

## Reply to Daungsupawong and Wiwanitkit



Dear Editor, Dr. Ray, and Dr. Wiwanitkit

I have carefully read your letter and the important issues raised in it regarding our work on the evaluation of custom ChatGPT agents in the management of pancreatic cystic lesions [1,2]. I agree with the limitations you have pointed out, keeping in mind that this study is intended as a proof of concept for a novel feature of large language models (LLMs). Although 60 scenarios is not a large sample size, most similar studies in the field of gastroenterology and LLMs studied a similar number of scenarios [3,4,5]. The issue of interrogation of discrepancies as part of a systemic approach to the development and training of LLMs should indeed be further implemented and improved. An example of such a systemic approach is described in a recent work that evaluated various setups using ChatGPT for management of hepatitis C infection [5]. The ethical considerations including bias and error handling and user responsibility should also be explored. Hopefully, these issues will be addressed gradually in future studies as this field of currently novel research progresses from generated scenarios to evaluation on clinical data and prospective interventional studies.

### Conflict of Interest

The authors declare that they have no conflict of interest.

### The authors

#### Yuri Gorelik<sup>1</sup>

- 1 Department of Internal Medicine D, Gastroenterology, Rambam Health Care Campus, Haifa, Israel

### Corresponding author

#### Dr. Yuri Gorelik, MD, MPH

Rambam Health Care Campus, Department of Internal Medicine D, Gastroenterology, Haaliya Hashnia, 31999 Haifa, Israel  
yurigorelik@gmail.com

### Publication note

Letters to the editor do not necessarily represent the opinion of the editor or publisher. The editor and publisher reserve the right to not publish letters to the editor, or to publish them abbreviated or in extracts.

### References

- [1] Daungsupawong H, Wiwanitkit V. Comment on "Using a customized GPT to provide guideline-based recommendations for management of pancreatic cystic lesions. *Endosc Int Open* 2024; 12: 810
- [2] Gorelik Y, Ghersin I, Arraf T et al. Using a customized GPT to provide guideline-based recommendations for management of pancreatic cystic lesions. *Endosc Int Open* 2024; 12: E600–E603 doi:10.1055/a-2289-9334
- [3] Pugliese N, Wai-Sun Wong V, Schattenberg JM et al. Accuracy, reliability, and comprehensibility of ChatGPT-generated medical responses for patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* 2024; 22: 886–889.e5
- [4] Levartovsky A, Ben-Horin S, Kopylov U et al. Towards AI-augmented clinical decision-making: An examination of ChatGPT's utility in acute ulcerative colitis presentations. *Am J Gastroenterol* 2023; 118: 2283–2289 doi:10.14309/ajg.0000000000002483
- [5] Kresevic S, Giuffrè M, Ajcevic M et al. Optimization of hepatological clinical guidelines interpretation by large language models: a retrieval augmented generation-based framework. *NPJ Digit Med* 2024 71 2024; 7: 1–9 doi:10.1038/s41746-024-01091-y

### Bibliography

*Endosc Int Open* 2024; 12: E811

DOI 10.1055/a-2335-8405

ISSN 2364-3722

© 2024. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

