

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 95-197





Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 4 August 28, 2021

OPINION REVIEW

- 95 Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease
Lo B, Burisch J
- 103 Robotic pancreaticoduodenectomy: Where do we stand?
Khachfe HH, Habib JR, Chahrour MA, Nassour I

MINIREVIEWS

- 110 Robotic surgery in colon cancer: current evidence and future perspectives – narrative review
Tagliabue F, Burati M, Chiarelli M, Cioffi U, Zago M
- 117 Artificial intelligence in endoscopy: The challenges and future directions
Gao X, Braden B
- 127 Deep learning applied to the imaging diagnosis of hepatocellular carcinoma
Ballotin VR, Bigarella LG, Soldera J, Soldera J
- 136 Role of capsule endoscopy in inflammatory bowel disease: Anything new?
Pérez de Arce E, Quera R, Núñez F P, Araya R
- 149 Role of optical coherence tomography in Barrett's esophagus
Gupta N, Yelamanchi R, Agrawal H, Agarwal N
- 157 Artificial intelligence and colonoscopy – enhancements and improvements
Yoo BS, D'Souza SM, Houston K, Patel A, Lau J, Elmahdi A, Parekh PJ, Johnson D
- 168 Impact of endoscopic ultrasound elastography in pancreatic lesion evaluation
Lesmana CRA, Paramitha MS
- 179 Artificial intelligence as a means to improve recognition of gastrointestinal angiodysplasia in video capsule endoscopy
Cox II GA, Jackson CS, Vega KJ
- 185 Early gastrointestinal cancer: The application of artificial intelligence
Yang H, Hu B

Contents

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 4 August 28, 2021

ABOUT COVER

Associate Editor of *Artificial Intelligence in Gastrointestinal Endoscopy*, Peter Bauerfeind, MD, Professor, Department of Gastroenterology and Hepatology, University hospital Zurich, Zurich 8091, Switzerland.
peter.bauerfeind@usz.ch

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (AIGE, *Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangio-pancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Lin-YuTong Wang*; Production Department Director: *Xiang Li*; Editorial Office Director: *Jin-Li Wang*.

NAME OF JOURNAL

Artificial Intelligence in Gastrointestinal Endoscopy

ISSN

ISSN 2689-7164 (online)

LAUNCH DATE

July 28, 2020

FREQUENCY

Bimonthly

EDITORS-IN-CHIEF

Fatih Altintoprak, Sahin Coban, Krish Ragunath

EDITORIAL BOARD MEMBERS

<https://www.wjgnet.com/2689-7164/editorialboard.htm>

PUBLICATION DATE

August 28, 2021

COPYRIGHT

© 2021 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/gerinfo/287>

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

PUBLICATION ETHICS

<https://www.wjgnet.com/bpg/gerinfo/288>

PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

ARTICLE PROCESSING CHARGE

<https://www.wjgnet.com/bpg/gerinfo/242>

STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/gerinfo/239>

ONLINE SUBMISSION

<https://www.f6publishing.com>

© 2021 Baishideng Publishing Group Inc. All rights reserved. 7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

E-mail: bpgoffice@wjgnet.com <https://www.wjgnet.com>



Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

Bobby Lo, Johan Burisch

ORCID number: Bobby Lo 0000-0002-0252-9341; Johan Burisch 0000-0002-3312-5139.

Author contributions: Lo B and Burisch J authors have made a significant contribution to the research described in this manuscript; all authors approved the final manuscript as well as the authorship list.

Conflict-of-interest statement: Lo B has received a lecture fee from Janssen-Cilag; Burisch J has received consulting fees from Celgene, Janssen-Cilag, AbbVie, Vifor Pharma, Jansen and Ferring; lecture fees from Abbvie, Pfizer, MSD, Pharmacosmos and Takeda Pharma, and unrestricted grant support from Takeda Pharma and Tillotts Pharma.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

Bobby Lo, Johan Burisch, Gastrounit, Medical Section, Copenhagen University Hospital Hvidovre, Hvidovre 2650, Denmark

Bobby Lo, Johan Burisch, Copenhagen Centre for Inflammatory Bowel Disease in Children, Adolescents and Adults, Copenhagen University Hospital Hvidovre, Hvidovre 2650, Denmark

Corresponding author: Bobby Lo, MD, Doctor, Gastrounit, Medical Section, Copenhagen University Hospital Hvidovre, Kettegård Alle 30, Hvidovre 2650, Denmark.
bobby_lo@protonmail.com

Abstract

Assessment of endoscopic disease activity can be difficult in patients with inflammatory bowel disease (IBD) [comprises Crohn's disease (CD) and ulcerative colitis (UC)]. Endoscopic assessment is currently the foundation of disease evaluation and the grading is pivotal for the initiation of certain treatments. Yet, disharmony is found among experts; even when reassessed by the same expert. Some studies have demonstrated that the evaluation is no better than flipping a coin. In UC, the greatest achieved consensus between physicians when assessing endoscopic disease activity only reached a Kappa value of 0.77 (or 77% agreement adjustment for chance/accident). This is unsatisfactory when dealing with patients at risk of surgery or disease progression without proper care. Lately, across all medical specialities, computer assistance has become increasingly interesting. Especially after the emanation of machine learning – colloquially referred to as artificial intelligence (AI). Compared to other data analysis methods, the strengths of AI lie in its capability to derive complex models from a relatively small dataset and its ability to learn and optimise its predictions from new inputs. It is therefore evident that with such a model, one hopes to be able to remove inconsistency among humans and standardise the results across educational levels, nationalities and resources. This has manifested in a handful of studies where AI is mainly applied to capsule endoscopy in CD and colonoscopy in UC. However, due to its recent place in IBD, there is a great inconsistency between the results, as well as the reporting of the same. In this opinion review, we will explore and evaluate the method and results of the published studies utilising AI within IBD (with examples), and discuss the future possibilities AI can offer within IBD.

Key Words: Inflammatory bowel disease; Artificial intelligence; Deep learning; Endoscopy; Disease severity; Machine learning

[p://creativecommons.org/licenses/by-nc/4.0/](https://creativecommons.org/licenses/by-nc/4.0/)

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: Denmark

Peer-review report's scientific quality classification

Grade A (Excellent): 0

Grade B (Very good): B

Grade C (Good): C, C, C, C

Grade D (Fair): 0

Grade E (Poor): 0

Received: April 28, 2021

Peer-review started: April 28, 2021

First decision: June 13, 2021

Revised: June 27, 2021

Accepted: August 16, 2021

Article in press: August 16, 2021

Published online: August 28, 2021

P-Reviewer: Chabrun F, El-Shabrawi MH, Kamran M, Li A

S-Editor: Liu M

L-Editor: Webster JR

P-Editor: Xing YX



©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Artificial intelligence (AI) is on the rise in inflammatory bowel diseases (IBD). Endoscopic evaluation is so far the most studied modality with promising results. Studies with others or the combination of several modalities have been carried out with moderate results leaving room for future research. Data availability and standardisation of the reporting of these new models seem to be the biggest challenges for AI's breakthrough within IBD. International consensus in the field is required to optimise research in AI.

Citation: Lo B, Burisch J. Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease. *Artif Intell Gastrointest Endosc* 2021; 2(4): 95-102

URL: <https://www.wjgnet.com/2689-7164/full/v2/i4/95.htm>

DOI: <https://dx.doi.org/10.37126/aige.v2.i4.95>

INTRODUCTION

The inflammatory bowel diseases (IBD), which mainly consist of Crohn's disease (CD) and ulcerative colitis (UC), are idiopathic immune-mediated diseases usually affecting young adults[1,2].

Currently, colonoscopy is considered the gold standard in the disease assessment of patients with UC as well as CD located in the terminal ileum and/or colon[3,4]. Disease activity of UC is assessed using scoring systems such as the Mayo Endoscopic Subscore (MES) or UC Endoscopic Index of Severity[5]. Despite their widespread use and being easy to use, both indices suffer from moderate to high inter-observer variation which reduces the credibility of the scores[6]. This has been demonstrated in clinical trials where up to one-third of patients deemed eligible for inclusion based on the MES did not live up to the inclusion criteria after reassessment[7]. Even central reading is associated with noteworthy inter-observer variation[7,8].

In CD, the CD Endoscopic Index of Severity and Simple Endoscopic Score for CD are currently the most used indices[4]. Both have demonstrated varying observer variance with central reading improving the inter-observer variation[9-11]. Capsule endoscopy (CE) for evaluating the small bowel can be scored using the Lewis score [12]. While widely used, the interobserver agreement between parameters in the index fluctuates widely (kappa 0.37-0.83)[13,14].

These interobserver variations and the risk of misclassification has led to the exploration of artificial intelligence (AI) assisted endoscopic assessment[15], especially in the field of colon cancer detection[16,17]. AI, depending on which method is used, mimics the human brain by having interconnected neurons that process the information given; however, in contrast to the human brain, AI can theoretically process an unlimited number of variables. In the field of IBD, the use of AI remains limited although it has received increasing attention. In the following review, we will discuss the use of AI-assisted assessment of endoscopic disease activity among CD and UC patients from a clinical perspective, the challenges the model faces and unexplored areas where AI has the potential to help patients and physicians.

CROHN'S DISEASE

CD can be examined using many modalities. Imaging has been an area of interest in terms of AI - especially CE[18]. A CE camera takes between 2-4 frames per second and has an approximate transit time of 250 min which can result in a total of approximately 60000 images[18]. One of the challenges CE entails is that it is a time-consuming process whereby a trained person must subsequently review all images. New AI has since assisted physicians and endoscopists in filtering out non-informative images, thereby leaving an image series where the computer believes there is an area of interest. Since the year 2000, AI has been used to identify polyps/tumours, ulcers, celiac disease, hookworms, angioectasia, and bleeding[18]. Among CD patients, special focus has been on small bowel lesions, erosions and ulceration[19]. The

majority of recent studies that have examined the listed parameters use a convolutional neural network - a deep learning method that has been shown to be effective in image recognition[18,20]. Overall, these studies have shown an accuracy of > 90% which must be considered close to perfect. However, the majority of these studies are conducted retrospectively and prospective results are wanted to demonstrate the models potential in clinical practice.

ULCERATIVE COLITIS

Due to UC only involving the colon it has been easier to categorise these patients than CD according to the extent and severity of inflammation[21]. Accordingly, most advances regarding AI in IBD has been done in UC and several clinical tools have been developed to assess the endoscopic disease severity. Such models have achieved an accuracy of 56%–77% in assessing the disease severity according to the MES or UC Endoscopic Index of Severity which was comparable to IBD experts[22–27]. The majority of studies have used methods such as the convolutional neural network to categorize images taken during a colonoscopy or sigmoidoscopy according to the MES. Recently, studies have also investigated the applicability of AI on videos; demonstrating a promising area under the receiver operating characteristic curve (AUROC or AUC) > 90%[24,26,27].

Currently, the available models are unable to distinguish between the different levels of the MES with sufficient accuracy. However, this is an area under great development and it is expected that within the coming years a model will be able to distinguish between the different MES levels with a satisfactory result and thereby eliminate the inter-observer variance, and standardize the clinical and academic evaluation of the endoscopic disease severity[28].

Few studies have further examined their model's MES score in relation to histological findings[29,30]. One study used endocytoscopy with a support vector machine and achieved an accuracy of approximately 90% in predicting histological findings which must be considered excellent results[29]. Endocytoscopy is, however, not an integral method in most clinics. Furthermore, although the study group utilized both a training and a test set, the training and optimizing process of the models is not described, leaving the reader with uncertainty with regard to *e.g.*, model selection and tuning of. Finally, samples were divided into active inflammation *vs* remission which might be too simplified a way of considering both the endoscopic and histological findings. Similar results were demonstrated by Takenaka *et al*[30] with white-light endoscopy, but with the same challenges. Ultimately, none of these studies validated the results on an independent cohort analyzed by independent experts, in order to test the performance of their model when compared to another population or to the point of view of different experts.

POTENTIAL AND DIFFICULTIES

As previously mentioned, AI has been shown to have great potential in the evaluation of endoscopic severity among patients with CD and UC. The models have shown to be at a level with or better than physicians to classify endoscopic disease severity; especially among UC patients[25]. Uniformity in the approach to the endoscopic procedure will make new clinical tools more credible and hopefully lead to less discrepancy between clinical and observational studies[31]. However, it is crucial that new models are developed for clinical purposes, which can be implemented more quickly, thereby reducing the gap between research and clinical practice.

Besides endoscopic evaluation, disease prediction in IBD has also been investigated using AI models. Waljee *et al*[32,33] used two clinical trial databases to predict C-reactive protein < 5 mg/L after 42 wk treatment with ustekinumab and steroid-free remission after 52 wk treatment with vedolizumab among CD patients, respectively. These studies used a combination of demographic, clinical, and biochemical data in a random forest model to predict patients' course after initiation of treatment. The models achieved an accuracy of 42% and 69%, respectively. Furthermore, the same study group investigated the treatment effect of vedolizumab in UC patients[34]. Using a random forest model, the model achieved an accuracy of 58% in predicting corticosteroid-free remission after 52 wk. When grouping UC and CD together, Biasci *et al*[35] used transcriptomics to identify a blood sample panel of 17 genes with sensitivity and specificity of approximately 73% to predict patients' risk of treatment

escalating within 1 year. A 5-year prediction study from Choi *et al*[36] demonstrated a sensitivity and specificity of 71% for predicting the risk of the use of biologics. In contrast to Biasci *et al*[35], this study utilized only demographical, clinical and common laboratory markers. Furthermore, Waljee *et al*[37,38] attempted twice to predict the treatment effect within 1 year, resulted in an AUC of 79% and 87% and accuracy of 72% and 80%, respectively. A limitation of these studies is that findings are only presented for IBD patients in total and not stratified according to the type of IBD. Despite these efforts, accuracies below at least 80% must be considered insufficient. Furthermore, even with accuracies above 80%, the results must be taken into perspective with the sensitivity, specificity and AUC to achieve an overall picture of the model's performance. Unfortunately, the majority of the studies have only reported some but not all measures of validity of which AUC is most commonly reported.

OTHER AREAS

It is not uncommon for some patients to undergo a lengthy diagnostic process before a definite diagnosis of CD or UC can be made[39]. This can be a challenge for both physicians and patients, and result in over or under treatments with major consequences for the patient. Recent studies using AI have attempted to use several modalities to better distinguish between these patients: endoscopy, histology, genetic markers, biochemical markers, clinical factors, omics, or a combination of one or more of these modalities[40-43]. These have shown acceptable results with AUC and accuracy of > 80%. It should be emphasized that these studies do not always report all results and many of the results are from validation data and not necessarily test data (unseen data) exposing the models to overfitting. However, to our knowledge, none of these models has been applied in clinical practice and real-life data are warranted to evaluate their efficacy.

To our knowledge, no other modalities explored in connection with AI have been published to date. In particular, the complexity of CD results in several challenges when developing new AI models. One area that remains untouched is the use of AI during colonoscopy in CD patients. This could be due to challenges in the endoscopic disease assessment of CD as the disease can be patchy and the severity varies between patches. Besides, indices for CD are difficult or time-consuming to use in clinical practice[4]. This could be accommodated by developing new scoring indices based on an evaluation from an AI model, allowing the possibility of assessing the gut as a whole rather than the segmented method currently being used.

In addition to endoscopies for both UC and CD, modalities such as ultrasound, magnetic resonance imaging, colon CE and computed tomography are obvious opportunities for the development of new clinical tools[44].

Unfortunately, this field is also challenged by several issues. First and foremost, a paradigm shift is needed; from a medical professional to a computer-aided assessment. This will first and foremost require doctors to accept the new technology [45] which can be difficult to understand as the latest AI architectures use deep learning where a black-box appears (the process between input and output)[46]. As it is not 100% possible to account for what happens in this black-box, mistrust might arise among the clinicians toward the models. Despite different ways of explaining the black-box, mathematically and illustratively, it is only possible to give an estimate of its process[46].

Secondly, medical education may need to be reorganized in the future to have more focus on interpretation and critical evaluation of the results of these new models. The medical field has experienced a similar paradigm shift before with the introduction of the World Wide Web[47]. This gave patients equal access to knowledge that physicians had and doctors went from being the ultimate definitive truth to now having to explain how the symptoms and the disease are connected and which diagnosis and disease courses are most likely[47]. However, a new organization of the medical education in connection with AI may require interdisciplinary involvement with, among others, bioinformatics and computer scientists to better equip doctors to interpret and critically evaluate the models' output.

Thirdly, larger amounts of data are needed – more than previously accustomed to developing these new models. However, the amount of data needed varies significantly in relation to the outcome and the methods used and no specific number of required data exists. As data is resource demanding, the estimate must be adjusted to what is clinically possible. In recent years, cross-border collaborations have been

Table 1 Recommendations for reporting of studies regarding artificial intelligence

Section	Requirements
Method	<p>Origin of dataset and description of the acquisition process</p> <p>Pre-processing methods</p> <p>Definition of ground truth</p> <p>Split of data set and should include a training, validation and test set. A clear statement that the test set is not used to tune hyperparameters or in the selection of the model</p> <p>Method and architecture used, whether it is pretrained or not, and what dataset it is pretrained on</p> <p>Full technical detail should be included in supplementary files</p> <p>Statement of post-selection analyses and why these are conducted</p>
Results	<p>A complete report of all results including but not restricted to AUC, sensitivity, specificity, accuracy and kappa value for the overall model's performance and not for selected tasks</p>
Discussion	<p>Risks of overfitting and bias</p> <p>Generalisability and cautions to take</p> <p>Clinical implementation</p>

AUC: Area under the receiver operating characteristic curve.

formed to make large amounts of data available. However, these are rarely freely available and the quality must also be critically evaluated when the workflow and equipment vary markedly between nations. We, therefore, encourage everyone to make their data at least partially accessible - a good example is The HyperKvasir dataset[48].

Finally, international reporting standards must be set within the field of IBD regarding AI studies. AI is still a relatively unexploited territory within IBD. This has led to great variation in the way the studies report both their methods and results, despite several calls for uniformity[49]. A good example is the endoscopic evaluation of disease severity in UC patients. Often, only AUC is reported, which can be misleading as sensitivity, specificity and accuracy may be only modest[25]. This is due to the fact that the AUC is a measure of how well the true positive can be separated from the rest, while measures of *e.g.*, accuracy hint at the actual performance of the models. Even when the studies report the wanted parameters, the reporting method can vary. For example, calculating the sensitivity, specificity and accuracy for each class rather than reporting the overall sensitivity, specificity and accuracy for the entire index. We, therefore, encourage that future articles as a minimum must report the information and parameters described in Table 1.

In addition, international journals should set standards for what is required of future AI studies within the field. The use of previous reporting methods, *e.g.*, STARD guidelines, seems outdated and should be updated to the new technological reality [50].

CONCLUSION

AI is on the rise in IBD. Endoscopic evaluation is so far the most studied modality with promising results. Studies with others or the combination of several modalities have been carried out with moderate results leaving room for future research. Data availability and standardization of the reporting of these new models seem to be the biggest challenges for the AI's breakthrough within IBD. International consensus in the field is required to optimize research in AI.

REFERENCES

- 1 **Torres J**, Mehandru S, Colombel JF, Peyrin-Biroulet L. Crohn's disease. *Lancet* 2017; **389**: 1741-1755 [PMID: [27914655](#) DOI: [10.1016/S0140-6736\(16\)31711-1](#)]
- 2 **Ungaro R**, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. *Lancet* 2017; **389**: 1756-1770 [PMID: [27914657](#) DOI: [10.1016/S0140-6736\(16\)32126-2](#)]
- 3 **Magro F**, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, Burisch J, Geese KB, Hart AL, Hindryckx P, Langner C, Limdi JK, Pellino G, Zagórowicz E, Raine T, Harbord M, Rieder F; European Crohn's and Colitis Organisation [ECCO]. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. *J Crohns Colitis* 2017; **11**: 649-670 [PMID: [28158501](#) DOI: [10.1093/ecco-jcc/jjx008](#)]
- 4 **Peyrin-Biroulet L**, Sandborn W, Sands BE, Reinisch W, Bemelman W, Bryant RV, D'Haens G, Dotan I, Dubinsky M, Feagan B, Fiorino G, Geary R, Krishnareddy S, Lakatos PL, Loftus EV Jr, Marteau P, Munkholm P, Murdoch TB, Ordás I, Panaccione R, Riddell RH, Ruel J, Rubin DT, Samaan M, Siegel CA, Silverberg MS, Stoker J, Schreiber S, Travis S, Van Assche G, Danese S, Panes J, Bouguen G, O'Donnell S, Pariente B, Winer S, Hanauer S, Colombel JF. Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): Determining Therapeutic Goals for Treat-to-Target. *Am J Gastroenterol* 2015; **110**: 1324-1338 [PMID: [26303131](#) DOI: [10.1038/ajg.2015.233](#)]
- 5 **Peyrin-Biroulet L**, Panés J, Sandborn WJ, Vermeire S, Danese S, Feagan BG, Colombel JF, Hanauer SB, Rycroft B. Defining Disease Severity in Inflammatory Bowel Diseases: Current and Future Directions. *Clin Gastroenterol Hepatol* 2016; **14**: 348-354.e17 [PMID: [26071941](#) DOI: [10.1016/j.cgh.2015.06.001](#)]
- 6 **Mohammed Vashist N**, Samaan M, Mosli MH, Parker CE, MacDonald JK, Nelson SA, Zou GY, Feagan BG, Khanna R, Jairath V. Endoscopic scoring indices for evaluation of disease activity in ulcerative colitis. *Cochrane Database Syst Rev* 2018; **1**: CD011450 [PMID: [29338066](#) DOI: [10.1002/14651858.CD011450.pub2](#)]
- 7 **Samaan MA**, Mosli MH, Sandborn WJ, Feagan BG, D'Haens GR, Dubcenco E, Baker KA, Levesque BG. A systematic review of the measurement of endoscopic healing in ulcerative colitis clinical trials: recommendations and implications for future research. *Inflamm Bowel Dis* 2014; **20**: 1465-1471 [PMID: [24831558](#) DOI: [10.1097/MIB.0000000000000046](#)]
- 8 **Feagan BG**, Sandborn WJ, D'Haens G, Pola S, McDonald JWD, Rutgeerts P, Munkholm P, Mittmann U, King D, Wong CJ, Zou G, Donner A, Shackelton LM, Gilgen D, Nelson S, Vandervoort MK, Fahmy M, Loftus EV Jr, Panaccione R, Travis SP, Van Assche GA, Vermeire S, Levesque BG. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013; **145**: 149-157.e2 [PMID: [23528626](#) DOI: [10.1053/j.gastro.2013.03.025](#)]
- 9 **Daperno M**, Comberlato M, Bossa F, Biancone L, Bonanomi AG, Cassinotti A, Cosentino R, Lombardi G, Mangiarotti R, Papa A, Pica R, Rizzello F, D'Inca R, Orlando A. Inter-observer agreement in endoscopic scoring systems: preliminary report of an ongoing study from the Italian Group for Inflammatory Bowel Disease (IG-IBD). *Dig Liver Dis* 2014; **46**: 969-973 [PMID: [25154049](#) DOI: [10.1016/j.dld.2014.07.010](#)]
- 10 **Daperno M**, Comberlato M, Bossa F, Armuzzi A, Biancone L, Bonanomi AG, Cassinotti A, Cosentino R, Lombardi G, Mangiarotti R, Papa A, Pica R, Grassano L, Pagana G, D'Inca R, Orlando A, Rizzello F; IGIBDendo Group. Training Programs on Endoscopic Scoring Systems for Inflammatory Bowel Disease Lead to a Significant Increase in Interobserver Agreement Among Community Gastroenterologists. *J Crohns Colitis* 2017; **11**: 556-561 [PMID: [28453758](#) DOI: [10.1093/ecco-jcc/jjw181](#)]
- 11 **Rutgeerts P**, Reinisch W, Colombel JF, Sandborn WJ, D'Haens G, Petersson J, Zhou Q, Iezzi A, Thakkar RB. Agreement of site and central readings of ileocolonoscopy scores in Crohn's disease: comparison using data from the EXTEND trial. *Gastrointest Endosc* 2016; **83**: 188-197 [PMID: [26234693](#) DOI: [10.1016/j.gie.2015.06.018](#)]
- 12 **Gralnek IM**, Defranchis R, Seidman E, Leighton JA, Legnani P, Lewis BS. Development of a capsule endoscopy scoring index for small bowel mucosal inflammatory change. *Aliment Pharmacol Ther* 2008; **27**: 146-154 [PMID: [17956598](#) DOI: [10.1111/j.1365-2036.2007.03556.x](#)]
- 13 **Cotter J**, Dias de Castro F, Magalhães J, Moreira MJ, Rosa B. Validation of the Lewis score for the evaluation of small-bowel Crohn's disease activity. *Endoscopy* 2015; **47**: 330-335 [PMID: [25412092](#) DOI: [10.1055/s-0034-1390894](#)]
- 14 **Esaki M**, Washio E, Morishita T, Sakamoto K, Fuyuno Y, Hirano A, Umeno J, Kitazono T, Matsumoto T, Suzuki Y. Su1262 inter- and intra-observer variation of capsule endoscopic findings for the diagnosis of crohn's disease: A case-control study. *Gastrointest Endosc* 2018; **87**: AB302 [DOI: [10.1016/j.gie.2018.04.1649](#)]
- 15 **Buch VH**, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities. *Br J Gen Pract* 2018; **68**: 143-144 [PMID: [29472224](#) DOI: [10.3399/bjgp18X695213](#)]
- 16 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADE-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: [31981517](#) DOI: [10.1016/S2468-1253\(19\)30411-X](#)]

- 17 **Wang Y**, He X, Nie H, Zhou J, Cao P, Ou C. Application of artificial intelligence to the diagnosis and therapy of colorectal cancer. *Am J Cancer Res* 2020; **10**: 3575-3598 [PMID: [33294256](#)]
- 18 **Yang YJ**. The Future of Capsule Endoscopy: The Role of Artificial Intelligence and Other Technical Advancements. *Clin Endosc* 2020; **53**: 387-394 [PMID: [32668529](#) DOI: [10.5946/ce.2020.133](#)]
- 19 **Klang E**, Barash Y, Margalit RY, Soffer S, Shimon O, Albshesh A, Ben-Horin S, Amitai MM, Eliakim R, Kopylov U. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc* 2020; **91**: 606-613 [PMID: [31743689](#) DOI: [10.1016/j.gie.2019.11.012](#)]
- 20 **Soffer S**, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, Kopylov U, Barash Y. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. *Gastrointest Endosc* 2020; **92**: 831-839.e8 [PMID: [32334015](#) DOI: [10.1016/j.gie.2020.04.039](#)]
- 21 **Satsangi J**, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 2006; **55**: 749-753 [PMID: [16698746](#) DOI: [10.1136/gut.2005.082909](#)]
- 22 **Stidham RW**, Liu W, Bishu S, Rice MD, Higgins PDR, Zhu J, Nallamothu BK, Waljee AK. 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 [PMID: [31099869](#) DOI: [10.1001/jamanetworkopen.2019.3963](#)]
- 23 **Ozawa T**, Ishihara S, Fujishiro M, Saito H, Kumagai Y, Shichijo S, Aoyama K, Tada T. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc* 2019; **89**: 416-421.e1 [PMID: [30367878](#) DOI: [10.1016/j.gie.2018.10.020](#)]
- 24 **Gottlieb K**, Requa J, Karnes W, Chandra Gudivada R, Shen J, Rael E, Arora V, Dao T, Ninh A, McGill J. Central Reading of Ulcerative Colitis Clinical Trial Videos Using Neural Networks. *Gastroenterology* 2021; **160**: 710-719 [PMID: [33098883](#) DOI: [10.1053/j.gastro.2020.10.024](#)]
- 25 **Gutierrez Becker B**, Arcadu F, Thalhammer A, Gamez Serna C, Feehan O, Drawnel F, Oh YS, Prunotto M. Training and deploying a deep learning model for endoscopic severity grading in ulcerative colitis using multicenter clinical trial data. *Ther Adv Gastrointest Endosc* 2021; **14**: 2631774521990623 [PMID: [33718871](#) DOI: [10.1177/2631774521990623](#)]
- 26 **Yao H**, Najarian K, Gryak J, Bishu S, Rice MD, Waljee AK, Wilkins HJ, Stidham RW. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2021; **93**: 728-736 [PMID: [32810479](#) DOI: [10.1016/j.gie.2020.08.011](#)]
- 27 **Alammari A**, Islam ABMR, Oh JH, Tavanapong W, Wong J, De Groen PC. Classification of ulcerative colitis severity in colonoscopy videos using CNN. ICIME 2017: Proceedings of the 9th International Conference on Information Management and Engineering; 2019 Oct 9-11; Barcelona, Spain. Barcelona Spain, 2017: 139-44 [DOI: [10.1145/3149572.3149613](#)]
- 28 **Lo B**, Liu Z, Bendtsen F, Igel C, Vind I, Burisch J. Deep Learning Surpasses Gastrointestinal Experts at Classifying Endoscopic Severity in Patients with Ulcerative Colitis. 2021 Preprint. Available from: SSRN Electron Journal [DOI: [10.2139/ssrn.3824683](#)]
- 29 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. 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 [PMID: [30268542](#) DOI: [10.1016/j.gie.2018.09.024](#)]
- 30 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: [32060000](#) DOI: [10.1053/j.gastro.2020.02.012](#)]
- 31 **Salleron J**, Danese S, D'Agay L, Peyrin-Biroulet L. Effectiveness Research in Inflammatory Bowel Disease: A Necessity and a Methodological Challenge. *J Crohns Colitis* 2016; **10**: 1096-1102 [PMID: [26944416](#) DOI: [10.1093/ecco-jcc/jjw068](#)]
- 32 **Waljee AK**, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, Stidham RW, Zhu J, Higgins PDR. Development and Validation of Machine Learning Models in Prediction of Remission in Patients With Moderate to Severe Crohn Disease. *JAMA Netw Open* 2019; **2**: e193721 [PMID: [31074823](#) DOI: [10.1001/jamanetworkopen.2019.3721](#)]
- 33 **Waljee AK**, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, Higgins PDR. Predicting Corticosteroid-Free Biologic Remission with Vedolizumab in Crohn's Disease. *Inflamm Bowel Dis* 2018; **24**: 1185-1192 [PMID: [29668915](#) DOI: [10.1093/ibd/izy031](#)]
- 34 **Waljee AK**, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, Higgins PDR. Predicting corticosteroid-free endoscopic remission with vedolizumab in ulcerative colitis. *Aliment Pharmacol Ther* 2018; **47**: 763-772 [PMID: [29359519](#) DOI: [10.1111/apt.14510](#)]
- 35 **Biasci D**, Lee JC, Noor NM, Pombal DR, Hou M, Lewis N, Ahmad T, Hart A, Parkes M, McKinney EF, Lyons PA, Smith KGC. A blood-based prognostic biomarker in IBD. *Gut* 2019; **68**: 1386-1395 [PMID: [31030191](#) DOI: [10.1136/gutjnl-2019-318343](#)]
- 36 **Choi YI**, Park SJ, Chung JW, Kim KO, Cho JH, Kim YJ, Lee KY, Kim KG, Park DK. Development of Machine Learning Model to Predict the 5-Year Risk of Starting Biologic Agents in Patients with Inflammatory Bowel Disease (IBD): K-CDM Network Study. *J Clin Med* 2020; **9**: 3427 [PMID: [33114505](#) DOI: [10.3390/jcm9113427](#)]
- 37 **Waljee AK**, Lipson R, Wiitala WL, Zhang Y, Liu B, Zhu J, Wallace B, Govani SM, Stidham RW, Hayward R, Higgins PDR. Predicting Hospitalization and Outpatient Corticosteroid Use in

- Inflammatory Bowel Disease Patients Using Machine Learning. *Inflamm Bowel Dis* 2017; **24**: 45-53 [PMID: 29272474 DOI: 10.1093/ibd/izx007]
- 38 **Waljee AK**, Joyce JC, Wang S, Saxena A, Hart M, Zhu J, Higgins PD. Algorithms outperform metabolite tests in predicting response of patients with inflammatory bowel disease to thiopurines. *Clin Gastroenterol Hepatol* 2010; **8**: 143-150 [PMID: 19835986 DOI: 10.1016/j.cgh.2009.09.031]
- 39 **Vind I**, Riis L, Jess T, Knudsen E, Pedersen N, Elkjaer M, Bak Andersen I, Wewer V, Nørregaard P, Moesgaard F, Bendtsen F, Munkholm P; DCCD study group. Increasing incidences of inflammatory bowel disease and decreasing surgery rates in Copenhagen City and County, 2003-2005: a population-based study from the Danish Crohn colitis database. *Am J Gastroenterol* 2006; **101**: 1274-1282 [PMID: 16771949 DOI: 10.1111/j.1572-0241.2006.00552.x]
- 40 **Li J**, Qian JM. Artificial intelligence in inflammatory bowel disease: current status and opportunities. *Chin Med J (Engl)* 2020; **133**: 757-759 [PMID: 32132365 DOI: 10.1097/CM9.0000000000000714]
- 41 **Mossotto E**, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep* 2017; **7**: 2427 [PMID: 28546534 DOI: 10.1038/s41598-017-02606-2]
- 42 **Douglas GM**, Hansen R, Jones CMA, Dunn KA, Comeau AM, Bielawski JP, Tayler R, El-Omar EM, Russell RK, Hold GL, Langille MGI, Van Limbergen J. Multi-omics differentially classify disease state and treatment outcome in pediatric Crohn's disease. *Microbiome* 2018; **6**: 13 [PMID: 29335008 DOI: 10.1186/s40168-018-0398-3]
- 43 **Duttagupta R**, DiRienzo S, Jiang R, Bowers J, Gollub J, Kao J, Kearney K, Rudolph D, Dawany NB, Showe MK, Stamato T, Getts RC, Jones KW. Genome-wide maps of circulating miRNA biomarkers for ulcerative colitis. *PLoS One* 2012; **7**: e31241 [PMID: 22359580 DOI: 10.1371/journal.pone.0031241]
- 44 **Sandberg K**, Yarger E, Saeed S. Updates in diagnosis and management of inflammatory bowel disease. *Curr Probl Pediatr Adolesc Health Care* 2020; **50**: 100785 [PMID: 32402535 DOI: 10.1016/j.cppeds.2020.100785]
- 45 **Oppenheim M**. Stephen Hawking: Artificial intelligence could be the greatest disaster in human history. [cited 2021 Apr 11]. Available from: <https://www.independent.co.uk/news/people/stephen-hawking-artificial-intelligence-disaster-human-history-leverhulme-centre-cambridge-a7371106.html>
- 46 **Buhrmester V**, Münch D, Arens M. Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey. 2019 Preprint. Available from: arXiv:1911.12116
- 47 **Hatcher M**, Heetebry I. Information technology in the future of health care. *J Med Syst* 2004; **28**: 673-688 [PMID: 15615295 DOI: 10.1023/b:joms.0000044969.66510.d5]
- 48 **Borgli H**, Thambawita V, Smedsrud PH, Hicks S, Jha D, Eskeland SL, Randel KR, Pogorelov K, Lux M, Nguyen DTD, Johansen D, Griwodz C, Stensland HK, Garcia-Ceja E, Schmidt PT, Hammer HL, Riegler MA, Halvorsen P, de Lange T. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Sci Data* 2020; **7**: 283 [PMID: 32859981 DOI: 10.1038/s41597-020-00622-y]
- 49 **van der Sommen F**, de Groof J, Struyvenberg M, van der Putten J, Boers T, Fockens K, Schoon EJ, Curvers W, de With P, Mori Y, Byrne M, Bergman JJGHM. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. *Gut* 2020; **69**: 2035-2045 [PMID: 32393540 DOI: 10.1136/gutjnl-2019-320466]
- 50 **Cohen JF**, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HC, Bossuyt PM. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; **6**: e012799 [PMID: 28137831 DOI: 10.1136/bmjopen-2016-012799]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

Telephone: +1-925-3991568

E-mail: bpgoffice@wjgnet.com

Help Desk: <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

