## Review

# Artificial intelligence-assisted endoscopy changes the definition of mucosal healing in ulcerative colitis

Hiroshi Nakase, ⓘ Takehiro Hirano, Kohei Wagatsuma, Tadashi Ichimiya, Tsukasa Yamakawa, Yoshihiro Yokoyama, Yuki Hayashi, Daisuke Hirayama, Tomoe Kazama, Shinji Yoshii ⓘ and Hiro-o Yamano

*Department of Gastroenterology and Hepatology, Sapporo Medical University School of Medicine, Hokkaido, Japan*

The relevance of endoscopic monitoring of ulcerative colitis (UC) has been translated into the new concept of "mucosal healing (MH)" as the therapeutic goal to achieve because a large amount of scientific data have revealed the favorable prognostic value of a healed mucosa in determining the clinical outcome of UC. Recent interest in MH has skewed toward not only endoscopic remission but also histological improvement (so called histological MH). However, we should recognize that there have been no prospectively validated endoscopic scoring systems of UC activity in previous clinical trials. Artificial intelligence (AI)-assisted endoscopy has been developed for gastrointestinal cancer surveillance. Recently, several AI-assisted endoscopic systems have been developed for assessment of MH in UC. In the future, the development of a new endoscopic scoring system based on AI might standardize the definition of MH. Therefore, "The road to an exact definition of MH in the treatment of UC has begun only now".

**Key words:** artificial intelligence, histological healing, mucosal healing, red density system, ulcerative colitis

## INTRODUCTION

I N flammatory bowel diseases (IBD) comprise the two major forms of chronic digestive inflammation: Crohn's disease (CD) and ulcerative colitis (UC).[1] Both are immunopathogenic complex diseases in which, on the basis of genetic susceptibility of the host, an excessive mucosal immune response against the complex enteric microbiota plays a role in the initiation and perpetuation of intestinal inflammation.[2–4] Gastrointestinal (GI) endoscopy is clinically important for the management of IBD patients because of its involvement in the diagnosis of UC and CD, in monitoring the response to anti-inflammatory therapy, and dysplasia surveillance. Many physicians have understood that mucosal healing (MH) is a key clinical treatment goal in IBD patients.[5] Since it has been shown that complete MH is one of the most crucial aspects that predicts sustained clinical remission, the need for hospitalization and the probability of surgery-free survival among patients, a precise and detailed real-time assessment of the mucosal surface has become more important than ever for the medical management of IBD patients.[6–8] In this context, advanced endoscopic imaging techniques including endocytoscopy (EC; Olympus, Tokyo, Japan) and confocal laser endomicroscopy (CLE) have been applied to allow precise, ultrastructural and even microscopic characterization of the inflammation of UC.[9] However, even if we used these modalities to assess MH in UC, the interpretation of the gained findings was operator-dependent. Artificial intelligence (AI)-assisted endoscopy has been developed for evaluation of endoscopic activity of UC. In this article, we review the current literature on AI-associated endoscopy regarding the assessment of MH with histological remission in UC patients.

## WHAT IS MACHINE LEARNING, DEEP LEARNING, AND ARTIFICIAL INTELLIGENCE?

### Machine learning

I N 1959, SAMUEL defined machine learning as the "field of study that gives computers the ability to learn without being explicitly programmed".[10] There are four types of algorithms involved in machine learning (Fig. 1). (i) Supervised learning: This is a method of training the computer by labeling the training data tagged with the correct answer. (ii) Unsupervised learning: This is a method

of training without marking the training data. (iii) Semisupervised learning: This method utilizes a large amount of unlabeled data based on a small amount of labeled data, making it easier to train the model (computer). (iv) Reinforcement learning: In this method, the network learns behaviors to maximize the "rewards" (scores) set as objectives in a given "environment".[11–13]

## Neural network system

A neural network is a system that imitates the mechanism of human nerve cells (neurons). By using neural networks in multiple layers, it becomes possible to learn deeper features contained in the data in a stepwise fashion. By inputting a large amount of image, text, voice and other kinds of data into a multilayered neural network (convolutional neural network, CNN), the computer model automatically learns the features contained in the data in each layer. Deep learning (DL) is based on neural network systems. These structures are unique to DL, making DL models accurate and sometimes more accurate than humans.

## Deep learning

Deep learning is a machine learning method that allows a computer to learn the tasks that humans naturally perform.[14] It is a technology that supports the rapid development of the application of artificial intelligence (AI) to practical tasks in many fields. For example, DL is the key to self-driving cars and voice recognition. DL has been gaining attention recently because DL is becoming capable of achieving high levels of performance. The difference between traditional machine learning methods and DL is whether or not a human feature extraction is required (Fig. 2).[15,16]

## Artificial intelligence

Artificial intelligence involves a computer performing intelligent tasks that are normally performed by humans, such as learning and problem-solving.[16,17] There are two types of AI; general-purpose AI and specialized AI. Briefly, AI that solves various problems in different domains is called general-purpose AI. AI specializing in individual domains to demonstrate its capabilities is called specialized AI. For example, there is AI that specializes in voice recognition, and AI specialized in Go or Shogi. Figure 3 shows the relationship between DL and AI/machine learning.

## MUCOSAL HEALING IN UC – THE SUBJECTIVE ENDOSCOPIC SCORE AS THE MAIN STAY OF MH ASSESSMENT

THE CONCEPT OF MH was introduced approximately 50 years ago.[18] In a clinical practice, physicians have accepted the definition of MH as "the complete resolution of the visible alterations or lesions, regardless of their severity and/or type at baseline colonoscopy". Although there have been 50 years of clinical trials and endoscopic scoring systems with different designs have been developed (Baron score, Mayo score, Sutherland, Powell-Tuck and Rachmilewitz indices, among others),[19–16] the definitions and scoring methods of these instruments have never been prospectively validated. Two new scoring systems were developed and prospectively validated: the Ulcerative Colitis Endoscopic Index of Severity (UCEIS) and the Ulcerative Colitis
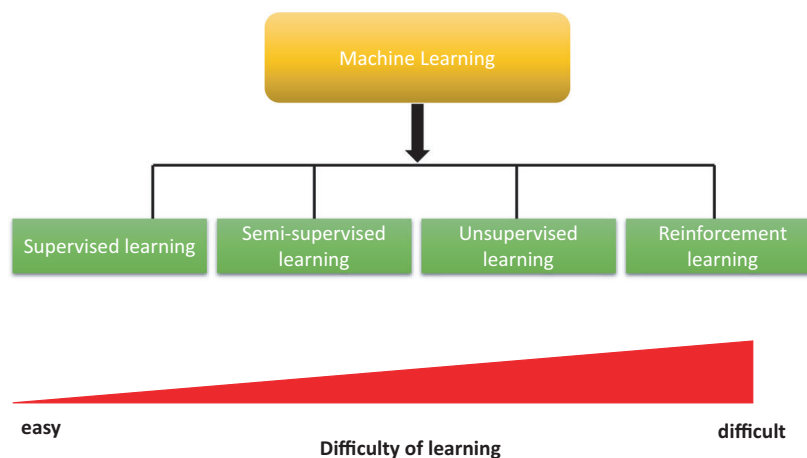


**Figure 1** Four kinds of algorithm regarding machine learning with the difficulty of learning.
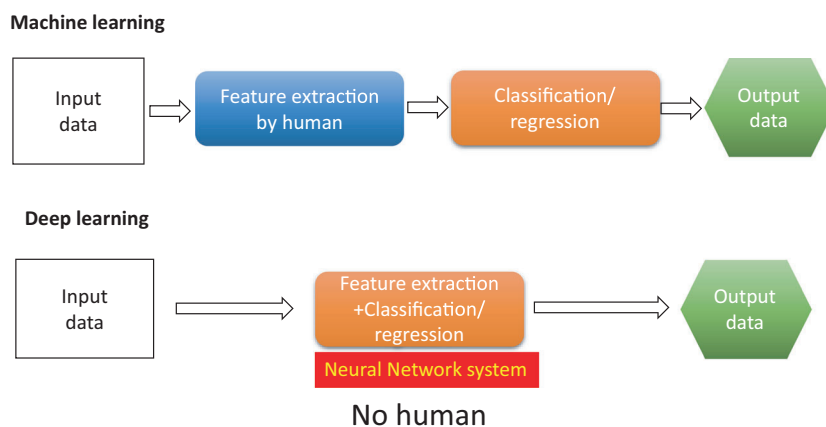
**Figure 2** The difference between traditional machine learning methods and deep learning.
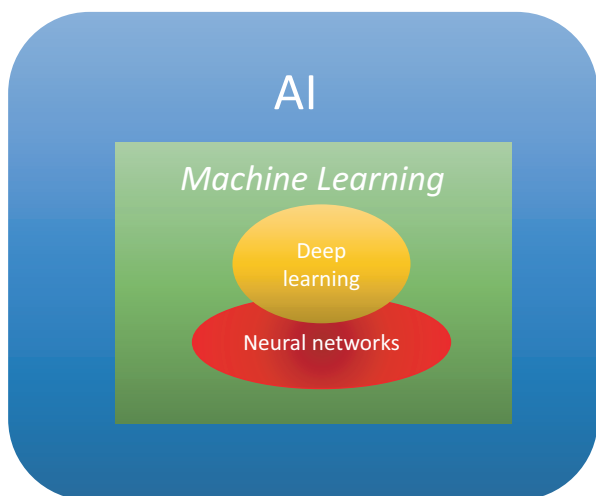


**Figure 3** The relationship between deep learning and machine learning/artificial intelligence.

Colonoscopic Index of Severity.[27,28] Nevertheless, no validated endoscopic score for the evaluation of UC activity evaluation reflects the complexity.

## HISTOLOGICAL MH – IS THIS AN IDEAL THERAPEUTIC GOAL?

THE NECESSITY OF the histological evaluation of the colonic mucosa of UC patients has been emphasized. Bessissow *et al.*[29] reported that the presence of basal plasmacytosis predicts UC clinical relapse in patients with complete MH. Peyrin-Biroulet *et al.*[30] indicated that data indicating a prognostically relevant role for histologic activity in the mucosa of UC patients, and that finding, in

addition to the macroscopic activity, has opened the door to the concept of "histological MH", with the complete absence of clinical, laboratory, endoscopic and histological features of active inflammation. However, it remains unclear whether the histological results of rectal biopsy reflect inflammatory conditions in the entire colonic mucosa of patients with extensive UC who have achieved endoscopic remission.

## EMERGENCE OF CONFOCAL LASER ENDOMICROSCOPY AND ENDOCYTOSCOPY FOR THE ASSESSMENT OF MUCOSAL INFLAMMATION IN UC

NEW ENDOSCOPIC TECHNIQUES such as in vivo confocal endomicroscopy (CLE), focus on histologic assessment. CLE allows the in vivo microscopic imaging of cellular and subcellular structures at approximately 1000×-fold magnification. This intriguing technique has the capability to obtain "real-time" histology-like images of the GI mucosa. Several reports have demonstrated that CLE enables us to identify the different features of colonic crypts in the inflamed and noninflamed rectal mucosa in UC patients and to visualize bacterial translocation and abnormal epithelial barrier function.[31–33] Thus, CLE can contribute to the understanding of the pathophysiology of UC. Endocytoscopy (EC) is based on the principle of contact light microscopy. EC systems are either integrated into the distal tip of a standard endoscope (iEC) or probe-based (pEC).[34,35] For example, pEC visualizes architectural details, cellular features, and vascular pattern morphology at a magnification of up to 1390×. Bessho *et al.*[36] reported the correlation of assessment of microscopic features in UC patients between EC and conventional histopathology.

Interestingly, a scoring system based on EC findings was strongly correlated with Matts' histological grades. Taken together, these findings show that both CLE and EC have strong potential for identifying subtle mucosal inflammation in real time by identifying single mucosal inflammatory cells in conjunction with architectural changes. On the other hand, we must acknowledge that the mucosal inflammation detected by CLE and EC only reflects a part of the histological condition in the colonic mucosa of UC alone because the observational range of intestinal mucosa for these modalities is very limited.

## IMPORTANCE OF AI-ASSISTED ENDOSCOPY FOR EVALUATING MUCOSAL INFLAMMATION

IMAGE RECOGNITION USING AI techniques can be used to grade endoscopic images and morphology. AI systems allow us to make assessments with lower bias and more objective interpretation methods not only for GI cancer but also for IBD. As mentioned earlier, objective treatment targets are required in a treat to target (T2T) approach for UC. Many physicians hope for the development of an automated endoscopic scoring system that correlates with histology because the AI-assisted evaluation of mucosal inflammation can serve as an objective predictor of sustained remission in UC. Currently, a computer-aided diagnosis (CAD) system, has been applied to the objective assessment of mucosal inflammation in UC. The characteristics of AI-assisted endoscopy are shown in Table 1.

## Computer-aided diagnosis of endoscopic activity in UC

Even if endoscopists experienced the evaluation of the inflammatory activity in UC, interobserver variability still exists. Ozawa *et al.*[37] reported the construction of a CAD system using a CNN, and the performance of the system was evaluated by tagging a large dataset of endoscopic images from patients with UC. They constructed this CNN-based CAD system based on GoogLeNet architecture. The data used to train the CNN were 26,304 colonoscopic images from a cumulative total of 841 patients with UC. These images were tagged with Mayo endoscopic scores (MESs). The performance of the trained CNN with regard to identifying normal mucosa (MES 0) and mucosal healing mucosa (MES 0-1) was evaluated in an independent test set of 3981 images from 114 patients with UC by calculating the areas under the receiver operating characteristic curves (AUROCs). As expected, this CNN-based CAD system had excellent performance, with AUROCs of 0.86 and 0.98 for the identification of MES 0 and 0-1, respectively. The authors stated that the CNN-based CAD system would support less-experienced endoscopists and reduce interobserver variability with regard to the MES. Unfortunately, this CAD system was not tagged with the histological findings.

## EndoBRAIN system (EndoBRAIN; Cybernet Systems, Tokyo, Japan)

Mori *et al.* reported the usefulness of a CAD system for EC imaging (EC-CAD) to allow untrained, nonexpert endoscopists to derive benefits from the high diagnostic performance of EC. This system provides fully automated instant classification of colorectal polyps during routine colonoscopy.[38] Maeda *et al.*[39] applied this CAD system to develop and evaluate the prediction of histologic inflammation in UC patients. They retrospectively reviewed the data of 187 patients with UC from whom biopsy samples were obtained after EC. All EC images were tagged with reference to the histologic activity observed in the biopsy samples. They validated 525 sets of 525 independent segments, which were collected from 100 patients. Machine learning was applied to 12,900 EC images from the

**Table 1**  The characteristics of artificial intelligence-assisted endoscopy

|  | Red density[44] | MAGIC score[43] | CAD system for Mayo score[37] | Endo brain[39] | Fujifilm research[40] (estimation mucosal capillary vessel) | Sony research[41] | Michigan research[42] |
|---|---|---|---|---|---|---|---|
| Range of assessment of mucosal inflammation[†] | Large | Large | Large | Small | Small | Large | Large |
| Scoring system | Yes | Yes | No | Yes | No | Yes | Yes |
| Correlation with histology | Yes | Yes | No | Yes | Yes | Yes | No |

[†]Large: whole assessment of inflammation in colonic mucosa on image. Small: partial assessment of inflammation in colonic mucosa on magnified image.

remaining 87 patients and were used to construct the CAD system. This CAD system was shown to have a diagnostic sensitivity, specificity, and accuracy of 74% (95% confidence interval (CI), 65–81%), 97% (95% CI, 95–99%), and 91% (95% CI, 83–95%), respectively. The authors concluded that the EC-CAD system could contribute to the fully automated identification of the persistent histologic inflammation associated with UC. However, we should keep in mind that the inflammation of the colonic mucosa in UC is not uniform, but heterogeneous. The images captured by the EC-CAD system does not necessarily reveal the complete circumferential colonic inflammatory status.

## Histological evaluation based on mucosal capillary pattern in UC by the CAD system

It is recognized that the infiltration of neutrophils is associated with irregularities in the pericryptal capillaries in the inflamed mucosa. Bossuyt et al.[40] developed an objective automated endoscopic tool to assess histological remission by focusing on the evaluation of the morphology of the pericryptal capillaries. The algorithm included two steps. First, bleeding (mucosal/luminal) was assessed by pattern recognition (number of pixels: bleeding (NPBL)). Samples with bleeding (high NPBL) were automatically classified as nonremission. Second, in the case of nonbleeding samples (low NPBL), the degree of congestion of the capillaries (number of pixels: high density (NPHD)) was measured to assess an ideal cut-off value to identify histological remission (Geboes score (GBS) <2B.1; no neutrophils in the lamina propria). This algorithm for automated image analysis detected histological remission with a higher performance (sensitivity 0.79, specificity 0.90) than the UCEIS (sensitivity 0.95, specificity 0.69) and MES (sensitivity 0.98, specificity 0.61), resulting in positive predictive values of 0.83, 0.65 and 0.59 for the algorithm of automated image analysis algorithm, UCEIS and MES, respectively. Mucosal capillary pattern recognition based on an automated image analysis with short-wavelength monochromatic light detects histological remission with high accuracy in UC patients. The authors concluded that this technique provides an objective and quantitative tool for the assessment of histological remission in UC. However, the NPHD value only reflects the grade of mucosal inflammation in a limited part of the colonic mucosa, as in the EndoBRAIN system.

## Deep neural network for grading the endoscopic disease severity of UC

Takenaka et al.[41] reported the development of a deep neural network system for the consistent, objective, and real-time analysis of endoscopic images from patients with UC. They constructed a deep neural network for the evaluation of UC (DNUC) algorithm with reference to 40,758 images of colonoscopies tagged with 6885 biopsy results. The accuracy of the DNUC algorithm was validated in a prospective study of 875 patients with UC who underwent colonoscopy from April 2018 to April 2019, with 4187 endoscopic images and 4104 biopsy specimens. They defined endoscopic remission (ER) as a UCEIS of 0; and histologic remission as a Geboes score of three points or less. The results of this study demonstrated that the accuracy ratio of identifying UC patients with ER by the DNUC algorithm was 90.1% accuracy (95% CI 89.2–90.9%) and the accuracy of identifying UC patients with histologic remission by the DNUC algorithm was 92.9% (95% CI 92.1–93.7%). Takenaka et al. concluded that the DNUC algorithm could provide objective, consistent and real-time endoscopic evaluation. Additionally, they emphasized that the DNUC algorithm could be not only an alternative to a central reader in clinical trials but also a useful tool in the training of gastroenterologists for the assessment of endoscopic activity of UC. This is a novel system of evaluation of circumferential mucosal inflammation in UC, therefore, we strongly hope that this system will become part of routine clinical practice. On the other hand, a multicenter prospective study is needed to confirm the accuracy of this system because the DNUC algorithm was constructed with data from a single center.

## Application of CNN to generate DL models for grading the endoscopic disease severity of UC

Stidham et al.[42] investigated the feasibility of generating a DL algorithm to grade the endoscopic severity of UC and applied it to full-motion video recordings of colonoscopies. Based on data from Michigan's endoscopic imaging database (16,514 colonoscopic images in 3082 UC patients) and the university's electronic health records, they constructed a 159-layer CNN as a DL model to train and categorize images into two groups: ER (defined by MES 0 or 1) and active disease (MES 2 or 3). They showed that this CNN could distinguish ER from active disease with an AUROC of 0.966 (95% CI, 0.967–0.972); a positive predictive value (PPV) of 0.87 (95% CI, 0.85–0.88) with a sensitivity of 83.0% (95% CI, 80.8–85.4%), and a specificity of 96.0% (95% CI, 95.1–97.1%); and a negative predictive value (NPV) of 0.94 (95% CI, 0.93–0.95). Applying the CNN to entire colonoscopy videos had similar accuracy for identifying moderate-to-severe disease (AUROC, 0.97; 95% CI, 0.963–0.969). Thus, the performance of DL models to approximate the disease assessment made by experts has excellent reproducibility, objectivity, and

speed. However, a limitation of this study was that they did not construct the system based on histological data. In addition, they did not differentiate between MES 0 and 1. Therefore, we hope that a more precise DL model will be developed that can differentiate MES 0 and 1 based on tagged histologic results because several studies have highlighted the distinct difference in the relapse rate between clinically quiescent UC patients with MES 0 and those with MES 1.

## Development of the MAGIC score for the assessment of mucosal inflammation

The assessment of a MES of 0 or 1 varies widely between endoscopists. Additionally, several studies have shown that endoscopic MH does not necessarily reflect quiescent microscopic UC disease activity. It should be emphasized that no scoring system based on the combination of endoscopic images and histology has been developed for clinical use. If we can develop a scoring system reflecting mucosal inflammation or specifically reflecting MH in UC, it will be ideal from the perspective of the current T2T concept. We retrospectively reviewed the data from 52 UC patients with clinical remission who had undergone routine colonoscopy with the i-scan TE-c system, which can enhance images of normal or inflamed colonic mucosa with improved color contrast.[43] With reference to images captured by the i-scan TE-c system, the degree of inflammation was quantified for the entire screen by correlating the value with the reference value for each pixel in the HSV color space. Based on the data, we developed the Mucosal Analysis of Inflammatory Gravity by i-scan TE-c Image (MAGIC) score by supervised machine learning, which was defined as the mean value of the quantified values for each pixel. Representative images of MAGIC score are shown in Figure 4. The associations among the MAGIC score, MES, and histologic activity (Geboes score) were investigated. The MAGIC score was significantly higher in the MES 1 group than in the MES 0 group (779.8 $\pm$ 488.4 vs. 487.2 $\pm$ 378.2, $P = 0.0034$) and was significantly correlated with the Geboes score ($P = 0.015$). Interestingly, even UC patients with MES 0 have widely ranging MAGIC scores, as do those with endoscopic MES 1. These data suggested that the stratification of mucosal inflammation in UC by the MAGIC scoring could be possible among UC patients with MESs of 0 or 1.

## Evolution of the MAGIC score: construction of the red density system

Using the MAGIC score as a foundation, we developed an operator-independent computer-based tool to determine UC disease activity based on endoscopic images, which we named the "Red density (RD) system".[44] First, a computer algorithm using data from 29 consecutive patients with UC and six healthy controls was constructed by supervised machine learning.

Red density was determined using normal images that were taken with a white light source and without magnification. The RD system was based on the red channel of the red-green-blue pixel values and pattern recognition from endoscopic images. The algorithm was refined in sequential steps to optimize the correlation with endoscopic and histological disease activity. Second, the operating properties were tested in patients with UC flares necessitating treatment escalation. To validate the algorithm, we tested the correlations between the RD system score and clinical, endoscopic and histological features in a validation cohort. Multiple regression analysis was used to assess the correlation of the data with Robarts histological index (RHI). The display of the RD system during endoscopy can be seen in the supplementary video (Fig 5, Video S1). We found that the RD score correlated with RHI ($r = 0.74$, $P < 0.0001$), the MES subscores ($r = 0.76$, $P < 0.0001$) and the UCEIS ($r = 0.74$, $P < 0.0001$). Thus, the RD system is a novel modality that provides an objective computer-based score that accurately assesses disease activity in UC. Currently, we are starting a larger, prospective study to confirm the predictive value and accuracy of the RD system score for MH. Currently, Pentax is seeking regulatory approval for the RD system and plan to hold meetings with the regulatory authorities before submission or registration.

## ROLE OF AI-ASSISTED ENDOSCOPY FOR THE SURVEILLANCE OF COLITIS-ASSOCIATED CANCER

IN ADDITION TO the assessment of MH in UC, the primary clinical issue is the surveillance of colitis-associated cancer (CAC). Chronically sustained inflammation triggers colitis-associated lesions, which tend to present in a flat or multifocal pattern.[45] Colitis-associated dysplasia and cancer have pathologically distinct patterns from sporadic adenoma/carcinoma. It is challenging to distinguish the morphological features of these lesions from the surrounding inflamed mucosa with standard white light endoscopy.[46] There are multiple new endoscopic imaging technologies, including standard chromoendoscopy with indigo carmine or methylene blue, virtual chromoendoscopy narrow-band imaging, Fujinon intelligent color enhancement, and i-Scan, autofluorescence imaging, CLE and EC, that have been applied to identify CAC.[47] However, there are few reports on the application of AI-assisted colonoscopy techniques for the diagnosis of CAC. Most recently,

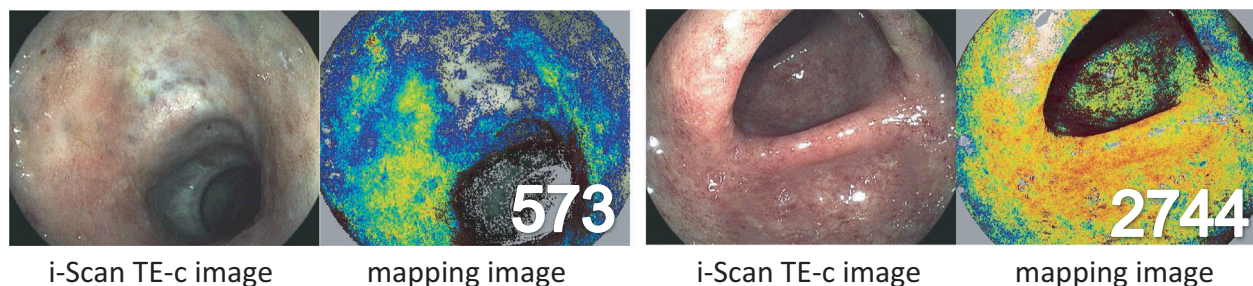i-Scan TE-c image    mapping image    i-Scan TE-c image    mapping image

**Figure 4** The representative images of MAGIC scores. i-Scan TE-c images (left) and mapping images (right). As the degree of inflammation increased, the calculated score of the mapping image increased.
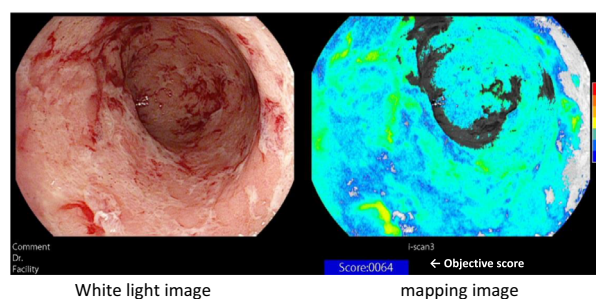


White light image    mapping image

**Figure 5** Red density (RD) image and score displayed on the monitor.

Fukunaga *et al.*[48] reported the usefulness of the Endo-BRAIN system for the diagnosis of high-grade dysplasia in a patient with a long-term history of UC. They concluded that the EndoBRAIN system could help nonexpert endoscopists identify CAC, thereby avoiding unnecessary biopsies. This case report strongly indicates the significance of EC for the diagnosis of neoplastic lesions associated with UC and sporadic neoplasia. However, there are still a small number of images of colitis-associated dysplasia and cancer in UC that have been histologically tagged. A large amount of imaging data on dysplasia and CAC in UC patients and validation is needed to improve the diagnostic performance of the EndoBRAIN system.

## LIMITATION OF AI FOR THE ASSESSMENT OF MUCOSAL INFLAMMATION IN UC

WE RECOGNIZE THAT a deeper model with a large number of parameters tends to be more accurate, and accuracy is also improved by combining the predictions of multiple models (ensemble) rather than using the predictions made by a single model. Based on these principles, a large number of images with an accurate diagnosis regarding MH is required. As mentioned earlier, many researchers have utilized a large number of endoscopic pictures evaluated by the MES, and UCEIS and tagged with the histologic results to develop an AI model for the assessment of intestinal inflammation in UC. However, there is still a problem because MH in UC lacks an exact, correct definition. Since DL is automatic, the content and direction of the learning change considerably depending on the data provided. Therefore, it is essential to select the data to be read carefully. In this regard, it remains unclear whether the present endoscopic images, endoscopic index and histologic score are suitable parameters for DL. At present, if we establish a more accurate AI model for the assessment of MH, endoscopic images combine with biological parameters such as the levels of fecal calprotectin and leucine-rich α2 glycoprotein should be used. The development of a graphic processing unit can contribute to the improvement of the recognition of details of endoscopic images; however, this may not be enough. The addition of molecular analysis data, such as cytokine and epithelial regenerative molecule expression in the intestinal mucosa of patients with long-term quiescent UC, as parameters for DL in the future may enable the precise identification of MH.

## IN SUMMARY

SINCE IT IS suggested that a T2T algorithm could improve outcomes in the long term in UC, the intra- and interobserver variations in endoscopic assessment of UC are a clinical issue. However, the development of AI-assisted endoscopy has led to growing interest in the use of new imaging technologies, and opened a new path toward not only the diagnosis of GI cancer but also the IBD field. Now, we are seeing AI being used in medical image interpretation to improve objectivity and reproducibility and there is a possibility that AI-assisted medical imaging can broaden access to expert-level of assessments or exceed it. There may even be a time when we do not grade or diagnose endoscopic images may come in near future. As AI systems contribute to the exact diagnosis and treatment of human

disease, we should continue to learn best practices in health care.

## ACKNOWLEDGMENT

THE AUTHORS WOULD like to thank Dr. Peter Bossuyt, Prof. Séverine Vermeire, Prof. Raf Bisshops and Yousuke Ikemoto, Takao Makino (Pentax Medical, Product development department, Tokyo, Japan)

## CONFLICT OF INTEREST

## FUNDING INFORMATION

## REFERENCES

1 Strober W, Fuss I, Mannon P. The fundamental basis of inflammatory bowel disease. *J Clin Invest* 2007; **117**: 514–21.

2 Cleynen I, Boucher G, Jostins L *et al*. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016; **387**: 156–67.

3 Jostins L, Ripke S, Weersma RK *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012; **491**: 119–24.

4 Parkes M. The genetics universe of Crohn's disease and ulcerative colitis. *Dig Dis* 2012; **30**: 78–81.

5 Nakase H. Optimizing the use of current treatments and emerging therapeutic approaches to achieve therapeutic success in patients with inflammatory bowel disease. *Gut Liv* 2020; **14**: 7–19.

6 Colombel JF, Rutgeerts P, Reinisch W *et al*. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. *Gastroenterology* 2011; **141**: 1194–201.

7 Frøslie KF, Jahnsen J, Moum BA *et al*. Mucosal healing in inflammatory bowel disease: results from a Norwegian population-based cohort. *Gastroenterology* 2007; **133**: 412–22.

8 Casellas F, Barreiro de Acosta M, Iglesias M *et al*. Mucosal healing restores normal health and quality of life in patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2012; **24**: 762–9.

9 Kim ES. Role of advanced endoscopic imaging techniques in the management of inflammatory bowel disease. *Clin Endosc* 2017; **50**: 424–8.

10 Samuel A. Some studies in machine learning using the game of checkers. *IBM J Res Dev* 1959; **3**: 210–29.

11 Mozaffari-Kermani M, Sur-Kolay S, Raghunathan A, Jha NK. Systematic poisoning attacks on and defenses for machine learning in healthcare. *IEEE J Biomed Health Inform.* 2015; **19**: 1893–905.

12 Abbasi B, Goldenholz DM. Machine learning applications in epilepsy. *Epilepsia* 2019; **60**: 2037–47.

13 Crites RH, Barto AG. Elevator group control using multiple reinforcement learning agents. *Mach Learn* 1998; **33**: 235–62.

14 Shen W, Zhou M, Yang F, Yang C, Tian J. Multi-scale convolutional neural networks for lung nodule classification. *Inf Process Med Imaging.* 2015; **24**: 588–99.

15 Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D. Deep learning for cellular image analysis. *Nat Methods.* 2019; **16**: 1233–46.

16 Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? *J Arthroplasty.* 2018; **33**: 2358–61.

17 Kohn LT, Corrigan J, Donaldson MS. *To Err is Human: Building a Safer Health System*. Washington: National Academy Press, 2000.

18 Wright R, Truelove SR. Serial rectal biopsy in ulcerative colitis during the course of a controlled therapeutic trial of various diets. *Am J Dig Dis* 1966; **11**: 847–57.

19 Neurath MF, Travis SP. Mucosal healing in inflammatory bowel diseases: a systematic review. *Gut* 2012; **61**: 1619–35.

20 Truelove SC, Witts LJ. Cortisone in ulcerative colitis; final report on a therapeutic trial. *Br Med J* 1955; **2**: 1041–8.

21 Baron JH, Connell AM, Lennard-jones JE. Variation between observers in describing mucosal appearances in proctocolitis. *Br Med J* 1964; **1**: 89–92.

22 Powell-Tuck J, Bown RL, Lennard-Jones JE. A comparison of oral prednisolone given as single or multiple daily doses for active proctocolitis. *Scand J Gastroenterol* 1978; **13**: 833–7.

23 Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med* 1987; **317**: 1625–9.

24 Sutherland LR, Martin F, Greer S *et al*. 5-Aminosalicylic acid enema in the treatment of distal ulcerative colitis, proctosigmoiditis, and proctitis. *Gastroenterology* 1987; **92**: 1894–8.

25 Rachmilewitz D. Coated mesalazine (5-aminosalicylic acid) versus sulphasalazine in the treatment of active ulcerative colitis: a randomised trial. *BMJ* 1989; **298**: 82–6.

26 Feagan BG, Greenberg GR, Wild G *et al* Treatment of ulcerative colitis with a humanized antibody to the alpha4beta7 integrin. *N Engl J Med*. 2005; **352**: 2499–507.

27 Travis SP, Schnell D, Krzeski P *et al*. Reliability and initial validation of the ulcerative colitis endoscopic index of severity. *Gastroenterology* 2013; **145**: 987–95.

28 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.

29 Bessissow T, Lemmens B, Ferrante M *et al*. Prognostic value of serologic and histologic markers on clinical relapse in ulcerative colitis patients with mucosal healing. *Am J Gastroenterol*. 2012; **107**: 1684–92.

30 Peyrin-Biroulet L, Bressenot A, Kampman W. Histologic remission: the ultimate therapeutic goal in ulcerative colitis? *Clin Gastroenterol Hepatol*. 2014; **12**: 929–34.e2.

31 Watanabe O, Ando T, Maeda O *et al*. Confocal endomicroscopy in patients with ulcerative colitis. *J Gastroenterol Hepatol*. 2008; **23**(Suppl 2): S286–90.

32 Li CQ, Xie XJ, Yu T *et al*. Classification of inflammation activity in ulcerative colitis by confocal laser endomicroscopy. *Am J Gastroenterol*. 2010; **105**: 1391–6.

33 Moussata D, Goetz M, Watson AJ *et al*. Confocal laser endomicroscopy is a new imaging modality for recognition of intramucosal bacteria in inflammatory bowel disease in vivo. *Gut* 2011; **60**: 26–33.

34 Neumann H, Fuchs FS, Vieth M *et al*. Review article: in vivo imaging by endocytoscopy. *Aliment Pharmacol Ther* 2011; **33**: 1183–93.

35 Minami H, Inoue H, Yokoyama A, *et al* Recent advancement of observing living cells in the esophagus using CM double staining: endocytoscopic atypia classification. *Dis Esophagus*. 2012; **25**: 235–41.

36 Bessho R, Kanai T, Hosoe N *et al*. Correlation between endocytoscopy and conventional histopathology in microstructural features of ulcerative colitis. *J Gastroenterol*. 2011; **46**: 1197–202.

37 Ozawa T, Ishihara S, Fujishiro M *et al*. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc*. 2019; **89**: 416–21.e1.

38 Mori Y, Kudo SE, Wakamura K *et al*. Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos). *Gastrointest Endosc* 2015; **81**: 621–9.

39 Maeda Y, Kudo SE, Mori Y *et al*. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc*. 2019; **89**: 408–15.

40 Bossuyt P, Hertogh GD, Eelbode T, Vermeire S, Bisschops R. P269 Mucosal capillary pattern recognition based on real-time computer-aided image analysis adequately detects histological remission in ulcerative colitis. *J Crohn's Colitis* 2020; **14**(Suppl 1): S284–5.

41 Takenaka K, Ohtsuka K, Fujii T *et al*. Development and validation of a deep neural network for grading endoscopic disease severity of UC. *Gastroenterology* 2020; **158**: 2150–7.

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

43 Honzawa Y, Matsuura M, Higuchi H, Sakurai T, Seno H, Nakase H. A Novel endoscopic imaging system for quantitative evaluation of colonic mucosal inflammation in patients with quiescent ulcerative colitis. *Endosc Int Open*. 2020; **8**: E41–E49.

44 Bossuyt P, Nakase H, Vermeire S *et al*. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut* 2020; **69**: 1778–86.

45 Odze RD, Farraye FA, Hecht JL, Hornick JL. Long-term follow-up after polypectomy treatment for adenoma-like dysplastic lesions in ulcerative colitis. *Clin. Gastroenterol. Hepatol*. 2004; **2**: 534–41.

46 Oka S, Tanaka S, Chayama K. Detection of nonpolypoid colorectal neoplasia using magnifying endoscopy in colonic inflammatory bowel disease. *Gastrointest Endoscopy Clin N Am*. 2014; **24**: 405–17.

47 Bessissow T, Bisschops R. Advanced endoscopic imaging for dysplasia surveillance in ulcerative colitis. *Expert Rev Gastroenterol Hepatol* 2013; **7**: 57–67.

48 Fukunaga S, Kusaba Y, Ohuchi A *et al*. Is artificial intelligence a superior diagnostician in ulcerative colitis? *Endoscopy* Published online: 26 Jun 2020; https://doi.org/10.1055/a-1195-1986

## SUPPORTING INFORMATION

ADDITIONAL SUPPORTING INFORMATION may be found in the online version of this article at the publisher's web site.

**Video S1** The RD image and score was seen in real time, along with the white light image.