POLISH JOURNAL *of* RADIOLOGY

Letter to the Editor

Correspondence to "Will ChatGPT pass the Polish specialty exam in radiology and diagnostic imaging?"

Hinpetch Daungsupawong[1,A,B,D,E], Viroj Wiwanitkit[2,A,D]

[1]Private Academic Consultant, Phonhong, Lao People's Democratic Republic
[2]Chandigarh University, India

Dear Editor,

We read with a great interest "Will ChatGPT pass the Polish specialty exam in radiology and diagnostic imaging? Insights into strengths and limitations [1]". The purpose of this paper is to investigate the performance of Chat-GPT, a cutting-edge language model, in relation to the pass rate of the national specialist examination (PES) in radiology and imaging diagnostics in the Polish educational system. To assess complexity, the researchers used a PES exam with 120 questions and classified them using Bloom's taxonomy. ChatGPT was utilized to deliver exam answers, and its confidence in each response was graded on a scale of 1 to 5. While ChatGPT did not meet the PES test pass rate criteria, it performed well in specific question categories, and no significant disparities in the percentage of right responses were noted across question types and sub-types.

The key weakness revealed in this study is that Chat-GPT did not meet the PES exam pass rate criteria. This implies that the model's performance fell short of the expected degree of expertise and comprehension in radiology and imaging diagnostics. It indicates a weakness in ChatGPT's capacity to answer sophisticated and specialized exam questions in this sector.

Modern methods and a large training set are needed to eliminate bias and errors from chatbots [2,3]. This is due to the risks associated with relying only on a large data source. The use of chatbots presents ethical concerns because some of their impacts may be unpleasant or unexpected. To prevent the spread of harmful ideas and incorrect information, ethical constraints and limitations must be introduced as AI language models advance. If there is not enough human monitoring or verification, a chatbot can give a false reference, which could lead to other issues [2,3].

Several future paths can be followed to solve the constraints and improve the performance of ChatGPT in the context of the national specialist examination (PES) in radiology and imaging diagnostics. To begin, domain-specific training and fine-tuning of ChatGPT are required to improve its comprehension and accuracy in answering test questions within the field. Incorporating specialist knowledge and resources into ChatGPT's training data would also help it give precise and insightful answers. Furthermore, performing a detailed examination of the specific types of questions or themes where ChatGPT faltered would aid in identifying areas for improvement and guiding the creation of tactics to increase its performance in those specific areas.

Conflict of interest

The authors report no conflict of interest.

References

1. Kufel J, Paszkiewicz I, Bielówka M, et al. Will ChatGPT pass the Polish specialty exam in radiology and diagnostic imaging? Insights into strengths and limitations. Pol J Radiol 2023; 88: e430-e434.
2. Kleebayoon A, Wiwanitkit V. Artificial intelligence, chatbots, plagiarism and basic honesty: comment. Cell Mol Bioeng 2023; 16: 173-174.
3. Kleebayoon A, Wiwanitkit V. Comment on "How may ChatGPT impact medical teaching?". Rev Assoc Med Bras (1992) 2023; 69: e20230593. doi: 10.1590/1806-9282.20230593.

Correspondence address:
Hinpetch Daungsupawong, Private Academic Consultant, Phonhong, Lao People's Democratic Republic, e-mail: hinpetchdaung@gmail.com

Authors' contribution:
A Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection