Artificial intelligence in inflammatory bowel disease endoscopy: current landscape and the road ahead

Suneha Sundaram iD**, Tenzin Choden, Mark C. Mattar, Sanjal Desai and Madhav Desai** iD

Abstract: Inflammatory bowel disease is a complex chronic inflammatory disorder with challenges in diagnosis, choosing appropriate therapy, determining individual responsiveness, and prediction of future disease course to guide appropriate management. Artificial intelligence has been examined in the field of inflammatory bowel disease endoscopy with promising data in different domains of inflammatory bowel disease, including diagnosis, assessment of mucosal activity, and prediction of recurrence and complications. Artificial intelligence use during endoscopy could be a step toward precision medicine in inflammatory bowel disease care pathways. We reviewed available data on use of artificial intelligence for diagnosis of inflammatory bowel disease, grading of severity, prediction of recurrence, and dysplasia detection. We examined the potential role of artificial intelligence enhanced endoscopy in various aspects of inflammatory bowel disease care and future perspectives in this review.

Keywords: artificial intelligence, endoscopy, inflammatory bowel disease

Correspondence to:
Madhav Desai
Assistant Professor of Clinical Medicine, University of Kansas School of Medicine, Kansas City VA Medical Center, 4801 Linwood Blvd, Kansas City, MO 64128, USA

Division of Gastroenterology, Hepatology and Motility, Department of Internal Medicine, University of Kansas School of Medicine, Kansas City, KS, USA

Department of Gastroenterology, Kansas City VA Medical Center, Kansas City, MO, USA
drmadhavdesai@gmail. com

Suneha Sundaram
Department of Gastroenterology, Kansas City VA Medical Center, Kansas City, MO, USA

Tenzin Choden
Mark C. Mattar
Division of Gastroenterology, Department of Medicine, MedStar Georgetown University Hospital, Washington, DC, USA

Sanjal Desai
Department of Hematology, Mayo Clinic, Rochester, MN, USA

Introduction

In their leading textbook, 'Artificial Intelligence: A Modern Approach', experts in artificial intelligence (AI), Stuart Russell and Peter Norvig, identify four possible goals of AI systems that think like humans, act like humans, think within the ideal concept of intelligence (i.e. rationally), or act rationally.[1] The scientific discipline of AI can be classified into 'Computer Vision' (image recognition and classification), 'Natural Language Processing' (speech-to-text generation and translation), and 'Machine Learning', which is the most commonly employed type in medicine.[2] The objective of machine learning (ML) is to build an algorithm based on inputted data with the power to predict new data, such as future outcomes. Processes of ML include supervised learning, unsupervised learning, and deep learning.[3,4] In supervised learning, the algorithm is trained on labeled data points. The algorithm learns the relationship between the known data and the correct outcomes. It then predicts future outcomes of unforeseen data. On other hand, unsupervised learning is evaluating for patterns based on available data. The algorithm learns the inherent structure/relationships within inputted data without the supervision of prescribed labels and outcomes.[1,4] Finally, the deep learning, specifically deep/convoluted neural networks (DNN/CNN), works by using layers of algorithms that are interconnected and receives weighted input from earlier layers to recognize patterns and ultimately capture complex relations within the data. In this way, these DNN function much like the mammalian brain.[4]

AI in medicine, specifically in gastroenterology, has been of growing interest with exciting developments in the last couple of decades. Two major applications of ML are computer-aided detection (CADe) for detection of a pathology (e.g. polyp identification) and computer-aided diagnosis (CADx) for predicting diagnosis (e.g. polyp classification).[4]
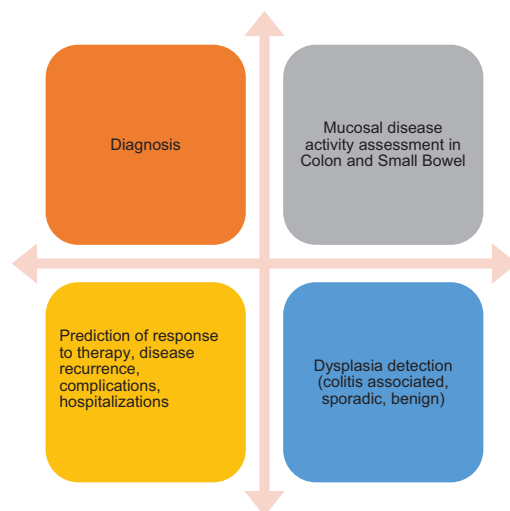
Figure 1. Potential AI Applications during Endoscopy in IBD.
AI, artificial intelligence; IBD, inflammatory bowel disease.

Previously, the use of AI in gastroenterology has largely focused on CADe with polyp detection during screening or surveillance colonoscopy with advancement to CADx in predicting polyp pathology to aid in cancer diagnosis.[2,5–7]

Another area of interest is endoscopy in inflammatory bowel disease (IBD) where there is significant interobserver variation in disease diagnosis, endoscopic activity assessment, dysplasia detection, and prognostic assessment of the disease course. Due to significant heterogeneity in presentation, disease course, treatment response, and provider assessment, AI offers objectifying the endoscopic assessment of inflammation, lesion detection, and possible role in prognostication of disease recurrence. Potential applications of AI in IBD include use in diagnosis, identifying mucosal disease activity assessment, prediction of response to therapy/recurrence/complications/hospitalizations, and dysplasia detection (Figure 1). This review will focus on the technologies currently being studied in the IBD and possible future applications. We will review published data on AI in disease diagnosis by endoscopy, staging or activity assessment, and lesion detection (summarized in Table 1). Given the evolving vastness of the role of AI in other aspects of IBD care including IBD patient education, electronic health record linkage using natural language processing (NLP), and pattern recognition during abdominal imaging, we have limited our discussion to AI application during IBD endoscopy for the purpose of this review.

Diagnosis of IBD

The diagnosis of IBD is a clinicopathological diagnosis that involves a combination of a thorough clinical history, imaging, endoscopy, histopathology as well as serum markers.[8,9] It can be a difficult diagnosis to make, even for gastrointestinal (GI) specialists and pathologists. The accuracy of endoscopic evaluation is based on the single subjective observation and biopsy retrieval of the endoscopist performing it. Most endoscopy technologies currently employ high-definition white light endoscopy (HD-WLE) with optical enhancing chromoendoscopy options. Therefore, the challenge in endoscopic diagnosis is primarily interpretation and less so visualization.[4] A poor-quality endoscopy can lead to a delayed diagnosis and severe complications.[10] There is documented interobserver variation in pathologist diagnosis of IBD and then furthermore in differentiating between ulcerative colitis (UC) and Crohn's disease (CD), even among specialized GI pathologists who are not available in most clinical centers.[9] For these reasons, serological, genetic, and inflammatory markers have become a focus of study for diagnosing IBD, differentiation of UC and CD, and grading severity of inflammation. A combination of all these techniques is the current practice for diagnosing both UC and CD.[8]

To aid in diagnosis for IBD and many other disorders, Thakkar and colleagues[11] developed a CADe system to detect the quality of endoscopy and present this interpretable feedback to the endoscopist in real time. The system yields a score for each of the four quality metrics (visible surface area, opened/distended colon, preparation conditions, and clarity of current view). When compared with scores given by expert endoscopists, this AI-driven system shows comparable performance. This system was used in 10 real-time colonoscopies to prove its feasibility. The quality of endoscopies is highly dependent on the expertise of the endoscopist performing them. Furthermore, the many AI systems aimed at improving diagnosis (e.g. polyp/lesion detection) and treatment (e.g. identifying inflammatory tissue to determine remission *versus* further treatment in UC) depend on the quality of a procedure to identify a specific abnormality. Systems like this can provide the clinical team with an objective measure of confidence in the quality of a procedure, ultimately, enhancing intraprocedural diagnostic precision.

Table 1. Brief Overview of Selected Prior Studies Examining Utility of AI in Various Aspects of IBD Management.

Primary author	Year	Technology	Potential application/ Objective	Findings
Thakkar	2020	Colonoscopy	Assessment of quality of colonoscopy examinations in real time	Produced score for 4 quality metrics (visible surface area, opened/distended colon, preparation conditions, and clarity of current view) and used in real-time endoscopy showing feasibility
Mossotto	2017	Endoscopic images and histology	Classification of pediatric IBD: UC *versus* CD at diagnosis	83.3% accuracy when classifying between CD and UC using endoscopic images and histology from diagnostic endoscopy
Kumar	2012	Capsule endoscopy images	Identify CD lesions	Identified active CD lesions with 91% precision but only 79% accuracy for classification of lesion severity
Girgis	2010	Capsule endoscopy images	Detection of CD inflammation	87% accuracy in identifying regions of inflammation compared with expert review
Barash	2021	Capsule endoscopy images	Classification of CD ulcer severity	System achieved 67% agreement for grade 1 and 91% agreement for identifying grade 3 ulcer compared with 2 expert endoscopists; classification accuracy of algorithm 0.91 (CI: 0.867–0.954) for grade 1 *versus* 3 ulcers.
Takenaka	2020	Endoscopic images and histology	Predicting UC remission using endomicroscopy	System predicted endoscopic remission with 90% accuracy, $\kappa = 0.80$ and histological remission with 93% accuracy, $\kappa = 0.86$
Bossuyt	2020	Endoscopic images and histology	Predicting UC remission/ inflammation using endomicroscopy	System produced a Red Density score with $r = 0.65$ correlation to the Robarts histological index
Ozawa	2019	Endoscopic images	Grading UC disease severity	System distinguished disease in remission *versus* moderate/severe disease with AUROC of 0.98 based on the Mayo score
Gottlieb	2020	Endoscopy video	Grading UC disease severity	Produced accurate Mayo scores and UCEIS scores with agreement/reproducibility of $\kappa = 0.84$ and 0.85, respectively
Stidham	2019	Endoscopic images	Grading UC disease severity	System distinguished disease in remission *versus* moderate/severe disease using Mayo scores with AUROC of 0.97 and agreement to human reviewer scores, $\kappa = 0.86$
Yao	2020	Endoscopic images	Grading UC disease severity	Overall system accuracy, 84%. Improved agreement with human scores using high-definition images
Bhambhvani	2020	Endoscopic images	Classification of UC severity	System accurately assigned Mayo scores of 1, 2, or 3 with an AUC of 0.96, 0.86, and 0.89, respectively; system performed with an average specificity of 85.7% and average sensitivity of 72.4%
Maeda	2019	Endoscopic images and histology	Detection of UC inflammation using endocytoscopy	Per segment classification, system performed with 91% accuracy, 97% specificity, and 74% sensitivity
Khandiah	2020	Colonoscopy and red light emission	Neoplasia detection using chromoendoscopy	No difference in neoplasia detection between high-definition white light endoscopy and new virtual chromoendoscopy system
Selaru	2002	Complementary DNA	Differentiating IBD-related dysplasia and spontaneous colorectal adenomas	System correctly diagnosed 12/12 cases in validation set with 0.999 regression factor

AI, artificial intelligence; AUC, area under the curve; AUROC, area under the receiver operating curve; CD, Crohn's disease; CI, confidence interval; IBD, inflammatory bowel disease; UC, ulcerative colitis; UCEIS, Ulcerative Colitis Endoscopic Index of Severity.

For the purpose of diagnosing and further differentiating between UC and CD, Mossotto and colleagues developed a CADx system using supervised and unsupervised learning in pediatric patients. The training data set used endoscopic images and histopathology data, from previously diagnosed patients, to build an algorithm that could classify CD *versus* UC. On the validation set, the combined DNN model classified CD and UC at diagnosis with 83.3% accuracy.[12] The importance of distinguishing CD and UC at diagnosis cannot be underemphasized because treatment options and follow-up differ between these diagnoses. A notable difference is the surveillance for dysplasia detection and subsequent intervention. Due to the significantly increased risk of colon cancer in UC, colectomy is often recommended if high-grade dysplasia is detected.[9] This system begins to show the potential for AI in diagnosis of IBD; however, integration of serum markers and application in real time will be future developments that are necessary for this complex diagnostic process.

Detection of inflammation

An early area of study was the use of ML for detection of inflammatory lesions in IBD patients including the areas that are beyond the reach of ileocolonoscopy. Systems that have been developed to do this use video capsule endoscopy and endomicroscopy technologies. While systems that can interpret live endoscopy images have not yet been developed, the DNN technology described here used with video capsule endoscopy and endomicroscopy may theoretically be used in the real-time setting.

Video capsule endoscopy

In the early 2010s, Kumar and colleagues used a supervised learning model to classify lesions on video capsule endoscopy images. Their system detected CD lesions with 91% precision but only 79% accuracy for classification of lesion severity.[13] Girgis and colleagues[14] also used capsule endoscopy to detect regions of inflammation in CD patients with an accuracy of 87%. Most recently, in 2021, Barash and colleagues demonstrated that their DNN technology which was trained first to identify CD lesions *versus* normal mucosa could also grade the severity of CD ulcers with impressive accuracy. The system performed with the highest classification accuracy [0.91, confidence interval (CI): 0.716–0.844] for distinguishing grade 1 *versus* grade 3 ulcers.[15,16] These CADe systems show evidence that AI can be used with video capsule endoscopy to help decrease manpower and time spent reviewing these images while improving subjective human accuracy.

Endomicroscopy

A novel frontier is the field of endomicroscopy that allows histologic observation during endoscopy without requiring biopsies. Using a multidimensional system that incorporates endoscopic visualization and histological evaluation is important for the diagnosis of IBD.

A DNN model developed by Takenaka and colleagues in Japan is built on data from both endoscopic images and histopathology from patients with UC. The team used the Ulcerative Colitis Endoscopic Index of Severity (UCEIS) and Geboes score to define endoscopic and histologic remission, respectively. In the validation phase, this study found that the model predicted endoscopic remission with 90% accuracy and had a kappa score of 0.80 when compared with experienced reviewer scores. Histological remission was identified with 93% accuracy and reproducibility kappa score of 0.86 compared with pathologist readings.[17] This innovative system shows the potential to avoid biopsies in UC patients in remission, decreasing healthcare expenditure in this chronic disease for both the system (procedural costs and pathologist compensation) and patient.

Bossuyt and colleagues created a Red Density (RD) score as an endomicroscopic evaluation of UC activity. This DNN was built on algorithmic data from redness color map and vascular patterns on endoscopic images that were then linked to the Robarts histological index (RHI). The system performed with statistically significant correlation to RHI in the validation set ($r = 0.65$).[18] As discussed, this endomicroscopy system allows for computer-aided evaluation of patient inflammation *versus* remission without the need for extensive biopsies, and the score produced presents this information in a functionally objective form.

Scoring disease severity in UC

Multiple studies have found promising evidence that deep learning algorithms have the potential to predict severity of disease in UC. Systems are being trained to produce Mayo Endoscopic Disease Activity Assessment score and UCEIS

scores based on endoscopic images. This has the clinical benefit of assessing the effectiveness of a current treatment for patients and determining appropriate future treatments. Given the poor interobserver agreement for these endoscopic scores, a DNN system would reduce human error, allow for more timely treatments, and improve a patient's quality of life.[19] We will discuss these systems based on the imaging technologies.

Endoscopic images

Ozawa and colleagues developed a complex neural network system to label endoscopic mucosal images with Mayo endoscopic scores to evaluate disease activity. The system labeled images with Mayo scores of 0, 1, or 2–3 with an area under the receiver operating curve (AUROC) of 0.98. This system can be used to differentiate if a patient with UC is in endoscopic remission (defined as Mayo score: 0–1) or has active inflammation (Mayo score: 2–3). Ozawa and colleagues[20] acknowledged the necessity of using this system in real time and claimed that the DNN processing time was less than 30 ms which is fast enough for use during a colonoscopy.

Bhambhvani and Zamora recently demonstrated that their DNN system could go further and accurately classify UC lesion severity by specific Mayo scores of 1, 2, or 3 with an area under the curve (AUC) of 0.96, 0.86, and 0.89, respectively. The system performed with an average specificity of 85.7% and average sensitivity of 72.4%.[21] This expansion of AI from identification of inflammation to classification of severity is a promising advancement for patients with complex and UC, requiring detailed distinction between their levels of active disease.

Numerous DNN models have also been built based on data from full endoscopy videos. Gottlieb and colleagues used both Mayo and UCEIS scores to compare the performance of a recurrent neural network model with human reader scores with very good reproducibility, $\kappa = 0.84$ for Mayo scoring and $\kappa = 0.85$ for UCEIS scoring. A DNN system by Stidham and colleagues[22] is capable of distinguishing UC in remission *versus* moderate/severe disease state using the Mayo score with an AUROC of 0.97. This model had comparable agreement to human reviewer scores, $\kappa = 0.86$. Yao and colleagues built a similar model and investigated its performance both on high- and

low-resolution endoscopic videos. Their system performed on par with the aforementioned studies with the same aims of grading UC severity. They found improved agreement of the model and human scores with high-resolution videos ($\kappa = 0.84$) and an overall system accuracy of 84%. This highlights the need for standardization of endoscopy digitization if AI is to be used more broadly in both clinical trial and clinical treatment settings.[23]

Maeda and colleagues[19] constructed a CADx system to predict persistent histological inflammation in patients with UC using endocytoscopy technique, which uses a specialized endocytoscope that have a forward-facing microscope capable of $1000\times$ real-time magnification compared with the traditional $50\times$ of a standard endoscope. In the cohort of 187 patients with UC, when scoring per segment for histological active inflammation *versus* healing, the system performed with 91% accuracy, 97% specificity, and perfect agreement with histologist Geboes scoring ($\kappa = 1$), but only 74% sensitivity.[24] Because additional indications such as dysplasia surveillance necessitate endoscopic biopsy specimens regardless, lower sensitivity may be tolerated, given the exceptional specificity and agreement regarding histological remission.

Dysplasia detection

Current CADe/CADx systems aimed at the detection and differentiation of colon polyps/lesions have shown remarkable data for colon cancer screening studies, but this technology remains to be applied as broadly for dysplasia surveillance in IBD patients. Theoretically, the same technology differentiating different types of polyps could be used to survey for visible dysplasia. In addition, virtual chromoendoscopy (VCE) is another technique that has recently been of interest for dysplasia surveillance in patients with IBD. Few recent studies have evaluated the potential role for VCE in the identification of dysplastic lesions. A multicenter randomized control trial by Khandiah and colleagues[25] compared the performance of high-definition VCE (i-scan OE mode 2) with HD-WLE, showing comparable detection of neoplasia. El-Dallal and colleagues recently published a meta-analysis looking at the comparative efficacy of VCE compared with HD-WLE or traditional chromoendoscopy for dysplasia detection in higher risk IBD patients. When looking at per patient analysis, there were no statistical differences between VCE, traditional

chromoendoscopy, and HD-WLE.[26] Although still being investigated, VCE could be an additional modality that could be used by AI systems to target dysplasia in IBD.

In an alternate approach to AI outside of endoscopy, Selaru and colleagues demonstrate the potential for AI in dysplasia detection for patients with long-standing colitis using tissue specimens. This group trained a DNN system using complementary DNA to distinguish between IBD-related dysplasia and sporadic colorectal adenomas.[27] Notably, as it requires tissue samples, this system still necessitates colonoscopy with biopsies of any suspicious lesions. The decrease in human error and manpower for tissue analysis and diagnosis are benefits such DNN-based technology improving patient outcomes. There are several challenges to address when implementing this technology in clinical practice for IBD patients including the ability to differentiate pseudopolyps from true polyps and detection of abnormal flat lesions that are common in long-standing colitis.

Limitations
While AI is not ready for prime time in IBD endoscopy and the field is overall evolving with no system ready to be used in daily practice, there are certainly growing interests in various systems from different organizations. While competition for the best system is important, incorporation of such DNN that can perform diagnosis, differentiation, and detection of lesions in the same system, although ideal, is challenging. Another concern with the use of AI in gastroenterology is in the form of ethics. How independently can and should these systems perform? What level of responsibility for negative outcomes can and should be placed on the system *versus* a physician (i.e. complications from procedures, missed diagnoses)? As science progresses, these questions will need to be answered for the safety and comfort of patients. Similarly, in IBD endoscopy, before availability of these systems at the bedside, multicenter prospective trials with comparison against current standard of care followed by rigorous regulatory approval are required.

While these AI systems show promising evidence in validation sets, many of these current AI technologies have not yet been integrated into current patient care. Therefore, there is a lack of data with regard to patient outcomes. AI technologies applied to detecting inflammation and grading severity should improve patient experience during a procedure and decrease procedure time. A major improvement in patient outcome would be that the use of AI would decrease time required for data interpretation/time to diagnosis and treatment significantly enough to improve patients' quality of life. Ultimately, the optimal goal of AI use in IBD patients is improving the clinical course of the disease by earlier diagnosis, leading to improved disease burden, quality of life, and possible decrease or early detection of dysplasia.

Potential applications and future
The benefits and potential for the use of AI in IBD diagnosis and treatment as well as clinical trials are vast. Computer systems processing is undeniably faster than that of its human counterparts. Decreasing the need for biopsies can contribute to overall savings in healthcare costs. Finally, the potential to improve interobserver agreement in endoscopic disease scoring between gastroenterologists is possibly the most exciting aim of ML. Coupling CAD systems with physician observation and clinical gestalt amplifies the positive impact and illuminates an exciting future in medicine.

What is on the horizon for AI in IBD? Endomicroscopy is clearly an exciting new technology with promising results; however, the system is not available widely and requires a steep learning curve. Future studies in AI systems that incorporate clinical information such as biomarkers and symptoms may further the use of DNN in the multifactorial diagnosis of IBD. Finally, the current and future technologies should be applied to live colonoscopies to fully utilise their potential. Live CADe and CADx systems will provide the most clinically significant benefits to patients and endoscopists, so the next step includes using these newest technologies in clinical trials., The ideal AI system would help improve all aspects of endoscopy-related care in IBD including detection and characterization of inflammation, mucosal healing, recurrence pattern, and detection of any concerning lesion in the same AI system helping endoscopists and patients from one-step assessment.

Conflict of interest statement
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References
1. Russell SJ and Norvig P. *Artificial intelligence: a modern approach*. London: Pearson, 2021.

2. Fernández-Esparrach G, Bernal J, López-Cerón M, *et al*. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy* 2016; 48: 837–842.

3. Deo RC. Machine learning in medicine: will this time be different? *Circulation* 2020; 142: 1521–1523.

4. van der Sommen F, de Groof J, Struyvenberg M, *et al*. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. *Gut* 2020; 69: 2035–2045.

5. Urban G, Tripathi P, Alkayali T, *et al*. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 2018; 155: 1069–1078.

6. Byrne MF, Chapados N, Soudan F, *et al*. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2019; 68: 94–100.

7. Mori Y, Kudo S-E, Misawa M, *et al*. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Ann Intern Med* 2018; 169: 357–366.

8. Plevy S, Silverberg MS, Lockton S, *et al*. Combined serological, genetic, and inflammatory markers differentiate non-IBD, Crohn's disease, and ulcerative colitis patients. *Inflamm Bowel Dis* 2013; 19: 1139–1148.

9. Farmer M, Petras RE, Hunt LE, *et al*. The importance of diagnostic accuracy in colonic inflammatory bowel disease. *Am J Gastroenterol* 2000; 95: 3184–3188.

10. van der Laan JJH, van der Waaij AM, Gabriëls RY, *et al*. Endoscopic imaging in inflammatory bowel disease: current developments and emerging strategies. *Expert Rev Gastroenterol Hepatol* 2021; 15: 115–126.

11. Thakkar S, Carleton NM, Rao B, *et al*. Use of artificial intelligence-based analytics from live colonoscopies to optimize the quality of the colonoscopy examination in real time: proof of concept. *Gastroenterology* 2020; 158: 1219–1221.

12. Mossotto E, Ashton JJ, Coelho T, *et al*. Classification of paediatric inflammatory bowel disease using machine learning. *Sci Rep* 2017; 7: 2427.

13. Kumar R, Zhao Q, Seshamani S, *et al*. Assessment of Crohn's disease lesions in wireless capsule endoscopy images. *IEEE Trans Biomed Eng* 2012; 59: 355–362.

14. Girgis HZ, Mitchell BR, Dassopoulos T, *et al*. An intelligent system to detect Crohn's disease inflammation in Wireless Capsule Endoscopy videos. In: *2010 IEEE international symposium on biomedical imaging: from Nano to macro*, Rotterdam, 14–17 April 2010, pp. 1373–1376. New York: IEEE.

15. Klang E, Barash Y, Margalit RY, *et al*. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc* 2020; 91: 606–613.

16. Barash Y, Azaria L, Soffer S, *et al*. Ulcer severity grading in video capsule images of patients with Crohn's disease: an ordinal neural network solution. *Gastrointest Endosc* 2021; 93: 187–192.

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. 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.

19. 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.

20. 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.

21. Bhambhvani HP and Zamora A. Deep learning enabled classification of Mayo endoscopic subscore in patients with ulcerative colitis. *Eur J Gastroenterol Hepatol* 2020; 33: 645–649.

22. 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.

23. Yao H, Najarian K, Gryak J, *et al*. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2020; 93: 728.e1–736.e1.

24. 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.

25. Kandiah K, Subramaniam S, Thayalasekaran S, *et al*. Multicentre randomised controlled trial on virtual chromoendoscopy in the detection of neoplasia during colitis surveillance high-definition colonoscopy (the VIRTUOSO trial). *Gut*. Epub ahead of print 19 November 2020. DOI: 10.1136/gutjnl-2020-320980.

26. El-Dallal M, Chen Y, Lin Q, *et al*. Meta-analysis of virtual-based chromoendoscopy compared with dye-spraying chromoendoscopy standard and high-definition white light endoscopy in patients with inflammatory bowel disease at increased risk of colon cancer. *Inflamm. Bowel Dis* 2020; 26: 1319–1329.

27. Selaru FM, Xu Y, Yin J, *et al*. Artificial neural networks distinguish among subtypes of neoplastic colorectal lesions. *Gastroenterology* 2002; 122: 606–613.