Exploratory Research and Hypothesis in Medicine **2024** vol. 000(000) | 000–000 Epub
DOI: 10.14218/ERHM.2024.00016

Letter to the Editor

ChatGPT in Radiology: Insights into Current Advantages and Limitations of Artificial Intelligence in Radiology Reporting

Shuai Ren[1#], Lina Song[1#], Marcus J. Daniels[2], Ying Tian[1*] and Zhongqiu Wang[1*]

[1]*Department of Radiology, Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China;* [2]*Department of Radiology, NYU Langone Health, New York, USA*

Received: April 05, 2024 | **Revised:** July 22, 2024 | **Accepted:** July 26, 2024 | **Published online:** August 22, 2024

Dear Editors,

In recent years, artificial intelligence (AI) has emerged as a valuable tool in radiology, promising to enhance diagnostic accuracy, efficiency, and predictions of patient outcomes. ChatGPT, a powerful large language model (LLM) developed by OpenAI, has garnered considerable attention worldwide.[1] ChatGPT has the potential to revolutionize radiology by providing a more streamlined and accurate approach to analyzing and interpreting medical images.[2]

We recently read with great interest the insightful article titled "Feasibility of differential diagnosis based on imaging patterns using a large language model" by Dr. Kottlors *et al*.,[3] published in the July 2023 issue of *Radiology*. The authors aimed to assess the performance of LLMs in identifying relevant differential diagnoses based on specific imaging patterns. Dr. Kottlors *et al*.[3] found that GPT-4 achieved a concordance rate of 68.8% (55 out of 80 cases) with expert consensus in generating the top differential diagnoses. Notably, 93.8% (75 out of 80 cases) of the differential diagnoses proposed by GPT-4 were considered acceptable alternatives.

Dr. Kottlors *et al*.[3] have conducted groundbreaking research demonstrating the potential of LLMs to generate relevant differential diagnoses based on imaging patterns. Their work serves as a proof-of-concept for enhancing diagnostic decision-making and significantly reducing the time and resources required for diagnosis by enabling real-time analysis and interpretation of images. In clinical practice, variability in observation and interpretation is common among radiologists due to individual differences in biases, training, and specialized knowledge. Remarkably, LLMs address these challenges by employing a fixed algorithm trained on extensive data, thereby offering consistent and accurate interpretations of medical images.[4] Moreover, LLMs contribute to streamlined workflows and enhance patient experience by enabling radiologists to analyze and interpret images more efficiently. Another significant advantage is their ability to generate code tailored for medical imaging research.[5] Additionally, LLMs can empower individuals with minimal or no coding experience to transform research concepts into practical code,[5] which is instrumental in developing machine learning models specifically designed for medical imaging research.

While the limitations of ChatGPT are acknowledged, a significant concern is its performance dependency on the quality and quantity of training data.[4] Data collection, analysis, and interpretation can be complex and time-consuming, with data often being noisy or biased, significantly impacting the accuracy and reliability of ChatGPT. ChatGPT may underperform in detecting rare or atypical cases for which it lacks specific training. Its performance may also be limited in specific subgroups or modalities if the training data predominantly represents other subgroups or different imaging modalities. Additionally, ethical considerations are crucial to ensure its reasonable and beneficial application. Challenges such as detecting plagiarism or publication fabrication with language models such as ChatGPT are notable.[6] Furthermore, new ethical challenges related to accountability, bias, and transparency may arise. Inaccurate diagnoses can profoundly impact patients when algorithms are biased or trained on misleading data, potentially leading to disparities in medical care. Copyright issues also need to be addressed, as algorithms are trained using essential information sourced from various origins.

Horiuchi *et al*.[7] assessed GPT's diagnostic performance in neuroradiology across various conditions, highlighting that despite its adaptability, GPT's diagnostic accuracy might still vary among specific diseases. They emphasized the need for further comparisons with radiologists to fully evaluate its reliability and effectiveness. In another study, Nakamura *et al*.[8] explored GPT's potential to automate lung cancer staging and TNM classification using CT radiology reports, suggesting significant promise in this area with recommendations to enhance numerical reasoning and domain-specific knowledge. Chung *et al*.[9] investigated GPT's capability to generate concise, patient-friendly MRI reports for prostate cancer patients, achieving readability at a sixth-grade level and enhancing physician satisfaction. Additionally, Nakaura *et al*.[10] investigated GPT-3.5 and GPT-4's ability to produce readable radiology reports from succinct imaging findings, emphasizing the need for radiologist verification of accuracy in clinical impressions and diagnoses.

The continuous evolution of language models, such as ChatGPT, is paving the way for extensive research and development

***Correspondence to:** Ying Tian and Zhongqiu Wang, Department of Radiology, Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, No. 155 Hanzhong Road, Nanjing, Jiangsu 210029, China. ORCID: https://orcid.org/0000-0002-6031-2335 (YT); https://orcid.org/0000-0001-6681-7345 (ZQW). Tel: +86-025-86617141, Fax: +86-025-86619843, E-mail: ty52111@sina.com (YT) and zhongqiuwang@njucm.edu.cn (ZQW)

[#]These authors have contributed equally to this work.

How to cite this article: Ren S, Song L, Daniels MJ, Tian Y, Wang Z. ChatGPT in Radiology: Insights into Current Advantages and Limitations of Artificial Intelligence in Radiology Reporting. *Explor Res Hypothesis Med* 2024;000(000):000–000. doi: 10.14218/ERHM.2024.00016.

in conversational AI. This section explores potential technical advancements and innovative directions aimed at enhancing ChatGPT's capabilities, addressing current limitations, and advancing conversational AI systems. Future research avenues for ChatGPT span a spectrum of technological challenges and opportunities. Looking ahead, the future of LLMs in radiology appears promising, poised to improve patient care, enhance outcomes, and empower radiologists with advanced capabilities.

In conclusion, efforts should be increased to advance ChatGPT, ensuring that the algorithm is trained on high-quality data and a diverse range of imaging findings in clinical conditions. It is crucial to address ethical challenges, including accountability, bias, transparency, and privacy concerns, in the integration of AI-generated decision aids into human decision-making processes.

Acknowledgments

None.

Funding

This study was funded by the National Natural Science Foundation of China (82371919, 82372017, 82202135, 82171925), China Postdoctoral Science Foundation (2023M741808), the Young Elite Scientists Sponsorship Program by Jiangsu Association for Science and Technology (JSTJ-2023-WJ027), Jiangsu Provincial Key Research and Development Program (BE2023789), the Foundation of Excellent Young Doctor of Jiangsu Province Hospital of Chinese Medicine (2023QB0112), Nanjing Postdoctoral Science Foundation, Natural Science Foundation of Nanjing University of Chinese Medicine (XZR2023036, XZR2021003), the Medical Imaging Artificial Intelligence Special Research Fund Project, Nanjing Medical Association Radiology Branch, and the Developing Program for High-level Academic Talent in Jiangsu Hospital of Chinese Medicine (y2021rc03 and y2021rc44).

Conflict of interest

None.

Author contributions

Study concept and design (SR, ZQW), funding acquisition (SR, YT, ZQW), drafting of the manuscript (SR, LNS), critical revision of the manuscript for important intellectual content (MJD, YT, ZQW), and study supervision (YT, ZQW). All authors have made significant contributions to this study and have approved the final manuscript.

References

[1] Elkassem AA, Smith AD. Potential Use Cases for ChatGPT in Radiology Reporting. AJR Am J Roentgenol 2023;221(3):373–376. doi:10.2214/AJR.23.29198, PMID:37095665.

[2] Temperley HC, O'Sullivan NJ, Mac Curtain BM, Corr A, Meaney JF, Kelly ME, *et al*. Current applications and future potential of ChatGPT in radiology: A systematic review. J Med Imaging Radiat Oncol 2024;68(3):257–264. doi:10.1111/1754-9485.13621, PMID:38243605.

[3] Kottlors J, Bratke G, Rauen P, Kabbasch C, Persigehl T, Schlamann M, *et al*. Feasibility of Differential Diagnosis Based on Imaging Patterns Using a Large Language Model. Radiology 2023;308(1):e231167. doi:10.1148/radiol.231167, PMID:37404149.

[4] Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, *et al*. ChatGPT and Other Large Language Models Are Double-edged Swords. Radiology 2023;307(2):e230163. doi:10.1148/radiol.230163, PMID:36700838.

[5] Akinci D'Antonoli T, Stanzione A, Bluethgen C, Vernuccio F, Ugga L, Klontzas ME, *et al*. Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions. Diagn Interv Radiol 2024;30(2):80–90. doi:10.4274/dir.2023.232417, PMID:37789676.

[6] Cao JJ, Kwon DH, Ghaziani TT, Kwo P, Tse G, Kesselman A, *et al*. Accuracy of Information Provided by ChatGPT Regarding Liver Cancer Surveillance and Diagnosis. AJR Am J Roentgenol 2023;221(4):556–559. doi:10.2214/AJR.23.29493, PMID:37222278.

[7] Horiuchi D, Tatekawa H, Shimono T, Walston SL, Takita H, Matsushita S, *et al*. Accuracy of ChatGPT generated diagnosis from patient's medical history and imaging findings in neuroradiology cases. Neuroradiology 2024;66(1):73–79. doi:10.1007/s00234-023-03252-4, PMID:37994939.

[8] Nakamura Y, Kikuchi T, Yamagishi Y, Hanaoka S, Nakao T, Miki S, *et al*. ChatGPT for automating lung cancer staging: feasibility study on open radiology report dataset. medRxiv 2023. doi:10.1101/2023.12.11.23299107.

[9] Chung EM, Zhang SC, Nguyen AT, Atkins KM, Sandler HM, Kamrava M. Feasibility and acceptability of ChatGPT generated radiology report summaries for cancer patients. Digit Health 2023;9:20552076231221620. doi:10.1177/20552076231221620, PMID:38130802.

[10] Nakaura T, Yoshida N, Kobayashi N, Shiraishi K, Nagayama Y, Uetani H, *et al*. Preliminary assessment of automated radiology report generation with generative pre-trained transformers: comparing results to radiologist-generated reports. Jpn J Radiol 2024;42(2):190–200. doi:10.1007/s11604-023-01487-y, PMID:37713022.