Fagbola, T. M., Dhiman, A., Mboli, J., & Mishra, B. Paper presented at 1st International Workshop on Responsible AI (RAI) for Healthcare and Net Zero, IIT Madras, Chennai, India, 16 Oct 2024.
https://www.responsibleaihull.com/events/responsibleai-chennai.

A Responsible AI Perspective to implementing Generative AI in Personalized Healthcare: Implications, Challenges and Future Directions

Abstract

Generative AI (GenAI) is transforming personalized healthcare by enabling customized treatment plans, advancing drug discovery, and offering targeted diagnostic support. While these advancements offer significant potential, they also present complex ethical and practical challenges. This paper explores the ethical implications and practical challenges associated with integrating GenAI into personalized healthcare, with a focus on the need for comprehensive Responsible AI frameworks. We critically assess existing frameworks, highlighting their limitations in the context of personalized healthcare. Key ethical concerns include algorithmic bias, threats to patient privacy, diminished patient autonomy, and the lack of accountability for AI-driven errors. On a practical level, challenges such as the integration of GenAI with current healthcare systems, the need for high-quality and diverse training data, and issues related to trust, transparency and explainability are examined. Our approach involves a systematic review of recent literature on personalized healthcare, AI ethics, healthcare GenAI applications, and international AI regulatory and governance standards. Our findings indicate that while GenAI holds great promise for improving personalized healthcare outcomes, current frameworks often fail to adequately address healthcare-specific challenges. These gaps include insufficient measures to mitigate bias, inadequate regulation of data privacy, and a lack of clear universally acceptable requirements for explainability in medical AI applications. This review contributes to the ongoing discussion by offering specific recommendations to enhance Responsible AI frameworks. These include fostering interdisciplinary collaboration, improving data governance strategies, and implementing stricter transparency standards as GenAI advancement continues to evolve. We call for continued research and policy development to ensure that GenAI integration in personalized healthcare remains ethical, equitable, and focused on promoting patient welfare without compromising ethical standards.

Keywords: Responsible AI, Generative AI, Conversational AI, Personalized medicine, Healthcare AI, Explainable healthcare, Inclusive Chatbot, Sustainable healthcare, AI Framework.

1. Introduction

Over the years, healthcare systems globally have been experiencing unprecedented crisis and multifaceted challenges due to increasing demand for healthcare services, aging population, global shortage of health professionals and the vicious cycle associated with it, widening health inequities, and poor funding [1}. For example, in England, as of July 2023, around 7.7 million people were on the non-emergency waiting list while over 350,000 people had been waiting for over a year [2]. Currently, huge concerns have been raised by over 74% health institutions in England regarding staff shortages. By implication, late diagnosis, widening health disparity gaps, and avoidable deaths are common outcomes [3]. These challenges have opened opportunities for proactive digital interventions, enabling the seamless integration of advanced Fourth Industrial Revolution technologies, such as Generative Artificial Intelligence (GenAI). These innovations aim to enhance healthcare practices, improve efficiency and effectiveness, boost productivity, and increase patient satisfaction [4,5,6]. For clarity, GenAI is a type of AI system that generates new information based on patterns it learns from large datasets. In healthcare, it utilizes relevant patient data, including genomic information, electronic health records, and clinical notes, to generate insights that support quick and effective medical decision-making [7,8]. GenAI systems are often deployed as conversational agents or chatbots, and their use in healthcare cut across various application domains. Major application areas of GenAI healthcare include personalized treatment, medical image analysis and diagnostics, patient monitoring, drug discovery and development and patient assistance [4]. It has also been deployed to downstream health services and tasks, such as screening and (follow-up) treatment support, medical training, administrative practice optimization, robot-assisted AI surgery, and virtual nursing assistants [8]. GenAI integration in healthcare is a huge and promising revolution. In 2023, the market size of healthcare GenAI in the U.S. grew to over $520.11 million, and globally stood at an excess of $1.5 Billion in revenue. By 2033, the market size is projected to reach over $8,131.58 million in the U.S. and $29.8 Billion globally [9,10,11].

Beyond the traditional one-size-fits-all treatment strategy, the healthcare industry is now drifting towards personalized healthcare [12], otherwise known as precise medicine. This trend is heightened by the growing demand of personalized healthcare interventions globally, which has now become a key determinant and driving factor of the global GenAI healthcare market [10,11]. Personalized healthcare in the GenAI era involves an intervention plan in which GenAI is used to predict and tailor medical treatment towards the unique biological characteristics, lifestyle and health history of a patient [12]. A key goal of personalized healthcare is to optimize clinical decision-making and improve patient outcomes [12, 13]. For instance, in oncology, it is being used to simulate patient-specific tumour behaviour and to predict health outcomes based on different treatment options [14]. In pharmacogenomics, GenAI uses individual patients' genetic makeup to personalize medical interventions, translating to reduced adverse drug effects and improved treatment effectiveness [7]. Similarly, GenAI models like GANs have seen successful application in creating novel drug compounds that are targeted to specific diseases and persons, thereby improving health outcomes [7]. These models have also assisted in personalized kidney care with the ability to predict potential acute failure 48 hours in advance, thereby enabling early intervention [15]. They are very effective and efficient in analyzing electronic health records of patients, and provide predictive insights, such as the likelihood of hospital readmission, disease progression rate or reoccurrence ratio towards early and targeted interventions [5,12,13]. GenAI has also demonstrated capacity to enhance diagnostic accuracy towards early disease detection, such as identifying tumours in mammograms and brain magnetic resonance images [16,17]. GenAI systems have natural language processing capabilities. Hence, they have been deployed as virtual health assistants to monitor patient symptoms, support chronic disease management and provide personalized health advice [18].

However, from a responsible AI standpoint, the goal of GenAI in personalized healthcare is to promote health equity, close disparity gaps and ensure that treatments are ethics-compliant and inclusive [22-25]. Despite the promising potential of GenAI systems in personalized healthcare, their implementation, as they directly interact with sensitive patient data, has raised serious ethical concerns including inadvertent data privacy and security breaches, and their potential to compromise patient autonomy [19]. Furthermore, lack of informed consent and satisfactory explanation, especially when GenAI systems use sensitive patient's data to make personalized recommendations, also exacerbate patients concerns. Many GenAI models are not interpretable, thereby complicating healthcare professional's ability to validate and trust AI-generated recommendations. Inherent bias in these systems can also lead to disparities in healthcare outcomes, undermining the principle of fairness in medical practice. In the remaining part of this paper, we discuss the ethical implications and practical challenges associated with implementing generative AI in personalized healthcare. We investigate existing responsible AI frameworks, their relevance in (personalized) healthcare, and how they can be adapted to birth healthcare applications that are ethically sound. We emphasize the need to develop responsible AI frameworks that balance technological innovation with ethical considerations, focusing on data privacy, patient autonomy, equity, transparency, safety and accountability. Furthermore, we discuss the growing need for inclusive design that involves all health stakeholders, including AI developers, AI researchers, health professionals, policymakers, regulatory bodies, patients and the public in general. We provide tailored recommendations such as fostering interdisciplinary collaboration, refining data governance strategies, and enforcing rigorous transparency standards to improve current responsible AI frameworks tailoring them to specific healthcare needs. We advocate for continued research and policy development to ensure that the integration of AI into healthcare is both ethical and equitable, promoting patient welfare without compromising ethical standards. The rest of this article is organized as follows; we discuss the ethical implications of healthcare GenAI in section 2 and the practical challenges of their implementation in section 3. In section 4, we present and discuss some Responsible AI frameworks for healthcare. We argue the need for interdisciplinary collaboration in section 5 and we conclude in section 6.

2. Ethical Implications of Generative AI in Healthcare

Despite the potential of GenAI to transform healthcare, its integration into healthcare systems has raised several key ethical concerns. These include bias in AI-generated recommendations, privacy and data security, patient autonomy and informed consent, transparency and explainability, safety and reliability, and responsibility and accountability. Bias in AI-generated recommendations is one of the most significant ethical concerns in healthcare. Biased outcomes often result from AI models trained on datasets that lack diversity, and are not representative of the broader population, which in turn can disproportionately affect minority groups [20]. For example, a healthcare AI model may exhibit racial bias, if it misrepresents the health needs of patients due to their race [21]. Bias in GenAI can manifest in terms of overdiagnosis or underdiagnosis in certain populations, unequal allocation of healthcare resources, and misrepresentation of disease risks. To mitigate these biases, diverse representative datasets should be employed in model training and the implementation of fairness checks be ensured [22]. Additionally, AI outputs should be consistently monitored to identify and correct biases as they emerge. With respect to privacy and security issues, when GenAI models are trained with sensitive patient data, concerns relating to data privacy, informed consent, potential breaches and unauthorized access are raised [23]. There is a potential risk of re-identifying individuals if GenAI models create synthetic patient data from original dataset that was not sufficiently anonymized [24]. An important question for GenAI healthcare stakeholders is if patients are aware of how their data will be used, its potential impact on their care, or the inherent risks of using their data for GenAI model training. In this regard, transparency in data usage, robust encryption, secure data storage and explicit consent procedures are crucial to address these concerns [24].

Patient autonomy and informed consent are also very pressing ethical issues especially as GenAI's recommendations can strongly influence clinical decision-making toward specific treatment options which may impart patient’s autonomy. A major concern is possible overreliance of clinicians on AI-generated recommendations while overlooking alternative and equally viable treatments [25]. Hence, a "human-in-the-loop" process is imperative [26], where GenAI outputs are reviewed by healthcare experts for relevance, appropriateness, and accuracy. These efforts ensure that GenAI supports rather than replaces clinical judgment. Patients must also play an active role in their treatment [27]. A shared decision-making approach is necessary to obtain patient consent and ensure they are fully informed about the integration of GenAI in their care, how it generates recommendations, and any potential risks involved. In terms of transparency and explainability, GenAI in healthcare must transition from a black-box to a white-box model. This shift would enable healthcare professionals to understand how the system arrives at its recommendations, ensuring its safe use in clinical practice. Without sufficient explanation, mistrust, skepticism, and errors in patient care become unavoidable [28]. Consequently, Explainable AI (XAI) methods are now being developed to make the decision-making process of GenAI systems more transparent to clinicians, for example, through visual explanations of their network activities [29]. Given that GenAI systems directly impact health outcomes, they must undergo rigorous testing, continuous validation, and regular updates based on clinical standards before deployment to ensure safety and reliability [30]. Validation processes should also confirm that GenAI’s performance remains consistent across diverse populations and dynamic healthcare settings [31]. Such evaluations will help end-users better understand the behaviour of healthcare

GenAI, thereby enhancing trust [11]. Healthcare stakeholders and policymakers should establish regulatory frameworks and safety guidelines as a minimum standard for GenAI evaluation and deployment in clinical practice. The safety of AI systems further involves regular updates and monitoring to incorporate new medical knowledge and data [32]. This is crucial for protecting the model's objective function from adversarial attacks. Regarding responsibility and accountability, a critical question remains: who is liable when a healthcare GenAI makes a faulty diagnosis or an ineffective treatment recommendation [33]? Reaching a consensus on this issue is essential to protect patient rights and ensure the ethical use of GenAI in healthcare. There is a growing movement toward adopting a shared responsibility framework, guided by relevant legal and regulatory guidelines, which involves clinicians, healthcare institutions, and GenAI developers [5]. Such a framework would not only promote the ethical deployment of GenAI but also help build public trust in AI-assisted healthcare.

3. Practical Challenges in Implementing Generative AI in Healthcare

Aside the ethical concerns related to the use of GenAI in healthcare as discussed in the previous section, there are also practical challenges associated with their implementation. Major challenges revolve around the quality and diversity of data used in training, difficulties integrating GenAI solutions with existing healthcare systems, as well as issues of poor regulation and compliance at putting emerging healthcare GenAI in check. Considering the austerity of healthcare, it becomes critical to ensure that training data for healthcare GenAI models are both diverse and of good quality. Tan et al. [34] argued that several real-world datasets employed in training these models are often imbalanced and suffer from limited size and quality. The imbalanced nature of these datasets emanates from issues like sampling, observation, measurement and algorithmic biases. When bias dataset is used to train a model, the output becomes unrepresentative and erroneous [35]. In a bid to ensure data quality, pre-processing and in-processing are the two commonly employed steps [34]. In a pre-processing step, data bias is targeted and eliminated during the model training process. This could be by collecting new (additional) data or balancing data with appropriate data selection methods. During an in-processing step, data quality is addressed by introducing alternative training objectives. Inappropriate contents such as discriminative, violent or offensive data also impede data quality [36]. As a result, training data should be subjected to scrutiny by healthcare stakeholders, AI experts, and diverse group of potential application users. Data equity must also be pursued to promote privacy, indigenous data autonomy and accessibility [37]. The inclusion of data equity and GenAI presents unique challenges as it requires proactively auditing data and algorithms and continuous intervention at every step of the GenAI application development process starting from data collection to model training, then to its implementation. To mitigate bias issues, dataset augmentation, user feedback and bias-aware algorithms are often employed [35]. Dataset augmentation implies the inclusion of additional data into the training dataset to foster wider representation and minimise bias. User feedback seeks to obtain constructive feedback from application users about their experience to spot inherent bias which can be used to make corrections in the model. Bias-aware algorithms are designed to become aware of representativeness, diversity and completeness issues in the training data to minimise their collective influence on the models' output.

The incessant proliferation of advanced GenAI models have raised the problem of regulation and compliance that could effectively govern AI implementation and use in healthcare settings. Lately, some proactive strategies are being proposed towards regulation and compliance enforcement [5]. An example is the EU AI Act (2024), which advocates for transparent development and use of AI to include a summarised publication of copyrighted dataset used for model training. There is also the Medicines and Healthcare Products Regulatory Agency (MHRA) in the UK that sees to ensuring that healthcare GenAI models and associated medical products are subjected to regular re-evaluation and re-approval. However, these Acts should be flexible to accommodate mitigations to potential loopholes that may suffice from new and advanced GenAI capabilities which are everyday emerging. Ethical, legal and social aspects are another concern in the regulation strategies. Every GenAI model is expected to comply with all social and ethical principles, as well as relevant laws and regulations to avoid any unintentional harm [38]. Another problem bothers on integrating healthcare GenAI models with existing healthcare systems. The potential integration challenges that may arise can be categorised as technological, organizational or environmental [39]. The technological component accounts for compatibility and complexity issues, including ease of use and user interface design associated with the existing system. The organisational component is focused on the readiness and absorptive capacity while the environmental component is concerned with regulatory support governing adaptation and dealing with uncertainty. While the benefits of GenAI are noticeable, it is crucial to explore how it can also accommodate and support the technical know-how of end-users of existing systems [36]. Generally, successful integration of new technologies with existing systems may depend upon their readiness, technology acceptance, availability of relevant interoperable hardware and technologies, initial cost of integration and maintenance, change management and appropriate risk mitigation strategies [40].

4. Responsible AI Frameworks for Healthcare

Responsible AI frameworks in healthcare are essential to ensure that AI systems are developed and deployed ethically, transparently, and effectively. These frameworks address concerns such as bias, fairness, privacy, and accountability. Governments and organizations worldwide are working to guarantee AI systems' fairness, trust, transparency, privacy, and robustness. The following sections explore some existing frameworks and case studies.

4.1 Overview of AI Frameworks

The rapid advancement of healthcare GenAI necessitates robust frameworks to ensure its ethical and beneficial application. There are several GenAI models that can be deployed for various applications in healthcare and other industries [41]. A few of these models, their descriptions, possible applications for healthcare and considerations for responsible AI is presented in Table 1 [42].

Table 1: Selected Generative AI Models

Generative AI Model	Description	Healthcare Applications	Considerations for Responsible AI
Generative Adversarial Networks (GANs) [43]	GANs consist of two neural networks, a generator and a discriminator, that compete against each other. Often used in image synthesis, super-resolution, and style transfer.	**Medical image synthesis**, such as generating realistic MRI, CT, or X-ray images for training or testing purposes, **data augmentation** for rare diseases, and **drug discovery** by simulating molecular structures.	- **Bias detection-** GANs must be trained with diverse datasets to avoid biased outputs. - **Data privacy-** Ensure synthetic data is not identifiable or reverts to real patient data. - **Transparency-** Outputs should be interpretable for healthcare professionals.
Variational Autoencoders (VAEs) [42]	VAEs are a type of autoencoder which generates structured representations, useful for tasks such as creating new images or other data points.	**Anomaly detection** in medical imaging (e.g., detecting tumours in MRI scans), **image denoising** for clearer diagnostics, and **medical data compression**.	- **Data integrity-** Ensure no medically relevant details are lost during data compression or denoising. - **Explainability-** VAEs need to be interpretable, especially when used in diagnostics. - **Fairness**: Avoid potential biases from skewed training data that could misrepresent certain demographics.
Autoregressive Models [44]	These models predict the next output in a sequence based on previous outputs. Extensively	**Electronic Health Record (EHR) predictions**, such as forecasting disease progression, **clinical note**	- **Patient safety-** Errors in generated clinical text could have serious consequences. Outputs must be verified by clinicians.

	used in text and time-series predictions.	**generation**, and **personalized treatment suggestions**.	- **Accountability-** Clear documentation of model decisions for liability and traceability. - **Data security-** Careful handling of patient data to ensure privacy compliance.
Flow-based Models [45]	These models use the change of variables formula to model complex distributions, allowing both generation and efficient inference.	**Anomaly detection** in vital signs monitoring or lab results, **image generation** for rare medical conditions, and **disease modelling**.	- **Robustness-** Ensure models are reliable across various patient groups and rare conditions. - **Explainability-** Results must be interpretable and easily understood by medical professionals. - **Security-** Safeguard against adversarial attacks that could mislead medical decisions.
Energy-based Models (EBMs) [46]	EBMs learn an energy function to assign low-energy values to data points from the distribution and higher values to others. Used in image restoration and structured prediction.	**Image restoration** (e.g., improving noisy MRI scans), **unsupervised learning** for identifying patient subtypes, and **structured prediction** for treatment outcomes.	- **Ethical use-** Ensure no harmful treatment recommendations are generated. - **Bias mitigation-** Use diverse data to ensure fair treatment recommendations for all patient demographics. - **Data fidelity-** Maintain high-quality outputs that do not distort medically relevant information.
Diffusion Models [47]	These models reverse a diffusion process to generate high-quality, diverse samples.	**Medical image generation** (e.g., enhancing or generating training data for rare conditions), **protein folding prediction**, and **drug molecule generation**.	- **Reliability** High-fidelity generation is critical in healthcare to avoid misleading data. - **Accountability**: Ensure transparency and auditability in molecular design for drug discovery. - **Regulatory compliance**: Generated molecular structures or drugs must adhere to strict healthcare regulations.

4.2 Enhancing Generative AI Frameworks for Healthcare-Specific Needs

As discussed in the previous sections, the benefits of implementing GenAI in healthcare are vast, including its potential to significantly enhance patient outcomes and optimize healthcare delivery. However, realizing these benefits requires careful consideration of the ethical concerns and challenges associated with GenAI deployment. In this regard, it is crucial to implement robust AI governance frameworks that address issues such as accuracy, fairness, and data privacy [24,34]. This involves establishing clear guidelines for the ethical use of AI, continuous monitoring for biases, and ensuring transparency in AI decision-making processes. Such frameworks are key to fostering responsible and trustworthy AI in healthcare. Moreover, the adoption of generative AI in healthcare must include comprehensive evaluation mechanisms to assess its impact and effectiveness. Recent studies highlight the importance of a "human-in-the-loop" approach, where healthcare practitioners actively oversee AI-generated outputs to mitigate risks associated with incorrect AI responses, which can be particularly hazardous in healthcare settings [26]. Additionally, there is an increasing need for regulatory frameworks that can keep pace with the rapid advancements in AI technologies [48]. These frameworks should prioritize safeguarding patient data, ensuring the reliability and safety of AI systems, and promoting equitable access to AI-driven healthcare solutions. By addressing these challenges, generative AI can be responsibly integrated into healthcare, ultimately leading to improved patient outcomes and more efficient healthcare delivery.

5. Interdisciplinary Collaboration

Currently, GenAI models in healthcare face several challenges, including a lack of quality medical data, concerns over data privacy, availability, and security, as well as difficulties in determining relevant clinical metrics and selecting appropriate methodologies [33-38]. When AI makes errors, questions about accountability and responsibility arise, requiring ethical and legal frameworks to manage such risks. Despite existing laws, patients often bear the consequences of AI-related failures [22]. Additionally, many AI models are developed without meaningful engagement with the communities they are designed to serve, which can lead to societal harm [21,49]. This emphasizes the need for more inclusive AI development that considers community input to avoid unintended consequences. In reality, AI researchers and developers make critical choices throughout the process—from data collection to deployment—imprinting their own ethics and values on the technology. Every decision must be responsibly considered to ensure the ethical use of AI [22-25]. Moreover, measuring fairness and effectiveness of machine learning (ML) models in complex societal contexts can often be misguided and even harmful if not carefully managed [34,38]. Dr. Chang, Dean of Harvard Medical School, argues that as AI advances, medical students will need to shift their focus away from data gathering and summarization and concentrate more on complex clinical decision-making. AI can assist with data processing and diagnostics, but critical thinking, compassionate communication, and physical examination remain irreplaceable skills [26-27]. A multi-disciplinary approach, involving cooperation between AI scientists, healthcare professionals, and clinicians, is essential for developing ethical and effective AI systems. By incorporating clinical expertise, selecting relevant features, testing for clinical validity, and improving human-computer interaction, this collaboration ensures that AI models are both clinically useful and ethically sound. Cross-institutional collaboration also facilitates data sharing, improves AI model verification, and addresses concerns about bias, privacy, and transparency, leading to greater trust and broader adoption of the technology [50].

In 2023, Google established the Impact Lab as part of its responsible AI team, focusing on the socioeconomic and human rights impacts of AI [51]. This initiative led to the development of the Equitable AI Research Roundtable (EARR), a coalition of experts in law, education, community engagement, social justice, and technology. In EARR, technical teams provide use cases, market insights, and early-stage AI prototypes, while external experts contribute knowledge on AI ethics and technology equity, focusing on the communities most affected by unfair predictions. Feedback from these experts helps technical teams adjust their models by diversifying datasets, improving user communication, and refining outreach strategies. This collaboration ensures that AI models become more representative of society, safer, more trustworthy, and reflective of global diversity [20,22,31,34,37]. Despite the transformative potential of AI in diagnostics, treatment, and personalized medicine, several barriers impede its integration into clinical practice. These challenges include workflow disruptions due to continuous documentation requirements, an overwhelming influx of alerts and notifications, over-reliance on AI, and a lack of interoperability between AI systems and existing technologies [25,39,40]. Moreover, the absence of effective human computer interaction (HCI) has been identified as a key reason for the failure of clinical decision support systems (CDSS) to be fully embraced by healthcare professionals [52]. A Human-Centered AI (HCAI) approach, which emphasizes clinician involvement throughout the design, development, and implementation phases, offers a viable solution to these issues [36,49]. Engaging clinicians from the beginning allows for two primary benefits: seamless integration of AI tools into clinical workflows, ensuring that the AI aligns with real-world challenges faced by healthcare professionals, and validation of AI models through comparison with clinician performance, leading to improved accuracy and trust in AI systems [53].

Incorporating clinician feedback throughout the AI lifecycle also addresses concerns about ethical biases, trust, and accountability. By involving end-users—such as physicians, nurses, and other healthcare workers—during the development phase, AI systems can better account for the practical and ethical complexities of patient care. This collaboration not only improves the functionality of AI tools but also fosters a sense of ownership and trust among clinicians, which is crucial for the widespread adoption of AI in healthcare. Despite these advantages, many AI development processes fail to consistently involve clinicians, leading to issues of trust, usability, and adoption. Studies have shown that nearly half of AI development projects do not evaluate or address the trust clinicians place in these tools, which further impedes their acceptance in clinical environments [53]. Clinicians bring essential insights that go beyond data-driven decision-making, incorporating social, institutional, and ethical factors that AI cannot fully replicate. Their involvement is crucial for ensuring that AI systems are not only technically robust but also capable of supporting the complex decision-making processes that occur in healthcare settings [54]. Moreover, AI's role in fields such as radiography [17] and precision medicine [12-13] has demonstrated promising results, but there remains significant variability in how AI impacts clinical outcomes. For instance, research on AI-assisted radiology has shown that experience and diagnostic skill do not consistently predict the extent to which AI improves radiologist performance, highlighting the need for AI systems to provide transparent explanations and detailed reports, rather than relying solely on probabilistic outputs [28-29].

However, HCAI does come with its own limitations. Human involvement in the development process introduces the potential for biases, which can stem from personal, technical, or socio-economic backgrounds and may inadvertently influence AI system outcomes. Additionally, engaging clinicians throughout the design and development phases can increase the time and cost of AI projects, requiring openness to diverse perspectives and extensive collaboration. While this investment is necessary for creating more effective and trustworthy AI systems, it can be a challenge in fast-paced healthcare environments. Furthermore, new biases may emerge as a result of diverse human input, including industry-specific, geographic, or ethical considerations, potentially complicating the development of universally applicable AI models. Therefore, while HCAI offers substantial benefits for aligning AI with clinical practice, its implementation must be carefully managed to mitigate these potential drawbacks [55]. Many AI companies, including major players like Google, Meta, and Amazon, have committed to following AI ethics principles centered around fairness, transparency, security, privacy, and accountability [56]. While some companies have established ethics advisory groups to guide responsible AI practices, regulations from governments are increasingly becoming the primary means of enforcing responsible AI usage. Existing regulations like the General Data Protection Regulation (GDPR) cover important aspects of AI, but as GenAI evolves, some laws may need revisions or new frameworks [57]. Countries like the U.S. [58] and China [59] have introduced new regulations to address AI-specific risks, such as the spread of misinformation, misuse, harmful content generation and other unknown technical dangers. These regulations aim to create a safer environment for innovation by holding companies accountable and reducing potential risks. For instance, the U.S. government has developed key frameworks like the AI Bill of Rights [60], AI risk management framework [61], and launched the National AI Research Institutes with significant funding [62]. These policies underscore the importance of balancing AI advancements with safeguards that protect individuals and society, while still fostering innovation in a responsible manner. National and local regulations, while important, are often limited in addressing key issues that transcend borders, such as legal inconsistencies, varying definitions of AI, and compliance across jurisdictions [63]. Since AI technologies are global in nature, operating on a transnational scale, with networks, computing resources, and entities spread across multiple countries, a unified international framework is necessary to ensure that AI is developed and used responsibly and ethically by all countries, not just a few powerful corporations or states [64]. Domestic regulations tend to prioritize local interests, and without global oversight, AI development may disproportionately benefit certain regions while excluding others. For example, most AI advancements and rewards are currently concentrated in a small number of private companies and states, leaving much of the global population without access to its benefits [65]. Global cooperation through organizations like the United Nations (UN) and international initiatives such as the AI

Alliance can ensure that AI is governed inclusively and benefits all of humanity [66]. These global bodies emphasize the need for ethical principles, human rights, and the rule of law in guiding AI development [63-66]. A global regulatory framework would address the transboundary nature of AI, ensuring accountability for both developers and users of AI technologies [67]. It would also support open-source data sharing and the development of AI models that can be adapted across different regions, similar to how generic medicines improve healthcare access globally [68]. By empowering the UN and involving diverse stakeholders, global AI governance can effectively manage AI's wide-ranging impacts on global economic, social, health, and security issues [69]. Thus, a global approach is necessary to address the gaps left by national regulations and to foster a fair, inclusive, and ethical AI ecosystem worldwide [63-64].

6. Conclusion and Future Direction

GenAI models have shown great potential in shaping the future of personalized healthcare globally. These advanced models can deliver personalized treatment plans tailored to the specific needs of each patient. They also play a significant role in drug discovery and development, targeting patient-specific diseases with greater efficacy and reduced costs. GenAI has been widely applied in targeted diagnostic support by identifying patterns and correlating insights from diverse medical databases, leading to more accurate and early detection tailored to individual patients' health conditions. Additionally, these models have proven effective in uncovering links between genetic mutations and diseases based on a patient's genetic profile. GenAI has also been adapted for use as virtual health assistants and clinical decision support agents in various real-world applications. However, GenAI models face several challenges, both ethical and practical. These challenges raise concerns among healthcare stakeholders, which in turn limit the implementation, adoption, and use of GenAI in real-world healthcare settings for providing or supporting critical personalized healthcare services. In this paper, we conduct a systematic review of the implications and physical challenges of implementing GenAI in personalized healthcare from a Responsible AI perspective. GenAI models are expected to provide specific, relevant answers to clinical questions and ensure transparency in their decision-making processes, and they should be designed accordingly. Trust and interpretability are key, and legal requirements like the EU's GDPR mandate explanations for decisions affecting individuals. Methods like LIME [70] can improve interpretability, but the level of explanation needed varies based on the impact of the decision, for example, allocation of potentially life-prolonging treatments requires detailed explanations to satisfy the affected individuals. This can be achieved by extensive collaborations and consultation with the target audience (patients), and end users (clinicians) [27-29]. The future of responsible AI in healthcare requires careful validation and verification of AI tools, as their perceived advanced technology can falsely limit the need for thorough testing. For instance, Babylon Health's symptom checker faced concerns during early trials about its potential misuse by patients. This shows that AI algorithms may not be used as intended in real-world settings. To address this, researchers should consider potential pitfalls early on and develop AI tools with pragmatic clinical trials to ensure their effectiveness and proper usage in clinical practice. Researchers should design AI models with built-in mechanisms for reassessment, planning these from the start of implementation rather than adding them later [71]. The future directions for implementing Responsible AI in healthcare emphasize the need for updated regulations, unified professional guidelines, and the creation of new roles to address the complexities introduced by AI-based Clinical Decision Support Systems (AI-CDSS) [72]. First, professional regulations must evolve to accommodate the integration of AICDSS in healthcare, addressing issues of responsibility and accountability. This requires collaboration with regulators such as NICE [73], NHS England [74], and NHS Digital[75] to establish enforceable standards and guidance that apply to all clinicians across various professions, ensuring consistency. Additionally, new roles like "digital and AI specialist clinicians" are proposed to bridge the gap between AI developers and clinical users, helping ensure that AI technologies are safely and effectively integrated into healthcare [70]. However, these roles will also require professional guidance to define the appropriate use of AI in clinical settings. Before deployment, robust technology assessment frameworks must be established to ensure the safety and accuracy of AI-CDSS tools. Labels that provide transparency, explainability, and safety information are essential to help users understand and trust these systems. Moreover, unified and standardized guidance is crucial to prevent inconsistent or potentially harmful practices across different healthcare professions. Collaborative guidance from regulatory bodies and professional organizations will enable better integration of AICDSS into healthcare, benefiting both clinicians and patients [72-73].

References

1. Khan Z. (2023). The Emerging Challenges and Strengths of the National Health Services: A Physician Perspective. Cureus, 15(5), e38617. https://doi.org/10.7759/cureus.38617
2. Nuffield Trust. NHS performance tracker. Nuffield Trust; 2023 (www.nuffieldtrust.org.uk/qualitywatch/nhs-performance-summary).
3. https://www.health.org.uk/publications/long-reads/nine-major-challenges-facing-health-and-care-in-england
4. The Lancet Regional Health-Europe. Embracing generative AI in health care. Lancet Reg Health Eur. 2023 Jul 3;30:100677. doi: 10.1016/j.lanepe.2023.100677. PMID: 37465322; PMCID: PMC10350845.
5. Reddy, S., 2024. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. *Implementation Science*, *19*(1), p.27.
6. Mulukuntala, S., 2022. Generative AI-Benefits, Limitations, Potential risks and Challenges in Healthcare Industry. *EPH-International Journal of Medical and Health Science*, *8*(4), pp.1-9.
7. Zhavoronkov A., et al. (2019). Artificial intelligence for ageing and longevity research: recent advances and perspectives. *Ageing Research Reviews*. *Elsevier* 49, 49–66. 10.1016/j.arr.2018.11.003
8. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artificial Intelligence in Healthcare. 2020:25–60. doi: 10.1016/B978-0-12-818438-7.00002-2. Epub 2020 Jun 26. PMCID: PMC7325854.
9. https://www.sphericalinsights.com/reports/generative-artificial-intelligence-ai-in-healthcare-market
10. https://www.polarismarketresearch.com/industry-analysis/us-healthcare-generative-ai-market
11. https://finance.yahoo.com/news/global-generative-artificial-intelligence-ai-050000828.html
12. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, Zhao J, Snowdon JL. Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci. 2021 Jan;14(1):86-93. doi: 10.1111/cts.12884.
13. August GJ et al. Moving Toward a Precision-Based, Personalized Framework for Prevention Science: Introduction to the Special Issue. Prev Sci. 2019 Jan;20(1):1-9. doi: 10.1007/s11121-018-0955-9.
14. Sparano, Joseph A. et al. "Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer." *The New England Journal of Medicine* 379 (2018): 111–121.
15. Tomašev, Nenad et al. "A clinically applicable approach to continuous prediction of future acute kidney injury." *Nature* vol. 572,7767 (2019): 116-119. doi:10.1038/s41586-019-1390-1
16. Oyelade, O.N., Ezugwu, A.E., Almutairi, M.S. *et al.* A generative adversarial network for synthetization of regions of interest based on digital mammograms. *Sci Rep* **12**, 6166 (2022). https://doi.org/10.1038/s41598-022-09929-9
17. Babaferi et al. DeepCAI-V3: Improved Brain Tumor Classification from Noisy Brain MR Images Using Convolutional Autoencoder and Inception-V3 Architecture. *2024 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, 1-7.

18. Samala, A. D., & Rawas, S. (2024). Generative AI as Virtual Healthcare Assistant for Enhancing Patient Care Quality. *International Journal of Online and Biomedical Engineering (iJOE)*, *20*(05), pp. 174–187. https://doi.org/10.3991/ijoe.v20i05.45937
19. Nagabhushana, G. & Boinodiris, P. 2024. *Delivering responsible AI in the healthcare and life sciences industry* [Online]. Available: https://www.ibm.com/think/topics/responsible-ai-healthcare
20. Mehrabi, N., et al. (2021). A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, 54(6), 1-35. https://doi.org/10.1145/3457607
21. Obermeyer, Z., et al. (2019). Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations. *Science*, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342
22. McCradden, M. D., et al. (2020). Ethical Concerns around Use of Artificial Intelligence in Health Care Research from the Patient Perspective: A Scoping Review. *BMC Medical Ethics*, 21(1), 1-11. https://doi.org/10.9778/cmajo.20190151
23. Jiang, F., et al. (2017). Artificial Intelligence in Healthcare: Past, Present, and Future. *Stroke and Vascular Neurology*, 2(4), 230-243. https://doi.org/10.1136/svn-2017-000101
24. Murdoch, T. B. (2021). Privacy and Artificial Intelligence: Challenges for Protecting Health Information in a New Era. *BMC Medical Ethics*, 22(1), 1-6. https://doi.org/10.1186/s12910-021-00687-3
25. Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare. *Artificial Intelligence in Healthcare*, 295-336. https://www.sciencedirect.com/science/article/pii/B9780128184387000125
26. Bhasker, S., Bruce, D., Lamb, J. & , A. G. S. 2023. *Tackling healthcare's biggest burdens with generative AI* [Online]. Available: https://www.mckinsey.com/industries/healthcare/our-insights/tackling-healthcares-biggest-burdens-with-generative-ai [Accessed 21/08/2024].
27. Wahl, B., et al. (2018). Artificial Intelligence (AI) and Global Health: How Can AI Contribute to Health in Resource-Poor Settings? *BMJ Global Health*, 3(4), e000798. https://doi.org/10.1136/bmjgh-2018-000798
28. Tonekaboni, S., et al. (2019). What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use. *ArXiv, abs/1905.05134*.
29. Arrieta, A. B., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Information Fusion*, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012
30. Wang, Changyu et al. "Ethical Considerations of Using ChatGPT in Health Care." *Journal of medical Internet research* vol. 25 e48009. 11 Aug. 2023, doi:10.2196/48009
31. Gulshan, V., et al. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 316(22), 2402-2410. https://doi.org/10.1001/jama.2016.17216
32. Haq IU, Chhatwal K, Sanaka K, Xu B. Artificial Intelligence in Cardiovascular Medicine: Current Insights and Future Prospects. Vasc Health Risk Manag. 2022 Jul 12;18:517-528. doi: 10.2147/VHRM.S279337
33. Morley, J., et al. (2020). The Ethics of AI in Health Care: A Mapping Review. *Social Science & Medicine*, 260, 113172. https://doi.org/10.1016/j.socscimed.2020.113172
34. Tan, S., Shen, Y. and Zhou, B., 2020. Improving the fairness of deep generative models without retraining. *arXiv preprint arXiv:2012.04842*.
35. Ferrara E. Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Sci*. 2024; 6(1):3. https://doi.org/10.3390/sci6010003
36. Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, *25*(3), 277–304. https://doi.org/10.1080/15228053.2023.2233814
37. Stonier, J. et al., 2023. Data Equity: Foundational Concepts for Generative AI. *arXiv preprint arXiv:2311.10741*.
38. Fedele, A., Punzi, C. and Tramacere, S., 2024. The ALTAI checklist as a tool to assess ethical and legal implications for a trustworthy AI development in education. *Computer Law & Security Review*, *53*, p.105986.

39. Prasad Agrawal, K. (2023). Towards Adoption of Generative AI in Organizational Settings. *Journal of Computer Information Systems*, *64*(5), 636–651. https://doi.org/10.1080/08874417.2023.2240744
40. Akin, Oyeleye Christopher et al. "The Impact and Challenges of Cloud Computing Adoption on Public Universities in Southwestern Nigeria." *International Journal of Advanced Computer Science and Applications* 5 (2014): n. pag.
41. Gozalo-brizuela, R. & Garrido-merchan, E. C. 2023. ChatGPT is not all you need. A State of the Art Review of large Generative AI models. *arXiv preprint arXiv:2301.04655*.
42. Kingma, D. P. & Welling, M. 2019. An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12**,** 307-392.
43. Creswell, A. et al. 2018. Generative adversarial networks: An overview. *IEEE signal processing magazine*, 35**,** 53-65.
44. Liu, T., Jiang, Y., Monath, N., Cotterell, R. & Sachan, M. 2022. Autoregressive structured prediction with language models. *arXiv preprint arXiv:2210.14698*.
45. Kumar, M. et al, 2019. Videoflow: A conditional flow-based model for stochastic video generation. *arXiv preprint arXiv:1903.01434*.
46. Du, Y. & Mordatch, I. 2019. Implicit generation and modeling with energy based models. *Advances in Neural Information Processing Systems*, 32.
47. Croitoru, F.-A. et al. 2023. Diffusion models in vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45**,** 10850-10869.
48. Nagabhushana, G. & Boinodiris, P. 2024. *Delivering responsible AI in the healthcare and life sciences industry* [Online]. Available: https://www.ibm.com/think/topics/responsible-ai-healthcare [Accessed 12/07/2024].
49. Emily Black et al."A call for universities to develop requirements for community engagement in ai research". In: The Fair and Responsible AI Workshop at the 2020 CHI Conference on Human Factors in Computing Systems. 2020.
50. Brian A. and Eric S. "Collective action for responsible AI in health". In: (2024).
51. Jamila Smith-Loud. Responsible AI at Google Research: The Impact Lab. Mar. 2023. url: https://blog.research.google/2023/03/responsible- ai- at- googleresearch.html.
52. Glyn Elwyn et al. "Many miles to go. . .": a systematic review of the implementation of patient decision support interventions into routine clinical practice". In: BMC medical informatics and decision making 13 (2013), pp. 1–10.
53. Stephanie Tulk Jesso et al. "Inclusion of clinicians in the development and evaluation of clinical artificial intelligence tools: a systematic literature review". In: Frontiers in Psychology 13 (2022), p. 830345.
54. Saif Khairat et al. "Reasons for physicians not adopting clinical decision support systems: critical analysis". In: JMIR medical informatics 6.2 (2018), e8912.
55. You Chen et al. "Human-centered design to address biases in artificial intelligence". In: Journal of medical Internet research 25 (2023), e43251.
56. Paul B de Laat. "Companies committed to responsible AI: From principles towards implementation and regulation?" In: Philosophy & technology 34.4 (2021), pp. 1135–1193.
57. As gen AI advances, regulators—and risk functions—rush to keep pace — mckinsey.com. https://www.mckinsey.com/capabilities/risk-and-resilience/ourinsights/as- gen- ai- advances- regulators- and- risk- functions- rush- tokeep-pace. [Accessed 12-09-2024].
58. FACT SHEET: Biden-Harris Administration Announces New Actions to Promote Responsible AI Innovation that Protects Americans' Rights and Safety—OSTP—The White House — whitehouse.gov. https://www.whitehouse.gov/ostp/newsupdates/2023/05/04/fact-sheet-biden-harris-administration-announcesnew-actions-to-promote-responsible-ai-innovation-that-protects-americansrights-and-safety/. [Accessed 12-09-2024].
59. Translation: Measures for the Management of Generative Artificial Intelligence Services (Draft for Comment) – April 2023 - DigiChina — digichina.stanford.edu.

https://digichina.stanford.edu/work/translation-measures-for-the-managementof- generative - artificial - intelligence - services - draft - for - comment -april-2023/. [Accessed 12-09-2024].
60. Blueprint for an AI Bill of Rights — OSTP — The White House — whitehouse.gov. https://www.whitehouse.gov/ostp/ai-bill-of-rights/. [Accessed 12-09-2024].
61. AI Risk Management Framework — nist.gov. https://www.nist.gov/itl/ai-riskmanagement- framework. [Accessed 12-09-2024].
62. ai.gov. https://www.ai.gov/wp-content/uploads/2023/01/NAIRR-TF-Final-Report-2023.pdf. [Accessed 12-09-2024].
63. Fan Yang. Who will write the rules for AI? How nations are racing to regulate artificial intelligence — theconversation.com. https://theconversation.com/who-will-write-the-rules-for-ai-how-nations-are-racing-to-regulate-artificialintelligence-216900. [Accessed 12-09-2024].
64. Nicholas Emery-Xu, Andrew Park, and Robert Trager. "Uncertainty, information, and risk in international technology races". In: Journal of Conflict Resolution (2023), p. 00220027231214996.
65. https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/ai_advisory_body_interim_report.pdf. [Accessed 12-09-2024].
66. Technical Difficulties— state.gov. https://www.state.gov/artificial-intelligencefor-accelerating-progress-on-the-sustainable-development-goals-addressingsocietys-greatest-challenges/. [Accessed 12-09-2024].
67. Robert Trager et al. "International governance of civilian AI: A jurisdictional certification approach". In: arXiv preprint arXiv:2308.15514 (2023).
68. Elizabeth D Gibbons. "Toward a more equal world: the human rights approach to extending the benefits of artificial intelligence". In: IEEE Technology and Society Magazine 40.1 (2021), pp. 25–30.
69. David Leslie et al. "'Frontier AI,' Power, and the Public Interest: Who benefits, who decides?" In: Harvard Data Science Review Special Issue 5 (2024).
70. Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ""Why should i trust you" Explaining the predictions of any classifier". In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016, pp. 1135–1144.
71. Sebastian Vollmer et al. "Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness". In: bmj 368 (2020).
72. Helen Smith, John Downer, and Jonathan Ives. "Clinicians and AI use: where is the professional guidance?" In: Journal of Medical Ethics 50.7 (2024), pp. 437–441.
73. https://www.facebook.com/NationalInstituteforHealthandCareExcellence. Find guidance— nice.org.uk. https://www.nice.org.uk/guidance. [Accessed 14-09-2024].
74. NHS England. NHS England — england.nhs.uk. https://www.england.nhs.uk/. [Accessed 14-09-2024].
75. https://digital.nhs.uk/. [Accessed 14-09-2024].