Article published online: 2024-04-19

Exploring the Potentials of Large Language Models in Vascular and Interventional Radiology: Opportunities and Challenges

Taofeeq Oluwatosin Togunwa[1,2] Abdulquddus Ajibade[1] Christabel Uche-Orji[1,2] Richard Olatunji[1]

[1] Department of Radiology, College of Medicine, University of Ibadan, Oyo, Nigeria
[2] College Research and Innovation Hub, University College Hospital, Ibadan, Oyo, Nigeria

Address for correspondence Taofeeq Oluwatosin Togunwa, MBBS, Department of Radiology, College of Medicine, University of Ibadan, Oyo 200212, Nigeria (e-mail: togunwataofeeq@gmail.com).

Arab J Intervent Radiol 2024;8:63–69.

Abstract

The increasing integration of artificial intelligence (AI) in healthcare, particularly in vascular and interventional radiology (VIR), has opened avenues for enhanced efficiency and precision. This narrative review delves into the potential applications of large language models (LLMs) in VIR, with a focus on Chat Generative Pre-Trained Transformer (ChatGPT) and similar models. LLMs, designed for natural language processing, exhibit promising capabilities in clinical decision-making, workflow optimization, education, and patient-centered care. The discussion highlights LLMs' ability to analyze extensive medical literature, aiding radiologists in making informed decisions. Moreover, their role in improving clinical workflow, automating report generation, and intelligent patient scheduling is explored. This article also examines LLMs' impact on VIR education, presenting them as valuable tools for trainees. Additionally, the integration of LLMs into patient education processes is examined, highlighting their potential to enhance patient-centered care through simplified and accurate medical information dissemination. Despite these potentials, this paper discusses challenges and ethical considerations, including AI over-reliance, potential misinformation, and biases. The scarcity of comprehensive VIR datasets and the need for ongoing monitoring and interdisciplinary collaboration are also emphasized. Advocating for a balanced approach, the combination of LLMs with computer vision AI models addresses the inherently visual nature of VIR. Overall, while the widespread implementation of LLMs in VIR may be premature, their potential to improve various aspects of the discipline is undeniable. Recognizing challenges and ethical considerations, fostering collaboration, and adhering to ethical standards are essential for unlocking the full potential of LLMs in VIR, ushering in a new era of healthcare delivery and innovation.

Keywords
- ► artificial intelligence
- ► vascular and interventional radiology
- ► large language models
- ► machine learning
- ► radiology
- ► patient-centered care

Introduction

Over the past decade, the popularity and utilization of artificial intelligence (AI) within the healthcare industry have experienced a significant surge.[1] This has led to the potential for substantial enhancements in healthcare efficiency and precision, ultimately resulting in improved patient care, more informed decision-making, and

article published online
April 19, 2024

DOI https://doi.org/10.1055/s-0044-1782663.
ISSN 2542-7075.

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India

considerable cost savings. AI facilitates the evaluation and analysis of various disease conditions, often using complex and rapidly expanding datasets. Many times, AI achieves remarkable accuracy and depth in these tasks by harnessing cutting-edge concepts and technologies such as machine learning (ML), neural networks (NNs), and large language models (LLM).[2]

The discipline of vascular and interventional radiology (VIR) has also seen significant advancements in recent years, with an increase in the number of procedures performed, subspecialization, and improved accuracy in interventions.[3] The demonstrated AI applications in various VIR use cases abounds including AI-assisted endovascular clot retrieval for acute ischemic stroke, predicting responses to transcatheter arterial chemoembolization in hepatocellular carcinoma, AI-guided ultrasound in echocardiography, and angiography-based ML algorithms for real-time estimation of fractional flow reserve.[4,5] So, as VIR techniques and procedures mature and achieve wider acceptance, it could prove more useful in multidisciplinary care with the incorporation of AI potentially leading to improved outcomes for all stakeholders.[6]

In this context, VIR is the frontier of the concept of Imaging 3.0, an initiative aimed at showcasing the expanded contributions of radiologists beyond conventional image interpretation.[7] In detailing key aspects, VIR shows its pivotal role in percutaneous, image-guided procedures such as abscess drainage and needle biopsy. This emphasizes the specialty's capacity to reduce morbidity, enhance patient outcomes, and offer cost-effective alternatives to other surgical interventions. Additionally, the significance of transjugular intrahepatic portosystemic shunts and the central role of interventional radiologists in delivering cost-effective central venous access services are emphasized, making substantial contributions to multidisciplinary cancer treatment.

While many of the AI-used cases in VIR primarily fall within the domains of computer vision and image classification/segmentation in AI, the impact of AI in VIR potentially extends beyond these conventional boundaries. LLMs are a promising domain in AI with opportunities for substantial applications in VIR. LLMs are advanced AI models, such as Open AI's Chat Generative Pre-Trained Transformer (ChatGPT), which are capable of understanding and generating human-like text.[8] LLMs have emerged as a revolutionary breakthrough in AI, as it pertains to natural language processing (NLP). Their ability to process and comprehend language has propelled them into diverse sectors, including finance, marketing, and healthcare.

In general healthcare, the LLM technology is now widely acclaimed for its role in medical language tasks, including automated report generation and integration with healthcare systems.[9] It enables the comprehensive extraction of patient data from electronic health records, laboratory results, and prior imaging studies, significantly enhancing the diagnostic process and patient care. The literature on the potential application of LLMs in VIR is, however, sparse relative to the broader discussions of opportunities and concerns provoked by AI generally but more importantly, the readily accessible LLM technologies like ChatGPT.

Hence, the aim of this article is to comprehensively explore the potentials of LLMs in VIR. By reviewing the current abilities of LLMs, the aim is to highlight the potentials of this technology, outlining its current and prospective contributions to advancing clinical practices, patient outcomes, and educational activities in VIR. The challenges, potential future directions, and advancements of LLMs in the field of VIR are also reviewed.

Overview of LLMs

LLMs are gaining significant traction as invaluable tools in the field of radiology.[10] These advanced AI models work by transforming text into numerical tokens, thus capturing contextual information during training.[11] Consequently, they predict and select plausible next tokens based on learned language patterns, creating the appearance of encyclopedic knowledge and reasoning. ChatGPT is one of the popular LLMs, developed by OpenAI and made publicly available in November 2022. It is trained on extensive text datasets in various languages and can produce human-like responses to text input. ChatGPT utilizes the GPT architecture to process natural language and generate context-based responses. Other publicly available LLMs include T5, Pythia, and LlaMA.[12–14] For an overview of additional open access LLMs suitable for personal and research applications, the reader is referred to a detailed Github repository.[15]

The scientific community has shown diverse reactions toward ChatGPT, reflecting the ongoing debate surrounding the benefits and risks of LLMs and generative AI technologies in general.[16] On one hand, ChatGPT and other LLMs have demonstrated usefulness in conversational and writing tasks in medicine, enhancing output efficiency and accuracy.[9] On the other hand, concerns have emerged regarding potential bias in its training datasets, leading to limitations and factual inaccuracies, a phenomenon referred to as "hallucination." Additionally, there are security concerns related to the spread of misinformation and the possibility of cyber-attacks utilizing LLMs.[17]

The Present and Promising Future of LLMs in VIR

In a notable experiment, ChatGPT 3.5 demonstrated impressive performance on 376 USMLE test questions from the June 2022 sample exam, achieving a passing or near-passing score threshold of 60%, while exhibiting high concordance and insightful responses, without specialized training.[18] A more recent study indicates that the latest version, ChatGPT 4 performed even better and demonstrated remarkable medical reasoning.[19,20] Additionally, Yan et al[21] introduced RadBERT, a language model fine-tuned for radiology, excelling in NLP tasks and promising automation in abnormal findings identification and report creation. These advances could alleviate healthcare workload and burnout, suggesting future developments in automated radiology reports and broader healthcare applications. It is fascinating that this technology can achieve such outcomes in its early stages, and without domain-specific training, more so the LLMs can reason through novel problems to a remarkable degree without specific training, a phenomenon known as "zero shot."[22]

Considering these factors, alongside the novel capabilities of LLMs, their potential applications are noteworthy. While the literature on this emerging subject is still limited, key areas where LLMs demonstrate promising applications in VIR are highlighted (►**Table 1**).

Supporting Clinical Decision-Making

LLMs, equipped with sophisticated NLP methods, possess the capability to analyze extensive volumes of medical literature, electronic health records, and patient data. By processing and understanding the intricate patterns and nuances within this information, LLMs can assist the interventional radiologist in making more informed and precise decisions regarding disease diagnosis, treatment planning, and prognostic predictions.

Shen et al[23] have shown that ChatGPT can use large knowledge bases to swiftly answer questions about the most suitable imaging study for specific clinical scenarios. A recent study assessed the performance of two LLMs, ChatGPT and Glass AI, in predicting optimal neuroradiology imaging modalities compared with an experienced neuroradiologist.[24] Both LLMs scored similarly at 1.75 and 1.83, respectively, out of a maximum possible 3 points, while the neuroradiologist outperformed with a score of 2.20. ChatGPT showed greater variability, suggesting room for improvement with targeted medical text training, unlike Glass AI, which has more precise training on medical literature. Furthermore, ChatGPT showed promising prospects to enhance diagnostic accuracy, streamline workflow, and improve patient care by providing evidence-based recommendations and facilitating personalized treatment strategies.[16]

Additionally, the ability of LLMs to continuously learn from new data ensures that their recommendations evolve with the dynamic landscape of VIR. The regenerative attribute of LLMs makes it amenable to keep pace with the prolific medical devices industry stocking the cath laboratories. Evidence indicates improved performance on clinical tasks when LLMs are trained on domain-specific clinical data.[9,18] Based on this, there is some optimism that LLMs trained on

Table 1 Summary of current and future uses of LLMs in VIR

S/N	Application	Specifics	References
1.	Supporting clinical-decision-making	• Analyzing medical literature, electronic health records, and patient data	23
		• Assisting in disease diagnosis, treatment planning, and prognostic predictions	24
		• Providing evidence-based recommendations and facilitating personalized treatment strategies	16
		• Continuous learning from new data for evolving recommendations in a dynamic VIR landscape	9,18
2.	Improving clinical workflow and patient scheduling	• Use in radiology report generation, reducing addendum requests and improving reporting processes	25,26
		• Intelligent patient scheduling for risk identification and preventive measures	27
		• Handling administrative duties like patient billings and extracting relevant summaries from patient records	27,28
		• Alleviating healthcare provider workload and reducing risks to patients	29
3.	Enhancing VIR education	• Assisting medical students and trainees in board-style examinations	30
		• Synergizing with attending physicians for a comprehensive learning experience	31
4.	Patient education and patient-centered care in VIR	• Simplifying medical reports for patient understanding	32
		• Providing patient education on VIR procedures with potential improvements in accuracy	33
		• Generating patient-friendly explanations of complex medical conditions, treatment options, and risks	35

Abbreviations: LLMs, large language models; VIR, vascular and interventional radiology.

VIR specific data will offer clinical decision support utility for the interventional radiologist, particularly in suggesting relevant procedures and appropriate treatment modalities including device choice and compatibility. However, it is crucial to acknowledge that the expertise of a trained interventional radiologist remains indispensable for interpreting and verifying the outputs from LLMs.

Improving Clinical Workflow and Patient Scheduling

AI has the potential to improve the interventional radiologist's daily practice in various ways. For instance, structured reporting has been shown to lead to a reduction in addendum requests for insufficient documentation, indicating a more comprehensive and clear reporting process.[25] Recognizing this unique need, recent studies have explored the use of LLMs in radiology report generation (R2Gen). R2GenGPT is an emerging innovative framework for R2Gen, which leverages LLMs for automated radiology reporting.[26] It demonstrates state-of-the-art performance and reduced computational complexity. Incorporation of LLMs similar to R2GenGPT as adjuncts for generating structured VIR report holds promise for further improving the clinical practice workflow.

Another significant aspect is intelligent patient scheduling, where AI can identify patients at high risk and take precautions to avoid potentially preventable morbidities or mortalities, while also reducing the chances of missing necessary care through smart scheduling and patient selection.[27] Though the core of these algorithms may be based on complex supervised learning models or advanced ML techniques other than LLMs, LLMs still possess the ability to be the front-end conversational interface. This interface would be able to take the output of the back-end AI models and present it in an intelligible and interactive manner.

Furthermore, LLMs can handle administrative duties such as patient billings, and extract relevant summaries from a patient's records such as problem lists, clinical notes, laboratory data, pathology reports, vital signs, prior treatments, and prior imaging reports.[27,28] These summaries provide the interventional radiologist with crucial contextual information for clinical uses. Patel and Lam[29] demonstrated the utility of ChatGPT in creating discharge summaries, allowing the clinicians to focus on more clinical commitments. Employing LLMs for laborious tasks like these also potentially reduces risks to the patient.

Enhancing VIR Education

The potential of LLMs in VIR education for medical students and trainees is promising. Recently, ChatGPT demonstrated impressive performance on a radiology-board style examination, correctly answering 69% of questions.[30] It excelled in lower-order thinking questions but faced challenges with higher-order thinking questions, particularly those related to describing imaging findings, calculations, classifications, and applying concepts. Another study compared ChatGPT-4 and Bard (developed by Google) in responding to questions from the American College of Radiology's Diagnostic Radiology In-Training (DXIT) examination. ChatGPT-4 exhibited a higher overall accuracy of 87.11% compared with Bard's 70.44%. Despite occasional failures in addressing questions accurately, the authors expressed cautious optimism, suggesting that LLMs like ChatGPT-4 could serve as valuable study tools for trainees in the future.

Consequently, the VIR-specific trained LLM can assume a crucial role in the learning curve of the VIR trainee within a learner-centered collaborative training framework. Within this framework, the LLM acts as an immediate repository, delivering a trove of updated literature, procedural guidelines, and case studies to enrich the learning experience. Proficient in evaluating lower-order thinking questions, it also becomes an invaluable tool in gauging the trainee's foundational knowledge. As the trainee confronts higher-order challenges, the attending physician and LLM synergize, addressing complexities and filling knowledge gaps. The LLM's identified limitations in imaging, procedural descriptions, and calculations are mitigated by the attending physician's expertise, creating a dynamic feedback loop for comprehensive learning. Such a personalized adaptive learning pathway aligns with the findings of Duong et al,[31] demonstrating the potential benefits of AI in achieving "precision education" within the field of radiology. The unique strengths and weaknesses of the trainee are harnessed to achieve superior learning experience facilitated by the personalized integration of LLMs thus heralding a new dawn in enhanced VIR education.

Patient Education and Patient-Centered Care in VIR

The integration of LLMs into VIR education extends beyond the training of the workforce. It holds some promise in enhancing patient-centered care through patient education as well. An exploratory case study conducted by radiologists revealed promising results in utilizing ChatGPT to simplify medical reports while maintaining factual accuracy, completeness, and patient safety.[32] Scheschenja et al[33] also explored the viability of using LLMs, specifically ChatGPT-3 and ChatGPT-4, for patient education in VIR. The authors designed hypothetical questions about common VIR procedures, comparing the accuracy of responses from the two models. While both models provided accurate information on general procedure details, preparation, risks, and postinterventional aftercare, ChatGPT-4 demonstrated better overall accuracy than ChatGPT-3 in answering questions related to Port Implantation, PTA, and TACE procedures. Recognizing the complexities associated with ensuring language clarity and response coherence, they concluded still that the LLMs exhibit feasibility for safe and relatively accurate patient education in VIR, with GPT-4 showing incremental improvements.

Other authors have reported similar results after investigating ChatGPT's performance on VIR knowledge.[34,35] McCarthy et al[35] evaluated the LLM's efficacy in delivering educational content on VIR, comparing it to standard material from the Society of Interventional Radiology Patient Care Web site. Despite occasional inaccuracies and a tendency to produce lengthy and somewhat complex responses, ChatGPT was generally deemed a reliable source for most VIR procedures.

By leveraging the capabilities of LLMs in this manner, interventional radiologists can generate patient-friendly explanations of complex medical conditions, treatment options, and potential risks associated with VIR procedures. This empowers patients with accessible information, fostering a better understanding of their health conditions and treatment plans. Moreover, the LLM's ability to produce human-like text can enhance communication between healthcare professionals and patients, fostering a more empathetic and transparent doctor–patient relationship. Hopefully, further improvements will minimize instances of incorrect information and lead to safer patient education.

Challenges, Ethics, and Recommendations in LLM Implementation for VIR

Although the incorporation of LLMs into VIR holds great promise, it is also fraught with numerous challenges and ethical considerations.[36] These include reduced human involvement, potential harm resulting from LLM reasoning weaknesses, limited availability of comprehensive datasets, the risk of biases leading to healthcare disparities, and cost constraints in low-resource settings. An overdependence on AI has the potential to diminish human involvement in decision-making processes.[37] To counteract this trend, it is essential for LLMs to augment human expertise rather than replace it, emphasizing the continued centrality of interventional radiologists.

The assessment of the generative capabilities of LLMs, particularly in the context of VIR, heavily relies on the availability of comprehensive datasets. The scarcity of medical data from VIR operating suites poses a significant obstacle to effective data collection.[38] Also, LLMs encounter challenges such as hallucinations and weak numerical reasoning.[36,39] When applied in patient care without due caution, these issues can lead to severe harm or even fatal consequences, underscoring the urgency of developing improved mitigation techniques.

To address these concerns, recent research has introduced effective strategies. These include the integration of external tools such as code interpreters, retrieval augmentation, knowledge graphs, and other mathematical tools.[40–42] These measures aim to enhance the reliability and safety of LLM applications in VIR, ensuring that they contribute positively to healthcare without compromising patient well-being.

Also, an issue of utmost concern revolves around bias and fairness, and addressing these are pivotal concerns in VIR. LLMs acquire knowledge from the data they are trained on. If this data contains inherent imbalances and biases against certain races or group of peoples, there is a risk of replicating these biases in its AI predictions.[32,43] This, in turn, could result in unfair outcomes, potentially worsening existing healthcare disparities. So, there must be emphasis on fairness-aware ML and transparent development,[44] where data are well balanced and representative of all groups, and any patient data are well protected to ensure privacy and security.

Moreover, the widespread adoption of LLMs faces challenges due to cost and resource constraints, particularly in low-resource settings.[45] Addressing this, efficient transfer models and leveraging cloud resources can enhance the accessibility of LLMs.[46] Fostering public–private collaboration can also help distribute costs and resources, facilitating broader adoption.

Given these challenges, additional research is needed to assess the performance of LLMs in clinical VIR. These investigations should be organized around the various phases of clinical interactions: preoperative, perioperative, and postoperative care. Evaluation criteria should include conventional AI metrics like specificity, sensitivity, and F1-score. Moreover, it is crucial to also employ metrics tailored to LLMs, such as BLEU, ROUGE, BERT Score, and LLM-EVAL.[47–49]

Ultimately, rigorous validation, ongoing monitoring, and collaboration with medical experts and VIR specialists are crucial on all these bases. Creating a regulatory framework that spans the multiple disciplines is crucial for the secure integration of LLM into VIR practice. This ensures transparency and accountability without impeding advancements. Addressing these challenges and ethical concerns can maximize LLMs' potential to improve healthcare outcomes.

Conclusion

The incorporation of LLMs into the domain of VIR signifies a promising frontier poised to enhance the discipline. Although the current landscape indicates that the widespread implementation of LLMs in VIR may be premature, their potential holds the promise of improving various aspects of the practice. These advanced AI tools have the capacity to improve clinical decision-making, streamline workflow, enhance medical education, and facilitate patient-centered care.

Moreover, full integration of these technologies into the clinical workflow of VIR necessitates further exploration of multi-modal AI. This involves leveraging the text and language capabilities of LLMs in conjunction with computer vision AI models, recognizing the inherently visual nature of VIR. Additional research in this direction is crucial to unlock the full spectrum of benefits and possibilities that LLMs can bring to the field.

Effective use of LLMs in VIR also requires recognizing challenges and ethical considerations, such as AI over-reliance, potential misinformation, and the need for rigorous validation. Collaboration among radiologists, AI researchers, and regulators is essential for balancing LLMs' potential with patient safety. Unlocking LLMs' full potential in VIR also requires training and refining for domain nuances, implementing robust frameworks, and adhering to ethical standards. This fosters a new era of medical practice, blending human expertise with advanced AI for patient-centered care and innovation.

Authors' Contribution
T.O.T. conceptualized the study. R.O. is the guarantor of the study. T.O.T., R.O., and A.A. were involved in

methodology. T.O.T., R.O. A.A., and C.U-O. helped in providing resources, writing—original draft preparation and editing. All authors have read and agreed to the final version of the manuscript.

Conflict of Interest
None declared.

References

1 Bohr A, Memarzadeh K. Chapter 2 - The rise of artificial intelligence in healthcare applications. In:Bohr A, Memarzadeh K, eds. Artificial Intelligence in Healthcare: Academic Press; 2020: 25–60
2 Parampreet K, Aaron Alexander M, Naitik P, et al. Unlocking the potential of artificial intelligence (AI) for healthcare. In: Stanislaw PS, ed. Artificial Intelligence in Medicine and Surgery - An Exploration of Current Trends, Potential Opportunities, and Evolving Threats. Rijeka: IntechOpen; 2023:Ch. 3
3 Frandon J, Beregi J-P. Special issue: present and future perspectives of vascular interventional radiology. J Pers Med 2023;13 (07):1131
4 Seah J, Boeken T, Sapoval M, Goh GS. Prime time for artificial intelligence in interventional radiology. Cardiovasc Intervent Radiol 2022;45(03):283–289
5 Waller J, O'Connor A, Rafaat E, et al. Applications and challenges of artificial intelligence in diagnostic and interventional radiology. Pol J Radiol 2022;87(01):e113–e117
6 Weiss CR, Hafezi-Nejad N. Interventional radiology: past, present, and future. Radiology 2023;308(01):e230809
7 Charalel RA, McGinty G, Brant-Zawadzki M, et al. Interventional radiology delivers high-value health care and is an imaging 3.0 vanguard. J Am Coll Radiol 2015;12(05):501–506
8 Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. Adv Neural Inf Process Syst 2020;33:1877–1901
9 Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med 2023;29 (08):1930–1940
10 Mallio CA, Sertorio AC, Bernetti C, Beomonte Zobel B. Large language models for structured reporting in radiology: performance of GPT-4, ChatGPT-3.5, Perplexity and Bing. Radiol Med (Torino) 2023;128(07):808–812
11 Browning J, LeCun Y. Language, common sense, and the Winograd schema challenge. Artif Intell 2023;325:104031
12 Biderman S, Schoelkopf H, Anthony QG, et al , Eds. Pythia: A suite for analyzing large language models across training and scaling. International Conference on Machine Learning. 2023: PMLR.
13 Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. J Mach Learn Res 2020;21(01):5485–5551
14 Touvron H, Lavril T, Izacard G, et al. Llama: open and efficient foundation language models. arXiv preprint arXiv:230213971. 2023
15 Open LLMs [Internet]. 2023. Accessed February 17, 2024 at: https://github.com/eugeneyan/open-llms
16 Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023;11(06):887
17 Eggmann F, Weiger R, Zitzmann NU, Blatz MB. Implications of large language models such as ChatGPT for dental medicine. J Esthet Restor Dent 2023;35(07):1098–1102(n/a)
18 Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023;2(02):e0000198
19 Nori H, King N, McKinney SM, Carignan D, Horvitz E. Capabilities of gpt-4 on medical challenge problems. arXiv preprint arXiv:230313375. 2023
20 Truhn D, Weber CD, Braun BJ, et al. A pilot study on the efficacy of GPT-4 in providing orthopedic treatment recommendations from MRI reports. Sci Rep 2023;13(01):20159
21 Yan A, McAuley J, Lu X, et al. RadBERT: adapting transformer-based language models to radiology. Radiol Artif Intell 2022;4 (04):e210258
22 Webb T, Holyoak KJ, Lu H. Emergent analogical reasoning in large language models. Nat Hum Behav 2023;7(09):1526–1541
23 Shen Y, Heacock L, Elias J, et al. ChatGPT and other large language models are double-edged swords. Radiology 2023;307(02): e230163
24 Nazario-Johnson L, Zaki HA, Tung GA. Use of large language models to predict neuroimaging. J Am Coll Radiol 2023;20(10): 1004–1009
25 Nguyen Q, Sarwar A, Luo M, Berkowitz S, Ahmed M, Brook OR. Structured reporting of IR procedures: effect on report compliance, accuracy, and satisfaction. J Vasc Interv Radiol 2018;29(03): 345–352
26 Wang Z, Liu L, Wang L, Zhou L. R2GenGPT: Radiology Report Generation with frozen LLMs. Meta-Radiology. 2023;1(3):100033
27 Gurgitano M, Angileri SA, Rodà GM, et al. Interventional Radiology ex-machina: impact of artificial intelligence on practice. Radiol Med (Torino) 2021;126(07):998–1006
28 Zheng Y, Wang L, Feng B, Zhao A, Wu Y. Innovating healthcare: the Role of ChatGPT in streamlining hospital workflow in the future. Ann Biomed Eng 2023
29 Patel SB, Lam K. ChatGPT: the future of discharge summaries? Lancet Digit Health 2023;5(03):e107–e108
30 Bhayana R, Krishna S, Bleakney RR. Performance of ChatGPT on a radiology board-style examination: insights into current strengths and limitations. Radiology 2023;307(05):e230582
31 Duong MT, Rauschecker AM, Rudie JD, et al. Artificial intelligence for precision education in radiology. Br J Radiol 2019;92(1103): 20190389
32 Jeblick K, Schachtner B, Dexl J, et al. ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. arXiv preprint arXiv:221214882. 2022
33 Scheschenja M, Viniol S, Bastian MB, Wessendorf J, König AM, Mahnken AH. Feasibility of GPT-3 and GPT-4 for in-depth patient education prior to interventional radiological procedures: a comparative analysis. Cardiovasc Intervent Radiol 2024;47(02): 245–250
34 Barat M, Soyer P, Dohan A. Appropriateness of recommendations provided by ChatGPT to interventional radiologists. Can Assoc Radiol J 2023;74(04):758–763
35 McCarthy CJ, Berkowitz S, Ramalingam V, Ahmed M. Evaluation of an artificial intelligence chatbot for delivery of IR patient education material: a comparison with societal website content. J Vasc Interv Radiol 2023;34(10):1760–1768.e32
36 Javan R, Kim T, Mostaghni N, Sarin S. ChatGPT's potential role in interventional radiology. Cardiovasc Intervent Radiol 2023;46 (06):821–822
37 Huang S-C, Chaudhari AS, Langlotz CP, Shah N, Yeung S, Lungren MP. Developing medical imaging AI for emerging infectious diseases. Nat Commun 2022;13(01):7060
38 Demir KC, May M, Schmid A, et al , Eds. PoCaP Corpus: A Multimodal Dataset for Smart Operating Room Speech Assistant Using Interventional Radiology Workflow Analysis. Cham: Springer International Publishing; 2022
39 De Angelis L, Baglivo F, Arzilli G, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front Public Health 2023;11:1166120
40 Gao L, Madaan A, Zhou S, et al. PAL: Program-aided language models. In:Andreas K, Emma B, Kyunghyun C, Barbara E, Sivan S, Jonathan S, eds. Proceedings of the 40th International Conference on Machine Learning; Proceedings of Machine Learning Research. : PMLR; 2023:10764–99

41 Yao S, Zhao J, Yu D, et al. React: Synergizing reasoning and acting in language models. arXiv preprint arXiv:221003629. 2022
42 Agrawal G, Kumarage T, Alghami Z, Liu H. Can knowledge graphs reduce hallucinations in LLMs?: a survey arXiv preprint arXiv:231107914. 2023
43 Yang J, Li HB, Wei D. The impact of ChatGPT and LLMs on medical imaging stakeholders: perspectives and use cases. Meta-Radiology. 2023;1(01):100007
44 Xu J, Xiao Y, Wang WH, et al. Algorithmic fairness in computational medicine. EBioMedicine 2022;84:104250
45 Köpf A, Kilcher Y, von Rütte D, et al. OpenAssistant Conversations–Democratizing Large Language Model Alignment. arXiv preprint arXiv:230407327. 2023
46 Niu S, Liu Y, Wang J, Song H. A decade survey of transfer learning (2010–2020). IEEE Trans Artif Intell 2020;1(02):151–166
47 Chan Y-H, Fan Y-C, Eds. A recurrent BERT-based model for question generation. Proceedings of the 2nd workshop on machine reading for question answering. Hong Kong, China: Association for Computational Linguistics;2019
48 Lin Y-T, Chen Y-N. LLM-Eval: unified multi-dimensional automatic evaluation for open-domain conversations with large language models. arXiv preprint arXiv:230513711. 2023
49 Bandi A, Adapa PVSR, Kuchi YEVPK. The power of generative AI: a review of requirements, models, input–output formats, evaluation metrics, and challenges. Future Internet 2023;15(08):260

