Dear Editor:

We greatly appreciate your letter and reviewers' comments concerning our manuscript entitled **"Acute pancreatitis: A review of diagnosis, severity prediction and prognosis assessment from imaging technology, scoring systems and artificial intelligence"** (Manuscript NO.: 85503, Review). All comments are valuable and helpful for improving our manuscript. Based on all reviewers' suggestions, we have carefully revised the manuscript and underline it in black, and we hope that the revised manuscript will meet the journal's standard. Below you will find our point-by-point responses to the reviewers' comments.

The corresponding author is Cun-Rong Chen, and his address and other information are as follow:

Address: Department of Critical Care Medicine, Fujian Medical University Union Hospital, 29 Xinquan Road, Gulou District, Fuzhou 350001, China

Email: chcr789@139.com

Thanks very much for your attention to our manuscript.

We are looking forward to hearing from you at your earliest convenience.

Sincerely yours

Jianxiong Hu and Chengfei Zhao

The itemized responses to all reviewers' comments are as follows:

Reviewer #1:

Scientific Quality: Grade B (Very good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Accept (General priority)

Specific Comments to Authors: I read this article and learned a lot for myself. I think it's pretty well written for general surgeons and gastroenterologists to update their basic knowledge of AP. My suggestions for this article: It would be great if the authors presented the scoring systems and prognostic features they presented in the article in the form of a few tables. However, I suggest that this article be published as soon as possible.

  Reply: First of all, we sincerely thank you very much for your recognition and guidance of our work. And thanks very much for your suggestions. According to your suggestion, we have added Table S1, which mainly shows the clinical common score systems used to predict the severity of acute pancreatitis, including SAP, mortality, organ failure, ICU admission, location complications, in-hospital adverse events, and pancreatic necrosis. Because of the amount of content in the table, we have included Table S1 in the Supplementary Material. In addition, according to Table S1, we have also made a proper summary.

  The corresponding content in the revised manuscript is as follows:

  Page 23, line 31-33, and page 24, line 1-13:

  "Drawing on the insights gleaned from the aforementioned body of literature, we have meticulously synthesized a detailed appraisal of the application of various clinical scoring systems in prognosticating severity, local complications, organ failure, and mortality rates associated with AP. These summarizations are comprehensively depicted in Table S1. Various scoring systems exhibited diverse levels of sensitivity, specificity, and accuracy in forecasting the severity, local complications, organ failure, and associated mortality. Further, it is noteworthy that numerous studies have indicated the existence of substantial differences among these scoring systems, highlighting their lack of uniform standards and, in some instances, a concerning

2

degree of inconsistency in their projections. Given the variability in accuracy among diverse scoring systems for predicting the severity, local complications, organ failure, and mortality associated with AP, there is a plausible need for further refinement and design optimization of each scoring system to enhance the precision of these predictions. Moreover, another potential area of research could be the amalgamation of multiple existing scoring systems to boost the predictive accuracy for AP through a more comprehensive scoring approach."

We hope that the revised manuscript will meet the requirements of the journal. And we cordially hope that you will give us guidance if the revised manuscript still has deficiency. Thank you very much again!

Reviewer #2:

Scientific Quality: Grade C (Good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Minor revision

Specific Comments to Authors: This is a detailed review of imaging, scoring, and AI diagnosis in acute pancreatitis. It is easier to understand if each reference is listed in a table.

Reply: Firstly, we sincerely thank you very much for your guidance and suggestions. we have added Table S1, which mainly shows the clinical common score systems used to predict the severity of acute pancreatitis, including SAP, mortality, organ failure, ICU admission, location complications, in-hospital adverse events, and pancreatic necrosis. Because of the amount of content in the table, we have included Table S1 in the Supplementary Material.

We hope that the revised manuscript will meet the requirements of the journal. And we cordially hope that you will give us guidance if the revised manuscript still has deficiency. Thank you very much again!

Reviewer #3:

Scientific Quality: Grade B (Very good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Major revision

Specific Comments to Authors: This is a very interesting paper on an important topic such as the evaluation of acute pancreatitis and the challenges involved in the prediction of its severity. Could the authors please respond to the following comments? 1) Although AI features prominently in the title and the abstract (and is really the significant part in terms of novelty), it is only a very small part of this paper towards the end. The authors provide us with a very thorough description of the different types of radiological means and scoring systems in the area of acute pancreatitis, but, as important as it may be, this part of the paper should be more concise, whereas the latter part with the discussion about AI, should be expanded and especially the part about the limitations and challenges of AI.

Reply: Firstly, we sincerely thank you very much for your recognition and guidance of our work. In response to your issues, we would answer them one by one.

a. 1) Although AI features prominently in the title and the abstract (and is really the significant part in terms of novelty), it is only a very small part of this paper towards the end.

Reply: You are absolutely right about that. Moreover, AI is also a hot topic at present, which is gradually being used in the field of clinical medicine and gradually emerging in clinical practice. In the title and abstract, AI really stands out. According to your suggestion, we have removed the "AI" in the title, which may be more in line with the structure of the overall content of the article.

At present, AI is gradually appearing in the diagnosis, severity prediction and prognosis assessment of acute pancreatitis, but the relevant studies are significantly less than the application of imaging and scoring systems in acute pancreatitis. As a result, AI is only a small part of this article. According to your suggestion, we have added the latest relevant research.

The added content in the revised manuscript is as follows:

Page 36, line 1-13:

"In a recent retrospective study involving a cohort of 460 AP patients, to predict ARDS in these patients at admission, Zhang and Pang constructed and optimized four machine learning models, including Support Vector Machine (SVM), Ensembles of

4

Decision Trees (EDTs), Bayesian Classifier (BC) and nomogram models, based on 31 features with significant differences between the groups with and without ARDS [161]. Among the four models, the BC algorithm exhibited superior predictive performance with the highest AUC (0.891), surpassing SVM (0.870), EDTs (0.813) and the nomogram (0.874) in the test set [161]. Concurrently, the EDT algorithm achieved the highest accuracy at 0.891, precision at 0.800 and F1 score at 0.615, but registered the lowest FDR at 0.200 and the second-highest NPV at 0.902 [161]. In terms of predictive performance for ARDS as a complication of AP, they concluded that BC was the superior predictive model in the test set and EDTs exhibited promising potential for predicting large samples [161]."

b. The authors provide us with a very thorough description of the different types of radiological means and scoring systems in the area of acute pancreatitis, but, as important as it may be, this part of the paper should be more concise, whereas the latter part with the discussion about AI, should be expanded and especially the part about the limitations and challenges of AI.

Reply: In this review, we discuss and analyze in detail the applications and studies of imaging, scoring systems, and artificial intelligence in the diagnosis, severity prediction, and prognosis assessment of acute pancreatitis. Moreover, in the review, the information related to imaging and scoring systems is largely derived from a number of recent references, which we have carefully selected to increase the attractiveness of the review to readers, and to provide clinicians with clinically valuable guidance.

According to your suggestion, appropriate modifications have been made to the content of the imaging and scoring system responses in the revised manuscript.

We specifically expanded on the existing limitations and possible challenges of the application of artificial intelligence in the diagnosis, severity prediction, and prognosis assessment of acute pancreatitis.

The corresponding text in the revised manuscript is as follows:

Page 36, line 16-32, and page 37, line 1-22:

"The application of AI in the diagnosis, severity prediction, and prognosis assessment

of AP represents an exciting development in the field of medicine. However, based on these current researches, we recognize several limitations and potential challenges that must be addressed to fully leverage the capabilities of AI in this context.

Data quality and availability: AI algorithms require high-quality, comprehensive, and diverse data to build robust and accurate models. In the context of AP, such data sets may not be readily available, especially for rare subtypes of the disease or patient populations with specific comorbidities. Furthermore, incomplete or inconsistent data can lead to biased or flawed results.

Interpretability: AI models, especially those employing complex algorithms like deep learning, often operate as 'black boxes', providing outputs without clear, understandable reasons for their decisions. This can limit their acceptance in the clinical setting, as healthcare professionals typically prefer to understand the reasoning behind a diagnosis or prediction.

Standardization: AI algorithms are typically designed and validated on specific datasets. Their generalizability to other populations or healthcare settings, especially those that are vastly different from the original context, is not guaranteed. This lack of standardization can lead to inconsistent results when the models are used in different settings.

Generalizability: Models trained on a specific set of data may not perform well when applied to different datasets, especially if there are demographic or geographical differences. For example, an AI model trained on data from a high-income country might not work as well in a low-income setting due to differences in healthcare infrastructure, disease prevalence, and patient characteristics.

Regulation: The use of patient data to develop and apply AI models raises significant concerns around data privacy, consent, and security. It's crucial that these concerns are addressed to ensure ethical usage and maintain public trust. For instance, who is responsible if an AI system makes an incorrect diagnosis or prognosis? How is patient data privacy ensured?

Implementation: The successful implementation of AI in healthcare settings requires clinicians to have a certain level of understanding and trust in the technology. This can

be challenging due to varying levels of digital literacy among healthcare providers, and resistance to change.

Given these challenges, ongoing research is critical to improve the reliability, interpretability, and generalizability of AI tools in healthcare, and to address the ethical, legal, and workflow integration issues associated with their use. It's important that as we move forward, these tools are developed and used in a manner that complements the expertise of healthcare professionals, rather than seeking to replace it."

We hope that the revised manuscript will meet the requirements of the journal. And we cordially hope that you will give us guidance if the revised manuscript still has deficiency. Thank you very much again!

Company editor-in-chief:

I have reviewed the Peer-Review Report, the full text of the manuscript, the relevant ethics documents, and the English Language Certificate, all of which have met the basic publishing requirements of the World Journal of Gastroenterology, and the manuscript is conditionally accepted. I have sent the manuscript to the author(s) for its revision according to the Peer-Review Report, Editorial Office's comments and the Criteria for Manuscript Revision by Authors. Before final acceptance, the author(s) must add a table/figure (medical imaging) to the manuscript. There are no restrictions on the figures (color, B/W). Before final acceptance, when revising the manuscript, the author must supplement and improve the highlights of the latest cutting-edge research results, thereby further improving the content of the manuscript. To this end, authors are advised to apply a new tool, the Reference Citation Analysis (RCA). RCA is an artificial intelligence technology-based open multidisciplinary citation analysis database. In it, upon obtaining search results from the keywords entered by the author, "Impact Index Per Article" under "Ranked by" should be selected to find the latest highlight articles, which can then be used to further improve an article under preparation/peer-review/revision. Please visit our RCA database for more information at: https://www.referencecitationanalysis.com/.

Reply: we firstly thank you very much for your guidance and suggestions. According to reviewers' suggestion, Editorial Office's comments, and the Criteria for Manuscript Revision, we have revised the full text of the manuscript, and added a table (Table S1). Table S1 mainly shows the clinical common score systems used to predict the severity of acute pancreatitis, including SAP, mortality, organ failure, ICU admission, location complications, in-hospital adverse events, and pancreatic necrosis. Because of the amount of content in the table, we have included Table S1 in the Supplementary Material.

In order to further improve the content and readability of the review article, we supplement the latest cutting-edge research results and improv the corresponding content. This article mainly includes three parts, namely imaging technology, imaging technology and artificial intelligence. According to the specific content of each part, we make a summary, and put forward the existing problems, as well as solutions and future development direction.

In the manuscript, the content of major changes has been underlined in black.

We actively used the Reference Citation Analysis (RCA) in the process of revising the manuscript. RAC is very useful for us because it helps us to search for the references we need.

We have recommended the RAC to our colleagues and a number of academics and have received positive feedback from them. They all praise the usefulness of the RAC.

Here, based on the actual situation in China, I suggest that more people who need RAC can know the existence of RAC through WeChat.

Because WeChat has an extremely large user population in China, it can make RAC well promoted.

Thank you very much for your guidance to improve our manuscript. We sincerely hope that the revised manuscript meet the requirements of the journal. If there are still deficiencies in the revised manuscript, we will endeavor to improve it.

We sincerely thank you for your guidance again! Wish you success in your work!