Applications of Generative AI in Healthcare: algorithmic, ethical, legal and societal considerations

Onyekachukwu R. Okonji¹, Kamol Yunusov¹, and Bonnie Gordon¹

¹Affiliation not available

May 09, 2024

Abstract

Generative AI is rapidly transforming medical imaging and text analysis, offering immense potential for enhanced diagnosis and personalized care. However, this transformative technology raises crucial ethical, societal, and legal questions. This paper delves into these complexities, examining issues of accuracy, informed consent, data privacy, and algorithmic limitations in the context of generative AI’s application to medical imaging and text. We explore the legal landscape surrounding liability and accountability, emphasizing the need for robust regulatory frameworks. Furthermore, we dissect the algorithmic challenges, including data biases, model limitations, and workflow integration. By critically analyzing these challenges and proposing responsible solutions, we aim to foster a roadmap for ethical and responsible implementation of generative AI in healthcare, ensuring its transformative potential serves humanity with utmost care and precision.

Generative AI, medical imaging, text analysis, ethics, informed consent, data privacy, algorithmic bias, model limitations, workflow integration, healthcare, regulation, liability, accountability.

Introduction

The landscape of medical diagnosis is undergoing a seismic shift. Generative AI, with its ability to analyze images and craft insightful text reports, holds immense promise for revolutionizing healthcare. Yet, beneath the surface of this technological marvel lie ethical, societal, and legal complexities that demand careful scrutiny.

This report embarks on a critical exploration of this transformative technology. We delve into the ethical considerations that underpin its application in medical imaging and text analysis, examining the delicate balance between accuracy, informed consent, and data privacy. We navigate the uncharted legal landscape, dissecting the challenges of liability and accountability in the face of AI-driven diagnoses. Finally, we turn a discerning lens to the algorithmic intricacies, highlighting concerns around data biases, model limitations, and the need for seamless workflow integration.

This is not merely an academic exercise. It is a call for responsible innovation, a plea to harness the power of AI while safeguarding the well-being of patients. By dissecting the complexities that lie at the heart of generative AI in medicine, we aim to pave the way for its ethical and responsible implementation, ensuring that this transformative technology serves humanity with the utmost care and precision.
Selection of application

This report will be looking at 4 critical aspects related to the application of Generative AI models in healthcare with a focus on medical imaging and text.

Justification of application

Generative AI’s burgeoning influence in medical imaging and text analysis demands a rigorous exploration of its ethical, societal, and legal implications. This report delves into the intricacies of this potent technology, dissecting its promising potential alongside the complexities that necessitate careful consideration. Ethical questions weave throughout the fabric of generative AI in medicine. Accuracy, a cornerstone of trust, necessitates rigorous testing and transparency about performance limitations. Informed consent becomes paramount when AI generates potentially life-altering reports, demanding clear communication about its role and limitations. Accountability rests not only with developers but also with healthcare institutions wielding this tool, requiring robust audit trails and clear channels for redressal. Data privacy, ever-sensitive in healthcare, demands stringent safeguards against unauthorized access and secondary use.

Societal considerations matter too. Will generative AI exacerbate existing healthcare disparities? How can we ensure equitable access to this technology for all? What are the public’s concerns about AI in healthcare? How can we build trust and address ethical concerns to ensure widespread adoption? How will generative AI impact healthcare professionals? What are the implications for job displacement and the need for retraining and reskilling?

Legally, Generative AI navigates uncharted waters. While existing regulations apply to data privacy and medical practice, specific frameworks for AI-generated diagnoses and reports are nascent. Issues of negligence and liability in the event of AI-based errors will require careful legal dissection, balancing accountability with fostering innovation. Ensuring transparency and explainability of AI algorithms becomes crucial for informed legal decisions and building public trust.

Algorithmic considerations cast an introspective lens on generative AI’s limitations. Concerns about insufficient training data loom large, as biases within this data can lead to discriminatory outputs. Hallucinations and inaccuracies in AI-generated images or text can have dire consequences, necessitating stringent filters and human oversight. Training models and running complex algorithms necessitate substantial computational resources, posing challenges for wider adoption in resource-constrained settings. Moreover, integrating AI seamlessly into existing clinical workflows requires careful consideration, ensuring minimal disruption and smooth human-machine collaboration. The real-time nature of medical decision-making further complicates matters, demanding swift AI inference and processing without compromising accuracy.

This report aims to untangle this intricately woven tapestry of ethical, societal, and legal threads, illuminating both the transformative potential and the challenges inherent in generative AI’s application to medical imaging and text. By fostering a comprehensive understanding of these complexities, we can pave the way for responsible innovation, harnessing AI’s power to serve humanity while safeguarding the well-being of patients and upholding the highest ethical standards.

Analysis

Ethical aspects

Artificial Intelligence (AI) has revolutionized the field of medical imaging, bringing forth numerous advancements and opportunities. However, with these advancements come important ethical considerations that must be addressed. This report explores the ethical aspects of AI in medical imaging, focusing on privacy,
bias, and accountability. AI has the potential to improve medical imaging by making them more efficient, affordable, and widely available. (AP & Neri E. Artificial intelligence in radiology: Ethical Considerations. Diagnostics Basel, 2020) With the use of Generative AI in medical imaging there has been an increase in the precision on diagnosis, treatment care plans and monitoring of numerous diseases. The use of specific techniques such as generative adversarial networks (GANs) and variational autoencoders (VAEs) strengthen medical imaging by generating synthetic images and improving reconstruction. (G et al., 2021) However, this is only going to be feasible if patients and medical professionals agree that AI medical devices (AIMDs) are reliable and trustworthy (G & Hawk KE, 2020). Ethical considerations must be made when using LLMs within medical radiography. While these models yield promising results, there are elements that must be considered when using these models.

Privacy

AI in medical imaging relies heavily on patient data, which raises concerns about privacy and data security. It is crucial to ensure that patient information is protected and used responsibly. As these models rely extensively on large datasets such as medical records, genetics and ethnical information, stricter regulations and robust security measures should be implemented to safeguard patient privacy and prevent unauthorized access or misuse of data (G et al., 2021). Severe consequences of unauthorised use or breaches within the data security can result in identity theft, privacy violations and misuse of personal information.

Bias

AI algorithms used in medical imaging have the potential to introduce biases that may impact healthcare outcomes. It is essential to address and mitigate these biases to ensure fairness and accuracy in diagnosis and treatment. Transparency in algorithm development, diverse training datasets, and continuous monitoring can help identify and rectify biases in AI systems.

Transparency and Accountability

As AI systems become more integrated into medical imaging, questions of accountability arise. Who is responsible when an AI system makes a wrong diagnosis or misses a critical finding? Clear guidelines and mechanisms should be established to assign accountability and address any errors or adverse outcomes resulting from AI-assisted diagnosis (G & Hawk KE, 2020). Additionally, healthcare professionals should be trained to understand the limitations and potential biases of AI systems.

Autonomy of use

Some concerns have stemmed from the impact on patient-doctor rapport and patient autonomy. Whilst LLMs provide effective insights and support to healthcare professionals, the over-reliance on the AI system can cause a wedge between patients trusting their care providers and lack of human touch. Patient relationship is built on empathy, care and understanding so to maintain such level of care, healthcare professionals need to use LLMs simply as a tool in supporting their decisions or care plan with the involvement of the patient. (JR et al., 2019)

Societal aspects

The advent of artificial intelligence (AI) technologies has revolutionized many sectors, including healthcare. Particularly, as mentioned, AI can be instrumental in medical imaging, enabling precision and accuracy in detecting and diagnosing diseases. However, this increased reliance on AI may impact the nature of social interaction in healthcare.
Human touch

Healthcare has traditionally been a high-touch field, with human interaction playing a critical role in patient care. The introduction of AI, while augmenting diagnostic efficiency, could potentially alter this dynamic. It may reduce direct interaction between healthcare providers and patients, as AI takes over tasks such as image interpretation and diagnosis. This could lead to a shift in the patient-provider relationship, posing questions about the importance of human touch and interaction in the healing process. (Blease et al., 2019) According to a study by Huang and Liang (2020), the potential lack of human touch in AI-based services may affect patient satisfaction and the perceived quality of care. (Huang et al., 2020) They argue that while AI improves efficiency, it may not be able to replicate the empathy, reassurance, and personal care that human healthcare providers offer. Patients often derive comfort from the human interaction they receive during their healthcare experience, which AI, at its current state, cannot provide. This lack of ‘human touch’ could lead to a perceived decrease in the quality of care and may affect patient satisfaction. The implementation of AI in healthcare settings also brings forth complex societal issues that need addressing. These issues span the spectrum from ethical considerations, such as informed consent and privacy concerns, all of which will be discussed in this paper, to social implications, like potential job displacement and the digital divide. (BD et al., 2053)

Job displacement

From a societal perspective, the increased use of AI could potentially displace jobs, particularly those that involve routine tasks. AI can automate certain tasks, freeing up radiologists and other imaging professionals to focus on areas requiring human expertise. For example, AI algorithms can be trained to accurately identify abnormalities in imaging scans, thus improving efficiency and reducing the workload of clinicians. (Liu & others, 2019) However, AI is also likely to create new roles within the field. The design, development, deployment and maintenance of AI systems will require new skills, leading to the creation of jobs such as AI specialists and data scientists. (Davenport & R Kalakota, 2019) This necessitates a shift in the education and training of healthcare workers. Reskilling will be key in helping these professionals adapt to the changing landscape. Medical education will need to incorporate AI literacy, including understanding the strengths and limitations of AI, data interpretation, and ethical considerations. (E Topol, 2019) Therefore, continuous professional development programs will need to focus on skill acquisition in AI-related areas. Furthermore, interdisciplinary collaboration between medical professionals, computer scientists, and engineers will become increasingly important. (Naylor & others, 2018) Efforts should also focus on developing AI applications that complement, rather than replace, the human aspects of healthcare. This could involve creating AI systems that still involve human healthcare providers in the decision-making process, or developing AI tools that assist healthcare providers in delivering more personalised care. (E Topol, 2019)

Digital divide

Moreover, there’s a risk of exacerbating the digital divide, as access to AI-powered healthcare may be limited to certain segments of society, potentially widening health disparities. It is accepted that AI algorithms can help analyse medical images more efficiently and accurately, potentially reducing diagnostic errors, enhancing patient care, and improving health outcomes. (E Topol, 2019) However, if not properly managed, AI could also exacerbate health disparities. It’s important to note that AI algorithms are often trained on datasets that might not be representative of diverse populations. If these datasets predominantly include data from certain groups and exclude others, this could lead to biased outcomes and widen existing health disparities. (Gianfrancesco & others, 2020) Additionally, AI can potentially improve access to care in underserved areas by allowing for remote diagnostics and consultations, thus overcoming geographical barriers. (Meskó & others, 2018) However, this opportunity also presents a risk: if AI advancements are primarily accessible to affluent communities due to cost or infrastructure requirements, it could increase health inequalities. Addressing health disparities in AI requires a careful consideration of legal frameworks. Issues such as privacy, data security, and informed consent are key. (W N Price & Bias’, 2019) Furthermore, regulatory mechanisms...
must ensure the quality and safety of AI systems and their algorithms. Anti-discrimination laws must also be enforced to prevent biased data collection and usage. (B et al., 2017)

Public trust

Lastly, while social aspect of AI is analysed, consideration of public trust is a crucial factor for the adoption of AI in healthcare, including in the realm of medical imaging. AI can potentially revolutionize healthcare by making diagnoses more accurate and efficient, but this relies on public acceptance and trust. (Minssen & Bias’, 2020) With the increasing use of AI in medical imaging, concerns about privacy, reliability, and fairness are paramount. (Lerberghe, 2008)

Public trust is the bedrock upon which the successful implementation and acceptance of AI applications in healthcare are built. (Amann, 2021) Transparency about how AI is used in medical imaging is vital to gain and maintain public trust. This involves clear communication about the AI decision-making process, its accuracy, and potential limitations. (n1, n.d.) Regulations that protect patient interests are also important. This includes laws to safeguard patient data, ensure the reliability and validity of AI algorithms, and uphold the principles of fairness and equity. (Haas & Bias’, 2020) Strong regulations thus serve as a guarantee to the public that their interests are protected, further fostering trust. (n1, n.d.) Societal frameworks play a fundamental role in fostering public trust in AI use in healthcare. These frameworks include educational initiatives to improve public understanding of AI, and mechanisms for public input and feedback on AI use. (Amann, 2021) Complex societal issues such as inequality and bias in AI algorithms must also be addressed. For instance, AI systems trained on data from one demographic might not work as well for another, leading to disparities in healthcare outcomes. (Haas & Bias’, 2020) Policies should be in place to ensure diversity in training data and to test AI systems across different groups. (n1, n.d.)

Legal aspects

Accuracy

Although AI is considered to have the potential to improve medical efficiency and precision, especially in medical imaging, its reliance on the quality of training data can lead to diagnostic errors, which could potentially harm patients. In conventional practices, incorrect diagnoses can result in legal liabilities, with the responsibility usually resting on the healthcare providers. However, in the context of AI, an important question emerges: “Who is accountable for AI negligence - the healthcare providers, the developers, or the institutions using AI?” (2020) Considering AI’s self-governing decision-making and learning capabilities could reshape conventional legal notions of liability and accountability, it is debated that distinctive legal frameworks need to be established. (Brownword R et al., 2019) From a legal viewpoint, accuracy isn’t solely about a correct diagnosis; it also relates to the rights and dignity of those availing these services.

Informed consent

One potential solution to eliminate this issue is to provide patients with an informed choice, making them aware of potential risks and benefits of proposed medical procedures. (NC & O’Neill O, 2007) Based on this informed consent, patients have the autonomy to make their own decisions, a principle that is legally mandated in healthcare systems globally. (Beauchamp TL & present challenges’ 20(4) Cambridge Quarterly of Healthcare Ethics 515.-523 2017, 2017) The significance of informed consent escalates when patient data is processed by AI. (2020) Since decisions or recommendations made by AI systems based on this data can significantly impact patient health, underscoring the critical nature of informed consent. (Char DS & Magnus D, 2018) However, the concept of informed consent poses a considerable challenge due to the complexities of understanding potential risks and benefits of AI integration in their medical procedures and the “decision making” process of AI. This lack of understanding could potentially undermine the legitimacy of informed consent. (Schönberger D & 2020, 2020) Understanding and defining the “decision making” process also brings forth complex legal concerns. This requires addressing issues such as data protection, transparency, and
accountability.(2017) However, the "black box" nature of AI, where decision-making processes are not easily understood, adds another layer of complexity.(M & Bryson J, 6298) Therefore, incorporating data protection, transparency, and accountability elements and adapting to evolving legal frameworks are considered crucial for understanding the "decision making" process of AI.(ME et al., 2020)

Transparency & Explainability

It’s essential to emphasize the significance of transparency and explainability in decisions made by AI, especially in areas like medical imaging where AI algorithms interpret intricate images.(B et al., 2017) Explainability allows healthcare professionals to assess AI’s decisions, enhancing confidence and fostering a better collaboration between AI systems, specialists,(R et al., 2015) and patients who are directly affected by these decisions.(Rajkomar A & Kohane I, 1347) Also, explainability empowers patients to make well-informed choices regarding their healthcare in AI-based diagnoses and treatments.(A & others, 1712) Similarly, transparency enables healthcare practitioners to understand specific diagnoses or treatment prescriptions made by AI, as well as to audit AI systems to ensure they are functioning as intended, thus maintaining accountability.(n17), n.d. Legally, the matters of transparency and explainability in healthcare are intricate and continuously evolving. This complexity stems from legal concerns such as liability, accountability, and patient rights associated with the medical imaging sector. The primary challenge is to determine who to hold accountable if an AI system provides an incorrect diagnosis or treatment recommendation: the healthcare provider, the AI developer, or the AI itself. The lack of specific laws and regulations governing AI in healthcare further complicates these legal issues.(2020) Therefore, it is crucial to ensure transparency in the development and application of AI algorithms to identify and mitigate potential risks or errors.(B et al., 2017)

Accountability

The concept of defining responsibility within the sphere of Artificial Intelligence (AI) presents significant challenges due to the inherent intricacies and lack of transparency of AI systems.(II & Cohen IG, 2019) To answer the question, "Where does the responsibility for AI-induced errors fall?" a thorough evaluation of legal liability, and notably, the aspect of responsibility, is necessitated. This is crucial as it underlines the demand for accountability and guarantees that appropriate measures are implemented in the event of damages.(Floridi & the governance of the digital. Philosophy and Technology; 2019) From a legal viewpoint, it could be proposed that developers who design the system in a manner that results in harm or fail to ensure its safety should bear the legal responsibility. Conversely, if users handle the AI system recklessly or improperly, they could be held accountable for negligence. However, if the AI system exhibits autonomous decision-making capacity and inflicts harm through its activities, it could be held responsible for those actions.(Calo & roadmap. Policy and Internet, 2017) The next logical question is how causation can be attributed to the non-human AI system. Firstly, it is essential to thoroughly comprehend the critical factors and processes that led to a particular outcome. Stakeholders should be able to determine whether the AI system’s actions were intentional, accidental, or the result of concealed implicit biases or flaws, a concept closely tied to the principles of transparency and explainability. Secondly, the technical challenge lies in identifying a single user responsible for the damage incurred since algorithms, data sources, and human input form an integral part of AI systems. This complexity raises the question of the extent of responsibility that should be assigned to developers, users, or even the AI system itself.(II & Cohen IG, 2019)

Data privacy & protection

In the domain of medical imaging, artificial intelligence (AI) methodologies frequently require access to a broad range of patient information, encompassing both medical imagery and related health records. This becomes particularly essential when AI systems scrutinize and learn from considerable quantities of confidential patient data. Thus, the elements of data privacy and data protection within legal considerations demand that data collection, storage, and usage are carried out in compliance with the relevant data protection laws.(El Emam K & Arbuckle L, 2807) The transfer and storage of sensitive patient information during
the usage of AI in medical imaging amplify the threat of data breaches. (Abouelmehdi et al., 2018) And any unauthorized intrusion into such data can instigate harm and raise privacy issues. (V Dignum & Use AI in a Responsible Way', 2019) Furthermore, the evaluation of medical imaging data by AI might risk patient re-identification, notwithstanding the confidentiality measures in place. (2019) Advanced techniques might enable de-identified data to be associated with other publicly accessible information, thereby endangering patient privacy. (El Emam K & Arbuckle L, 2807) Also, AI algorithms employed in medical imaging may harbour biases, (6464) potentially leading to inconsistencies in patient care, erroneous diagnoses, or unequal treatment. (Rajkomar A & Kohane I, 1347) This practice not only infringes on privacy by amplifying existing biases but also jeopardizes the confidentiality of certain groups. (6464)

**Intellectual property rights**

As previously highlighted, Intellectual Property Rights (IPR) can indeed intersect with, and occasionally conflict with, the need for transparency in legal aspects. Nevertheless, this area presents its unique set of challenges and plays a pivotal role in the advancement of novel algorithms and AI technologies in healthcare. These rights offer legal safeguards to the originators of such innovations, thus fostering innovation and investment in the field. By gaining exclusive rights to their creations, inventors may disclose and commercialise their inventions, which in turn propels progress in medical imaging and other healthcare applications. (2019)

The matter becomes considerably more intricate when new algorithms and AI technologies are conceived by an AI without direct human intervention. Does this imply that the possession and entitlement to intellectual property are transferred to the AI? Conventionally, legal frameworks predominantly acknowledge human inventors as the legitimate owners of intellectual property. However, when it comes to AI-generated innovations, the issues of authorship and ownership are not straightforward. Some scholars have observed that resolving this dilemma necessitates meticulous scrutiny of existing IP laws, along with potential revisions to accommodate the distinct nature of AI-generated inventions. (Pasquale & machine learning for the law. Harvard Journal of Law and Technology, 2019) From a legal standpoint, if the allocation of ownership and entitlement to intellectual property is indeterminate, a significant challenge arises in establishing liability for AI-generated inventions and their potential infringement upon existing patents. Moreover, issues emerge concerning the extent of human participation in AI-generated inventions, the repercussions on conventional notions of creativity and inventiveness, and the demand for transparency and accountability in dealing with autonomous AI systems. (Ohly & the question of originality in the European Union. European Intellectual Property Review, 2018)

The development of comprehensive legal frameworks that tackle these complexities is pivotal to stimulating innovation while safeguarding intellectual property rights. (Geiger et al., 2018) However, it’s debatable how far the fundamental principles of intellectual property law and the intersection of privacy and intellectual property can accommodate these complexities. Some scholars propose that both human and AI contributions should be considered, along with the threshold of originality, as an approach to ascertain originality and copyright protection for AI-generated content. (Van Overwalle et al., 2009) Others recommend a re-evaluation of the core principles and notions of intellectual property law, (Dinwoodie, 2012) or suggest viewing privacy itself as a form of intellectual property. (2011)

Overall, these discussions emphasize the intricacy of determining originality and copyright protection in AI-produced works, along with the implications of AI on legal aspects of intellectual property.

**Regulatory compliance**

The final aspect of legal considerations surrounding medical imaging pertains to the assessment of existing regulations in the UK and the element of regulatory compliance. Regulated and compliant AI systems is integral to ensuring safe and effective patient care, upholding ethical standards, and maintaining public trust. Therefore, aligning AI systems with existing regulations such as the Data Protection Act 2018 is of paramount importance. However, the swift evolution of AI technologies could potentially outpace these regulatory measures, necessitating continual updates to the legal framework. (Wachter S et al., 2017) There are numerous other measures implemented to address transparency and explainability in AI, such as published guidelines and frameworks like the "Code of Conduct for Data-Driven Health and Care Technology"
and the "NHS AI Lab AI Ethics Framework", (of Health et al., 2019) are viewed as significant steps in addressing related concerns. These measures are perceived to serve as tools for transparency, explainability, and accountability, thereby maintaining public trust and ethical standards. (2019) Meanwhile, the Medicines and Healthcare Products Regulatory Agency (MHRA) supervises the regulation of AI and medical devices, including software. But there is an ongoing discourse regarding the efficiency of these regulations in keeping pace with the rapid advancement of AI and the complex issues it introduces, (Cohen et al., 2020)

To tackle issues of accountability, the UK government has instituted the Centre for Data Ethics and Innovation (CDEI) to offer guidance on the ethical application of AI, (Wachter S et al., 2017) and the protection of individuals’ data rights. (2010) The Information Commissioner’s Office (ICO)(intelligence & intellectual property 2021, 2021) is another regulatory body tasked with enforcing data protection laws within the UK. They dispense guidance on data protection practices and probe into any instances of breaches or non-compliance. (2015) In contrast, the Intellectual Property Office (IPO) offers guidance concerning the patentability of AI inventions and the ownership of AI-produced works. Moreover, The Copyright, Designs and Patents Act 1988 has been revised to elucidate aspects related to AI-generated works. (Copyright & Patents Act 1988, 1988) Nevertheless, the effectiveness and efficiency of these laws in dealing with the intricacies of AI and intellectual property necessitate continual appraisal and potential modification to keep abreast of technological progressions and legal challenges. (2020; 2017; Kosta et al., 1448; Laurie & Sethi, 2018)

**Algorithmic and Technical aspects**

With respect to algorithmic aspects of the application of Generative AI in Healthcare, there are several ways to look at it from. This report looks at the subject from 6 angles namely:

- Unavailability of sufficient data for training.
- Issues with accuracy and hallucinations.
- Model training and compute limitations.
- Issues with interpretability and explainability.
- Issues with workflow integration.
- The need for real-time inference and processing.

**Unavailability of sufficient data for training**

The healthcare domain faces unique challenges in applying generative AI models due to limited availability of data, sensitive patient information, and scarcity of well-posed discriminative tasks. Additionally, the datasets used for benchmarking often fail to represent real-world data due to spectrum bias and measurement or imaging device discrepancies. To address these challenges, the AI for Health Imaging Initiative (AI4HI)(2023) was formed, bringing together five EU projects to develop Big Data infrastructure while adhering to GDPR.

Individually, CHAIMELEON collects data in local or central databases, EuCaNImage builds a federated cancer imaging platform, INCISIVE aims for a pan-European repository of multi-modal data, ProCaNcer-1 focuses on prostate cancer imaging, and PRIMAGE develops a cloud-based platform for pediatric cancer decision-making. Additionally, OpenAI has announced partnerships to produce public and private datasets for enhanced AI training.

**Issues with accuracy and hallucination**

Hallucination refers to the generation by a LLM seemingly correct sentences by virtue of its innate probabilistic generation, which aren’t entirely based on facts and can be largely incorrect. This can lead to inaccurate and misleading outputs that could have serious consequences in healthcare. There are several sources of hallucinations in LLMs, some of which includes:
- Unreliable sources of information: The training data for LLMs can be unreliable, such as if it contains false or inaccurate information.
- Probabilistic generation: LLMs generate text based on probabilities, which means that there is a chance that they will generate incorrect or misleading content.
- Biased training data: If the training data is biased, the model will be more likely to generate biased content.
- Insufficient context: LLMs may not be able to understand the context of a query, which can lead to inaccurate or misleading responses.

To measure and prevent hallucinations in, improve accuracy and foster workflow inclusion of LLMs, the following strategies are currently being used:

- Med-HALT(2023): A medical dataset framework designed to measure and evaluate hallucinations in LLMs.
- Human evaluation(2023): Having human experts evaluate the output of LLMs.
- Metrics(2023): Using metrics such as perplexity and cross-entropy to assess the quality of LLM output.
- Human-in-the-loop: Having humans involved in the development and deployment of LLMs.
- Algorithmic corrections: Using machine learning techniques to correct hallucinations.
- Fine-tuning GPTs on healthcare data: Training GPT models on healthcare data.
- Improving prompts: Using prompts that encourage the LLM to generate accurate and relevant content. Recently, OpenAI released this guide(OpenAI. Prompt Engineering, 2023) on prompt engineering.
- Instruction fine-tuning, as was reasonably effectively used in developing Google’s QA LLM, Flan-PaLM(2210) which achieved state-of-the-art performance in Medical Question and Answering tasks.
- The introduction of more robust medical datasets e.g. MultiMedQA, a new dataset which comprises 6 popular medical QA datasets (MedQA(Jin et al., 2009), MedMCQA(2203), PubMedQA(1909), LiveQA(2010), MedicationQA(Aug 21;264:25-29. Doi: 10.3233/shti19 Pmid: 31437878, 2019), MMLU Clinical Topics(Hendrycks, 8550) and finally the newer HealthSearchQA(K. Singhal et al. Large Language Models Encode Clinical Knowledge, 2212))
- Chain of prompt prompting(2023): Using a series of prompts to help the LLM stay on track and avoid generating hallucinations. Commonly used alongside this in recent state-of-the-art Medical LLMs is self-consistency(2023).
- Retrieval-augmented generation (RAG)(K. Martineau, 2023): Using vector databases to provide the LLM with additional context.

Model training and compute

There are challenges of training and deploying large language models (LLMs) in healthcare. One of the major challenges is the high cost of training LLMs, which can be prohibitive for healthcare facilities. The cost of training LLMs is due to the need for powerful hardware, such as GPUs, CPUs, and RAM storage. Additionally, data gathering and annotation is a time-consuming and labor-intensive process. To address these challenges, researchers have developed techniques such as model, data, and tensor parallelism to reduce the training time and optimize the compute requirements of LLMs. However, these techniques require effective communication between GPU nodes and high-speed internet connectivity, which may not be available to all healthcare facilities. Retrieval Augmented Generation (RAG), which is a technique that can be used to improve the performance of LLMs without the need for extensive training involving the use of vector databases to provide LLMs with additional context, can help them to generate more accurate and relevant
responses could be a valuable tool for healthcare facilities that are looking to use LLMs without the high cost of training them. However, there is not yet much research into the application of RAG in healthcare.

**Issues with interpretability and explainability**

In 2021, a study led by Michael Roberts et al. (2021) highlighted issues in machine learning (ML) algorithms for COVID-19 detection from chest radiographs and CT scans. Algorithms wrongly associated disease severity with imaging views rather than features, hindering early-stage disease detection. In healthcare, minimizing wrong predictions and ensuring interpretable ML models is crucial. Vera Liao et al. research (2023) on Large Language Models (LLMs) identified challenges in transparency due to complex architectures and proprietary technologies. Addressing interpretability, existing approaches include model reporting, evaluation result publication, explanation provision, and communicating uncertainty. Hallucinations in LLMs pose challenges, with research suggesting issues in training datasets and model reliance on memorization instead of robust learning. Open-source tools like EvidentlyAI (EvidentlyAI, 2023) and Arize (Arize, 2023) aim to enhance LLM observability and interpretability by modifying prompts, tracking data changes, and fine-tuning models. However, the effectiveness of these strategies in improving trust and LLM adoption in healthcare requires further examination.

**Issues with workflow integration**

In addition to aforementioned issues such as hallucinations, inaccuracies and lack of transparency, healthcare practitioners (HCPs) draw insights from a diverse array of data sources, encompassing patient records, prescriptions, diagnostic images, and test results as such, single-mode LLMs i.e. those that deal with only text input, tend to have limited functionality in clinical settings for example. Recently, there has been the growing development of multi-modal LMMs comprising input from images/videos with text and/or sound input, the goal being to facilitate workflow integration and simulate the activities of physicians and other healthcare practitioners when making clinical decisions. Projections from a comprehensive EIT Health and McKinsey report (2020) anticipate the automation of approximately 15% of healthcare workflows by 2030, potentially streamlining 20% to 8% of tasks performed by physicians and nurses through AI. Despite the considerable potential of AI in healthcare, its widespread adoption remains limited, with 44% of European healthcare professionals reporting minimal involvement in AI development and implementation. In this pursuit of broader AI integration in healthcare, there is an immediate need for innovative applications. For instance, the successful integration of Electronic Medical Records (EMR) (2022), as highlighted in the 2022 Brookings report, underscores the pivotal role of innovation in software systems. Among the persisting challenges, the lackluster impact of existing AI models on workflow optimization is a notable concern for healthcare practitioners. Unimodal AI models, primarily employed in healthcare, particularly in radiology, fall short of significantly enhancing workflow. For instance, in cardiac diagnostics, cardiologists still heavily rely on traditional text-based medical records and lab reports in conjunction with medical images. In contrast, a comprehensive multimodal AI system adept at processing various data formats can liberate the cardiologist to focus on tasks demanding human expertise, representing a transformative shift in healthcare AI adoption.

**The need for real-time inference and processing**

Large Language Models (LLMs) and Large Multimodal Models (LMMs) pose several challenges when it comes to real-time inference in healthcare applications. These models, such as GPT-3 (May 2020, 4165) for LLMs and models like CLIP (Radford A. et al. Learning Transferable Visual Models From Language Supervision, 20AD) for LMMs, have demonstrated impressive capabilities in understanding and generating human-like text and handling multimodal inputs. However, deploying them in real-time healthcare scenarios comes with specific challenges:

- LLMs and LMMs are computationally expensive and resource intensive. Real-time inference requires quick processing, and these models may struggle to meet the time constraints of healthcare applications.
The sheer size of the models and the number of parameters they possess make it challenging to deploy them on resource-constrained devices, like those commonly found in healthcare environments.

- Real-time healthcare applications demand low latency to provide timely responses. The inference time of LLMs and LMMs may be too high for certain use cases, impacting the overall responsiveness of the system. High latency can be critical in healthcare scenarios where quick decision-making is essential, such as in emergency situations or during surgery.

- Storing and loading large models can be challenging in healthcare environments with limited storage capacity. Deploying such models on edge devices may require compromises in terms of model size and capabilities. Model updates and maintenance can be cumbersome, especially if frequent updates are necessary to keep the model’s knowledge up to date.

- Healthcare data is highly sensitive, and using cloud-based services for real-time inference with LLMs and LMMs may raise concerns about data privacy and security. On-device deployment might mitigate some of these concerns, but it brings challenges related to the computational power and storage capacity of the devices.

To address these challenges, ongoing research and development are necessary to optimize the efficiency, interpretability, and privacy aspects of LLMs and LMMs for real-time healthcare applications. Additionally, tailored training on healthcare-specific data and collaboration with domain experts are crucial to enhance the models’ performance and relevance in medical contexts.

**Conclusion**

Generative AI in medical imaging and text analysis stands at a crossroads. While its potential to revolutionize healthcare is undeniable, navigating the ethical, legal, societal, and algorithmic complexities is paramount to ensure its responsible implementation. Addressing ethical concerns around accuracy, informed consent, and data privacy necessitates robust frameworks and ongoing public dialogue. The legal landscape needs clear regulations on liability and accountability, fostering innovation while safeguarding patient well-being. Societally, ensuring equitable access to AI-powered healthcare and mitigating potential biases within algorithms are crucial for building trust and inclusivity. Algorithmically, overcoming limitations in training data, model capabilities, and workflow integration requires continuous research and development, prioritizing real-time inference and explainability. By acknowledging these interconnected challenges and actively seeking solutions, we can unlock the transformative potential of generative AI in medicine, paving the way for a future where technology enhances human expertise, leading to a brighter horizon for healthcare and human well-being.

**References**

(2020).
(2021).
(2020).
(2021).
(2020).
(2019).
(2019).
(2020).
(2021).
(2019).
(2019).
(2019).
(2020).
(2019).


(2019).


(2020).
(2019).
(2007).
(2017).
(2018).
(2020).
(2017).


(2020).
(1347).
(1712).
(2019).
(1999).
(2019).
(2017).
(2807).
(2018).
(2019).
(2019).