

## PEER-REVIEW REPORT

**Name of journal:** Artificial Intelligence in Gastrointestinal Endoscopy

**Manuscript NO:** 67597

**Title:** Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

**Reviewer's code:** 05709186

**Position:** Peer Reviewer

**Academic degree:** MD

**Professional title:** Doctor

**Reviewer's Country/Territory:** France

**Author's Country/Territory:** Denmark

**Manuscript submission date:** 2021-04-28

**Reviewer chosen by:** AI Technique

**Reviewer accepted review:** 2021-04-29 08:38

**Reviewer performed review:** 2021-04-29 15:42

**Review time:** 7 Hours

<b>Scientific quality</b>	<input type="checkbox"/> Grade A: Excellent <input type="checkbox"/> Grade B: Very good <input checked="" type="checkbox"/> Grade C: Good <input type="checkbox"/> Grade D: Fair <input type="checkbox"/> Grade E: Do not publish
<b>Language quality</b>	<input checked="" type="checkbox"/> Grade A: Priority publishing <input type="checkbox"/> Grade B: Minor language polishing <input type="checkbox"/> Grade C: A great deal of language polishing <input type="checkbox"/> Grade D: Rejection
<b>Conclusion</b>	<input type="checkbox"/> Accept (High priority) <input type="checkbox"/> Accept (General priority) <input checked="" type="checkbox"/> Minor revision <input type="checkbox"/> Major revision <input type="checkbox"/> Rejection
<b>Re-review</b>	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
<b>Peer-reviewer statements</b>	Peer-Review: <input checked="" type="checkbox"/> Anonymous <input type="checkbox"/> Onymous Conflicts-of-Interest: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No



**Baishideng  
Publishing  
Group**

7041 Koll Center Parkway, Suite  
160, Pleasanton, CA 94566, USA  
**Telephone:** +1-925-399-1568  
**E-mail:** bpgoffice@wjgnet.com  
**https://**www.wjgnet.com

## **SPECIFIC COMMENTS TO AUTHORS**

Please find below my comments regarding this manuscript: Line 115: “In addition, the group used only a training and a validation set but not a test set to assess whether the algorithm was overfitted”. I do not completely agree with this statement. From what I understand from this manuscript, Maeda et al. have trained a model on a dataset denoted as the “training set”, and have validated its results on a dataset, denoted as the “validation set”, consisting of data unseen by the model during training. I agree that the term “validation set” commonly denotes the dataset used for optimizing a model (e.g. hyperparameters grid search), while the “test set” commonly denotes unseen data used only for assessing the performance of the final model. However, the way they describe their study makes me think that the dataset they denote as a “validation set” is actually a “test set”, since they do not mention any data used for comparing architectures or hyperparameters configurations. However, it may be interesting and pertinent to highlight the fact that this study, as many others, did not validate its 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. Line 205: “Often, only AUC is reported, which can be misleading as sensitivity, specificity and accuracy may be only modest.” I believe this sentence does not exactly point out why the presentation of such metrics (sensitivity, specificity, accuracy...) is important, rather than presenting only the ROC-AUC. Usually, this is important because the ROC-AUC evaluates the performance of a model’s output regardless of any threshold, and thus does not allow assessing the consistency of this output on different datasets. For instance, if a case-control prediction model consistently outputs 0% for control samples, but outputs 100% for case samples from the training set, and 50% for case samples from the test set: this model will have a ROC-AUC of 1 for both datasets,

but its sensitivity will decrease from 100% on the training set to 0% on the test set. Maybe the authors could better highlight the fact that the ROC-AUC only allows assessing the performance of a model on a single cohort, and does not allow assessing the consistency of its results throughout multiple datasets. Minor comments: The exact formulation may be improved for points listed below. Please note that I am not a native English speaker. Therefore, please ignore my suggestions concerning English issues if those are not appropriate. Line 31: maybe prefer “data analysis methods” to “data analyzing methods” Line 33: “[...] its ability to learn and optimize its models/predictions from new inputs.” sounds a little strange. Maybe replace by “[...] its ability to learn and optimize its predictions from new inputs.” Line 100: “[...] methods such as the convolutional network [...]” sounds really strange to me, since convolutional neural networks are rather a type of architecture which belongs to the field of machine learning (or more frequently deep learning, even if these arbitrary definitions are a bit vague) than a unique and homogeneous method. Prefer perhaps: “[...] methods such as convolutional neural networks” or even “[...] methods such as deep convolutional neural networks [...]”. Line 111: “support vector machine” instead of “support vector machine learning” Table 1: maybe replace “Generalisability” by “Generalizability”

## PEER-REVIEW REPORT

**Name of journal:** Artificial Intelligence in Gastrointestinal Endoscopy

**Manuscript NO:** 67597

**Title:** Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

**Reviewer's code:** 05088096

**Position:** Editorial Board

**Academic degree:** FAASLD, MD

**Professional title:** Professor

**Reviewer's Country/Territory:** Egypt

**Author's Country/Territory:** Denmark

**Manuscript submission date:** 2021-04-28

**Reviewer chosen by:** AI Technique

**Reviewer accepted review:** 2021-04-29 05:06

**Reviewer performed review:** 2021-04-30 09:52

**Review time:** 1 Day and 4 Hours

<b>Scientific quality</b>	<input type="checkbox"/> Grade A: Excellent <input type="checkbox"/> Grade B: Very good <input checked="" type="checkbox"/> Grade C: Good <input type="checkbox"/> Grade D: Fair <input type="checkbox"/> Grade E: Do not publish
<b>Language quality</b>	<input type="checkbox"/> Grade A: Priority publishing <input checked="" type="checkbox"/> Grade B: Minor language polishing <input type="checkbox"/> Grade C: A great deal of language polishing <input type="checkbox"/> Grade D: Rejection
<b>Conclusion</b>	<input type="checkbox"/> Accept (High priority) <input type="checkbox"/> Accept (General priority) <input type="checkbox"/> Minor revision <input checked="" type="checkbox"/> Major revision <input type="checkbox"/> Rejection
<b>Re-review</b>	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
<b>Peer-reviewer statements</b>	Peer-Review: <input type="checkbox"/> Anonymous <input checked="" type="checkbox"/> Onymous Conflicts-of-Interest: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

#### **SPECIFIC COMMENTS TO AUTHORS**

The Abstract is very defective in describing the aim and the main contents of the article.

What do you mean by the abbreviation CTI?

## PEER-REVIEW REPORT

**Name of journal:** Artificial Intelligence in Gastrointestinal Endoscopy

**Manuscript NO:** 67597

**Title:** Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

**Reviewer's code:** 05909344

**Position:** Peer Reviewer

**Academic degree:** MD

**Professional title:** Assistant Professor

**Reviewer's Country/Territory:** Pakistan

**Author's Country/Territory:** Denmark

**Manuscript submission date:** 2021-04-28

**Reviewer chosen by:** AI Technique

**Reviewer accepted review:** 2021-04-29 04:04

**Reviewer performed review:** 2021-05-07 15:15

**Review time:** 8 Days and 11 Hours

<b>Scientific quality</b>	<input type="checkbox"/> Grade A: Excellent <input type="checkbox"/> Grade B: Very good <input checked="" type="checkbox"/> Grade C: Good <input type="checkbox"/> Grade D: Fair <input type="checkbox"/> Grade E: Do not publish
<b>Language quality</b>	<input type="checkbox"/> Grade A: Priority publishing <input checked="" type="checkbox"/> Grade B: Minor language polishing <input type="checkbox"/> Grade C: A great deal of language polishing <input type="checkbox"/> Grade D: Rejection
<b>Conclusion</b>	<input type="checkbox"/> Accept (High priority) <input checked="" type="checkbox"/> Accept (General priority) <input type="checkbox"/> Minor revision <input type="checkbox"/> Major revision <input type="checkbox"/> Rejection
<b>Re-review</b>	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
<b>Peer-reviewer statements</b>	Peer-Review: <input checked="" type="checkbox"/> Anonymous <input type="checkbox"/> Onymous Conflicts-of-Interest: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No



**Baishideng  
Publishing  
Group**

7041 Koll Center Parkway, Suite  
160, Pleasanton, CA 94566, USA  
**Telephone:** +1-925-399-1568  
**E-mail:** bpgoffice@wjgnet.com  
<https://www.wjgnet.com>

#### **SPECIFIC COMMENTS TO AUTHORS**

This opinion review presents an important and upcoming concept of utilizing artificial intelligence for the assessment of inflammatory bowel disease. This disorder has unclear etiology and heterogeneous diagnostic tools with various confounding factors including skill of the endoscopist. Therefore diagnostic accuracy is required in order to stage and subsequently manage the condition to improve patients' quality of life as well as limit disease complications. However, to provide useful information, artificial intelligence models need to be user friendly and easy to apply, and hence this is an area where much more prospective research with larger data set is needed. The table proposed by the authors can also serve as a useful guide and standard for future articles on the subject.

## PEER-REVIEW REPORT

**Name of journal:** Artificial Intelligence in Gastrointestinal Endoscopy

**Manuscript NO:** 67597

**Title:** Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

**Reviewer's code:** 05885746

**Position:** Editorial Board

**Academic degree:** PhD

**Professional title:** Research Assistant Professor

**Reviewer's Country/Territory:** United States

**Author's Country/Territory:** Denmark

**Manuscript submission date:** 2021-04-28

**Reviewer chosen by:** AI Technique

**Reviewer accepted review:** 2021-04-30 16:32

**Reviewer performed review:** 2021-05-08 14:53

**Review time:** 7 Days and 22 Hours

<b>Scientific quality</b>	<input type="checkbox"/> Grade A: Excellent <input type="checkbox"/> Grade B: Very good <input checked="" type="checkbox"/> Grade C: Good <input type="checkbox"/> Grade D: Fair <input type="checkbox"/> Grade E: Do not publish
<b>Language quality</b>	<input type="checkbox"/> Grade A: Priority publishing <input checked="" type="checkbox"/> Grade B: Minor language polishing <input type="checkbox"/> Grade C: A great deal of language polishing <input type="checkbox"/> Grade D: Rejection
<b>Conclusion</b>	<input type="checkbox"/> Accept (High priority) <input type="checkbox"/> Accept (General priority) <input checked="" type="checkbox"/> Minor revision <input type="checkbox"/> Major revision <input type="checkbox"/> Rejection
<b>Re-review</b>	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
<b>Peer-reviewer statements</b>	Peer-Review: <input checked="" type="checkbox"/> Anonymous <input type="checkbox"/> Onymous Conflicts-of-Interest: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No





**Baishideng  
Publishing  
Group**

7041 Koll Center Parkway, Suite  
160, Pleasanton, CA 94566, USA  
**Telephone:** +1-925-399-1568  
**E-mail:** [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
<https://www.wjgnet.com>

#### **SPECIFIC COMMENTS TO AUTHORS**

1. Please include machine learning as a key word 2. What criteria were used to select literature to review?

## RE-REVIEW REPORT OF REVISED MANUSCRIPT

**Name of journal:** Artificial Intelligence in Gastrointestinal Endoscopy

**Manuscript NO:** 67597

**Title:** Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

**Reviewer's code:** 05088096

**Position:** Editorial Board

**Academic degree:** FAASLD, MD

**Professional title:** Professor

**Reviewer's Country/Territory:** Egypt

**Author's Country/Territory:** Denmark

**Manuscript submission date:** 2021-04-28

**Reviewer chosen by:** Man Liu

**Reviewer accepted review:** 2021-07-28 18:13

**Reviewer performed review:** 2021-07-31 19:26

**Review time:** 3 Days and 1 Hour

<b>Scientific quality</b>	<input type="checkbox"/> Grade A: Excellent <input checked="" type="checkbox"/> Grade B: Very good <input type="checkbox"/> Grade C: Good <input type="checkbox"/> Grade D: Fair <input type="checkbox"/> Grade E: Do not publish
<b>Language quality</b>	<input type="checkbox"/> Grade A: Priority publishing <input checked="" type="checkbox"/> Grade B: Minor language polishing <input type="checkbox"/> Grade C: A great deal of language polishing <input type="checkbox"/> Grade D: Rejection
<b>Conclusion</b>	<input type="checkbox"/> Accept (High priority) <input checked="" type="checkbox"/> Accept (General priority) <input type="checkbox"/> Minor revision <input type="checkbox"/> Major revision <input type="checkbox"/> Rejection
<b>Peer-reviewer statements</b>	Peer-Review: <input type="checkbox"/> Anonymous <input checked="" type="checkbox"/> Onymous Conflicts-of-Interest: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

## SPECIFIC COMMENTS TO AUTHORS



7041 Koll Center Parkway, Suite  
160, Pleasanton, CA 94566, USA  
**Telephone:** +1-925-399-1568  
**E-mail:** [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
**https://**[www.wjgnet.com](http://www.wjgnet.com)

Thanks for addressing all the points