

Peer Review File

Article information: <https://dx.doi.org/10.21037/tcr-23-2234>

Review Comments

Reviewer A

The manuscript addresses a timely and relevant topic - the application of AI, specifically ChatGPT, in medical education for renal cell carcinoma. The focus on both patient and trainee education is a unique aspect that adds value to the current literature. I have a few suggestions:

Major:

Comment 1: The currently stated aim in line 98-102 contains “to identify dangerous or misleading information in its responses”. Please explain how the results support this aim or please revise this sub-aim.

Reply 1: **Addressed in the text. This sub-aim was removed as an explicit goal in lines 98-102.**

Comment 2: It would be helpful to actually see the 15 questions and ChatGPT answers as a supplemental material. Also there is a little discrepancy in the results - Line 199: What is question 19 when there are 15 questions?

Reply 2: **Added as supplementary material and re-numbered question 19 in text. Question 19 in our survey corresponded to Question 9 of the content questions (numbering included pre-assessment questions). Revised for clarity.**

Comment 3: It would also be very useful to include the pre- and post- assessment questionnaires as a supplemental material.

Reply 3: **Added as supplementary material.**

Comment 4: Why Fig 1 is change in responses? I did not fully understand which one is before which one is after. There seems to be only one time point in this figure.

Reply 4: **I believe the Figure legends and images were swapped in error. This has been corrected and double-checked to show the appropriate legend for the labeled figure.**

Comment 5: Please provide a proper conclusion based on the results before going into its potential meaning.

Reply 5: **Provided in the text.**

Changes in the Text:

We found that clinicians and lay assessors consistently rated ChatGPT highly on the accuracy and usefulness of information provided in response to questions regarding the management of SRMs and RCC. Completing content assessment improved confidence in the accuracy of ChatGPT's information and increased agreement that it should be used for medical education. These results are an early, informal evaluation of the capabilities of evolving AI tools but show great promise for this new technology.

Minor:

Comment 6: Introduction line 70, please define limited AI (in comparison to general AI) for readers.

Reply: **This is provided in the text.**

Changes in the text: "Limited or narrow AI are programs with a specific set of constraints and output types and can only handle tasks within their programming or training parameters. General AI has not yet been achieved but would be capable of learning and solving problems of any kind or format, on the level of a human mind and without concrete limits to its abilities."

Comment 7: It would be interesting and more relevant to see the performance of ChatGPT 4.0. It would be good to mention it in the discussion.

Reply 7: **4.0 is now mentioned explicitly and the importance of ongoing assessments highlighted even more.**

Changes in the Text: "Since this work was completed, ChatGPT 4.0 and countless other AI tools have been released and will need similar assessments prior to widespread adoption in a medical context."

Overall, the manuscript provides valuable insights into the use of AI in medical education, specifically in the context of renal cell carcinoma. Its emphasis on both patients and trainees is commendable, and the methodology is sound. With minor improvements, the manuscript could make a significant contribution to the field.

Reviewer B

Comment:

Dear authors,

I reviewed the paper "Integrating Artificial Intelligence in Renal Cell Carcinoma: Evaluating ChatGPT's Performance in Educating Patients and Trainees" with interest.

The paper is interesting and fluent, the topic is appealing. The study has a simple but effective design where statistical analysis is trivial.

I would only suggest the authors discussing about other technological tools that can be useful for patients and trainees. In particular telemedicine is increasing after COVID pandemics and I believe that it should be mentioned in the discussion section. In this light, I would suggest citing these two relevant papers on the topic

1. doi: 10.1016/j.urology.2020.09.015

2. doi: 10.1016/j.acuro.2020.06.008

Reply: **Included an additional paragraph in the discussion section with this point and citation of the 2nd of these two papers, as I did not feel the first represented the point we are trying to make in our discussion. However, I do feel that contextualizing AI in the broader schema of technological advancements was helpful.**

Changes to text, Page 7: "The recent focus on AI is in many ways just one facet of rapid technologic advancement impacting medicine, with telehealth being the iteration prior. Particularly during the COVID-19 pandemic, this technology allowed physicians to continue to achieve patient care goals safely and effectively. Patients on the whole welcomed the change, with a majority satisfied with its use for their care and not feeling depersonalization from its use.³⁸ Younger and more tech-savvy

patients are even more likely to embrace these technological changes in medicine, and physicians should take note.³⁰ While adoption of telehealth was difficult for some, its utility was immense and its introduction into healthcare has expanded our options for interacting with and treating patients in ways that meet their needs. AI will be a similarly impactful tool, but the burden is on physicians to use it safely and well. ”