

AI and endoscopy/histology in UC: the rise of machine

Olga Maria Nardone*^{ID}, Yasuharu Maeda* and Marietta Iacucci^{ID}

Ther Adv Gastroenterol

2024, Vol. 17: 1–16

DOI: 10.1177/
17562848241275294

© The Author(s), 2024.
Article reuse guidelines:
[sagepub.com/journals-
permissions](https://sagepub.com/journals-permissions)

Abstract: The gap between endoscopy and histology is getting closer with the introduction of sophisticated endoscopic technologies. Furthermore, unprecedented advances in artificial intelligence (AI) have enabled objective assessment of endoscopy and digital pathology, providing accurate, consistent, and reproducible evaluations of endoscopic appearance and histologic activity. These advancements result in improved disease management by predicting treatment response and long-term outcomes. AI will also support endoscopy in raising the standard of clinical trial study design by facilitating patient recruitment and improving the validity of endoscopic readings and endoscopy quality, thus overcoming the subjective variability in scoring. Accordingly, AI will be an ideal adjunct tool for enhancing, complementing, and improving our understanding of ulcerative colitis course. This review explores promising AI applications enabled by endoscopy and histology techniques. We further discuss future directions, envisioning a bright future where AI technology extends the frontiers beyond human limits and boundaries.

Keywords: artificial intelligence, computer-aided diagnosis, convolutional neural network, digital pathology, endo-histologic mucosal healing, machine learning, ulcerative colitis

Received: 14 February 2024; revised manuscript accepted: 13 May 2024.



Correspondence to:
Marietta Iacucci
Mercy/Cork University
Hospitals, Room 1.07,
Clinical Sciences Building,
Cork, Ireland

APC Microbiome Ireland,
College of Medicine and
Health, University College
Cork, Cork T12YT20,
Ireland
miaucucci@ucc.ie

Olga Maria Nardone
Division of
Gastroenterology,
Department of Public
Health, University Federico
II of Naples, Naples, Italy

Yasuharu Maeda
APC Microbiome Ireland,
College of Medicine and
Health, University College
Cork, Cork, Ireland

*These authors
contributed equally

Video Abstract Please click on the image to play the video (also available as supplemental material).

Introduction

Endoscopic remission remains a primary endpoint in the management of inflammatory bowel disease (IBD), but it is not enough, as a remarkable proportion of patients continue to relapse despite a normal-appearing mucosa. This implies that endoscopy may underestimate the true activity of the disease, especially in ulcerative colitis (UC), because subtle microscopic alterations in the context of apparently normal mucosa are associated with an increased risk of clinical relapse in up to 30% of patients. Conversely, patients who achieve histological remission have a greater likelihood of improved long-term clinical outcomes.^{1,2}

Accordingly, treatment goals for UC are evolving to deeper level of healing that incorporates both endoscopic and histologic healing.³ Hence, recently a composite endpoint of endo-histology mucosal healing (MH) encompassing both endoscopic remission/improvement and histologic remission has been proposed as a more “complete” and stringent outcome measure in clinical trials.⁴⁻⁶ A recent meta-analysis has further confirmed the magnitude of the potential benefit of the composite endpoint of endo-histologic MH, which is associated with a significantly reduced risk of clinical relapse.⁷ However, significant barriers and limitations remain, including the heterogeneity of histologic definitions in the various published studies, the absence of a standardized endoscopic and histologic scoring system, variability in biopsy interpretation, and differences in the duration of follow-up.

Emerging advanced endoscopic techniques have reduced the discrepancies between endoscopy and histology, moving toward a deep definition of MH aiming to restore the function of the colon’s mucosal barrier and disease clearance.^{8,9} A recent international multicenter study showed that a single measure of endoscopic remission defined using electronic virtual chromoendoscopy performed similarly to combined endoscopic and histologic remission for predicting specified clinical outcomes at 12 months.³ Hence, electronic virtual chromoendoscopy can aid in simplifying the assessment of histo-endoscopic remission, facilitating better understanding and reducing complexity in evaluation.

Furthermore, advancements in artificial intelligence (AI) technologies hold promise for

enhancing endoscopic image interpretation and aiding in the quantification of vascular and mucosal patterns to provide a deeper definition of remission. AI moves beyond human vision of endoscopy and ultimately assists in real-time histologic assessment during endoscopy.¹⁰ Remarkably, utilization of AI as an ideal adjunct to existing tools can revolutionize the way therapeutic decisions are made by supporting diagnosis, monitoring, and predicting the course of disease in patients with UC.⁸

In this current narrative review, we explore promising AI applications in IBD endoscopy and histology. We further discuss future directions, envisioning a bright future in which AI technology adds value to precision medicine.

AI for endoscopy scoring disease activity assessment and prediction of clinical outcomes

AI for scoring disease activity assessment

Mayo endoscopic subscore (MES) and Ulcerative Colitis Endoscopic Index of Severity (UCEIS) are the most widely used endoscopy scoring for disease activity assessment in clinical trials and in practice.^{11,12} However, in both cases, it is known that there are variabilities among examiners imposing significant costs on image interpretation for clinical trials.¹³ Additionally, in the context of the treat-to-target (T2T) approach^{11,12} which has gained recent traction, achieving endoscopic remission as a treatment goal varies among examiners and institutions, thereby posing a challenge in establishing a common treatment objective. To the best of our knowledge, the first physician-initiated study of a computer-aided diagnosis (CAD) system in patients with UC was conducted by Sasaki et al.¹⁴ in 2003. Based on Bayesian decision theory, they established a computer-aided grading system for endoscopic severity (Matts grade).

The emergence of convolution neural network (CNN) and machine learning (ML) models have been attracting much attention to effectively learn patterns from extensive image datasets, automatically identify abnormal regions, and provide corresponding scores.¹⁵⁻¹⁸ This technology is expected to streamline the labor-intensive and costly process of central image interpretation in clinical trials,¹⁹ diminish variations in

interpretation among examiners, and promote more consistent evaluations.

Similarly, in 2019, Ozawa *et al.*¹⁵ succeeded in using a CNN powered by deep learning to develop CAD. Using expert diagnosis as the gold standard, their retrospective analysis showed that their CAD system correctly classified 73% of the MES0 images, 70% of the MES 1 images, and 63% of the MES 2–3 images with an appropriate MES. At the same time, Stidham *et al.*¹⁶ also developed a CNN model to output 4-class prediction (MES 0–3). In a retrospective comparison with experts' diagnosis, their CAD system had good predictive value for differentiating an MES of 0–1 from an MES of 2–3 as shown by sensitivity of 83%, specificity of 96%, positive predictive value of 86%, and negative predictive value of 94%.

Further, Takenaka *et al.*¹⁷ constructed a deep neural network for evaluation of UC (DNUC) algorithm, which was designed to extract features from conventional white-light still images and predict endoscopic remission (defined as a UCEIS of 0) with 90% accuracy.

However, these were validations for still images, and there was concern about selection bias in selecting images for validation. To address some of these issues, video instead of still images is recommended for assessments to reduce subjectivity related to image acquisition sites.^{20,21}

Subsequently, Yao *et al.*,²² from the same group as Stidham, investigated the application of their CAD system into unaltered full motion videos and still images and then automatically output a summary MES for each colonoscopy. Notably, their new algorithm automatically excluded inadequate frames (including those in proximity to the mucosa and those with light reflection, debris and blurring, and motion blur) from the analysis for a final score of endoscopic severity.

Likewise, Takenaka *et al.*²¹ updated their DNUC for application to video-based analysis. In their multicenter cross-sectional study, the DNUC correctly evaluated the presence or absence of histological inflammation in 81% of cases. They also showed that the DNUC result was strongly correlated with the UCEIS provided by central readers: with an intraclass correlation coefficient of 0.93 and good accuracy in determining

endoscopic remission, defined as a UCEIS of 0 or 1 (sensitivity of 82% and specificity of 95%). Notably, through this study, they found that the main reason for discrepancy of two or more points between the DNUC and the central evaluator was poor bowel preparation or the presence of inflammatory polyps. Further improvement in accuracy is expected by incorporating an algorithm that automatically excludes such images.

Since then, several AI models have been reported that employ the ability to automatically exclude inappropriate images of similar concepts from the analysis.^{20,23} However, the accuracy with which inappropriate images can be extracted has not yet been verified. Furthermore, the extent to which these additional functions ameliorate differences in performance among examiners remains to be tested.

AI has been expected to address the issue of requiring human central readers for endoscopic disease activity scoring, which makes clinical trials costly and time-consuming. Gottlieb *et al.*¹⁹ used the data set from a phase II Mirikizumab clinical trial²⁴ as the training and validation of a CAD system. Their CAD output a final score representing the endoscopic severity (MES of 0–3 and UCEIS of 0–8) for each full-length video in a fully autonomous manner. In their retrospective analysis, their CAD system achieved excellent accuracy of 96% for endoscopic remission when defined as an MES of 0 and 97% when defined as a UCEIS of 0.

AI-enabling inflammatory distribution

AI-assisted colonoscopy has also been used to visualize and evaluate the severity and distribution of inflammation more accurately in the entire colon. Currently, endoscopic scores, such as MES and UCEIS, evaluate the grade of inflammation in the most severe segment; the distribution of inflammation is not considered, even though it also significantly impacts drug selection. While there have been efforts to develop scoring systems that account for the extent of inflammation, they have not gained widespread adoption in clinical practice.^{25,26}

Takabayashi *et al.*²⁷ developed a user interface in which Ulcerative Colitis Endoscopic Gradation Scale using a ranking-CNN, results diagnosed were incorporated into the schema for the colon.

Visualization in their novel user interface facilitates the comprehension of the extent and grading of inflammation throughout the colon. Fan et al.²⁸ have introduced a new AI-driven scoring system that incorporates the distribution of inflammation in the output obtained from video-based AI-driven scoring. They divided the colon into a fixed number of “areas” (cecum, 20; transverse colon, 20; descending colon, 20; sigmoid colon, 15; rectum, 10). The scoring system automatically assessed the inflammatory severity of 85 areas from each video and generated a visualized result depicting the full-length intestinal inflammatory activity.

Most recently, Stidham et al.²⁹ introduced the automated Cumulative Disease Score (CDS) system which sums the MES-squared values for all 50 evenly spaced increments in the left colon, including the rectum. Notably, the CDS focuses solely on the left-sided colon and does not provide an evaluation of the entire colon. In a clinical trial cohort including 748 induction and 348 maintenance patients (UNIFI study⁵), CDS exhibited a significant correlation with the MES ($p < 0.001$) and all clinical components of the partial Mayo score ($p < 0.001$). When stratified by pretreatment CDS, it was observed that ustekinumab was more effective than placebo ($p < 0.001$), with a more pronounced effect in severe versus mild disease (-85.0 vs -55.4 ; $p < 0.001$). Furthermore, compared to the MES, the CDS revealed endoscopic differences between ustekinumab and placebo (Hedges’ g , 0.743 vs 0.460) and demonstrated greater sensitivity to change, requiring 50% fewer participants to show endoscopic differences between ustekinumab and placebo (Hedges’ g , 0.743 vs 0.460).

The AI-assisted assessment of inflammation distribution may offer more comprehensive descriptions of UC activity, contributing to a detailed evaluation of treatment effectiveness. This new concept enhances, complements, and improves our understanding of the disease.

AI-enabled advanced endoscopic technology

Another highlight in this field is the integration of AI and advanced endoscopic technology.^{8,10,30,31} Specifically, the use of advanced endoscopic technology to predict histologic remission is anticipated to offer a more precise assessment of inflammation and activity, addressing the unmet

need for biopsies. Currently, histology stands as the most reliable predictor of sustained clinical remission in UC. However, it has not been widely applied in daily practice³² because of the burden of taking biopsies, the additional time required for a pathological diagnosis, and the increased cost.

Bossuyt et al. utilized a red density algorithm (Pentax Medical, HOYA Corp., Tokyo, Japan) to predict histological activity. When the red density score cutoff was ≤ 60 , this approach exhibited a 96% sensitivity and 80% specificity in predicting histologic remission, defined as a Robarts histopathology index (RHI) of ≤ 6 .³³

Likewise, the same group explored in a pilot study the CAD system that utilized a single short-wavelength monochromatic Light Emitting Diode light illumination (Fujifilm, Tokyo, Japan).³⁴ This system enabled real-time evaluation of mucosal architecture, including the crypts, pericryptal capillaries, and bleeding. The algorithm successfully identified histologic remission, defined as a Geboes score of $< 2B.1$, with high performance, boasting a sensitivity of 79% and specificity of 90%.

Similarly, Iacucci et al.^{35,36} have developed an AI tool capable of generating the Paddington International Virtual Chromoendoscopy Score using iSCAN (Pentax) short-length videos. They conducted their analysis based on histologic remission criteria, specifically RHI ≤ 3 , Nancy histological index (NHI) ≤ 1 , and PICaSSO Histological Remission Index (PHRI) = 0. The accuracy values for these criteria were reported as 83%, 81%, and 83%, respectively.³⁷

Furthermore, a CAD system was developed to predict histological healing using a commercially available 520-fold ultra-magnifying contact microscope (Endocryo: CF-H290EC; Olympus Corp., Tokyo, Japan) images, which allows real-time evaluation of the microvessels, crypts, and goblet cells when used with narrow-band imaging.^{38–42} When histological remission was defined as a Geboes score of < 3.1 , the diagnostic accuracy of the CAD system was 74% for sensitivity, 97% for specificity, and 91%.⁴³ The updated version of this CAD system received regulatory approval in Japan and has been on the market since February 2021 (EndoBRAIN-UC; Cybernet Systems Corp., Tokyo, Japan). The

system is the only commercially available CAD system powered by AI during colonoscopy in patients with UC. However, it is not widely used because the system requires specific expertise in using the ultra-high magnified colonoscope. To address this issue, the same team developed another system that can be applied using a wide range of endoscopes to provide an objective two-class diagnosis of “AI-based vascular-healing” versus “AI-based vascular-active.”⁴⁴

With the most recent, Akiyama *et al.*⁴⁵ reported a novel hypoxia imaging algorithm (Fujifilm), which provides colonic tissue oxygen saturation and correlates with clinical, endoscopic, and histological activities. Notably, they have shown that rectal oxygen saturation levels correlate with bowel urgency, which may have opened the door to a new diagnostic of objective endoscopic assessment of dysfunction in patients with UC.

AI-assisted endoscopy to predict clinical outcomes

AI-enabled precise scoring has also been demonstrated to predict future relapse. Stratifying relapse risk through objective measures is anticipated to facilitate early therapeutic intervention and assist patients in sustaining long-term remission.

Takenaka *et al.*¹⁷ conducted a prospective follow-up study involving 875 patients, as outlined in their initial research. The patients were categorized into groups diagnosed by AI: the AI-diagnosed mucosal activity group and the AI-diagnosed MH group. The hazard ratios yielded by their algorithm for various outcomes, such as hospitalization, colectomy, steroid use, and relapse (defined as a partial Mayo score of ≥ 3 and C-reactive protein/calprotectin positivity (≥ 3 mg/L/ ≥ 250 μ g/g)), were 48.4, 46.4, 10.2, and 8.8, respectively.⁴⁶

Moreover, Maeda *et al.* performed colonoscopy using real-time AI in patients with UC who were in clinical remission. They prospectively monitored these patients ($n = 134$) for 12 months following colonoscopy. Based on the AI outputs, the patients were categorized into the AI-active and AI-healing groups. Clinical relapse, defined as a partial Mayo score of ≥ 3 , occurred in 28.4% (21/74) of the patients in the AI-active group and in 4.9% (3/61) in the AI-healing group.⁴⁷

Encouraging results stem from their recently CAD-developed system to identify “vascular healing” using commercially available magnifying colonoscopy. In the study, the clinical relapse rate during 12 months after colonoscopy was significantly higher in the AI-based vascular-active group (23.9% (16/67)) compared with the AI-based vascular-healing group (3.0% (1/33)) ($p = 0.01$).⁴⁴

Similarly, Iacucci *et al.*³⁷ followed up with 232 patients after AI-assisted colonoscopy and showed that the hazard ratios of their AI-assisted algorithms for adverse clinical outcomes (UC-related hospitalization, colectomy, and change in UC treatment owing to relapse) were 2.9 for the high-definition white-light endoscopy model and 4.0 for the iSCAN model (Table 1).

In conclusion, AI-assisted endoscopic scoring can offer not only endoscopist-independent objective tool analysis but also automatically exclude inadequate frames, the distribution of inflammation, optical prediction of histological remission, and on-site prediction of clinical prognosis (Figure 1). Offering objective and precise treatment targets would pave the way for personalized medicine, as well as enhance the efficiency of clinical trials.

AI systems for digital histological assessment of UC and therapeutic response

A potential application of AI is to assist in real-time histological evaluation. ML approaches are increasingly being developed to aid pathologists in accurately and reproducibly scoring histology, enabling precise quantification of clinically relevant features. Furthermore, promising applications in digital pathology can include the quantification of the expression of molecular targets.

The advent of digital pathology with the integration of digital slides into whole-slide images (WSIs) has extended the frontiers of the pathologist’s view beyond a microscopic slide and enabled the true utilization and integration of knowledge that in the future could surpass human limits and boundaries.⁴⁸ Noteworthy, the application of computer vision methods involves converting glass slides containing tissue samples into high-resolution digital images that can be viewed and analyzed using computer software. However, in the field of UC pathology, few AI models have been developed so far.⁴⁸

Table 1. Summary of artificial intelligence for endoscopy in ulcerative colitis.

Author	Ref.	Study design	Modality	No. of training samples	No. of validation samples	Outcome measures	Results
Artificial intelligence for disease activity							
Ozawa et al., 2019	15	Retrospective single-center	White-light endoscopy	26,304 images (841 pts)	3981 images (114 pts)	ER (MES=0 or 1)	CAD system correctly classified 73% of the MES0 images, 70% of the MES 1 images, and 63% of the MES 2–3 images with an appropriate MES
Stidham et al., 2019	16	Retrospective single-center	White-light endoscopy	14,862 images (2778 pts)	1652 images (304 pts) 11,432 frames (30 videos)	ER (MES=0 or 1)	CAD predictive value for differentiating an MES of 0–1 from an MES of 2–3 as shown by sensitivity of 83%, specificity of 96%, positive predictive value of 86%, and negative predictive value of 94%
Takenaka et al., 2020	17	Prospective single-center	White-light endoscopy	40,758 images	4187 images (875 pts)	ER (UCEIS=0) HR (Geboes < 3.1)	DNUC accuracy of 90% for ER and 93% accuracy for HR
Yao et al., 2021	20	Prospective multicenter	White-light endoscopy	N/A	51 videos (internal data) 264 videos (external data)	ER (MES=0 or 1)	CAD correctly predicted the exact MES in 78% (internal data) and 57% (external data)
Takenaka et al., 2022	21	Prospective multicenter	White-light endoscopy	N/A	900 segments (180 pts) 590 pts	ER (UCEIS=0 or 1)	DNUC sensitivity of 82%, and specificity of 95%
Byrne et al., 2023	23	Prospective single-center	White-light endoscopy	1,550,030 frames (134 UC videos)	100 videos	Grading of MES and UCEIS	The quadratic weighted kappa between experts' labels and the model's predictions were 0.87, 0.88 at frame level, 0.88, 0.90 at section level and 0.90, 0.78 at video level, for MES and UCEIS, respectively
Gottlieb et al., 2021	19	Prospective multicenter	White-light endoscopy	N/A	249 videos	ER (MES=0, and UCEIS=0)	Accuracies of 96% for MES and 97% for UCEIS.
Lo et al., 2022	18	Retrospective single-center	White-light endoscopy	1484 images (467 pts)	Five-fold cross-validation	Grading of MES categories (0 vs 1–3 and 0–1 vs 2–3)	94% and 93% in distinguishing MES 0 vs 1–3 and 0–1 vs 2–3.
Artificial intelligence-enabling inflammatory distribution							
Takabayashi et al., 2023	27	Retrospective multicenter	White-light endoscopy	14,208 images	1479 images	The correlation coefficients between IBD expert endoscopists and the AI	Spearman's correlation coefficients were all higher than 0.95 ($p < 0.01$).
Fan et al., 2022	28	Retrospective single-center	White-light endoscopy	5875 images (332 pts)	20 full-length videos (18 pts)	Grading of MES and single UCEIS items in WLE	86.5% accuracy in the MES ($\kappa=0.81$), UCEIS items with accuracies of 90.7%, 84.6%, and 77.7% for vascular pattern ($\kappa=0.82$), erosions and ulcers ($\kappa=0.78$), and bleeding ($\kappa=0.70$).
Stidham et al., 2024	29	Prospective multicenter	White-light endoscopy	N/A	748 pts	Cumulative disease score performance	Cumulative disease score had better sensitivity for detecting endoscopic changes than MES.

(Continued)

Table 1. (Continued)

Author	Ref.	Study design	Modality	No. of training samples	No. of validation samples	Outcome measures	Results
Artificial intelligence-enabled advanced endoscopic technology							
Bossuyt et al., 2020	33	Prospective multicenter	Red density	N/A	29 UC patients and 6 healthy controls	Correlation of red density score with RHI, MES, and UCEIS	Red density score correlated with RHI ($r=0.74$), MES ($r=0.76$), and UCEIS ($r=0.74$).
Bossuyt et al., 2023	34	Prospective single-center	Single short-wavelength monochromatic LED light illumination	N/A	113 segments (58 pts)	HR (Geboes <2 B.1)	79% sensitivity, and 90% specificity.
Iacucci et al., 2023	37	Prospective multicenter	iSCAN	67,280 frames (283 pts)	242 videos	ER (UCEIS ≤ 1 , and PICaSSO ≤ 3) HR (RHI ≤ 3 , NHI ≤ 1 , and PHRI ≤ 1)	ER of UCEIS ≤ 1 with a sensitivity of 72%, specificity of 87%, and PICaSSO ≤ 3 with a sensitivity of 79%, specificity of 95%. Accuracies ranging in HR from 80% to 85%.
Maeda et al., 2019	43	Retrospective single-center	Endocytoscope-narrow-band imaging	12,900 images (87 pts)	525 segments (100 pts)	HR (Geboes <3.1)	74% sensitivity, and 97% specificity.
Akiyama et al., 2024	45	Retrospective single-center	Hypoxia imaging	N/A	490 images (100 pts)	Correlation of colonic oxygen saturation, and MES and Geboes score.	At a colonic oxygen saturation cutoff of 45.5%, AUCs for endoscopically and histologically active diseases were 0.79 and 0.72.
Artificial intelligence-assisted endoscopy to predict clinical outcomes							
Takenaka et al., 2021	46	Prospective single-center	White-light endoscopy	40,758 images	875 pts	Future hospitalization, colectomy, steroid use, and relapse	Hazard ratios were 48.4, 46.4, 10.2, and 8.8 for hospitalization, colectomy, steroid use, and relapse.
Maeda et al., 2022	47	Prospective single-center	Endocytoscope-narrow-band imaging	44,097 images	135 pts	Clinical relapse during 12 months after colonoscopy	The relapse rate was significantly higher in the AI-active group (28%) than in the AI-healing group (5%; $p < 0.001$).
Kuroki et al., 2024	44	Prospective single-center	Narrow-band imaging	8853 images (167 pts)	104 pts	Clinical relapse during 12 months after colonoscopy	The relapse rate was significantly higher in the vascular-active group (28%) than in the vascular-healing group (5%; $p < 0.001$).
Iacucci et al., 2023	37	Prospective multicenter	iSCAN	67,280 frames (283 pts)	232 patients	Adverse clinical outcomes (UC-related hospitalization, colectomy, and change in UC treatment owing to relapse)	Hazard ratios were 2.9 for the high-definition white-light endoscopy model and 4.0 for the iSCAN mode.
AI, artificial intelligence; AUC, area under the curve; CAD, computer-aided diagnosis; DNUC, deep neural network for evaluation of ulcerative colitis; ER, endoscopic remission; HR, histological remission; IBD, inflammatory bowel disease; LED, light emitting diode; MES, Mayo endoscopic subscore; NHI, Nancy histological index; PHRI, PICaSSO Histological Remission Index; PICaSSO, Paddington International Virtual Chromoendoscopy Score; pts, patients; RHI, Roberts histopathology index; UC, ulcerative colitis; UCEIS, ulcerative colitis endoscopic index of severity; WLE, white light endoscopy.							

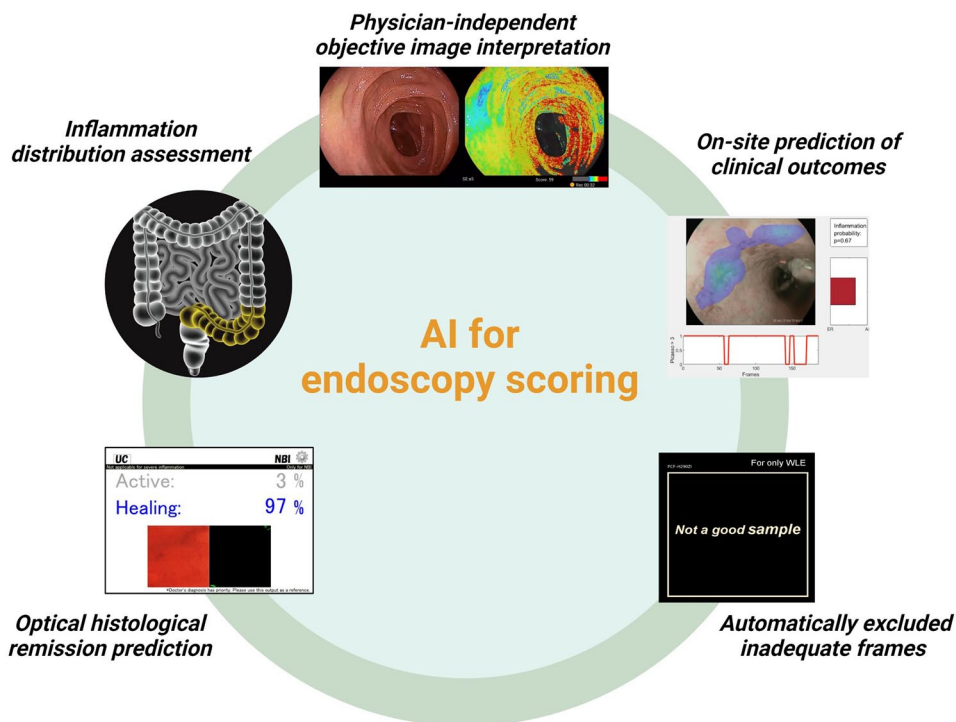


Figure 1. The role and potential of artificial intelligence for endoscopy scoring disease activity. Source: Created with BioRender.com.

In a pioneering study, Vande Casteele et al.⁴⁹ developed the first deep learning algorithm designed to identify eosinophils as markers of active UC. Eosinophils ($N=3480$) were manually annotated in digitized images of colonic biopsies to train a CNN. When applied to colonic biopsies from 88 patients with active UC, this algorithm exhibited perfect agreement with manual eosinophil counts performed by pathologists (intraclass correlation coefficients = 0.81–0.92). Nevertheless, eosinophil density did not show a good correlation with histological activity.

However, several papers have proposed that goblet cells play important roles in the onset and progression of UC. In a retrospective cohort study, Ohara et al.⁵⁰ developed a deep learning-based model to quantitatively assess goblet cell mucus in patients with UC who were in clinical and endoscopic remission. This model was designed to detect goblet cell mucus and epithelial cells in digital histological images of biopsy specimens and calculate their area ratio. Remarkably, rectal goblet cell ratio could be a valuable marker for assessing UC activity, and the use of deep learning-based models to quantify goblet cell mucus

area holds promise for predicting the future risk of clinical relapse.

In a recent study, Najdawi et al.⁵¹ developed an innovative ML-based strategy employing CNN models for segmenting tissue and cells, aimed at analyzing histologic aspects of UC from hematoxylin and eosin-stained WSIs. This model accurately assessed tissue and cell characteristics, exhibiting strong correlation with disease severity and pathologist-assigned scores. Notably, the extracted features closely mirrored pathologist annotations and strongly correlated with both disease severity and NHI scores. Based on the use of a random forest classifier on 13 human-interpretable features (HIF) extracted from these models, NHI score was accurately predicted, with substantial alignment demonstrated with pathologist consensus scores. Of significance, certain HIFs directly measured neutrophil features, acknowledged as pivotal indicators of disease activity, while the absence of neutrophil infiltration defined histologic remission.

Correspondingly, the Paddington International Virtual Chromoendoscopy Score (PICaSSO)

team pioneered a multiple instance learning approach employing CNNs to identify neutrophils in WSIs and classify them as indicative of either histological remission or predictive of adverse outcomes.⁵² Subsequently, the same group developed and validated an AI system CAD system trained and tested on a large set of digitized biopsies to detect UC disease activity as defined by different histologic indices and predict prognosis.⁵³ In this study, the CAD system trained and tested on a large set of digitized biopsies had a strong diagnostic performance to detect disease activity assessed by using the PICaSSO Histologic Remission Index (PHRI),⁵⁴ based on the presence of neutrophils from all areas (PHRI > 0) with an overall Area Under the Receiver Operating Characteristic Curve of 0.87, a sensitivity of 89%, and a specificity of 84%. This means that this new system exhibited high sensitivity and specificity in distinguishing and predicting histologic activity and remission, as well as stratifying the risk of flare at 12 months, similarly to human pathologists (hazard ratio 4.6 vs 3.6, respectively).⁵³

Furthermore, Rymarczyk *et al.*⁵⁵ developed an AI/ML model designed to automate the assessment of histological disease activity in UC. These models were trained using a vast collection of imaging data obtained from a company's late-stage clinical trial. They were specifically trained on scanned histology slides taken from the intestinal mucosa of adult patients and are capable of predicting the Geboes histopathology score for UC. Despite certain limitations such as imbalanced training data for certain Geboes subscores, as well as variations in the quality of histopathology slides, the models perform well in predicting Geboes scores.

More recently, Liu *et al.*⁵⁶ trained on 1,87,571 informative patches from rectal hematoxylin and eosin biopsy samples from 292 treatment-naïve pediatric patients with UC in a multicenter inception cohort (PROTECT2) study with an external validation test cohort included 113 pediatric patients followed in the Canadian Children Inflammatory Bowel Disease Network inception cohort study at the Hospital for Sick Children (SickKids) to predict treatment response in pediatric UC. The authors identified the set of 18 histomic features that, when incorporated in a ML model and digital histopathology, can be used to identify, before the initiation

of treatment, pediatric patients who cannot achieve steroid-free long-term remission with mesalamine alone. This model represents a promising approach to integrating digital histopathology-based histomic features and ML algorithms aiming at identifying therapeutic responses (Table 2).

In conclusion, AI-assisted histologic analysis is poised to become an integral component in advancing precision medicine for patients with IBD. In the near future, AI can serve as an assisting tool, holding the potential to facilitate decentralized trials, similar to the demonstrated utility in the evaluation of endoscopic images. We are approaching a new era of the shift to routinely incorporate biopsy collection and histologic readout for most UC drug development programs (Figure 2).

The role of AI in clinical trials

Histo-endoscopic mucosal improvement and remission are becoming crucial therapeutic endpoint in UC trials since they may provide a more accurate and deep evaluation of the disease healing. Yet, factors such as interobserver variability, endoscopist inexperience, and the variance in definitions can contribute to misjudging disease severity, potentially leading to inappropriate patient selection or incorrect assignment to treatment groups. Consequently, recruitment challenges may arise, potentially diminishing the pool of eligible participants and necessitating larger cohorts, thereby elevating the costs of clinical trials.⁵⁷

One potential strategy to break the therapeutic ceiling and reach 100% target could be to change the way to perform clinical trials in IBD.⁵⁸ In this context, AI is expected to raise the bar on study design, including patient selection, enrollment, and improving endoscopy quality through validated site reads, as well as facilitating stratification and re-randomization.⁵⁸

Firstly, central reading of endoscopic video plays a critical role in IBD clinical trials, to overcome the subjective variability in endoscopic scoring. Indeed, local readers tend to overscore the screening endoscopy and underscore the outcome of endoscopy.⁵⁷ Feagan *et al.*,⁵⁹ investigated the impact of centralized image review on mitigating discrepancies in placebo response rates in UC

Table 2. Summary of artificial intelligence for histology in ulcerative colitis.

Author	Ref.	Study design	No. of training samples	No. of validation samples	Outcome measures	Results
Gui et al., 2022	54	Prospective multicenter	97 biopsies	41 biopsies	HR (PHRI < 1)	78% sensitivity, 91.7% specificity, and 86.0% accuracy.
Ohara et al., 2022	50	Retrospective single-center	2300 images	114 patients	Automated calculation of goblet cell ratio, and prediction of clinical relapse	Patients with a GCR of $\leq 12\%$ had a significantly higher relapse rate than those with a GCR of $> 12\%$ (45% vs 6.5%; $p < 0.01$).
Vande Castele et al., 2022	49	Retrospective single-center	88 pts	20 tissue regions	Identification of eosinophil counts.	Almost perfect agreement with 4 pathologists (ICCs: 0.81–0.92).
Iacucci et al., 2023	53	Prospective multicenter	118 biopsies	A: 375 biopsies B: 154 biopsies (58 patients)	A: HR (PHRI ≤ 1 , RHI ≤ 3 , and NHI ≤ 1) B: Prediction of prognosis.	Sensitive and specific at 89% and 85% (PHRI), 94% and 76% (RHI), and 89% and 79% (NHI). B: the hazard ratio for disease flare-up was 4.6.
Najdawi et al., 2023	51	Retrospective single-center	512 WSI (334 pts)	308 WSI (243 pts)	Prediction of NHI and HR	NHI scores, with a weighted kappa ($\kappa = 0.91$) and Spearman correlation ($\rho = 0.89$), and HR, with 97% accuracy.
Rymarczyk et al., 2023	55	Retrospective multicenter	2696 biopsies	800 biopsies	Grading of Geboes score determined by experts.	Accuracy was 65% to 85% (kappa range: 0.44–0.68).
Liu et al., 2024	56	Inception cohort study multicenter	1,87,571 informative patches from rectal hematoxylin and eosin biopsy samples (292 pts); 250 histomic features	113 pediatric patients (external cohort)	Corticosteroid-free clinical remission on mesalamine maintenance therapy alone at 1 year	250 histomic features at the patch level achieved AUROC of 0.87 (95% CI, 0.73–1.00) and an accuracy of 0.90 (95% CI, 0.80–1.00) at the WSI level in predicting treatment response. Eighteen histomic features exhibited comparable performance with an AUROC of 0.89 (95% CI, 0.71–0.96) and accuracy of 0.90 (95% CI, 0.80–1.00); a set of 18 histomic features demonstrated comparable performance on the external cohort

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; HR, histological remission; ICC, interclass correlation coefficient; NHI, Nancy histological index; PHRI, PICaSSO Histological Remission Index; PICaSSO, *Paddington International Virtual Chromoendoscopy Score*; pts, patients; RHI, Robarts histopathology index; WSI, whole-slide images.

trials. Interestingly, 31% of participants initially deemed eligible based on a UCDAI sigmoidoscopy score ≥ 2 by site investigators were subsequently deemed ineligible by the central reader's independent assessment due to insufficient endoscopic disease. Removing these patients who had been entered based on discrepantly high sigmoidoscopic scores not only increased the difference between remission rates for mesalamine 4.8g/day and placebo (29% and 13.8%, respectively, with a now widened difference of 15%; 95% confidence interval, 3.5%–26.0%; $p = 0.011$), but also led to the trial achieving its primary and all secondary endpoints.

Subsequently, Gottlieb et al.¹⁹ assessed the ability of a recurrent neural network (RNN) for predicting scores assigned by central readers and compared its performance with MES and UCEIS scoring systems. Remarkably, they revealed that an RNN algorithm, trained on full-length videos from an industry-sponsored phase II clinical trial in UC where a human central reader assigned only one video-level MES and UCEIS score, achieved an accuracy of 97.0% for UCEIS and 95.5% for MES.

Hence, AI has the potential to transform the assessment of clinical trials by providing more

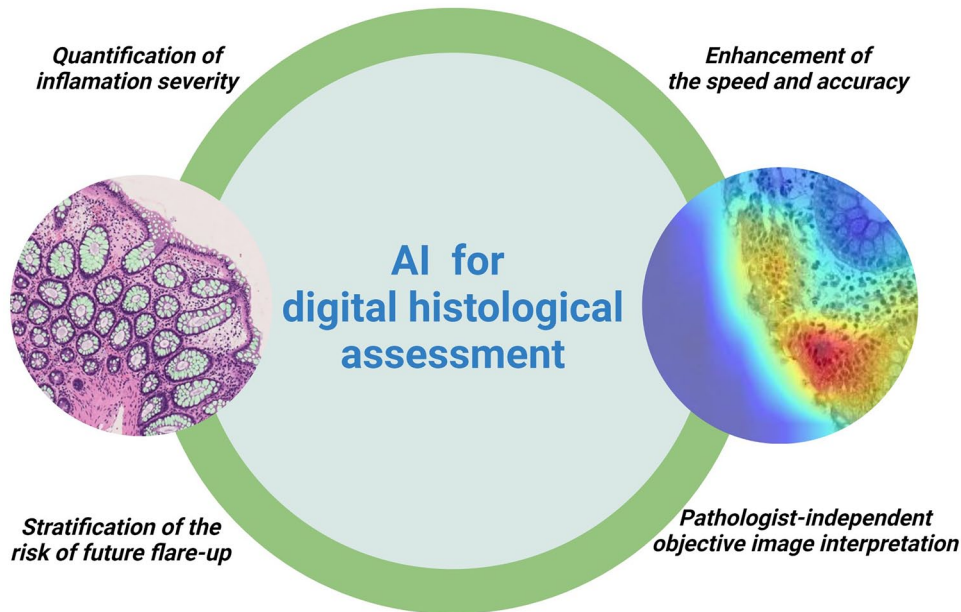


Figure 2. The role and potential of artificial intelligence for digital histological assessment.
Source: Created with BioRender.com.

reliable and efficient endpoint assessment through image-based endpoint detection. By implementing standardized imaging reading and scoring methods, AI can significantly enhance the speed and accuracy of disease assessment, offering a more precise identification of treatment response and non-response. Additionally, AI-assisted disease activity and remission assessment are expected to decrease variability and minimize the need for second reader/adjudication. Of note, cost savings are a further consideration to bear in mind, with AI potentially reducing central reading costs, a major component of trial budgets.⁵⁷

In conclusion, AI has the potential to revolutionize various aspects of IBD clinical trials, from expediting central reading to improving endpoint assessment and defining remission, ultimately leading to more effective decision-making and efficient management of IBD. Integration of data from multiple sources, including clinical symptoms, endoscopic readouts, histopathology, and gene expression values, through AI can hold potential in predicting treatment outcomes and informing prognosis and therapeutic response.

The role of AI in clinical practice

The role of AI in clinical practice is expected to encompass education, enhance quality, and facilitate visualization. With the global rise in the number of patients with UC, colonoscopists are not always sufficiently equipped with expertise in UC. Furthermore, given the evolution of new endoscopic scoring systems and the growing sophistication of endoscopic technology, the proficiency expected from endoscopists has also increased. Histological scoring is in a similar situation. As the concept of T2T becomes more prevalent in routine practice, the existence of inter-examiner differences in endoscopic and pathological scores highlights the necessity of establishing definitive treatment goals. To tackle these issues, AI is anticipated to play a pivotal role as an educator, helping to improve physicians' ability.

In CAD systems for colorectal neoplasia, which have been already undergoing clinical implementation, reports have shed both positive and negative light on the use of AI in its characterization,^{60–62} and no conclusions have yet been reached. Unfortunately, the effectiveness of AI as an educator in the UC field has not yet been tested.

Next, the utilization of AI is anticipated for the quality improvement of colonoscopy.⁶³ Even if the AI serves as an accurate image reader, its capabilities may not be effectively demonstrated if the images acquired by the endoscopist are of low quality. The function to automatically extract scoring ineligible images has already been reported in several papers. Verification of whether real-time use of that function improves the quality of colonoscopy is required.

Lastly, regarding the visualization of endoscopic examinations: in conventional clinical decisions, physicians typically review reports provided by endoscopists, mentally visualize the distribution of inflammation, and formulate a treatment plan. However, there is a risk that both parties' interpretations may not always align. If computer vision, facilitated by AI, enables the visualization of the degree and distribution of inflammation, it will not only simplify the comprehension of pathophysiology but also serve as a robust tool for sharing the disease status with the entire medical team.

In conclusion, AI holds the potential to address several unmet needs in UC clinical practice, resulting in increased efficacy and objective, precise management of UC. Moreover, by integrating endoscopic, pathological, and clinical data through AI, it is anticipated that more detailed predictions of future relapse and assessments of treatment efficacy will become feasible. There is a strong hope for the expeditious launch of a product suitable for clinical practice.

Currently, the limitations and challenges of AI

Although numerous reports exist in the research field, only a few of them have received regulatory approval, resulting in a limited level of clinical implementation. Consequently, there is a scarcity of real-time validation for AI. The question of whether AI can achieve the anticipated accuracy in actual clinical practice remains unclear. Furthermore, the value added by the use of AI to clinician diagnoses has not been substantiated. For instance, it remains to be determined if employing pathology AI enhances the accuracy of pathologists' scoring. Additionally, the impact of AI on clinicians' judgments, who bear ultimate responsibility, has not been verified.

Crucially, it is imperative to ascertain whether AI contributes to improvements in patients' quality of life and long-term prognosis and whether it possesses the potential to influence the standard of care.

The primary concern associated with the utilization of AI in endoscopy and histology for UC lies in the interpretation of AI's diagnostic output by physicians. There is apprehension that physicians lacking an integrated understanding of various factors, including clinical symptoms, laboratory results, and patient history, may overly depend on AI's diagnostic output, potentially resulting in inadequate care. While AI has the potential to assist physicians in offering a more standardized diagnosis, especially in the case of UC patients, shared decision-making tailored to individual preferences and lifestyle habits is crucial. AI-based practices, masquerading under the banner of standardization, need to be vigilantly monitored to mitigate the risk of interference with optimal personalized medicine. Therefore, striking a balance between leveraging AI for uniform diagnoses and ensuring personalized care for UC patients remains a critical consideration.

In conclusion, the utilization of AI in the management of UC patients is an expanding field of research with high expectations. However, it is crucial to objectively acknowledge both the positive and negative implications that AI may introduce to UC management.

Future directions and conclusion: new algorithms of AI endoscopy and histology to assess and manage UC

The potential of automating the extraction of endoscopic disease activity, histologic indices, and disease phenotyping using ML image analysis applications is expected to revolutionize the standard of care for UC in the near future. AI can play a crucial role in defining remission by providing comprehensive assessments beyond human capabilities, such as quantifying vascular and mucosal patterns and subtle details. It can also assist in real-time histological evaluation, predicting histological remission, with positive implications for clinical trials and practice. Therefore, integrating AI with endoscopy and histology extends beyond diagnosis to predict disease trajectories and clinical outcomes in colitis and offers the promise of identifying new treatment target

goals leading to better long-term outcomes for patients.

Most studies employing AI in UC have been post hoc evaluations or single-center, prospective studies. While one of exception is the PICaSSO multicenter, international, prospective study in which PICaSSO endoscopy and histology AI were developed based on numerous AI endoscopy videos and histological slides in a large patient cohort.⁵³ In the future, the developed AI needs to be used in real-time in real clinical practice and randomized controlled trials are required to better clarify the benefits and limitations. In the context of clinical trials, an AI/ML algorithm could be used to select patients accurately and thus reduce reliance on human central readers. However, since the training of CNNs for specific findings depends on human judgments, AI currently cannot surpass humans, as humans serve as the point of reference. Besides, AI systems that integrate disease activity beyond human cognitive limitations would need validation through disease outcomes, both short-, mid-, and long-term. While AI has the potential to assist physicians in offering a more standardized diagnosis, especially in the case of UC patients, shared decision-making tailored to individual preferences and lifestyle habits remains crucial. Integrating large datasets from multiple sources, including clinical symptoms, endoscopic read-outs, histopathology, and gene expression values, through AI can provide further insight into IBD and lead to the discovery of novel biomarkers. This fusion of AI with new advanced endoscopy technologies and histology has the potential to redefine our understanding of disease progression, response assessment, and treatment planning. This will require the cooperation and commitment of endoscopists, specialists, and societies. Thus, in the future, new AI algorithms for the clinical management of IBD patients by integrating innovative endoscopic tools with AI systems and molecular endoscopy hold the potential to deliver precision medicine in IBD, turning what was once a vision into an achievable reality.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Author contributions

Olga Maria Nardone: Conceptualization; Validation; Visualization; Writing – original draft; Writing – review & editing.

Yasuharu Maeda: Conceptualization; Validation; Visualization; Writing – original draft; Writing – review & editing.

Marietta Iacucci: Conceptualization; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

Acknowledgements

None.

Disclaimer

The article processing charges for this publication have been waived under an educational grant from Takeda Pharmaceuticals Limited.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Competing interests

O.M.N.: lecture fees from Ferring, Abbvie, Janssen, Pfizer; Advisory Board from Nestle. M.I.: research grants and equipment loans from Pentax USA, Olympus, and Fujifilm. Y.M.: nothing to declare.

Availability of data and materials

Data sharing does not apply to this article as no new data were created or analyzed in this study.

ORCID iDs

Olga Maria Nardone  <https://orcid.org/0000-0002-9554-4785>

Marietta Iacucci  <https://orcid.org/0000-0002-3142-9550>

Supplemental material

Supplemental material for this article is available online.

References

1. Nardone OM, Cannatelli R, Zardo D, et al. Can advanced endoscopic techniques for assessment of mucosal inflammation and healing approximate histology in inflammatory bowel disease? *Therap Adv Gastroenterol* 2019; 12: 1756284819863015.

2. Gupta A, Yu A, Peyrin-Biroulet L, et al. Treat to target: the role of histologic healing in inflammatory bowel diseases: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2021; 19: 1800–1813.e4.
3. Nardone OM, Bazarova A, Bhandari P, et al. PICaSSO virtual electronic chromendoscopy accurately reflects combined endoscopic and histological assessment for prediction of clinical outcomes in ulcerative colitis. *United European Gastroenterol J* 2022; 10: 147–159.
4. Magro F, Pai RK, Kobayashi T, et al. Resolving histological inflammation in ulcerative colitis with mirikizumab in the LUCENT induction and maintenance trial programmes. *J Crohns Colitis* 2023; 17: 1457–1470.
5. Sands BE, Sandborn WJ, Panaccione R, et al. Ustekinumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2019; 381: 1201–1214.
6. Yarur AJ, Chiorean MV, Panés J, et al. Achievement of clinical, endoscopic, and histologic outcomes in patients with ulcerative colitis treated with etrasimod, and association with fecal calprotectin and C-reactive protein: results from the phase 2 OASIS trial. *J Crohns Colitis* 2024; 18(6): 885–894.
7. Yoon H, Jangi S, Dulai PS, et al. Incremental benefit of achieving endoscopic and histologic remission in patients with ulcerative colitis: a systematic review and meta-analysis. *Gastroenterology* 2020; 159: 1262–1275.e7.
8. Nardone OM, Cannatelli R, Ghosh S, et al. New endoscopic tools in inflammatory bowel disease. *United European Gastroenterol J* 2022; 10: 1103–1112.
9. D’Amico F, Peyrin-Biroulet L and Danese S. Disease clearance in ulcerative colitis: is the ultimate therapeutic target? *United European Gastroenterol J* 2023; 11: 717–719.
10. Zammarchi I, Santacroce G and Iacucci M. Next-generation endoscopy in inflammatory bowel disease. *Diagnostics (Basel)* 2023; 13: 2547.
11. Le Berre C and Peyrin-Biroulet L; SPIRIT-IOIBD Study Group. Selecting end points for disease-modification trials in inflammatory bowel disease: the SPIRIT consensus from the IOIBD. *Gastroenterology* 2021; 160: 1452–1460.e21.
12. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the selecting therapeutic targets in inflammatory bowel disease (STRIDE) initiative of the international organization for the study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. *Gastroenterology* 2020; 160: 1570–1583.
13. Travis SP, Schnell D, Krzeski P, et al. Reliability and initial validation of the ulcerative colitis endoscopic index of severity. *Gastroenterology* 2013; 145: 987–995.
14. Sasaki Y, Hada R and Munakata A. Computer-aided grading system for endoscopic severity in patients with ulcerative colitis. *Dig Endosc* 2003; 15: 206–209.
15. Ozawa T, Ishihara S, Fujishiro M, et al. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc* 2019; 89: 416–421.e1.
16. 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.
17. Takenaka K, Ohtsuka K, Fujii T, et al. Development and validation of a deep neural network for accurate evaluation of endoscopic images from patients with ulcerative colitis. *Gastroenterology* 2020; 158: 2150–2157.
18. Lo B, Liu Z, Bendtsen F, et al. High accuracy in classifying endoscopic severity in ulcerative colitis using convolutional neural network. *Am J Gastroenterol* 2022; 117: 1648–1654.
19. Gottlieb K, Requa J, Karnes W, et al. Central reading of ulcerative colitis clinical trial videos using neural networks. *Gastroenterology* 2021; 160: 710–719.e2.
20. Yao H, Najarian K, Gryak J, et al. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2021; 93: 728–736.e1.
21. Takenaka K, Fujii T, Kawamoto A, et al. Deep neural network for video colonoscopy of ulcerative colitis: a cross-sectional study. *Lancet Gastroenterol Hepatol* 2021; 7: 230–237.
22. Yao H, Stidham RW, Soroushmehr R, et al. Automated detection of non-informative frames for colonoscopy through a combination of deep learning and feature extraction. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019 Jul 23 (pp. 2402–2406). IEEE.
23. Byrne MF, Panaccione R, East JE, et al. Application of deep learning models to improve ulcerative colitis endoscopic disease activity scoring under multiple scoring systems. *J Crohns Colitis* 2023; 17: 463–471.
24. Sandborn WJ, Ferrante M, Bhandari BR, et al. Efficacy and safety of mirikizumab in a randomized phase 2 study of patients with

- ulcerative colitis. *Gastroenterology* 2020; 158: 537–549.e10.
25. Lobaton T, Bessissow T, De Hertogh G, et al. The Modified Mayo Endoscopic Score (MMES): a new index for the assessment of extension and severity of endoscopic activity in ulcerative colitis patients. *J Crohns Colitis* 2015; 9: 846–852.
 26. Kaneshiro M, Takenaka K, Suzuki K, et al. Pancolonoscopic and histologic evaluation for relapse prediction in patients with ulcerative colitis in clinical remission. *Aliment Pharmacol Ther* 2021; 53: 900–907.
 27. Takabayashi K, Kobayashi T, Matsuoka K, et al. Artificial intelligence quantifying endoscopic severity of ulcerative colitis in gradation scale. *Dig Endosc* 2023; 36: 582–590.
 28. Fan Y, Mu R, Xu H, et al. Novel deep learning-based computer-aided diagnosis system for predicting inflammatory activity in ulcerative colitis. *Gastrointest Endosc* 2023; 97: 335–346.
 29. Stidham RW, Cai L, Cheng S, et al. Using computer vision to improve endoscopic disease quantification in therapeutic clinical trials of ulcerative colitis. *Gastroenterology* 2024; 166: 155–167.e2.
 30. Maeda Y, Kudo SE, Ogata N, et al. Use of advanced endoscopic technology for optical characterization of neoplasia in patients with ulcerative colitis: systematic review. *Dig Endosc* 2022; 34: 1297–1310.
 31. Santacrose G, Zammarchi I, Tan CK, et al. Present and future of endoscopy precision for inflammatory bowel disease. *Dig Endosc* 2024; 36: 292–304.
 32. Nardone OM, Iacucci M, Villanacci V, et al. Real-world use of endoscopic and histological indices in ulcerative colitis: results of a global survey. *United European Gastroenterol J* 2023; 11: 514–519.
 33. Bossuyt P, Nakase H, Vermeire S, et al. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut* 2020; 69: 1778–1786.
 34. Bossuyt P, De Hertogh G, Eelbode T, et al. Computer-aided diagnosis with monochromatic light endoscopy for scoring histologic remission in ulcerative colitis. *Gastroenterology* 2021; 160: 23–25.
 35. Iacucci M, Smith SC, Bazarova A, et al. An international multicenter real-life prospective study of electronic chromoendoscopy score PICaSSO in ulcerative colitis. *Gastroenterology* 2021; 160: 1558–1569.e8.
 36. Iacucci M, Daperno M, Lazarev M, et al. Development and reliability of the new endoscopic virtual chromoendoscopy score: the PICaSSO (Paddington International Virtual ChromoendoScopy ScOre) in ulcerative colitis. *Gastrointest Endosc* 2017; 86: 1118–1127.e5.
 37. Iacucci M, Cannatelli R, Parigi TL, et al. A virtual chromoendoscopy artificial intelligence system to detect endoscopic and histologic activity/remission and predict clinical outcomes in ulcerative colitis. *Endoscopy* 2023; 55: 332–341.
 38. Maeda Y, Ohtsuka K, Kudo SE, et al. Endocytoscopic narrow-band imaging efficiency for evaluation of inflammatory activity in ulcerative colitis. *World J Gastroenterol* 2015; 21: 2108–2115.
 39. Kudo SE, Misawa M, Wada Y, et al. Endocytoscopic microvasculature evaluation is a reliable new diagnostic method for colorectal lesions (with video). *Gastrointest Endosc* 2015; 82: 912–923.
 40. 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.
 41. Maeda Y, Kudo S-E, Ogata N, et al. Endocytoscopic intramucosal capillary network changes and crypt architecture abnormalities can predict relapse in patients with an ulcerative colitis Mayo endoscopic score of 1. *Dig Endosc* 2020; 32: 1082–1091.
 42. Takishima K, Maeda Y, Ogata N, et al. Beyond complete endoscopic healing: goblet appearance using an endocytoscope to predict future sustained clinical remission in ulcerative colitis. *Dig Endosc* 2022; 34: 1030–1039.
 43. Maeda Y, Kudo SE, Mori Y, et al. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; 89: 408–415.
 44. Kuroki T, Maeda Y, Kudo S-E, et al. A novel artificial intelligence-assisted “vascular-healing” diagnosis for prediction of future clinical relapse in patients with ulcerative colitis: a prospective cohort study. *Gastrointest Endosc* 2024; 100: 97–108.
 45. Akiyama S, Sakamoto T, Kobayashi M, et al. Clinical usefulness of hypoxia imaging colonoscopy for the objective measurement of ulcerative colitis disease activity. *Gastrointest Endosc* 2024; 99: 1006–1016.e4.
 46. Takenaka K, Ohtsuka K, Fujii T, et al. Deep neural network accurately predicts prognosis

- of ulcerative colitis using endoscopic images. *Gastroenterology* 2021; 160: 2175–2177.e3.
47. Maeda Y, Kudo SE, Ogata N, et al. Evaluation in real-time use of artificial intelligence during colonoscopy to predict relapse of ulcerative colitis: a prospective study. *Gastrointest Endosc* 2022; 95: 747–756.e2.
 48. Niazi MKK, Parwani AV and Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol* 2019; 20: e253–e261.
 49. Vande Castele N, Leighton JA, Pasha SF, et al. Utilizing deep learning to analyze whole slide images of colonic biopsies for associations between eosinophil density and clinicopathologic features in active ulcerative colitis. *Inflamm Bowel Dis* 2022; 28: 539–546.
 50. Ohara J, Nemoto T, Maeda Y, et al. Deep learning-based automated quantification of goblet cell mucus using histological images as a predictor of clinical relapse of ulcerative colitis with endoscopic remission. *J Gastroenterol* 2022; 57: 962–970.
 51. Najdawi F, Sucipto K, Mistry P, et al. Artificial intelligence enables quantitative assessment of ulcerative colitis histology. *Mod Pathol* 2023; 36: 100124.
 52. Del Amor R, Meseguer P, Parigi TL, et al. Constrained multiple instance learning for ulcerative colitis prediction using histological images. *Comput Methods Programs Biomed* 2022; 224: 107012.
 53. Iacucci M, Parigi TL, Del Amor R, et al. Artificial intelligence enabled histological prediction of remission or activity and clinical outcomes in ulcerative colitis. *Gastroenterology* 2023; 164: 1180–1188.e2.
 54. Gui X, Bazarova A, Del Amor R, et al. PICaSSO Histologic Remission Index (PHRI) in ulcerative colitis: development of a novel simplified histological score for monitoring mucosal healing and predicting clinical outcomes and its applicability in an artificial intelligence system. *Gut* 2022; 71: 889–898.
 55. Rymarczyk D, Schultz W, Borowa A, et al. Deep learning models capture histological disease activity in Crohn's disease and ulcerative colitis with high fidelity. *J Crohns Colitis* 2024; 18: 604–614.
 56. Liu X, Prasath S, Siddiqui I, et al. Machine learning-based prediction of pediatric ulcerative colitis treatment response using diagnostic histopathology. *Gastroenterology* 2024; 166: 921–924.e4.
 57. Ahmad HA, East JE, Panaccione R, et al. Artificial intelligence in inflammatory bowel disease endoscopy: implications for clinical trials. *J Crohns Colitis* 2023; 17: 1342–1353.
 58. Raine T and Danese S. Breaking through the therapeutic ceiling: what will it take? *Gastroenterology* 2022; 162: 1507–1511.
 59. Feagan BG, Sandborn WJ, D'Haens G, et al. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013; 145: 149–157.e2.
 60. Barua I, Wieszczy P, Kudo S-E, et al. Real-time artificial intelligence-based optical diagnosis of neoplastic polyps during colonoscopy. *NEJM Evid* 2022; 1: EVID0a2200003.
 61. Rondonotti E, Hassan C, Tamanini G, et al. Artificial intelligence-assisted optical diagnosis for the resect-and-discard strategy in clinical practice: the Artificial intelligence BLI Characterization (ABC) study. *Endoscopy* 2023; 55: 14–22.
 62. Kato S, Kudo S-E, Minegishi Y, et al.; Collaborators. Impact of computer-aided characterization for diagnosis of colorectal lesions, including sessile serrated lesions: multireader, multicase study. *Dig Endosc* 2024; 36: 341–350.
 63. Dekker E, Nass KJ, Iacucci M, et al. Performance measures for colonoscopy in inflammatory bowel disease patients: European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscopy* 2022; 54: 904–915.

Appendix

Abbreviations

AI	artificial intelligence
CAD	computer-aided diagnosis
CDS	cumulative disease score
CNN	convolution neural network
DNUC	deep neural network for evaluation of ulcerative colitis
IBD	inflammatory bowel disease
MES	Mayo endoscopic subscore
MH	mucosal healing
ML	machine learning
T2T	treat-to-target
UC	ulcerative colitis
UCEIS	ulcerative colitis endoscopic index of severity
WSIs	whole-slide images