March 28, 2021

Dear Editor-in-Chief

We thank Artificial Intelligence in Gastroenterology for considering our manuscript entitled, "Artificial Intelligence for Pancreatic Cancer Detection: Recent Development and Future Direction." by Laoveeravat et al. We have reviewed the comments and have considered them carefully. We have responded to each comment or request in detail. We do appreciate the Journal's recognition of the importance of the study. We trust our responses will prove acceptable.

The point-by-point response to reviewers' and editors' comments can be found below.

Major issue 1: Limited novelty and no discussion about other survey papers There are several recent survey papers on exactly this topic already and several are published within the last year. The authors should have referenced at least some of the good ones and then discussed what new aspects this manuscript brings to us.

Example survey articles include (but are not limited to):

[1] Lin et al., Application of artificial intelligence for the diagnosis, treatment, and prognosis of pancreatic cancer, Artificial Intelligence in Gastroenterology, 2020

[2] Gorris et al., Artificial intelligence for the management of pancreatic diseases, Digestive endoscopy, 2020

[3] Tonozuka et al., The Role of Artificial Intelligence in Endoscopic Ultrasound for Pancreatic Disorders, 2020

[4] Lakshmi et al., A Survey on Detection of Pancreatic Cancer using Deep Learning Techniques, International Journal of Grid and Distributed Computing Vol. 13, No. 1, (2020), pp. 2753– 2763

[5] Cazacu et al., Artificial intelligence in pancreatic cancer: Toward precision diagnosis, Endoscopic Ultrasound 2019

[6] Kuwahara et al., Current status of artificial intelligence analysis for endoscopic ultrasonography, Digestive endoscopy 2020

[7] Pereira et al., Early detection of pancreatic cancer, Lancet Gastroenterol Hepatol, 2020

[8] Akshintala et al., Artificial intelligence in pancreaticobiliary endoscopy, Gastroenterology and hepatology, 2021

These works cover very similar grounds as the manuscript.

For CT and MR, [2] does a better job in terms of these aspects. [7] includes a review of detection biomarkers in addition to using social media for detection and predictive models using electronic health records.

It is not clear what value this manuscript brings in light of these survey papers. The authors should, after conducting a systematic review of survey papers/meta-reviews, be clear about what is missing from these previous (but very recent) reviews and how the manuscript fills the gap.

Response

We have reviewed the prior review articles on artificial intelligence and pancreatic cancer and edited our manuscript to fill the gap that have not been addressed before.

Our work addresses the application of AI in medical imaging studies and biomarkers which were not mentioned in prior studies. We emphasize both aspects and provide the limitations and future directions which can be the idea of further research studies.

Major issue 2: Poor organization The manuscript is organized by imaging modality CT, MR, EUS which is fine. But fundamental aspects are not clarified.

The distinction of computer-assisted detection (CADe) and computer-assisted diagnosis (CADx) is not covered. From the FDA definitions: a CADe device is "intended to identify, mark, highlight or otherwise direct attention to portions of an image that may reveal abnormalities during interpretation of images by the clinician." A CADx device is "intended to provide information beyond identifying abnormalities, such as an assessment of disease."

It is very important to classify research and AI systems as CADe or CADx because they serve different purposes, performance requirements for real clinical use can be substantially different and patient risk changes. The manuscript mixes everything together and 'detection' is often confused with 'diagnostics'. Indeed the review of AI for EUS pancreatic cancer analysis in [2] separates works into CADe and CADx, and it is more comprehensive and systematic than the manuscript. EUS AI systems can also be categorized according to the modality used e.g. elastography, colour doppler, contrast enhancement, standard b-mode or combinations.

Contrast enhancement and elastography is mentioned but it is not clear surveyed papers use only b-mode.

The section on CT and MR is also quite poor with few references. Radiomics is not even mentioned but it is a large topic for CADe and CADx in CT and MR. Most of the discussion in CT is on pancreas segmentation, which is not very relevant for CADe or CADx for which there are many works see e.g. [2] and other survey papers.

Response

The discussion on CADe and CADx and radiomics were added to page 6-10.

We have separated the EUS studies into different modes: b-mode and elastography. No data is available on contrast enhancement. We added a sentence explaining more about the performance of elastography in the EUS section.

Major issue 3: Survey methodology is not described The manuscript is missing a section on survey methodology. How were papers discovered? Why were they included?

Response

We added the sentences on the survey method into the introduction as 'we conducted the systematic review on AI and pancreatic cancer with keywords of 'aritficial intelligenece' and 'pancreatic cancer' from PUBMED and IEEE databases.'

Major issue 4: Performance of papers are reported with metrics (usually sensitivity and specificity) but other details are missed. We know that the quality of the evaluation methodology of an AI system is fundamental. Results from papers are presented with performance metrics but we lack evaluation details and there is no real comment on the quality of the evaluation (a critical aspect). This manuscript would have much better value in my opinion (and differentiate it from other survey papers) if it also covered aspects such as: 1. Dataset size (test and training) 2. How datasets were divided into training and test sets 3. Dataset access (open or private) 4. Study type (retrospective/prospective, single-centre/multi-centre etc.) 5. AI models Note that [6] did something like this already for some papers for AI+EUS.

Response

We have edited the tables of EUS and added a table of CT and MRI per reviewer's comment.

Major issue 5: Abdominal US was not discussed There exists several works for pancreatic cancer detection with abdominal US see e.g. [6]. Why was abdominal US not

included? Unlike EUS, it is non-invasive and it is considered one of the first imaging modalities to consider.

Response

No available studies in AI-guided US and pancreatic cancer. We added the sentence emphasizing this point in the ultrasonography section.

Major issue 6: Future directions was quite weak For EUS, the suggested future directions are quite light: to develop AI systems capable of real-time performance and evaluation on larger datasets. But there are so many other important future directions and unsolved challenges that are not mentioned. I was hoping to read about the authors' insights into them such as: 1. How to systematically record and standardize EUS data, required for large-scale deep learning? Unlike CT and MR, this is much harder for EUS. 2. How do we motivate centres to systematically record and share EUS data? 3. How to train AI systems that handle the inherent operator-dependence of EUS and variable image/procedure quality? 4. How to make AI models interpretable and explainable, so that they are not just working in a 'black box' manner. This is a known issue in AI-based CADx and CADe systems today. What is particularly relevant when it comes to pancreatic cancer? 5. What are likely to be the fundamental limits of AI and EUS? We cannot expect AI to work magically. It has limits on diagnostic capabilities because of limits in the information contained in an EUS video/image. e.g. highly accurate differential diagnostics of IPMNs, SCNs and MCNs in b-mode may not be possible. There are no comments on the limits of AI. 6. What EUS modalities are likely to be important for use for CADe and CADx? What could we expect to achieve with b-mode compared to b-mode + elastography, or b-mode + elastography + contrast enhancement? 7. What other factors are preventing good results presented in papers from translating to clinical use? 8. How to combine data with CT, MR, biomarkers and electronic health records to make better predictions? Are there examples for pancreatic cancer or can we draw inspiration from work in other cancer types.

Response

We added additional aspects in Future Prospect as follows.

The modes of EUS: b-mode and elastography do not provide different accuracy and predictive value for pancreatic cancer. However, no data is available for EUS with contrast enhancement. B-mode which is generally used among centers can be the first step of AI implication. Ultimately, the data of imaging studies, biomarkers, and clinical parameters will be combined to build the sophisticated algorithm and implemented in the electronic medical records where clinicians use it as the predictive tool. There are a few limitations of AI application for EUS. First, the collection of EUS images as the big data is difficult. The collaboration of gastroenterologists, radiologists, and hospital

administration will help facilitate the retrieval of images into the system. Multicenter participation is required to create the large dataset of EUS images of which it will optimize the efficiency of AI. The platform of dataset in one institution can be the good example that other centers can adopt and join the group. Second, the root of clinical decision based on AI results is possibly affected by the black box issue (inability to identify the ground of decision). Although there are ways that enable AI to be more interpretable, it is still an active area of research in computer science.

Major issue 7: No figures or tables. Figures in previous survey papers can convey useful information such as EUS images of pancreatic lesions, diagnostic trees, decision processes (both by clinician and AI systems). The fact that this manuscript has no figures is a major weakness. Furthermore, the fact that this manuscript has no tables to systematically summarize works is also a major limitation.

Response

This is a review article, in which we did not request the permission to reuse any images from any publisher. Therefore, we have no available image or other figures, rather the one we created ourselves (Figure 1).

There are two tables: AI guided (1) CT and MRI and (2) EUS and one figure of the AI process.

We believe that our responses and manuscript modifications will prove satisfactory upon review. We thank the editors and reviewers for their insightful comments. We believe the manuscript is stronger for them.

Sincerely yours,

Amporn Atsawarungruengkit, MD

Department of xxx

Address:

Phone:

Email address: