

Reviewer #1:

Scientific Quality: Grade D (Fair)

Language Quality: Grade B (Minor language polishing)

Conclusion: Major revision

Specific Comments to Authors: Thank you for letting me review the present article entitled " ARTIFICIAL INTELLIGENCE in RECTAL CANCER " The paper may have merit but :

Thank you for taking the time to evaluate the article and for your valuable suggestions and comments.

1. In the INTRODUCTION section, please give the abbreviation where " Machine learning " first appears.

Author: The abbreviation was added in line with your suggestion.

2. In the ARTIFICIAL INTELLIGENCE IN DIAGNOSIS IN RECTAL CANCER section, the Detection of Lymph Node Metastasis was reviewed, why not review the studies of T stage or differentiation?

Author: In the literature, studies for predicting T stage and differentiation with artificial intelligence in rectal cancer are quite limited. This issue has been added to the review.

3. In the Artificial Intelligence in Detection of Distant Metastasis section, the MLM was reviewed, why not review the studies of SLM (if there is no definite CRLM, can we assess radiomics derived from the primary colorectal tumor to predict synchronous liver metastasis?)

Author: Synchronous liver metastases are already diagnosed during staging at the time of diagnosis. This section has been written to assess whether cases diagnosed with liver metastasis (metachronous) during follow-up can be predicted in advance.

4. This article did not include enough relevant advanced researches, such as molecular biological indicators, RAS mutations, MSI/MMR, or analysis of tumor microenvironmental features.

Author: There are very few studies on this subject in the literature. A title on the KRAS mutation has been added.

5. In the 3.B section: what is "PET-BT", is it a spelling mistake?

Author: There is spelling mistake, corrected.

Reviewer #2:

Scientific Quality: Grade C (Good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Accept (General priority)

Specific Comments to Authors: Determining the optimal treatment plan for a patient with rectal cancer is a complex process and the oncological results and toxicity are not the same in every patient with the same treatment at the same stage. This article reviews the application of deep learning in the diagnosis of rectal cancer. Multicenter studies with large data sets can provide algorithms with higher accuracy rates. Prediction algorithms can be standardized by sharing data between centers, data diversity, and creating big data. It can provide the better information for the relevant researchers. Thank you for taking the time to evaluate the article and for your valuable suggestions and comments.

Reviewer #3:

Scientific Quality: Grade D (Fair)

Language Quality: Grade B (Minor language polishing)

Conclusion: Major revision

Specific Comments to Authors:

Thank you for taking the time to evaluate the article and for your valuable suggestions and comments.

1.Please show specific methodology of AI in rectal cancer? 2.Without PET/CT or PET/MR radiomic data ? 3.Without radiogenomics of rectal cancer?

Author: Prediction studies with artificial intelligence using radiogenomics and radiomic properties are very rare in rectal cancer. A study on this subject has been added to the review.

4.Please give us the problems and prospects of AI in rectal cancer.

Author: Although artificial intelligence can take place in every step from patient consultation to patient follow-up in rectal cancer and can contribute to the clinician and the society, there are still many challenges and problems to be solved. Big data sets should be created for AI first and these data sets should be improved. The development of prediction tools with a wide variety of variables and models limits the comparability of existing studies and the use of standards. Prediction algorithms can be standardized by sharing data between centers, data diversity and creating big data. In addition, the models can be made clinically applicable by updating the models by entering new data into the models. Today, the accuracy and quality of the data is also of great importance, as no AI algorithm can fix the problems in training data.

It has been added to the article in line with your suggestion.

5.It is worth discussing on P5 "In the literature, the number of studies evaluating LN metastasis with AI in rectal cancer is limited. "?

Author: "In the literature, the number of studies evaluating LN metastasis with AI in rectal cancer is limited" has been deleted.

6.The citation of references in the paper is not standard.

Author: It was reviewed and edited in line with your suggestion.

Reviewer #4:

Scientific Quality: Grade E (Do not publish)

Language Quality: Grade C (A great deal of language polishing)

Conclusion: Major revision

Specific Comments to Authors: Dear editors and authors, Below are my comments on this review article titled “Artificial intelligence in rectal cancer”.

Thank you for taking the time to evaluate the article and for your valuable suggestions and comments.

Title and Abstract: The abstract should be revised extensively since it does not give any important information to readers on what can they learn from this review articles. It is inappropriate to use abstract for defining AI, machine learning, and deep learning.

Author: Changes have been made in line with your suggestion.

The authors mentioned about colorectal cancer here. Although colorectal cancer and rectal cancer are related, they are not the same. Clinically, the authors should establish a linkage from colorectal cancer to rectal cancer here. Otherwise, they can entirely ignore colorectal cancer and focus on rectal cancer instead since it is clearly a topic of focus based on the predefined title.

Author: Artificial intelligence studies are still not very common in all areas. There are no studies specific to rectum cancer on some topics. Therefore, studies on colorectal cancers are presented.

Core tip: First, there is a typo in the “core tip” itself. Second, I thought I was reading the core tip of AI in general oncology. There is nothing to do with rectal cancer here.

Author: Changes have been made in line with your suggestion.

Introduction: It is appropriate to define AI, machine learning, and deep learning here. However, the schematic representation in Figure 1 is more like a figure for a PowerPoint presentation, but really too simple for publishing to a scientific article for the current state of AI. Plus, the labels of AI, Machine Learning, and Deep Learning in Figure 1 are not in English. I would recommend removing this figure and replace it with something that benefit readers more (if possible), for examples, how deep learning or convolutional neural networks (CNN) works.

Author: Figure-1 has been deleted.

Once again, I have the same impression with the abstract such that colorectal cancer was emphasized here without a linkage to rectal cancer.

Author: Changes have been made in line with your suggestion.

Main content and tables: I like the organization that the authors divide the sections into the roles of AI in rectal cancer (i.e., diagnosis, treatment, and follow-up). However, there are so many times that I felt like the authors might not have sufficient statistical or technical background on AI. Here are some examples: • Faster Region-based Convolutional Neural Network (Faster R-CNN) was defined as a new AI system. In fact, many things in AI are new. Faster R-CNN is not a system, but it is an algorithm, framework, or architect. • AUC and ROC are separately defined. However, ROC was mentioned in the number of 0.912, which is basically an AUC. After a fact check, it supposed to be AUC or AUROC, not ROC alone. • “Using the ResNet-3D algorithm + SVM algorithm, which includes deep learning” was a bit confusing. ResNet doesn’t include deep learning, but it is a well-known deep learning architect.

Author: The review is written with a clinician's perspective and knowledge, and its statistical parts are not strong. Corrections have been made on your suggestions.

The tables do not show a clear summary of standardized results. I can not see the problem that the AI models in each study trying to solve (classification or regression). Normally, ML and DL are used to create predictive model, in which the common performance metrics included accuracy, AUC, Sensitivity, Specificity, RMSE (for regression), and F1. The presentation here are not well-organized. Plus, there are many typos here (e.g., %##,##) Important information like the number of training data and testing data are not clearly presented. Note that the number of cases are not the same as number of images used for training and testing. I also have no idea if the performance are all on the testing dataset.

Author: Changes have been made in line with your suggestion.

Conclusion: I am not convinced with the ending statement “prediction algorithms can be standardized by sharing data between centers, data diversity, and creating big data”. It was not entirely wrong, but more explanation is necessary before this section. The authors should also include more insight on the future direction in this section.

Author: Changes have been made in line with your suggestion.