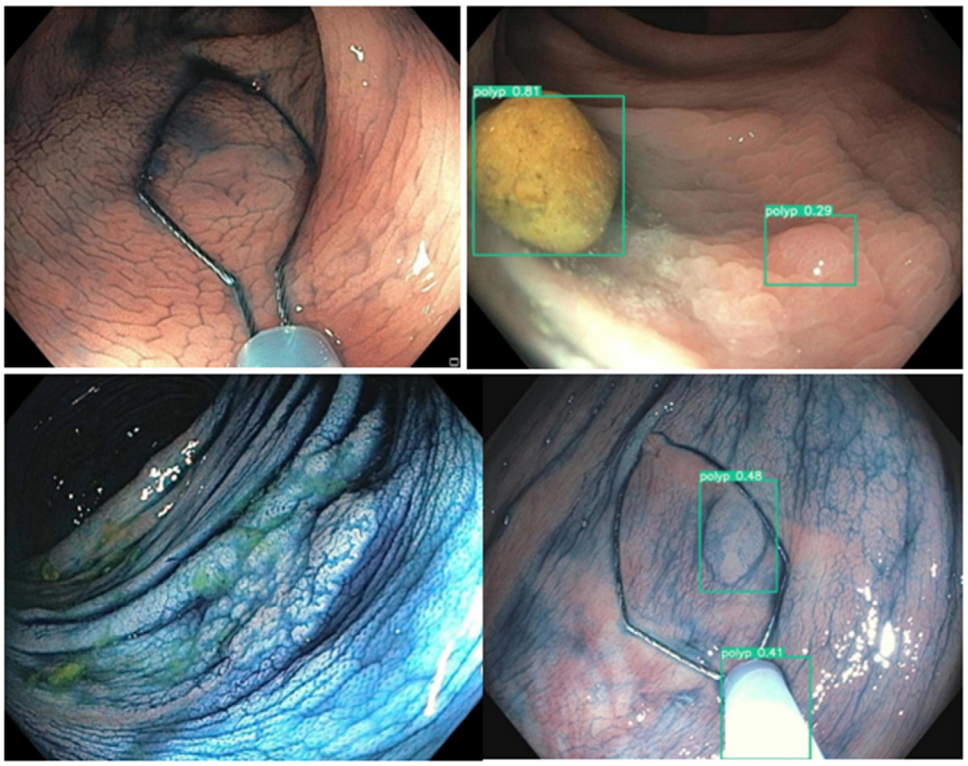
Presented at Digestive Disease Week, May 22-28, 2022, San Diego, California.



Supplementary Figure 1. Precision-recall curves for performance of the inflammatory bowel disease computer-aided detection system. **A,** Performance with high-density white-light endoscopy. **B,** Performance with chromoendoscopy. *AUC*, Area under the curve; Chromo, chromoendoscopy; *PR*, precision-recall.



Supplementary Figure 2. Examples of flat lesions detected by the inflammatory bowel disease computer-aided detection system.



Supplementary Figure 3. Examples of missed lesions and false positives by the inflammatory bowel disease computer-aided detection system.

SUPPLEMENTARY TABLE 1. Number of inflammatory bowel disease colorectal lesions included in the training, validation, and testing sets for high-definition white-light endoscopy and dye-based chromoendoscopy modalities

	High-definition white-light endoscopy		Chromoendoscopy	
Polyp type	**No. of polyps**	**No. of images**	**No. of polyps**	**No. of images**
Training set				
Dysplastic	246	204	97	75
Nondysplastic	170	154	108	81
Pseudopolyps	1121	388	372	84
Serrated epithelial changes	53	47	43	29
Serrated adenomas	108	103	41	37
Total	1698	896	661	306
Validation set				
Dysplastic	31	30	11	10
Nondysplastic	22	25	12	10
Pseudopolyps	140	80	41	5
Serrated epithelial changes	6	5	5	5
Serrated adenomas	14	10	5	5
Total	213	150	74	34
Testing set				
Dysplastic	31	30	27	21
Nondysplastic	21	21	30	23
Pseudopolyps	140	108	103	22
Serrated epithelial changes	7	7	12	8
Serrated adenomas	13	13	12	11
Total	212	179	184	85

SUPPLEMENTARY TABLE 2. Metrics used to analyze the performance of the inflammatory bowel disease computer-aided detection system

Metric	Formula or description
True positive	When the CNN bounding box overlapped with the annotated bounding box with a confidence threshold $\geq$.25
False positive	When the CNN drew a box in an area with no polypoid lesion that instead contained, for example, stool, bubbles, instruments, folds, glare, and inflamed mucosa
False negative	When the CNN failed to output a box in a region of interest containing a polypoid lesion
True negative	When the CNN did not output a box in an image that did not contain a polyp
Sensitivity (recall)	True positives / (true positives + false negatives)
Specificity	True negatives / (true negatives + false positives)
Positive predictive value (precision)	True positives / (true positives + false positives)
Negative predictive value	True negatives / (true negatives + false negatives)
False-positive rate	False positives / (true negatives + false positives)
False-negative rate	False negatives / (true positives + false negatives)
F1 score	2 × (precision × recall) / (precision + recall)
Accuracy	(True positives + true negatives) / all sample

CNN, Convolutional neural network.

SUPPLEMENTARY TABLE 3. Performance of the inflammatory bowel disease computer-aided detection system for high-definition white-light endoscopy and chromoendoscopy image datasets

Parameter	High-definition white-light endoscopy	Chromoendoscopy
Sensitivity, %	95.1	67.4
Specificity, %	98.8	88.0
Positive predictive value, %	98.9	83.3
Negative predictive value, %	94.7	74.3
False-positive rate, %	1.2	1.3
False-negative rate, %	4.9	3.3
F1 score, %	96.4	73.3
Accuracy,* %	96.8	77.8

*Confidence threshold selected at .25.

SUPPLEMENTARY TABLE 4. Characteristics of polypoid dysplastic lesions included in the test sets

Case no.	Type of inflammatory bowel disease	Size (mm)	Location	Mayo Score	Pathology	Paris classification	Detected by the inflammatory bowel disease computer-aided detection system	TP or FN
High-definition white-light endoscopy								
1	UC	>50	Ascending	2	Polypoid LGD	Is	Yes	TP
2	UC	7	Ascending	1	Polypoid LGD	IIa	Yes	TP
3	CD	5	Transverse	1	Polypoid LGD	Is	Yes	TP
4	CD	8	Ileum	1	Tubulovillous adenoma	Is	Yes	TP
5	UC	10	Ascending	0	Polypoid LGD	IIb	No	FN
6	UC	2	Transverse	1	Tubular adenoma	Is + IIc	Yes	TP
7	UC	5	Cecum	0	Tubular adenoma	Is	Yes	TP
8	UC	2	Rectosigmoid	3	Polypoid LGD	IIb	No	FN
9	UC	2	Ascending	1	Tubular adenoma	Is	Yes	TP
10	UC	25	Cecum	3	Polypoid LGD	IIa + Is	Yes	TP
11	UC	35	Hepatic flexure	2	Tubular adenoma	Is + IIa	Yes	TP
12	UC	10	Transverse	0	Polypoid LGD	Is	Yes	TP
13	UC	10	Cecum	1	Polypoid LGD	Is	Yes	TP
14	UC	4	Ascending	1	Polypoid LGD	Is	Yes	TP
15	CD	5	Sigmoid	0	Tubular adenoma	IIa	Yes	TP
16	IC	10	Transverse	1	Tubular adenoma	IIa	Yes	TP
17	UC	15	Ascending	2	Focal epithelial atypia and indefinite for dysplasia	IIb	Yes	TP
18	UC	30	Ascending	2	Polypoid LGD	Is + IIa	No	FN
19	UC	5	Sigmoid	1	Tubular adenoma	IIa	Yes	TP
20	UC	20	Rectosigmoid	1	Polypoid LGD	Is	Yes	TP
21	UC	20	Rectosigmoid	1	Polypoid LGD	Is	Yes	TP
22	CD	4	Rectum	0	Polypoid LGD	Is	Yes	TP
23	UC	5	Sigmoid	1	Polypoid LGD	IIa	Yes	TP
24	UC	4	Ascending	1	Polypoid LGD	IIa	Yes	TP
25	UC	2	Ascending	2	Tubular adenoma	Is	Yes	TP
26	UC	20	Ascending	2	Polypoid LGD	IIa + Is	Yes	TP
27	UC	15	Rectum	2	Polypoid LGD	IIb	Yes	TP
28	UC	3	Splenic flexure	1	Polypoid LGD	Is	Yes	TP
29	UC	6	Sigmoid	1	Polypoid LGD	Is	Yes	TP
30	UC	4	Cecum	2	Tubular adenoma	IIa	Yes	TP
Chromoendoscopy								
1	UC	2	Ascending	N/A	Tubular adenoma	Is	Yes	TP
2	UC	4	Transverse	N/A	Tubular adenoma	Is	Yes	TP
3	UC	4	Splenic flexure	N/A	Tubular adenoma	IIa	Yes	TP
4	UC	35	Hepatic flexure	N/A	Tubular adenoma	Is + IIa	Yes	TP
5	UC	10	Sigmoid	N/A	Polypoid LGD	IIa	Yes	TP
6	UC	2	Ascending	N/A	Tubular adenoma	Is	Yes	TP
7	UC	10	Rectum	N/A	NET	Is	Yes	TP
8	CD	20	Transverse	N/A	Polypoid LGD	Is + IIa	Yes	TP

(*continued on the next page*)

SUPPLEMENTARY TABLE 4. Continued

Case no.	Type of inflammatory bowel disease	Size (mm)	Location	Mayo Score	Pathology	Paris classification	Detected by the inflammatory bowel disease computer-aided detection system	TP or FN
9	UC	15	Ascending	N/A	Adenomatous change	IIb	Yes	TP
10	CD	6	Sigmoid	N/A	Polypoid LGD	Is	Yes	TP
11	CD	4	Descending	N/A	Tubular adenoma	Is	Yes	TP
12	CD	4	Descending	N/A	Tubular adenoma	Is	Yes	TP
13	UC	2	Sigmoid	N/A	Tubular adenoma	Is	Yes	TP
14	UC	20	Ascending	N/A	Polypoid LGD	IIa	Yes	TP
15	UC	2	Ascending	N/A	Tubular adenoma	Is	Yes	TP
16	CD	2	Splenic flexure	N/A	Tubular adenoma	Is	No	FN
17	UC	6	Sigmoid	N/A	Polypoid LGD	Is	No	FN
18	UC	30	Ascending	N/A	Polypoid LGD	IIa	No	FN
19	UC	2	Rectum	N/A	Tubular adenoma	IIb	No	FN
20	UC	2	Rectum	N/A	Tubular adenoma	IIb	No	FN
21	UC	2	Rectum	N/A	Polypoid LGD	IIa	No	FN
22	CD	20	Descending	N/A	Polypoid LGD	Is + IIa	No	FN

TP, True positive; *FN*, false negative; *UC*, ulcerative colitis; *CD*, Crohn's disease; *LGD*, low-grade dysplasia; *N/A*, not applicable Mayo Score with chromoendoscopy.

SUPPLEMENTARY TABLE 5. Performance of the inflammatory bowel disease computer-aided detection system in high-definition white-light endoscopy video clips

Video no.	Frames with polyps	video length (s)	No. (%) of true positives	No. (%) of false negatives	No. of false positives	Size (mm)	Paris classification	Mayo endoscopic subscore	Pathology	Reason for false positive	Observations
1	334	11.1	289 (87)	45 (13)	1	10	Ip	2	PSP	Irregular mucosa (Mayo 2)	
2	192	6.4	137 (71)	55 (29)	7	30	Iia+c	1	Colorectal cancer	Glare	Polyp covered with stool
3	190	6.3	184 (97)	6 (3)	38	8	Ip	2	PSP	Snare	Snare detected as polyp in 38 frames
4	400	13.3	399 (99.7)	1 (.3)	0	3	IIa	0	Serrated epithelial change		
5	275	9.16	249 (91)	26 (9)	3	5	Is	1	TA	Prominent fold and snare	
6	150	5	145 (97)	5 (3)	0	5	IIa	1	Sessile serrated adenoma		Sessile serrated adenoma covered in mucous cap
7	168	5.6	168 (100)	0 (0)	8	3	Is	0	Hyperplastic polyp	Scope in retroflexion	
8	198	6.6	83 (42)	115 (58)	52	2	IIa	0	TA	Biopsy forceps and vascular pattern	Blurry camera
9	170	5.7	170 (100)	0 (0)	0	4	Is	1	TA		Best video performance
10	162	5.4	149 (92)	13 (8)	85	6	Is	1	TA	Bubbles, glare, and mucus	

PSP, Pseudopolyp; *TA*, tubular adenoma.