

Principal Investigator

First Name: Bobby

Last Name: Lo

Degree: M.D.

Primary Affiliation: Gastrounit, medical section, Copenhagen University Hospital Hvidovre

E-mail: bobby.zhao.sheng.lo@gmail.com

Phone number: +4528580410

Address: Kaffevej 48

Kettegård alle 30

City: Hvidovre

State or Province: Denmark

Zip or Postal Code: 2650

Country: Denmark

General Information

Key Personnel (in addition to PI):

First Name: Bobby

Last name: Lo

Degree: M.D.

Primary Affiliation: Gastrounit, medical section, Copenhagen University Hospital Hvidovre

First Name: Johan

Last name: Burisch

Degree: M.D., PhD, Med.Sc.D.

Primary Affiliation: Gastrounit, medical section, Copenhagen University Hospital Hvidovre

First Name: Christian

Last name: Igel

Degree: Professor, Dr. habil.

Primary Affiliation: Copenhagen University, Department of Computer Science

First Name: Bulat

Last name: Ibragimov

Degree: Associate Professor

Primary Affiliation: Copenhagen University, Department of Computer Science

First Name: Bjørn

Last name: Møller

Degree: MSc, PhD student

Primary Affiliation: Copenhagen University, Department of Computer Science

Are external grants or funds being used to support this research?: No external grants or funds are being used to support this research.

How did you learn about the YODA Project?: Data Holder (Company)

Conflict of Interest

https://yoda.yale.edu/system/files/coi_7.pdf

https://yoda.yale.edu/system/files/sv_6m4tghhxg7w7uxe-r_27us65wzy9zuthm.pdf

<https://yoda.yale.edu/system/files/jb.pdf>

https://yoda.yale.edu/system/files/sv_6m4tghhxg7w7uxe-r_22ai5z4fh2wvrzi.pdf

.misc-fixes { display: none; } #admin-region { z-index: 9999999; } #admin-menu { z-index: 9999999; }

https://yoda.yale.edu/system/files/sv_6m4tghhxg7w7uxe-r_2tmt5a4sxq2ldl7.pdf

Certification

Certification: All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.

Data Use Agreement Training: As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

1. [NCT00036439 - C0168T37 - A Randomized, Placebo-controlled, Double-blind Trial to Evaluate the Safety and Efficacy of Infliximab in Patients With Active Ulcerative Colitis](#)
2. [NCT00096655 - C0168T46 - A Randomized, Placebo-controlled, Double-blind Trial to Evaluate the Safety and Efficacy of Infliximab in Patients With Active Ulcerative Colitis](#)
3. [NCT00336492 - C0168T72 - A Phase 3, Randomized, Open-label, Parallel-group, Multicenter Trial to Evaluate the Safety and Efficacy of Infliximab \(REMICADE\) in Pediatric Subjects With Moderately to Severely Active Ulcerative Colitis](#)
4. [NCT00487539 - C0524T17 - A Phase 2/3 Multicenter, Randomized, Placebo-controlled, Double blind Study to Evaluate the Safety and Efficacy of Golimumab Induction Therapy, Administered Subcutaneously, in Subjects with Moderately to Severely Active Ulcerative Colitis](#)
5. [NCT01551290 - CR018769: REMICADEUCO3001 - A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of Infliximab in Chinese Subjects With Active Ulcerative Colitis](#)
6. [NCT00488631 - C0524T18 - A Phase 3 Multicenter, Randomized, Placebo-controlled, Double-blind Study to Evaluate the Safety and Efficacy of Golimumab Maintenance Therapy, Administered Subcutaneously, in Subjects With Moderately to Severely Active Ulcerative Colitis](#)
7. [NCT00488774 - C0524T16 - A Phase 2/3 Multicenter, Randomized, Placebo-controlled, Double-blind Study to Evaluate the Safety and Efficacy of Golimumab Induction Therapy, Administered Intravenously, in Subjects With Moderately to Severely Active Ulcerative Colitis](#)
8. [NCT01863771 - CNTO148UCO3001 - A Safety and Effectiveness Study of Golimumab in Japanese Patients With Moderately to Severely Active Ulcerative Colitis](#)
9. [NCT01988961 - CNTO148UCO2001 - A Study to Evaluate the Accuracy of a Subset of the Length-109 Probe Set Panel \(a Genetic Test\) in Predicting Response to Golimumab in Participants With Moderately to Severely Active Ulcerative Colitis](#)
10. [NCT02407236 - CNTO1275UCO3001 - A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Protocol to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects With Moderately to Severely Active Ulcerative Colitis](#)

What type of data are you looking for?: Individual Participant-Level Data, which includes Full CSR and all supporting documentation

Research Proposal

Project Title

Validation of an artificial intelligence model on central read endoscopic disease severity in patients with ulcerative colitis

Narrative Summary:

Ulcerative colitis is a chronic disease that is often diagnosed in youth. A relapsing disease course characterises the disease despite treatment. Endoscopic evaluation is regarded as the gold standard for objectively measuring disease severity. However, it has a high observer variance; even with central reading, which is the gold standard in clinical trials. We have developed an artificial intelligence (AI) model that can classify endoscopic severity on par with experts in the field and aim to further validate it on images and videos from clinical trials scored by central reading. If successful, central reading can, in the future, be conducted by our AI model, saving time and resources.

Scientific Abstract:

Background

Endoscopic evaluation of disease severity of ulcerative colitis has shown high inter- and intra-observer variance. However, it remains an essential inclusion criterion and an outcome in many medical trials. Central reading has therefore become necessary to unify the assessment of disease severity, ensuring unbiased evaluation and reducing the observer variabilities. However, even central reading has demonstrated observer variance. The cost associated with central reading is high, with the risk of patients not being included in trials due to delayed assessment. We have developed an artificial intelligence (AI) model that can reliably score endoscopic disease severity during live endoscopy and on recordings or pictures from procedures.

Objective

We aim to evaluate endoscopic videos or images from all trials conducted in ulcerative colitis patients and compare the model's disease severity classification with the central read scores.

Study Design

Endoscopic videos or images from 10 trials (4,679 patients) will be requested. All videos or images will be analysed by the AI model and compared to the original scores.

Participants

Patients diagnosed with ulcerative colitis

Primary endpoint

The area under the receiver operating characteristic (AUROC) >0.90 & accuracy >0.85.

Secondary endpoints

Sensitivity and specificity will be reported as secondary evaluation metrics. We will also investigate the false positives and negatives of the AI model to evaluate its pitfalls.

Statistical analysis

We will apply the following metrics/analysis: AUROC, accuracy, sensitivity, specificity, Cohens Kappa, and adjusted Cohens Kappa.

Brief Project Background and Statement of Project Significance:

A key component in the care of UC patients is the monitoring of disease activity and timely response with treatment. Disease activity is defined by a combination of clinical and biochemical markers as well as endoscopy to evaluate mucosal inflammation. Current guidelines recommend mucosal healing, most commonly defined as a Mayo endoscopic subscore (MES) of 0–1, as the treatment target(2). However, endoscopic assessment suffers from substantial intra- and inter-observer variation limiting the reliability of individual assessments(3).

Studies of clinicians scoring endoscopic images have shown inter-observer kappa values between 0.45–0.75[1–3]. In light of these results, central reading emerged to standardize endoscopic evaluation for clinical trials and other research projects. However, although central reading improved the evaluation process, the results showed inter-observer reliability of only 0.79 and intra-observer reliability to be 0.89, measured using weighted kappa values [1,4].

We have developed a new automated way of evaluating endoscopic images from UC patients for both clinical and academic purposes. We have demonstrated that our deep learning (DL) model is highly capable of distinguishing MES 0 from 1–3, MES 0–1 from 2–3, and differentiating between all four MES levels of activity. We believe our model can support both clinical and academic assessment of disease severity in UC patients with an accuracy of 88. An application example could be in an RCT study using only the AI model as a central reader. Such a protocol reduces the time and expenditures associated with evaluating the endoscopic disease severity compared to a human central reader.

Furthermore, this can help standardize the field and help future clinical trials with inclusion and outcome measurements by discounting ineligible patients and making studies more easily and accurately comparable[5].

Specific Aims of the Project:

The aim of this study is to run our newly developed AI model on endoscopic images or videos from all clinical trials that have been conducted on patients with ulcerative colitis. The aim is to evaluate the accuracy of the AI model compared to centrally read images and assess whether such a tool can be used as a replacement for central reading.

What is your Study Design?:

Methodological research

What is the purpose of the analysis being proposed? Please select all that apply.

Confirm or validate previously conducted research on treatment effectiveness

Research on clinical trial methods

Other

Research Methods

Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:

Data source: NCT00487539, NCT00488631, NCT00488774, NCT01863771, NCT01988961, NCT00036439, NCT00096655, NCT00336492, NCT01551290, NCT02407236

Inclusion criteria: patients with endoscopic images/videos

Exclusion criteria: none

Primary and Secondary Outcome Measure(s) and how they will be categorized/defined for your study:

Primary endpoint:

The area under the receiver operating characteristic (AUROC) >0.90 & accuracy >0.85 will be deemed satisfactory. AUROC < 0.90 but >0.80 & accuracy <0.85 but >0.75 will be deemed moderate. Everything below will be deemed unsatisfactory.

Secondary endpoints

Sensitivity and specificity will be reported as secondary evaluation metrics. We will also investigate the false positives and negatives of the AI model to evaluate its pitfalls using machine learning techniques.

Main Predictor/Independent Variable and how it will be categorized/defined for your study:

Mayo endoscopic subscore (MES) will be the main variable in which the AI model shall correctly classify. The MES is a categorical score from 0-3.

Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:

Metadata/demographical and treatment data from before the assessment of the endoscopic disease severity will be required to get insight into images that are either false positively or negatively classified by the AI model.

Statistical Analysis Plan:

We will apply the following metrics/analysis: AUROC, accuracy, sensitivity, specificity, Cohens Kappa, and adjusted Cohens Kappa. Statistical regression models and machine learning models will be used to assess metadata/demographical and treatment data in which could have resulted in wrongly classified images, and get an

insight into pitfalls our AI model might have.

Software Used:

Python

Project Timeline:

The first study report is expected 6 months from data access. The manuscript will be submitted the same time as the study report.

Dissemination Plan:

Positive, negative and inconclusive results will be submitted for publication in a peer-reviewed international and/or Danish journal or published for public access if rejected. Results will be presented at national and international conferences. Bobby Lo will be the first author. The remaining authorship will be delegated according to the Vancouver guidelines.

Bibliography:

[1] Samaan MA, Mosli MH, Sandborn WJ, Feagan BG, D'Haens GR, Dubcenco E, et al. A systematic review of the measurement of endoscopic healing in ulcerative colitis clinical trials: Recommendations and implications for future research. *Inflamm Bowel Dis* 2014;20:1465–71. <https://doi.org/10.1097/MIB.000000000000046>.

[2] Fernandes SR, Pinto JSLD, Marques Da Costa P, Correia L. Disagreement among Gastroenterologists Using the Mayo and Rutgeerts Endoscopic Scores. *Inflamm Bowel Dis* 2018;24:254–60. <https://doi.org/10.1093/ibd/izx066>.

[3] Mohammed Vashist N, Samaan M, Mosli MH, Parker CE, Macdonald JK, Nelson SA, et al. Endoscopic scoring indices for evaluation of disease activity in ulcerative colitis. *Cochrane Database Syst Rev* 2018;2018. <https://doi.org/10.1002/14651858.CD011450.pub2>.

[4] Feagan BG, Sandborn WJ, D'Haens G, Pola S, McDonald JWD, Rutgeerts P, et al. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013;145. <https://doi.org/10.1053/j.gastro.2013.03.025>.

[5] Ket SN, Palmer R, Travis S. Endoscopic Disease Activity in Inflammatory Bowel Disease 2015. <https://doi.org/10.1007/s11894-015-0470-0>.

Supplementary Material

Lo, Bobby MD^{1,2}; Liu, ZhuoYuan MSc^{2,3}; Bendtsen, Flemming MD, DMSc^{1,2}; Igel, Christian Dr. Habil³; Vind, Ida MD, PhD^{1,2}; Burisch, Johan MD, PhD^{1,2}. High Accuracy in Classifying Endoscopic Severity in Ulcerative Colitis Using Convolutional Neural Network. *The American Journal of Gastroenterology*: October 2022 - Volume 117 - Issue 10 - p 1648-1654 doi: 10.14309/ajg.0000000000001904

https://journals.lww.com/ajg/Citation/2022/10000/High_Accuracy_in_Classi...

Supplementary Material:

https://yoda.yale.edu/sites/default/files/confirmation_yoda_project_data_use_agreement_training_completed_1.pdf