The histological detection of ulcerative colitis using a no-code artificial intelligence model

Yuichiro Hamamoto[1,2], Michihiro Kawamura[3], Hiroki Uchida[3], Kazuhiro Hiramatsu[3], Chiaki Katori[3], Hinako Asai[3], Shigeki Shimizu[4], Satoshi Egawa[5], Kyotaro Yoshida[3]

1 Department of Diagnostic Pathology, Kinki Central Hospital, Itami, Hyogo.
2 Department of Pathology, Osaka University Graduate School of Medicine, Suita.
3 Department of Clinical Laboratory, Kinki Central Hospital, Itami, Hyogo.
4 Department of Clinical Laboratory, National Hospital Organization Kinki-Chuo Chest Medical Center, Kita-ku, Sakai, Osaka, Japan
5 Department of Gastroenterology, Kinki Central Hospital, Itami, Hyogo.

Corresponding author:
Yuichiro Hamamoto, MD, PhD
Department of Diagnostic Pathology, Kinki central hospital, 3-1 Kurumaduka, Itami, Hyogo 664-8533, Japan.
Tel: +81-72-781-3712, Fax: +81-72-779-1567
E-mail: kyokui090088@gmail.com

Key word
Artificial intelligence, Machine learning, Ulcerative colitis, Inflammatory bowel disease, Adenocarcinoma

Running head
Artificial intelligence in ulcerative colitis

Conflict of interest
The authors declare no conflict of interest.

Funding disclosures

This work was financially supported by a research grant from the Kinki Central Hospital.

Acknowledgement

The authors thank Yumiko Yasuhara (Department of Pathology, Sakai City Medical Center, Sakai, Osaka, Japan) and all gastroenterologists of Kinki Center Hospital. The authors also thank Editage (http://www.editage.jp) for English language editing.

Ethics approval

This study was approved by the Ethical Review Board of the Kinki Central Hospital (Proposal No. 454, 467, 477). All procedures were performed in accordance with the guidelines and regulations of the committee. We have stated an opt-out policy on the website of the Kinki Central Hospital.

Abstract

Ulcerative colitis (UC) is an intractable disease that affects young adults. Histological findings are essential for its diagnosis; however, the number of diagnostic pathologists is limited. Herein, we used a no-code artificial intelligence (AI) platform "Teachable Machine" to train a model that could distinguish between histological images of UC, non-UC coloproctitis, adenocarcinoma, and control. A total of 5,100 histological images for training and 900 histological images for testing were prepared by pathologists. Our model showed accuracies of 0.99, 1.00, 0.99, and 0.99, for UC, non-UC coloproctitis, adenocarcinoma, and control, respectively. This is the first report in which a no-code easy AI platform has been able to comprehensively recognize the distinctive histologic patterns of UC.

Introduction

Ulcerative colitis (UC) is a type of chronic inflammatory bowel disease (IBD) that occurs frequently in young adults. However, its pathogenesis is not completely known[1]. The diagnosis of this disease is based on clinical, endoscopic, and histological findings. Histological characteristics, such as basal plasmacytosis, crypt distortion, and crypt abscesses, are commonly observed in UC and help make a definite diagnosis. Differentiation from Crohn's disease (CD), a form of IBD, and other forms of enteritis is important.

Since the number of diagnostic pathologists is limited, the utilization of artificial intelligence (AI) holds a significant potential for the diagnosis of UC. Numerous AI studies have been conducted in the fields of endoscopy and histology[2, 3]. However, medical staff often encounter challenges in adopting and utilizing AI technology. It would be ideal to implement a no-code easy AI, trained in pathology; however, only a few studies have been reported[4-7]. To the best of our knowledge, AI that can comprehensively recognize histological characteristics of UC has not been reported. In this study, we demonstrated for the first time, a no-code pathological AI system that can accurately detect UC.

Materials and Methods

Patients

We collected colorectal specimens from 150 patients from the pathology database (2010–2023) of our hospital. Fifty patients were clinically diagnosed with colorectal adenocarcinoma, 50 patients were diagnosed with UC, 20 patients were diagnosed with CD, 15 patients were diagnosed with ischemic colitis (IC), 10 patients were diagnosed with diverticulitis, and 5 patients were diagnosed with collagenous colitis (CC). The group of patients with adenocarcinoma comprised 30 men and 20 women (age: 41–94 years; mean age: 74.3 years). Nonlesioned specimens from the group of patients with adenocarcinoma were used as control. The patients with UC comprised 27 men and 23 women (age: 22–82 years; mean age: 51.6 years). The patients with CD comprised 16 men and 4 women (age: 31–90 years; mean age: 53.7 years). The patients with IC comprised 10 men and 5 women (age: 40–86 years; mean age: 68.8 years). The patients with diverticulitis comprised 7 men and 3 women (age: 46–81 years; mean age: 63.4 years). The patients with CC comprised 2 men and 3 women (age: 64–85 years; mean age: 76.4 years). CD, IC, diverticulitis, and CC were treated collectively as non-UC coloproctitis. Finally, 200 specimens were analyzed, including 50 specimens for adenocarcinoma, control, UC, and non-UC coloproctitis each.

Collecting images

Microscopic images were captured using a Bx53 microscope (EVIDENT CORPORATION, Shinjuku-ku, Tokyo, Japan) and a DP27 camera (EVIDENT CORPORATION, Shinjuku-ku, Tokyo, Japan). The images were set to 2,448 × 1,920 pixels and saved as .JPG files. We captured one microscopic image each at magnifications of 40x, 100x, and 200x and two microscopic images at a magnification of 400x for each sample. Eventually, we captured 1,000 images of the entire set of samples. The number of 1,000 images was increased six times to 6,000 images through argumentation. Argumentation was performed using a vertical flip, horizontal flip, and clockwise rotations at 90-degree, 180-degree, and 270-degree. We captured microscopic images of crypt distortion, goblet cell depletion, basal plasmacytosis, and crypt abscesses in UC specimens. For CD cases, areas of interest were selected in the part with lymphocyte aggregation. For IC cases, areas of interest were selected in the part with superficial

mucosa necrosis and lamina propria hyalinization. For diverticulitis cases, areas of interest were selected in the part with pseudodiverticulum and inflammation. For CC cases, areas of interest were selected in the part with subepithelial collagen band. For control cases, areas of interest with few artifacts were randomly selected. For colorectal adenocarcinoma cases, areas of interest were selected in the advanced part of the cancer (the part with the highest degree of invasion).

Machine learning

Subsequently, we used the machine-learning software, Teachable Machine (TM), to create a machine-learning model that could distinguish between UC, non-UC coloproctitis, adenocarcinoma, and control. TM, provided by Google (Google Inc., Mountain View, CA, USA), is a web-based tool that allows anyone to create machine-learning models without coding. The version of TensorFlow.js that supported Teachable Machine was 1.3.1. The number of training sets was 5,100, and the number of test sets was 900. The epoch number was 50, batch size was 16, and learning rate was set to 0.001.

Results

Representative histological images of adenocarcinoma, control, and UC are shown in Fig. 1. Histological and endoscopic findings of UC and adenocarcinoma are summarized in Table 1. Representative histological images of non-UC coloproctitis are shown in Fig. 2. The model performances of the four classes (UC, non-UC coloproctitis, adenocarcinoma, and control) are shown in Table 2, and the confusion matrix is shown in Fig. 3. We used accuracy as an indicator of AI performance. Anyone can access this model using the following uniform resource locator (URL): https://teachablemachine.withgoogle.com/models/xcMSOUU1T/.

Discussion

This is the first report in which a no-code AI comprehensively recognizes the characteristic histological patterns of UC. This AI model could also differentiate UC from non-UC coloproctitis and adenocarcinoma with high accuracy.

Several studies have reported the use of no-code AI to predict medical images

and perform scoring[4-7]. For example, histological scoring of the testes has been demonstrated using a no-code deep-learning software (Google Cloud AutoML Vision)[4]. In that study, a total of 9,822 testicular histological images were collected from 264 patients with azoospermia. The histological "Johnsen score" for spermatogenic activity was predicted using an automated machine-learning platform without coding. The average precision, precision, and recall values were 99.5, 96.29, and 96.23%, respectively. Additionally, Zeng Y et al. reported 378,215 histological images of mammary invasive ductal carcinoma, and Google Cloud AutoML Vision demonstrated an average accuracy of 91.6% in the diagnosis[5]. In contrast, Forchhammer S et al. included 831 patients with malignant melanoma and reported that TM could predict prognosis with an area under the curve of 0.694 in receiver operating characteristic analysis[6]. Furthermore, a TM, which can detect macroscopic abnormalities in the tympanic membrane, has been reported[7]. In that study, a total of 3,024 images of the tympanic membrane were collected, and the TM showed an accuracy of 90.1% in detecting the abnormalities. Google Cloud AutoML Vision and TM are both no-code effective AI platforms. Owing to its simple browser, the latter is easier to manage for many medical staffs.

Numerous studies have reported that AI can detect active inflammation, histological scoring, and/or prognostic prediction in UC[8-12]. Del Amor R et al. reported an AI that recognizes active UC[8]. In their study, whole-slide images of 230 colorectal specimens were obtained. Their AI model focused on neutrophils and demonstrated high accuracy of 0.96 with multiple instance learning. Additionally, Gui et al. demonstrated the utilization of Paddington International virtual ChromoendoScopy Score Histologic Remission Index (PHRI) based on AI to recognize neutrophils. The PHRI correlated with endoscopic scores, such as the Mayo endoscopic score and clinical outcomes, including hospitalization and surgery[9]. Interestingly, Ohara J et al. reported that the goblet cell ratio could be calculated in specimens of patients with UC using AI. This ratio was associated with disease relapse[10]. These reports are beneficial, but medical staff find the technology difficult to manage.

We demonstrated that a no-code effective AI model can be constructed by preparing high-quality data annotated by pathologists. However, it had a few limitations. First, the confusion matrix showed a small number of cases in which judgments regarding

malignancy were incorrect. Therefore, it is advisable to refrain from relying solely on AI for diagnostic confirmation. Second, the study included only four classes, namely, adenocarcinoma, control, UC, and non-UC coloproctitis. Further studies are required to identify other differential diagnoses. Third, care should be taken when grouping images into a single category is not appropriate, such as when adenocarcinoma coexists with ulcerative colitis. Fourth, this study was a single-institutional study and was based on our own hematoxylin and eosin staining protocol. This study lays a foundation for future research in this field.

References

1 Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. Lancet. 2017 Apr 29;389(10080):1756-1770.

2 Da Rio L, Spadaccini M, Parigi TL, et al. Artificial intelligence and inflammatory bowel disease: Where are we going? World J Gastroenterol. 2023 Jan 21;29(3):508-520.

3 Ahmad HA, East JE, Panaccione R, et al. Artificial intelligence in inflammatory bowel disease: implications for clinical practice and future directions. Intest Res. 2023 Apr 20.

4 Ito Y, Unagami M, Yamabe F, et al. A method for utilizing automated machine learning for histopathological classification of testis based on Johnsen scores. Sci Rep. 2021 May 10;11(1):9962.

5 Zeng Y, Zhang J. A machine learning model for detecting invasive ductal carcinoma with Google Cloud AutoML Vision. Comput Biol Med. 2020 Jul;122:103861.

6 Forchhammer S, Abu-Ghazaleh A, Metzler G, Garbe C, Eigentler T. Development of an Image Analysis-Based Prognosis Score Using Google's Teachable Machine in Melanoma. Cancers (Basel). 2022 Apr 29;14(9):2243.

7 Byun H, Lee SH, Kim TH, Oh J, Chung JH. Feasibility of the Machine Learning Network to Diagnose Tympanic Membrane Lesions without Coding Experience. J Pers Med. 2022 Nov 7;12(11):1855.

8 Del Amor R, Meseguer P, Parigi TL, et al. Constrained multiple instance learning for ulcerative colitis prediction using histological images. Comput Methods Programs Biomed. 2022 Sep;224:107012.

9 Gui X, Bazarova A, Del Amor R, et al. PICaSSO Histologic Remission Index (PHRI) in ulcerative colitis: development of a novel simplified histological score for monitoring

mucosal healing and predicting clinical outcomes and its applicability in an artificial intelligence system. Gut. 2022 May;71(5):889-898.

10 Ohara J, Nemoto T, Maeda Y, Ogata N, Kudo SE, Yamochi T. Deep learning-based automated quantification of goblet cell mucus using histological images as a predictor of clinical relapse of ulcerative colitis with endoscopic remission. J Gastroenterol. 2022 Dec;57(12):962-970.

11 Iacucci M, Parigi TL, Del Amor R, et al. Artificial Intelligence Enabled Histological Prediction of Remission or Activity and Clinical Outcomes in Ulcerative Colitis. Gastroenterology. 2023 Mar 4:S0016-5085(23)00216-0.

12 Najdawi F, Sucipto K, Mistry P, et al. Artificial Intelligence Enables Quantitative Assessment of Ulcerative Colitis Histology. Mod Pathol. 2023 Feb 15;36(6):100124.

Figure legends

Fig. 1 Representative captured images with hematoxylin and eosin stain（Original magnification 400×). (a) Adenocarcinoma. (b) Control. (c) Crypt abscess in ulcerative colitis (UC). (d) Crypt distortion and goblet cell depletion in UC.

Fig. 2 Representative captured images with hematoxylin and eosin stain. (a) Crohn's disease（Original magnification 200×). (b) Ischemic colitis（Original magnification 200×). (c) Diverticulitis（Original magnification 20×). (d) Collagenous colitis（Original magnification 200×).

Fig. 3 Confusion matrix that showed the correspondence between model predictions and correct classes.