Large Language Models: Could They Be the Next Generation of Clinical Decision Support Systems in Cardiovascular Diseases?

To the Editor,

We read with great interest the article by Bozyel et al,[1] published in the February 2024 issue of *the Anatolian Journal of Cardiology*, which assessed artificial intelligence (AI)-based clinical decision support systems in cardiovascular diseases. This article provides detailed information about the current status and potential application of AI in cardiology and cardiac imaging. Despite the insightful coverage of this topic by Bozyel et al,[1] we have not seen any details about the potential usage of large language models (LLMs) in cardiovascular diseases, which is a recently evolving area. Therefore, we would like to offer different perspectives on the potential of LLMs in cardiovascular diseases as clinical decision support systems.

There is an increasing interest in employing LLMs in cardiovascular diseases such as the assessment of electrocardiography and answering questions about acute coronary syndrome regarding European Society of Cardiology Guidelines.[2,3] Additionally, LLMs have been utilized for simplifying cardiovascular magnetic resonance reports in cardiovascular diseases.[4] Although AI can help in cardiovascular diseases, it is vital to evaluate the diagnostic performance of these LLMs.

The European Board of Cardiovascular Radiology (EBCR) Diploma has been conducting board examinations since 2011.[5] The examination comprises 2 parts: a written and an oral examination. The written examination consists of 30 multiple-choice questions on all aspects of cardiovascular radiology. A total score of a minimum of 60% from the written part and solving 2 cases during the oral examination are required to pass the examination. We assessed 10 sample questions from the sample written examination of the EBCR Diploma, which are openly accessible on the European Society of Cardiovascular Radiology website (https://www.escr.org/diploma). We initiated the input prompt as follows: "As a highly experienced professor of cardiovascular diseases with 30 years of imaging expertise, you assisted in cardiovascular radiology cases. Your role is to analyze the questions, patient histories, and imaging findings to determine the correct answer. There is only one correct answer." This prompt was presented in March 2024 on 5 distinct platforms with default hyperparameters: OpenAI's ChatGPT 3.5 and 4 (https://chat.openai.com), Google Gemini (https://gemini.google.com), Microsoft Copilot (https://copilot.microsoft.com) (Balanced), Perplexity (https://www.perplexity.ai). Our results indicated that ChatGPT 3.5 had 80.0% accuracy (8/10 questions), ChatGPT 4 also 60.0% (6/10 questions), Perplexity 50.0% (5/10 questions), Microsoft Copilot 90.0% (9/10 questions), and Google Gemini 60.0% (6/10 questions). This letter demonstrates that Microsoft Copilot outperforms other LLMs in terms of accuracy. The disparities in performance among Microsoft Copilot, Google Gemini, Perplexity, and ChatGPT 3.5 and 4 in text-based knowledge assessment can be attributed to differences in their designs. Future research on the capabilities of LLMs in cardiovascular imaging is necessary to demonstrate their diagnostic performance.

Yasin Celal Güneş[1] [ID]

Turay Cesur[2] [ID]

[1]*Department of Radiology, Kırıkkale Yüksek İhtisas Hospital, Kırıkkale, Türkiye*
[2]*Department of Radiology, Mamak State Hospital, Ankara, Türkiye*

Corresponding author:
Yasin Celal Güneş
✉ gunesyasincelal@gmail.com

REFERENCES

1. Bozyel S, Şimşek E, Koçyiğit Burunkaya D, et al. Artificial intelligence-based clinical decision support systems in cardiovascular diseases. *Anatol J Cardiol*. 2024;28(2):74-86. [CrossRef]
2. Günay S, Öztürk A, Özerol H, Yiğit Y, Erenler AK. Comparison of emergency medicine specialist, cardiologist, and Chat-GPT in electrocardiography assessment. *Am J Emerg Med*. 2024;80:51-60. [CrossRef]
3. Gurbuz DC, Varis E. Is ChatGPT knowledgeable of acute coronary syndromes and pertinent European Society of Cardiology Guidelines? *Minerva Cardiol Angiol*. 2024;72(3):299-303. [CrossRef]
4. Salam B, Kravchenko D, Nowak S, et al. GPT-4 makes cardiovascular magnetic resonance reports easy to understand. *J Cardiovasc Magn Reson*. 2024;26(1):101035. [CrossRef]
5. Natale L, Vliegenthart R, Salgado R, et al. Cardiac radiology in Europe: status and vision by the European Society of Cardiovascular Radiology (ESCR) and the European Society of Radiology (ESR). *Eur Radiol*. 2023;33(8):5489-5497. [CrossRef]