THERAPEUTIC ADVANCES in
Gastroenterology

Review

The impact of artificial intelligence on the endoscopic assessment of inflammatory bowel disease-related neoplasia

Ther Adv Gastroenterol

2025, Vol. 18: 1–14

DOI: 10.1177/17562848251348574

Siri A. Urquhart, Michael Christof and Nayantara Coelho-Prabhu

***Abstract*:** Inflammatory bowel disease (IBD) is a group of chronic inflammatory conditions of the gastrointestinal tract resulting from an inappropriate immune response to an altered gut microbiome in genetically predisposed individuals. Endoscopy plays a central role in IBD management, aiding in diagnosis, disease staging, monitoring, and therapeutic guidance. Patients with IBD face an increased risk of colorectal neoplasia due to chronic inflammation. Artificial intelligence (AI)-based systems show promise in detecting and classifying dysplasia and neoplasia during endoscopic evaluation. While there have been several studies on the application of AI to detect and diagnose various types of neoplasia in the non-IBD population, the literature in patients with IBD is limited. We aim to summarize the current evidence on the application of AI technologies to detect IBD-associated neoplasia, highlighting potential benefits, limitations, and future directions. A comprehensive literature search was performed using the PubMed database to identify relevant studies from January 2010 to February 2025. Additional references were identified from the relevant articles' bibliographies. AI-assisted endoscopy, particularly using machine learning and deep learning techniques, has shown promise in improving lesion detection rates and supporting real-time decision-making. Computer-aided detection systems may increase the sensitivity of dysplasia identification, while computer-aided diagnosis tools can aid in lesion characterization. Early studies suggest that AI can reduce interobserver variability, improve targeting of biopsies, and potentially lead to more personalized surveillance strategies. Although clinical data specific to IBD-related neoplasia remain limited compared to sporadic colorectal neoplasia, the integration of AI into endoscopic practice holds significant potential to enhance dysplasia detection and improve patient outcomes. Continued research, validation in IBD-specific cohorts, and integration with clinical workflows are essential to realize the full impact of AI in this setting.

Plain language summary

Impact of AI on IBD-related neoplasia

Patients with IBD have a higher risk of developing abnormal growths in the colon called dysplasia. These growths can sometimes turn into colorectal cancer if they are not found and treated early. Colonoscopy is the main tool healthcare providers use to look for these changes. However, it can be difficult to detect dysplasia during a colonoscopy as the signs are often subtle or hidden in inflamed tissue.

AI is a powerful new technology that can help healthcare providers during colonoscopy exams. AI uses computer programs that learn from thousands of images to recognize patterns and highlight suspicious areas in real time. When used during colonoscopy, AI can help healthcare providers detect more abnormal areas that may otherwise be missed.

Correspondence to:
Nayantara Coelho-Prabhu
Division of Gastroenterology and Hepatology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA
CoelhoPrabhu.Nayantara@mayo.edu

Siri A. Urquhart
Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA

Michael Christof
University of Rochester School of Medicine & Dentistry, Rochester, NY, USA

This means AI can improve the accuracy of detecting dysplasia, especially flat or subtle lesions that are more difficult to see. Early studies have shown that AI-assisted colonoscopy may lead to higher detection rates and more precise targeting of biopsies, which could reduce the chance of missing dangerous changes.

AI is also being used to review endoscopy videos and pathology images after procedures. This can support more consistent assessments between different healthcare providers and help identify which patients may need more frequent monitoring.

Overall, AI is becoming a helpful tool in managing IBD. It offers the potential to find pre-cancerous changes earlier and more reliably, helping healthcare providers make better decisions and giving patients greater peace of mind. While AI is still being studied and improved, it is likely to play a growing role in the future of IBD care.

Keywords: artificial intelligence, colorectal cancer, inflammatory bowel disease

Received: 26 February 2025; revised manuscript accepted: 23 May 2025.

Introduction

Inflammatory bowel disease (IBD), which consists of Crohn's disease (CD) and ulcerative colitis (UC), is a group of chronic inflammatory conditions of the gastrointestinal (GI) tract driven by inappropriate immune response to an altered gut microbiome in genetically susceptible individuals.[1] IBD affects at least 0.4% of Europeans and North Americans, with a rising prevalence worldwide, including in areas such as the Asian-Pacific and developing countries, where IBD was not previously described.[2–4]

Endoscopy is a cornerstone in the management of IBD, playing a crucial role in diagnosis, disease staging, and monitoring of disease activity. Endoscopy is an essential tool for diagnosing IBD, distinguishing between CD and UC, and excluding other etiologies. It allows for direct visualization of the mucosa and collection of biopsy specimens for histologic evaluation.[5,6] Endoscopic assessment aids in the definition of the extent and severity of mucosal inflammation, which is crucial for staging disease. This information influences medical and surgical decision-making and aids in targeting therapies.[5–7]

Endoscopic scoring systems play a crucial role in the management of IBD, offering objective measures of mucosal disease activity, which are essential for diagnosis, prognosis, and monitoring response to therapy. Endoscopic remission, or mucosal healing, has emerged as a key therapeutic goal in the management of IBD. Achieving endoscopic remission is associated with improved clinical outcomes, including prolonged remission, reduced hospitalization, and decreased risk of surgery and colorectal cancer (CRC).[8–11]

Several endoscopic scoring systems are used to quantify disease activity. For UC, commonly used scores include the Mayo Endoscopic Score (MES) and Ulcerative Colitis Endoscopic Index of Severity (UCEIS). The Toronto Inflammatory Bowel Disease Global Endoscopic Reporting score has also shown promise in improving the accuracy of clinical severity assessment and guiding therapeutic decisions.[12] For CD, the Simple Endoscopic Score for Crohn's Disease (SES-CD) and the Crohn's Disease Endoscopic Index of Severity (CDEIS) are frequently utilized.[13] Endoscopic scoring systems standardize the reporting of mucosal appearance, thereby enhancing clinical decision-making and facilitating a treat-to-target approach. This approach aims to treat patients early to prevent or limit intestinal injury and disability.[10] Despite the variability in inter- and intraobserver agreement rates, these scoring systems remain integral to the effective management of IBD.[14] Furthermore, standardization of endoscopic reporting in IBD clinical trials is essential for ensuring consistency, objectivity, and reliability in assessing therapeutic efficacy, reducing bias, meeting regulatory requirements, and improving trial design.[15–19]

Individuals with IBD are at increased risk of developing intestinal neoplasia—particularly colorectal neoplasia, colorectal dysplasia, and

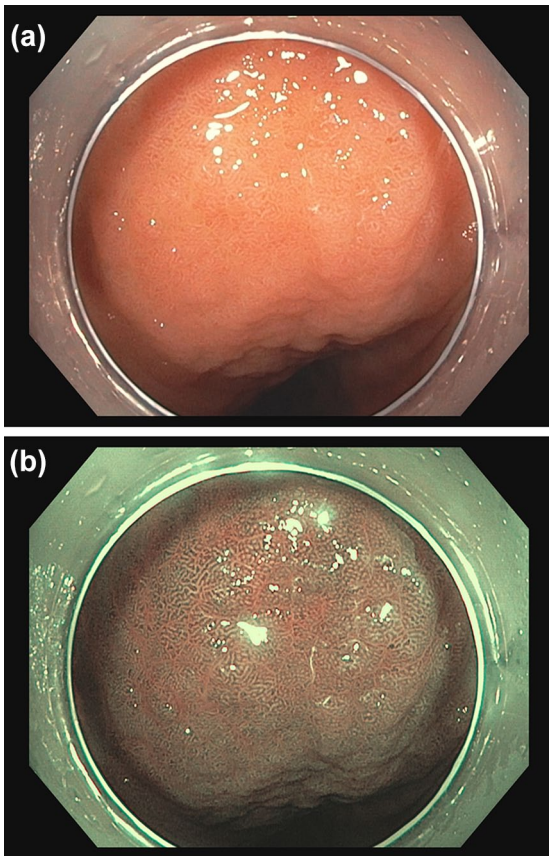

Figure 1. Representative endoscopic image of an elevated lesion in a patient with IBD, captured using high-definition white light endoscopy (a). The lesion appears subtly nodular with surface irregularities, features that may suggest dysplasia. Such lesions can be difficult to detect and characterize due to their often flat morphology and inflammation-related mucosal changes. Adjunct imaging techniques such as narrow band imaging (b), alongside emerging AI tools, may aid in improving lesion detection and differentiation during surveillance colonoscopy.
Source: Image used with permission from the authors.
AI, artificial intelligence; IBD, inflammatory bowel disease.

colorectal cancer—as a primary consequence of chronic colonic inflammation as well as intermittent episodes of acute or subacute inflammation as part of the natural disease course.[20–22] Regular endoscopic surveillance is critical for early detection of dysplasia and CRC in patients with long-standing IBD (Figure 1(a)). Advanced techniques, including chromoendoscopy, can enhance the detection of dysplasia[8,23] (Figure 1(b)).

Artificial intelligence (AI) is increasingly being utilized in medical imaging to enhance diagnostic accuracy, efficiency, and workflow. AI applications in medical imaging primarily involve machine learning (ML) and deep learning (DL) techniques, which have shown significant promise in various imaging modalities. Multiple technologies are used in this field, including convolutional neural networks (CNN),[24] generative adversarial networks (GAN),[25] computer-aided detection (CADe), and computer-aided diagnosis (CADx),[26] U-Net and variational autoencoder-UNet,[27] and automated image reconstruction and noise reduction.[28] In colonoscopy, AI-based CADe systems assist in detecting polyps (Figure 2), while CADx systems help predict polyp histology, improving adenoma detection rates and reducing the need for biopsies and/or endoscopic resection[26] (Table 1).

In this review, we aim to summarize the current evidence on the application of AI technologies to detect IBD-associated neoplasia, highlighting potential benefits, limitations, and future directions. While there have been several studies on the application of AI to detect and diagnose various types of neoplasia in the non-IBD population, the literature in patients with IBD is limited.

Methods

A structured literature review was conducted to identify and synthesize relevant studies evaluating the role of AI in the endoscopic assessment of neoplasia in IBD. The search was performed using the PubMed database. Our goal was to include more recent data pertaining to this topic, and thus articles published from January 2010 up to February 2025. Search terms included combinations of keywords such as "inflammatory bowel disease," "ulcerative colitis," "Crohn's disease," "dysplasia," "neoplasia," "colorectal cancer," "artificial intelligence," "machine learning," "deep learning," "computer-aided detection," and "endoscopy."

Studies were included if they reported original data or substantial reviews involving the application of AI in the detection, characterization, or risk stratification of IBD-associated dysplasia or neoplasia through endoscopic techniques. Articles focusing solely on sporadic CRC without reference to IBD, non-endoscopic applications of AI, or those not published in English were excluded.

Additional references were identified by manually screening the bibliographies of relevant articles. Both peer-reviewed research studies and key

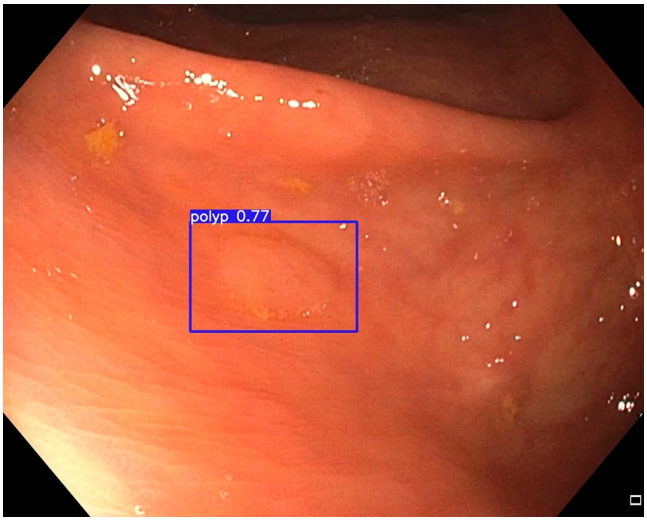

Figure 2. Endoscopic image of a lesion in a patient with IBD, captured using high-definition white light endoscopy with a CADe system. The CADe overlay highlights the lesion with a blue bounding box and an alert score of 0.77, indicating a high level of suspicion for neoplasia. Such real-time visual cues can assist endoscopists in detecting subtle or flat lesions that may otherwise be missed, supporting more effective dysplasia surveillance in IBD.
Source: Image used with permission from the authors.
CADe, computer-aided detection; IBD, inflammatory bowel disease.

review articles were considered to provide a comprehensive overview of the current state of the field. Data from eligible studies were extracted and qualitatively synthesized to highlight emerging AI technologies, clinical applications, performance metrics, and existing gaps in research related to IBD surveillance.

This review did not involve human subjects or original clinical data collection; therefore, ethical approval and patient consent were not required.

AI for disease activity scoring

Multiple endoscopic scores have been developed to assess disease activity in patients with IBD, including CDEIS and SES-CD, MES, UCEIS, and the UC colonoscopic index.[29–33] These scoring systems have been developed to define endoscopic disease activity through specific parameters and thereby limit interobserver variability.[34] ML algorithms, particularly DL models, have shown significant promise in standardizing endoscopic disease activity scoring in IBD. These algorithms aim to address the limitations of traditional endoscopic assessments, such as interobserver variability and the time-consuming nature of manual scoring.

Several studies have demonstrated the efficacy of ML models in this context. Gottlieb et al. developed a DL algorithm that predicted the MES and UCEIS from full-length endoscopy videos with high agreement metrics (quadratic weighted kappa (QWK) of 0.844 for MES and 0.855 for UCEIS).[35] Byrne et al. created a DL model that improved the accuracy and speed of UC detection and scoring, showing strong agreement with expert labels (QWK of 0.87 for MES and 0.88 for UCEIS).[36] Polat et al. utilized a regression-based DL approach to enhance the reliability of MES assessments, achieving excellent agreement with expert annotations.[37] Maeda et al. developed an endoscopic AI model used in real time on 135 patients with UC in clinical remission, underscoring the potential use of AI applications to predict clinical relapse of UC with statistical significance.[38]

Furthermore, Takenaka et al. created a deep neural network (DNUC) for real-time analysis of endoscopic images, achieving 90.1% accuracy in identifying endoscopic remission and 92.9% accuracy in identifying histologic remission. The intraclass correlation coefficient between the DNUC and endoscopists for UC endoscopic index of severity scoring was 0.917.[39] Lo et al. developed a CNN model to classify UC severity using endoscopic images. The model achieved a test accuracy of 0.84 for distinguishing all categories of MES and higher accuracies for binary tasks.[40]

Various articles have reported on the use of ML algorithms to also standardize disease activity scoring in CD utilizing radiographic imaging techniques. Guez et al. developed a multimodal ML fusion model to noninvasively assess ileal CD endoscopic activity by integrating magnetic resonance enterography (MRE) and biochemical biomarkers. This model showed improved performance over traditional linear-regression MRE models, with a better median test mean-squared-error and higher area under the curve for active disease classification.[41] Enchakalody et al. presented a hybrid model combining DL, 3-D CNN, and Random Forest to classify disease severity in the small bowel using Computerised Tomography (CT) enterography scans. The model demonstrated

Table 1. Summary of artificial intelligence tools in IBD.

Algorithm type	Training data	Performance metrics	Unique features	Limitations
CADe	Endoscopic image sets (white light and chromoendoscopy) of colitic bowel with and without lesions, neoplastic, and non-neoplastic IBD lesions	Sensitivity Specificity PPV NPV Accuracy AUC F1 score	IBD-specific training Real-time detection Increased accuracy compared to general endoscopists Improved ADR	Training data limitations Overfitting Heterogeneity of IBD Limited availability of CADe systems Algorithmic bias
CADx	Endoscopy images Electronic health records Multi-omics data Administrative databases	Sensitivity Specificity PPV Accuracy AUC	Longitudinal data analysis Predicting adverse outcomes Personalized treatment strategies Identifying molecular targets	Data bias Data integration challenges Ethical concerns Regulatory issues Generalizability
ML	Clinical data Imaging data Omics data	AUC Sensitivity Specificity Precision	Predicting therapeutic response Identifying IBD subtypes Assessing disease activity Identifying potential biomarkers	Data availability Data bias Generalization to new cohorts Clinical utility Lack of interpretability
DL	Imaging data Clinical data Multi-omics data	Accuracy Sensitivity Specificity AUC MCC Brier Score MAP DCA	Automatic feature extraction Multi-modal data integration Complex model architectures End-to-end learning	Data requirements Data quality Lack of interpretability Generalizability Bias Ethical concerns Regulatory challenges
CNN	Annotated endoscopic images from patients with IBD	Accuracy Sensitivity Specificity PPV NPV F1-score Kappa statistic	Convolutional layers Pooling layers Shared weights	Data requirements Data quality Bias Computational resources Interpretability Generalization Need for expert validation
GAN	Generates synthetic data similar to real data, allowing for data augmentation. For example, in IBD, microbiome datasets often have a greater number of features than samples, leading to overfitting. GANs can help mitigate this issue and improve the performance of ML models	FID	Conditional GANs Data augmentation	Mode collapse Non-convergence and instability Vanishing gradients
U-Net	Trained end-to-end, leveraging data augmentation to maximize use of limited data	Precision and recall IoU MCC Accuracy	Encoder-Decoder architecture Skip connections Efficient training Generalization	Blurred boundaries and irregular shapes Semantic gap Dependency on training data Potential bias Computational resources
VAE-UNet	Incorporates VAE module for encoding and decoding, aiming to learn meaningful representations of the data	Similar to U-Net	Generative model component	Computational complexity Complex data distributions

(Continued)

Table 1. (Continued)

Algorithm type	Training data	Performance metrics	Unique features	Limitations
Random forest	Endoscopic videos with histopathological confirmation, predicting IBD-related hospitalizations, initiation of biologics, identifying biomarkers, etc.	AUC Sensitivity Specificity	Handles high-dimensional data Features importance Ensemble of decision trees	Overfitting Parameter tuning Black box nature
SVM	Patient data including endoscopy and pathological results	Accuracy AUC Sensitivity (Recall) Specificity F1-score	Handling high-dimensional data Effective for IBD subtype discrimination Feature selection	Overfitting Computational complexity Interpretability Sensitivity to parameter tuning
Ensemble learning	Combination of clinical data and endoscopic image sets	AUC Accuracy	Cross-modal integration (clinical + image data)	Overfitting risk with small sample sizes

ADR, adenoma detection rate; AUC, area under the ROC curve; CADe, computer-aided detection; CADx, computer-aided diagnosis; CNN, convolutional neural network; DCA, decision curve analysis; DL, deep learning; FID, Fréchet inception distance; GAN, generative adversarial network; IBD, inflammatory bowel disease; IoU, intersection over union; MAP, mean average precision; MCC, Matthews correlation coefficient; ML, machine learning; NPV, negative predictive value; PPV, positive predictive value; SVM, support vector machine; VAE, variational autoencoder.

high precision, sensitivity, and accuracy, comparable to interobserver agreement between experienced radiologists.[42]

Xie et al. utilized a DL model, EfficientNet-B5, to detect and grade small bowel CD ulcers from double-balloon endoscopy images. The model achieved high accuracies in detecting ulcers, noninflammatory stenosis, and inflammatory stenosis, as well as in grading ulcerated surfaces, sizes, and depths.[43] Kumar et al. explored supervised classification for CD lesions using wireless capsule endoscopy images. Their methods showed high agreement with expert severity ratings and good precision and recall for lesion detection.[44] These studies collectively highlight the potential of ML algorithms to enhance the accuracy, consistency, and efficiency of endoscopic disease activity scoring in UC and CD.

Standardization of endoscopic disease activity scoring using AI significantly enhances the consistency of trial results in IBD clinical trials. AI algorithms, such as those described by Gottlieb et al. have demonstrated high agreement with human central readers in scoring endoscopic disease activity.[35] This high level of agreement reduces interobserver variability, which is a common issue in IBD trials, thereby improving the reliability and reproducibility of trial outcomes.

The use of AI in clinical trials can also increase efficiency by automating the scoring process, which traditionally requires human central readers. Ahmad et al. highlighted that AI could streamline the process by providing site-based quality evaluations and potentially eliminating the need for central readers in some contexts, while still allowing for AI-assisted central reading for follow-up assessments.[45] This can expedite the trial process and reduce costs.

However, central reading remains important in this context because it ensures an unbiased and consistent assessment of endoscopic images across different sites and readers. As noted by Gottlieb et al., central reading is crucial for maintaining the integrity of trial data, especially when AI is used to assist rather than replace human readers. Central reading provides a standardized reference that can validate and calibrate AI algorithms, ensuring automated scores are accurate and reliable.[15]

The integration of AI for standardizing endoscopic disease activity scoring enhances the consistency and efficiency of IBD clinical trials, while central reading remains essential for maintaining data integrity and validating AI performance.

Automated disease reporting

The use of automated disease scoring during endoscopy note writing significantly impacts the accuracy and consistency of endoscopic disease activity assessments in IBD clinical trials.

Automated systems, such as those developed by Yao et al., utilize CNNs to predict MES from endoscopic videos, achieving high agreement with human reviewers ($\kappa = 0.84$).[46] This reduces interobserver variability, a common issue in IBD trials, thereby enhancing the reliability and reproducibility of trial outcomes.

Stidham et al. demonstrated that computer vision methods could better quantify mucosal injury detail, offering a more sensitive measure than traditional scoring systems. Their Cumulative Disease Score was more sensitive to change and required 50% fewer participants to demonstrate endoscopic differences between treatments compared to MES.[47] This increased sensitivity and reduced sample size requirements can improve the statistical power of clinical trials.

Fan et al. developed a DL-based system that achieved 86.54% accuracy in predicting MES and 90.7% accuracy for UCEIS, further supporting the potential for AI to provide consistent and objective grading of endoscopic disease activity.[48]

Byrne et al. also highlighted the potential of DL models to improve the accuracy of MES and UCEIS, with strong agreement between AI predictions and expert labels ($\kappa = 0.87$ for MES and 0.88 for UCEIS).[36]

Gu et al. provided a comprehensive review on the role of AI in endoscopy, histology, and imaging for IBD. They highlighted the potential of AI to streamline diagnosis, reduce intra- and interobserver variability, and improve clinical decision-making. The review emphasized the promising results of early studies using DL and radiomics approaches for detecting and characterizing IBD-associated neoplasia.[49]

Stidham and Takenaka discussed the advancements in AI for disease assessment in IBD, including the interpretation of endoscopic images. AI has shown significant potential in replicating expert judgment and improving the accuracy of neoplasia detection, which is crucial for early intervention and management.[50]

Automated disease scoring enhances the accuracy and consistency of endoscopic disease activity assessments in IBD clinical trials by reducing interobserver variability, increasing sensitivity to changes, and improving the reproducibility of trial outcomes.

AI for IBD-associated neoplasia

AI plays a crucial role in detecting and managing IBD-associated neoplasia during endoscopy by enhancing the accuracy, consistency, and efficiency of lesion detection and classification.

AI systems, particularly those utilizing DL algorithms, have demonstrated superior diagnostic capabilities compared to human endoscopists. The very first case reports of the utilization of ML algorithms to identify IBD-associated dysplasia came from Japan. The EndoBrain technology, which was already in use as a CADe device for non-IBD screening, was applied to IBD surveillance. Both case series were able to identify low- and high-grade dysplasia in real time in videos.[51,52]

The first dedicated AI algorithm developed focusing on IBD dysplastic lesions was published by Vinsard et al. in 2023 and was notable for being trained on over a thousand images of IBD-associated dysplasia.[53] Even this algorithm demonstrated lower rates of success for lesions with flat morphology, suggesting that larger and varied datasets are needed to refine these algorithms. In 2024, a similar study was performed where the authors showed that the sensitivity and specificity of the algorithms improved when it was trained on IBD-specific lesions.[54]

Yamamoto et al. evaluated a deep CNN, EfficientNet-B3, for classifying IBD-associated neoplasia. The AI system demonstrated higher diagnostic accuracy compared to both expert and non-expert endoscopists, with a sensitivity of 72.5% and specificity of 82.9% for differentiating between high-grade dysplasia/adenocarcinoma and low-grade dysplasia/sporadic adenoma/normal mucosa.[55]

Okagawa et al. reviewed the rapid development of AI in GI endoscopy, including its application in IBD. AI has the potential to revolutionize the field by improving the detection and classification of neoplastic lesions, thereby enhancing surveillance and therapeutic strategies.[56] AI can assist in the characterization of detected lesions, aiding in decision-making regarding endoscopic resection

or surveillance. Santacroce et al. emphasized that AI-enabled models can provide objective and reproducible assessments of dysplasia, supporting endoscopic mucosal resection and submucosal dissection, which are crucial for a surgery-sparing approach in managing IBD-associated neoplasia.[57] Furthermore, AI-assisted real-time detection systems, such as those described by Lui et al. can significantly reduce the rate of missed lesions during colonoscopy. Their study demonstrated that AI detected 79.1% of missed proximal adenomas in tandem colonoscopy videos and increased the total number of adenomas detected by 23.6% in prospective colonoscopies.[58] This highlights the potential of AI to improve the detection of neoplastic lesions that might otherwise be overlooked, thereby enhancing surveillance and early intervention strategies.

However, there are inherent challenges with applying non-IBD trained AI algorithms to the detection of dysplasia and neoplasia in patients with IBD. AI models trained on non-IBD datasets may not generalize well to IBD patients due to differences in disease presentation and background inflammation. Rajpurkar and Lungren highlighted that AI performance can degrade when applied to different patient populations or clinical settings, a phenomenon known as "data set shift."[59] AI models developed without specific IBD training often exhibit a high risk of bias. Liu et al. found many AI models for IBD image-based prediction had high bias due to insufficient sample sizes, unreported missing data, and lack of external validation cohorts, making them less reliable for clinical implementation.[60] Non-IBD-trained AI algorithms may lack the specificity required to accurately differentiate between dysplasia and neoplasia in the context of chronic inflammation seen in IBD. Yamamoto et al. demonstrated that AI systems specifically trained on IBD-associated neoplasia achieved higher diagnostic accuracy compared to general AI models.[55]

Moreover, integrating AI into clinical practice requires overcoming technical challenges, such as ensuring compatibility with existing endoscopic equipment and electronic health record systems. Syed and Stidham emphasized the need for extensive multicenter evaluation and technical standardization to ensure AI models are robust and generalizable.[61] The application of AI in IBD raises ethical and regulatory issues, including data privacy, informed consent, and algorithmic transparency. Iacucci et al. stressed the importance of developing standardized guidelines and interdisciplinary collaboration to address these challenges.[62] Addressing these issues is crucial for the effective application of AI in detecting dysplasia and neoplasia in IBD patients.

Future directions

Future directions for the application of AI in detecting and managing IBD-associated dysplasia and neoplasia during endoscopy include several promising advancements that hold the potential to significantly improve the accuracy, efficiency, and personalization of IBD management.

Histological remission is now considered a target of treatment in UC and may be the most stringent way to assess disease remission.[63] There is growing evidence demonstrating that persistent histologic disease activity, even in the absence of endoscopic inflammation, is associated with worse clinical outcomes and risk of relapse.[63] Over 30 histologic scores have been proposed to grade histological activity in UC, but their application in clinical practice is minimal, primarily due to the impracticality of the scores.[64,65] Even when these scores are applied, such as in clinical trials, the interobserver variability is high, limiting comparison and reproducibility. As a result, clinical trials often utilize central reading systems in order for all biopsies to be evaluated by a few expert GI pathologists to reduce variability. Therefore, AI-based systems to automatically read UC biopsies would aid in standardizing assessment and decreasing interobserver variability.[66,67]

The first attempt to develop a CADe model to assess UC biopsies focused on eosinophils. The system had good agreement compared to the manual count performed by human pathologists (interclass correlation coefficient = 0.81–0.92) but did not demonstrate an association between eosinophils counted and overall inflammatory activity.[68] Gui et al. proposed to assess UC activity by taking into consideration the presence or absence of neutrophils, which are a hallmark of active inflammation. They proposed a simplified score, the PICaSSO histologic remission index that was then embedded into a CADe system that was able to distinguish histologic activity from remission in biopsies of UC with good accuracy.[69]

Further improvements to the same CADe have been presented, showing neutrophil-only assessment by CADe is largely consistent with prior scoring systems, such as Robarts and Nancy histological indexes.[70] Rymarczyk et al. demonstrated that DL models can effectively automate histological assessment in IBD, with accuracies in the colon in both CD and UC ranging from 87% to 94% and for the ileum in CD ranging from 76% to 83%.[71] AI can also enhance histopathological evaluations by predicting molecular markers such as p53 mutations from hematoxylin and eosin-stained slides, which can streamline the diagnostic process and reduce pathology workload.[72] This capability could lead to more timely and accurate diagnoses of colitis-associated neoplasia.

The ability to predict disease course as it relates to severity and progression is imperative in the field of IBD in order to implement specific, personalized approaches to management.[73] Development of AI-powered clinical decision support tools can aid in personalized treatment planning and risk stratification, improving patient outcomes and optimizing resource utilization.[74] Several studies have shown with use of natural language processing and ML algorithms trained on data from the electronic medical record, such as demographics, laboratory tests, and endoscopy reports, that it was possible to predict disease severity and surrogate markers of disease flare such as outpatient corticosteroid use or hospitalization.[75–78]

Lack of response to advanced therapies or progressive decline in response over time is an issue of utmost importance in the field of IBD, with significant downstream economic effects on total healthcare costs.[79] There is great interest in the possibility to anticipate the likelihood of a patient responding to a specific advanced therapy prior to its start may be a cost-effective approach to personalized treatment in patients with IBD (Da Rio). Through the use of ML on clinical data from trials, random forests were developed to predict response to biologics such as infliximab, vedolizumab, and ustekinumab, both in CD and UC with favorable results.[80–82] The integration and analysis of individualized patient molecular and clinical features via ML algorithms appears to be a promising field with significant potential for improvement in patient management and optimization of healthcare costs.[83]

Overall, the integration of AI in the detection and management of IBD-associated neoplasia holds the potential to revolutionize clinical practice by improving diagnostic accuracy, reducing variability, and enabling personalized care. Future research should focus on prospective validation studies and addressing ethical, legal, and regulatory challenges to facilitate the clinical implementation of these technologies.[84]

Conclusion

AI has shown significant promise in the endoscopic assessment of IBD-associated neoplasia. It has the potential to improve early detection of IBD-associated neoplasia, thereby reducing CRC rates in this patient population. AI has the ability to increase diagnostic confidence, particularly in identifying lesions in the setting of chronic inflammation and can enhance the quality and has the potential to enhance the quality and efficiency of endoscopic surveillance programs, benefiting both patients and the healthcare system.

Nonetheless, with future applications of AI in the field of IBD-associated neoplasia, there are several issues that will need to be addressed. There are inherent challenges with the application of non-IBD trained AI algorithms to the detection of dysplasia and neoplasia in patients with IBD. AI models trained on non-IBD datasets may not generalize well to IBD patients due to differences in disease presentation and background inflammation. Therefore, it is imperative to validate existing models in real-world settings in the IBD patient population to ensure robustness, and further studies are needed to build a more comprehensive foundation for future analyses. Given the numerous recent advances in the field of IBD, AI will likely gain a greater role in the practice and management of patients with IBD, ranging from evaluating the risk of developing IBD to assistance in the assessment of mucosal activity or detection of dysplasia to support in histopathological reporting, prediction of disease course, and treatment efficacy.[34] Bringing AI systems into routine clinical use will require robust, prospective validation in diverse IBD populations, development of interoperable platforms for seamless integration into endoscopy and electronic health record systems, and collaboration with regulatory bodies to ensure safety, transparency, and trust.

AI holds significant promise for enhancing the endoscopic assessment of IBD-related neoplasia, with demonstrated potential to improve lesion detection, reduce interobserver variability, and streamline documentation. However, key challenges remain, including dataset variability, integration into clinical workflows, and the need for regulatory guidance and clinician acceptance. To bridge the gap between innovation and implementation, future research must focus on prospective validation in IBD-specific cohorts, development of interoperable systems, and multidisciplinary collaboration across clinical, technical, and regulatory domains. Ultimately, the successful integration of AI into IBD surveillance will depend not only on technological performance but also on thoughtful infrastructure, clinical relevance, and trust, bringing us closer to more precise, personalized care for patients with IBD.

Declarations

Ethics approval and consent to participate

Not applicable. This article is a review of previously published literature and does not involve any studies with human participants or animals conducted by the authors.

Consent for publication

Not applicable. This article is a review of previously published literature and does not involve any studies with human participants or animals conducted by the authors.

Author contributions

Siri A. Urquhart: Conceptualization; Data curation; Writing – original draft; Writing – review & editing.

Michael Christof: Writing – original draft.

Nayantara Coelho-Prabhu: Conceptualization; Supervision; Writing – original draft; Writing – review & editing.

Acknowledgements

None.

Funding

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

Competing interests

S.A.U: Discloses no conflicts. M.C.: Discloses no conflicts. N.C.-P.: Research funding from Fujifilm, Cook Endoscopy, and Alexion Pharma.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID iD

Nayantara Coelho-Prabhu https://orcid.org/0000-0003-0775-4385

References

1. Kaser A, Zeissig S and Blumberg RS. Inflammatory bowel disease. *Annu Rev Immunol* 2010; 28: 573–621.
2. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2017; 390: 2769–2778.
3. Kaplan GG. The global burden of IBD: from 2015 to 2025. *Nat Rev Gastroenterol Hepatol* 2015; 12: 720–727.
4. Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012; 142: 46–54.e42; quiz e30.
5. Fefferman DS and Farrell RJ. Endoscopy in inflammatory bowel disease: indications, surveillance, and use in clinical practice. *Clin Gastroenterol Hepatol* 2005; 3: 11–24.
6. Spiceland CM and Lodhia N. Endoscopy in inflammatory bowel disease: Role in diagnosis, management, and treatment. *World J Gastroenterol* 2018; 24: 4014–4020.
7. Shergill AK, Lightdale JR, Bruining DH, et al. The role of endoscopy in inflammatory bowel disease. *Gastrointest Endosc* 2015; 81: 1101–1121.e13.
8. Núñez F P, Krugliak Cleveland N, Quera R, et al. Evolving role of endoscopy in inflammatory bowel disease: going beyond diagnosis. *World J Gastroenterol* 2021; 27: 2521–2530.
9. Christensen B and Rubin DT. Understanding endoscopic disease activity in IBD: how to

incorporate it into practice. *Curr Gastroenterol Rep*; 18. Epub ahead of print January 2016. DOI: 10.1007/s11894-015-0477-6.

10. Limdi JK, Picco M and Farraye FA. A review of endoscopic scoring systems and their importance in a "treat to target" approach in inflammatory bowel disease (with videos). *Gastrointest Endosc*; 91. Epub ahead of print November 2019. DOI: 10.1016/j.gie.2019.11.032.
11. Rubin DT, Ananthakrishnan AN, Siegel CA, et al. ACG clinical guideline: ulcerative colitis in adults. *Am J Gastroenterol* 2019; 114: 384–413.
12. Liu X-Y, Tian Z-B, Zhang L-J, et al. Clinical value of the Toronto inflammatory bowel disease global endoscopic reporting score in ulcerative colitis. *World J Gastroenterol* 2023; 29: 6208–6221.
13. Lichtenstein GR, Loftus EV, Isaacs KL, et al. ACG clinical guideline: management of Crohn's disease in adults. *Am J Gastroenterol* 2018; 113: 481–517.
14. Hashash JG, Yu Ci Ng F, Farraye FA, et al. Inter- and intraobserver variability on endoscopic scoring systems in Crohn's disease and ulcerative colitis: a systematic review and meta-analysis. *Inflamm Bowel Dis* 2024; 30: 2217–2226.
15. Gottlieb K, Daperno M, Usiskin K, et al. Endoscopy and central reading in inflammatory bowel disease clinical trials: achievements, challenges and future developments. *Gut* 2020; 70: gutjnl-320690.
16. Gottlieb K, Travis S, Feagan B, et al. Central reading of endoscopy endpoints in inflammatory bowel disease trials. *Inflamm Bowel Dis* 2015; 21: 1–1.
17. Vuitton L, Peyrin-Biroulet L, Colombel JF, et al. Defining endoscopic response and remission in ulcerative colitis clinical trials: an international consensus. *Aliment Pharmacol Ther* 2017; 45: 801–813.
18. Ma C, Hanzel J, Panaccione R, et al. CORE-IBD: a multidisciplinary international consensus initiative to develop a core outcome set for randomized controlled trials in inflammatory bowel disease. *Gastroenterology* 2022; 163: 950–964.
19. Khanna R, Ma C, Vipul Jairath, et al. Endoscopic assessment of inflammatory bowel disease activity in clinical trials. *Clin Gastroenterol Hepatol* 2020; 20: 727–736.e2.
20. Rutter MD, Saunders BP, Wilkinson KH, et al. Thirty-year analysis of a colonoscopic surveillance program for neoplasia in ulcerative colitis. *Gastroenterology* 2006; 130: 1030–1038.
21. Ullman TA and Itzkowitz SH. Intestinal inflammation and cancer. *Gastroenterology* 2011; 140: 1807–1816.e1.
22. Beaugerie L and Itzkowitz SH. Cancers complicating inflammatory bowel disease. *N Engl J Med* 2015; 372: 1441–1452.
23. Tharian B, George N and Navaneethan U. Endoscopy in the diagnosis and management of complications of inflammatory bowel disease. *Inflamm Bowel Dis* 2016; 22: 1184–1197.
24. Le EPV, Wang Y, Huang Y, et al. Artificial intelligence in breast imaging. *Clin Radiol* 2019; 74: 357–366.
25. Balaji V, Song T-A, Malekzadeh M, et al. Artificial intelligence for PET and SPECT image enhancement. *J Nucl Med*; 65. Epub ahead of print 9 November 2023. DOI: 10.2967/jnumed.122.265000.
26. Samarasena J, Yang D and Berzin TM. AGA clinical practice update on the role of artificial intelligence in colon polyp diagnosis and management: commentary. *Gastroenterology* 2023; 165: 1568–1573.
27. Zhou R, Kozlov A, Chen S-T, et al. Harnessing artificial intelligence to automate delineation of volumetric breast cancers from magnetic resonance imaging to improve tumor characterization. *J Clin Oncol* 2022; 40: 597–597.
28. Langlotz CP, Allen B, Erickson BJ, et al. A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019; 291: 781–791.
29. Mary JY and Modigliani R. Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Thérapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). *Gut* 1989; 30: 983–989.
30. Daperno M, D'Haens G, Van Assche G, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. *Gastrointest Endosc* 2004; 60: 505–512.
31. Schroeder KW, Tremaine WJ and Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. *N Engl J Med* 1987; 317: 1625–1629.
32. Travis SPL, Schnell D, Krzeski P, et al. Developing an instrument to assess the endoscopic severity of ulcerative colitis: the

Ulcerative Colitis Endoscopic Index of Severity (UCEIS). *Gut* 2011; 61: 535–542.

33. Samuel S, Bruining DH, Loftus EV, et al. Validation of the ulcerative colitis colonoscopic index of severity and its correlation with disease activity measures. *Clin Gastroenterol Hepatol* 2013; 11: 49–54.e1.
34. Da Rio L, Spadaccini M, Parigi TL, et al. Artificial intelligence and inflammatory bowel disease: where are we going? *World J Gastroenterol* 2023; 29: 508–520.
35. 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.
36. 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* 2022; 17: 463–471.
37. Polat G, Kani HT, Ergenc I, et al. Improving the computer-aided estimation of ulcerative colitis severity according to mayo endoscopic score by using regression-based deep learning. *Inflamm Bowel Dis* 2022; 29: 1431–1439.
38. Maeda Y, Kudo S, 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.
39. 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.
40. 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.
41. Guez I, Focht G, Greer M-LC, et al. Development of a multimodal machine-learning fusion model to non-invasively assess ileal Crohn's disease endoscopic activity. *Comput Methods Programs Biomed* 2022; 227: 107207–107207.
42. Enchakalody BE, Wasnik AP, Al-Hawary MM, et al. Local assessment and small bowel Crohn's disease severity scoring using AI. *Acad Radiol* 2024; 31: 4045–4056.
43. Xie W, Hu J, Liang P, et al. Deep learning–based lesion detection and severity grading of small-bowel Crohn's disease ulcers on double-balloon endoscopy images. *Gastrointest Endosc* 2024; 99: 767–777.e5.
44. 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.
45. 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.
46. Yao H, Najarian K, Gryak J, et al. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2021; 93: 728–736.e1.
47. Stidham RW, Cai L, Cheng S, et al. Using computer vision to improve endoscopic disease quantification in therapeutic clinical trials of ulcerative colitis. *Gastroenterology*; 166. Epub ahead of print 1 October 2023. DOI: 10.1053/j.gastro.2023.09.049.
48. 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.
49. Gu P, Mendonca O, Carter D, et al. AI-luminating artificial intelligence in inflammatory bowel diseases: a narrative review on the role of AI in endoscopy, histology, and imaging for IBD. *Inflamm Bowel Dis* 2024; 30: 2467–2485.
50. Stidham RW and Takenaka K. Artificial intelligence for disease assessment in IBD: how will it change our practice? *Gastroenterology*; 162. Epub ahead of print January 2022. DOI: 10.1053/j.gastro.2021.12.238
51. Maeda Y, Kudo S, Ogata N, et al. Can artificial intelligence help to detect dysplasia in patients with ulcerative colitis? *Endoscopy*; 53. Epub ahead of print 1 October 2020. DOI: 10.1055/a-1261-2944.
52. Fukunaga S, Kusaba Y, Ohuchi A, et al. Is artificial intelligence a superior diagnostician in ulcerative colitis? *Endoscopy* 2020; 53: E75–E76.
53. Vinsard DG, Fetzer JR, Agrawal U, et al. Development of an artificial intelligence tool for detecting colorectal lesions in inflammatory bowel disease. *iGIE* 2023; 2: 91–101.e6.
54. Abdelrahim M, Siggens K, Iwadate Y, et al. New AI model for neoplasia detection and characterisation in inflammatory bowel disease. *Gut* 2024; 73: gutjnl-330718.

55. Yamamoto S, Kinugasa H, Hamada K, et al. The diagnostic ability to classify neoplasias occurring in inflammatory bowel disease by artificial intelligence and endoscopists: a pilot study. *J Gastroenterol Hepatol* 2022; 37: 1610–1616.
56. Okagawa Y, Abe S, Yamada M, et al. Artificial intelligence in endoscopy. *Dig Dis Sci* 2021; 67: 1553–1572.
57. Santacroce G, Zammarchi I, Tan CK, et al. Present and future of endoscopy precision for inflammatory bowel disease. *Dig Endosc* 2023; 36: 292–304.
58. Lui TKL, Hui CKY, Tsui VWM, et al. New insights on missed colonic lesions during colonoscopy through artificial intelligence–assisted real-time detection (with video). *Gastrointest Endosc* 2021; 93: 193–200.e1.
59. Pranav Rajpurkar and Lungren MP. The current and future state of AI interpretation of medical images. *N Engl J Med* 2023; 388: 1981–1990.
60. Liu X, Reigle J, V.B. Surya Prasath, et al. Artificial intelligence image-based prediction models in IBD exhibit high risk of bias: a systematic review. *Comput Biol Med* 2024; 171: 108093.
61. Syed S and Stidham RW. Potential for standardization and automation for pathology and endoscopy in inflammatory bowel disease. *Inflamm Bowel Dis* 2020; 26: 1490–1497.
62. Iacucci M, Santacroce G, Zammarchi I, et al. Artificial intelligence and endo-histo-omics: new dimensions of precision endoscopy and histology in inflammatory bowel disease. *Lancet Gastroenterol Hepatol* 2024; 9: 758–772.
63. 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* 2021; 160: 1570–1583.
64. Mosli MH, Feagan BG, Sandborn WJ, et al. Histologic evaluation of ulcerative colitis. *Inflamm Bowel Dis* 2014; 20: 564–575.
65. Mojtahed A, Khanna R, Sandborn WJ, et al. Assessment of histologic disease activity in Crohn's disease. *Inflamm Bowel Dis* 2014; 20: 2092–2103.
66. Chateau T, Feakins R, Marchal-Bressenot A, et al. Histological remission in ulcerative colitis: under the microscope is the cure. *Am J Gastroenterol* 2019; 115: 179–189.
67. Rimondi A, Gottlieb K, Despott EJ, et al. Can artificial intelligence replace endoscopists when assessing mucosal healing in ulcerative colitis? A systematic review and diagnostic test accuracy meta-analysis. *Dig Liver Dis* 2024; 56: 1164–1172.
68. Vande Casteele 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* 2021; 28: 539–546.
69. Gui X, Bazarova A, Rocío del Amor, 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.
70. Villanacci V, Parigi TL, Del Amor R, et al. OP15 A new simplified histology artificial intelligence system for accurate assessment of remission in ulcerative colitis. *J Crohns Colitis* 2022; 16: i015–i017.
71. 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* 2023; 18: 604–614.
72. Noguchi T, Ando T, Emoto S, et al. Artificial intelligence program to predict p53 mutations in ulcerative colitis–associated cancer or dysplasia. *Inflamm Bowel Dis* 2022; 28: 1072–1080.
73. Marlicz W, Skonieczna-Żydecka K, Dabos KJ, et al. Emerging concepts in non-invasive monitoring of Crohn's disease. *Therap Adv Gastroenterol* 2018; 11: 175628481876907.
74. Silverman AL, Shung D, Stidham RW, et al. How artificial intelligence will transform clinical care, research, and trials for inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2025; 23: 428–439.e4.
75. Ananthakrishnan AN, Cai T, Savova G, et al. Improving case definition of Crohn's disease and ulcerative colitis in electronic medical records using natural language processing. *Inflamm Bowel Dis* 2013; 19: 1411–1420.
76. Stidham RW, Yu D, Zhao X, et al. Identifying the presence, activity, and status of extraintestinal manifestations of inflammatory bowel disease using natural language processing of clinical notes. *Inflamm Bowel Dis* 2022; 29: 503–510.
77. Reddy BK, Delen D and Agrawal RK. Predicting and explaining inflammation in Crohn's disease

patients using predictive analytics methods and electronic medical record data. *Health Informatics J* 2018; 25: 1201–1218.

78. Cai W, Xu J, Chen Y, et al. Performance of machine learning algorithms for predicting disease activity in inflammatory bowel disease. *Inflammation* 2023; 46: 1561–1574.
79. Argollo M, Kotze PG, Kakkadasam P, et al. Optimizing biologic therapy in IBD: how essential is therapeutic drug monitoring? *Nat Rev Gastroenterol Hepatol* 2020; 17: 702–710.
80. Li Y, Pan J, Zhou N, et al. A random forest model predicts responses to infliximab in Crohn's disease based on clinical and serological parameters. *Scand J Gastroenterol* 2021; 56: 1030–1039.
81. Waljee AK, Liu B, Sauder K, et al. Predicting corticosteroid-free biologic remission with vedolizumab in Crohn's disease. *Inflamm Bowel Dis* 2018; 24: 1185–1192.
82. Doherty MK, Ding T, Koumpouras C, et al. Fecal microbiota signatures are associated with response to ustekinumab therapy among Crohn's disease patients. *mBio*; 9. Epub ahead of print 2 May 2018. DOI: 10.1128/mbio.02120-17.
83. Pinton P. Computational models in inflammatory bowel disease. *Clin Transl Sci* 2022; 15: 824–830.
84. Kawamoto A, Takenaka K, Okamoto R, et al. A systematic review of artificial intelligence-based image diagnosis for inflammatory bowel disease. *Dig Endosc*; 34. Epub ahead of print 19 April 2022. DOI: 10.1111/den.14334.