Background

Artificial intelligence (AI), particularly generative AI, has emerged as a transformative tool in healthcare, with the potential to revolutionize clinical decision-making and improve health outcomes. Generative AI, capable of generating new data such as text and images, holds promise in enhancing patient care, revolutionizing disease diagnosis and expanding treatment options. However, the utility and impact of generative AI in healthcare remain poorly understood, with concerns around ethical and medico-legal implications, integration into healthcare service delivery and workforce utilisation. Also, there is not a clear pathway to implement and integrate generative AI in healthcare delivery.

- Methods

This article aims to provide a comprehensive overview of the use of generative AI in healthcare, focusing on the utility of the technology in healthcare and its translational application highlighting the need for careful planning, execution and management of expectations in adopting generative AI in clinical medicine. Key considerations include factors such as data privacy, security and the irreplaceable role of clinicians' expertise. Frameworks like the technology acceptance model (TAM) and the Non-Adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) model are considered to promote responsible integration. These frameworks allow anticipating and proactively addressing barriers to adoption, facilitating stakeholder participation and responsibly transitioning care systems to harness generative AI's potential.

- Results

Generative AI has the potential to transform healthcare through automated systems, enhanced clinical decision-making and democratization of expertise with diagnostic support tools providing timely, personalized suggestions. Generative AI applications across billing, diagnosis, treatment and research can also make healthcare delivery more efficient, equitable and effective. However, integration of generative AI necessitates meticulous change management and risk mitigation.
strategies. Technological capabilities alone cannot shift complex care ecosystems overnight; rather, structured adoption programs grounded in implementation science are imperative.

- Conclusions
It is strongly argued in this article that generative AI can usher in tremendous healthcare progress, if introduced responsibly. Strategic adoption based on implementation science, incremental deployment and balanced messaging around opportunities versus limitations helps promote safe, ethical generative AI integration. Extensive real-world piloting and iteration aligned to clinical priorities should drive development. With conscientious governance centred on human wellbeing over technological novelty, generative AI can enhance accessibility, affordability and quality of care. As these models continue advancing rapidly, ongoing reassessment and transparent communication around their strengths and weaknesses remain vital to restoring trust, realizing positive potential and, most importantly, improving patient outcomes.

- Contributions to the literature
• Generative AI has the potential to revolutionize clinical decision-making and improve health outcomes, but its utility and impact in healthcare remain poorly understood.
• The article outlines the vast capacity of generative AI to revolutionize healthcare system operations, scientific investigation and patient care, contending that if applied conscientiously, generative AI could enhance medical care quality, fairness and efficiency.

1. Though generative AI holds promise, successfully integrating it into the intricate healthcare system requires carefully planned approaches. The article provides methodical integration plans informed by implementation science principles, which are vital for gradually and effectively transforming complex care environments with new technologies over time.

2. Background
Artificial intelligence (AI) has become an increasingly popular tool in a variety of fields, including healthcare, with the potential to transform clinical decision-making and improve health outcomes [1,2,3]. Generative AI is one area of AI that has gained attention recently for its ability to use machine learning algorithms to generate new data, such as text, images and music [4,5,6,7]. Generative AI is proving to be a change catalyst across various industries, and the healthcare sector is no exception [8]. With its remarkable ability to analyse extensive datasets and generate valuable insights, generative AI has emerged as a powerful tool in enhancing patient care [9], revolutionizing disease diagnosis [10] and expanding treatment options [11]. By harnessing the potential of this cutting-edge technology, healthcare professionals can now access unprecedented levels of accuracy, efficiency and innovation in their practices.

Despite the potential benefits, the utility and impact of generative AI in healthcare remain poorly understood [12, 13]. The application of generative AI in healthcare raises ethical and medico-
legal concerns [14]. Moreover, it is unclear how generative AI applications can be integrated into healthcare service delivery and how the healthcare workforce can utilise them appropriately [15]. Furthermore, it is uncertain how far generative AI can improve patient outcomes and how this can be assessed. Finally, the value of generative AI beyond augmenting clinical and administrative tasks needs to be explored.

Realizing generative AI's vast potential in healthcare requires translational approaches rooted in implementation science. Such approaches recognize technological progress alone will not revolutionize healthcare overnight [16, 17]. Real change requires carefully orchestrated sociotechnical transitions that put people first. Implementation science-based approaches provide generalizable roadmaps grounded in empirical evidence from prior health IT deployments [16]. As such, healthcare leaders pioneering generative AI integration would be well served in leveraging these models to reinforce patient safety, trust and impact [17]. To facilitate the appropriate incorporation and application of generative AI in healthcare, this article aims to provide an overview of the use of generative AI in healthcare followed by guidance on its translational application.

3. Generative AI

Generative AI is a class of machine learning technology that learns to generate new data from training data [18, 19]. Generative models generate data that is similar to the original data. This can be useful in a variety of applications such as image and speech synthesis. Another unique capability is that they can be used to perform unsupervised learning, which means that they can learn from data without explicit labels [8]. This can be useful in situations where labelled data is scarce or expensive to obtain. Furthermore, generative AI models can generate synthetic data by learning the underlying data distributions from real data and then generating new data that is statistically similar to the real data. Generative models differ from other types of machine learning models in that they aim to endow machines with the ability to synthesise new entities [8]. They are designed to learn the underlying structure of a dataset and generate new samples that are like the original data. This contrasts with discriminative models, which are designed to learn the boundary between different classes of data. These models focus on tasks such as classification, regression or reinforcement learning, where the goal is to make predictions or take actions based on existing data.

There are several categories of generative AI, as outlined in

While there are several generative AI models, this article will mainly focus on two models, which are popular in the healthcare context: generative adversarial networks and large language models.

4. Generative adversarial networks

Generative adversarial networks (GANs) differ from traditional generative modelling techniques in their approach to learning [24]. GANs use a game-theoretic framework with competing networks. GANs consist of two neural networks, a generator and a discriminator, that compete against each
other. The generator creates fake data to pass to the discriminator. The discriminator then decides if the data it received is like the known, real data. Over time, the generator gets better at producing data that looks real, while the discriminator gets better at telling the difference between real and fake data. This adversarial training process allows GANs to learn representations in an unsupervised and semi-supervised fashion. In contrast, traditional generative modelling techniques often rely on explicit probabilistic models or variational inference methods.

Recent developments in GANs relating to representation learning include advancements in learning latent space representations [24]. These developments focus on improving the ability of GANs to transform vectors of generated noise into synthetic samples that resemble data from the training set. Some specific examples of recent developments in this area include GANs applied to image generation, semi-supervised learning, domain adaptation, generation controlled by attention and compression [5]. These advancements aim to enhance the representation learning capabilities of GANs and enable them to generate more realistic and diverse samples.

GANs have been used to generate realistic images [24]. These models can learn the underlying distribution of a dataset and generate new images that resemble the original data. This has applications in areas like computer graphics, art and entertainment. Moreover, GANs can be used to augment training data by generating synthetic samples. This can help in cases where the original dataset is small or imbalanced, improving the performance of machine learning models. Synthetic data, created by machine learning algorithms or neural networks, can retain the statistical relationships of real data while offering privacy protection. Synthetic data is also being considered for enhancing privacy.

5. Large language models

Large language models (LLMs) are powerful AI models that have shown promise in various natural language processing (NLP) tasks [25]. In particular, the availability of OpenAI’s GPT-4 [26], Anthropic’s Claude [27] and Google’s PaLM2 [28] has significantly galvanised the progress of not just NLP but the field of AI in general, whereby commentators are discussing achievement of human-level performance by AI [10, 29]. LLMs like OpenAI’s GPT-4 are based on the autoregressive model. An autoregressive model is used to generate sequences, such as sentences in natural language, by predicting a next item based on previous ones [30]. The difference between LLMs and traditional language models lies in their capabilities and training methods [25]. LLMs, like GPT-4, utilise the Transformers architecture, which has proven to be effective for understanding the context of words in a sentence. A transformer uses a mechanism called ‘attention’ to weigh the importance of words when making predictions [31]. This mechanism allows the model to consider the entire history of a sentence, making it a powerful tool for sequence prediction tasks. LLMs are trained on a large corpus of text data. During
training, the model learns to predict the next word in a sentence given the previous words. It does this by adjusting its internal parameters to minimise the difference between its predictions and the actual words that follow in the training data.

One of the key advantages of LLMs is their ability to perform many language processing tasks without the need for additional training data [32]. This is because they have already been trained on a vast corpus of text, allowing them to generate coherent and contextually relevant responses based on the input they receive. This makes them particularly useful as references or oracles for text summarization models. Text summarization is a complex task that involves understanding the main points of a piece of text and then condensing these points into a shorter form. LLMs can be used to generate summaries of text, which can then be used as a reference or ‘gold standard’ for other summarization models [25]. This can help to improve the performance of these models by providing them with high-quality summaries to learn from.

In addition to text summarizations, LLMs have also been used in a variety of other applications [15]. In the realm of text classification, LLMs can be used to automatically categorise pieces of text into predefined categories. This can be useful in a variety of applications, from spam detection in email filters to sentiment analysis in customer reviews. Finally, LLMs have been used for the automatic evaluation of attribution. This involves determining the source or author of a piece of text. For example, an LLM could be used to determine whether a particular tweet was written by a specific person, or to identify the author of an anonymous piece of writing.

It is important to note that while LLMs are powerful, they have limitations [15]. Because they generate sequences one component at a time, they are inherently sequential and cannot be parallelised. Moreover, they are causal, meaning that they can only use information from the past, not the future, when making predictions [33, 34]. They can struggle to capture long-range dependencies because of the vanishing gradient problem, although architectures like Transformers help mitigate this issue.

6. Application of generative AI in healthcare

Generative AI models that facilitate the creation of text and images are seen as a promising tool in the healthcare context [26, 35, 36]. Generative AI can transform healthcare by enabling improvements in diagnosis, reducing the cost and time required to deliver healthcare and improving patient outcomes (Fig. 1).

Use cases of generative AI in healthcare.

Generative AI models like generative adversarial
networks (GANs) and large language models (LLMs) are used to generate various data modalities including text and image data, which are then used for various scenarios including drug discovery, medical diagnosis, clinical documentation, patient education, personalized medicine, healthcare administration and medical education amongst other use cases.

7. Drug discovery as well as Synthetic data generation and data augmentation

Generative AI models are also being used to generate novel small molecules, nucleic acid sequences and proteins with a desired structure or function, thus aiding in drug discovery [11]. By analysing the chemical structure of successful drugs and simulating variations, generative AI can produce potential drug candidates at a much faster rate than traditional drug discovery methods. This not only saves time and resources but can also help to identify drugs that may have gone unnoticed using traditional methods. Moreover, the use of generative AI can also aid in predicting the efficacy and safety of new drugs, which is a crucial step in the drug development process. By analysing vast amounts of data, generative AI can help to identify potential issues that may arise during clinical trials, which can ultimately reduce the time and cost of drug development [11, 38].

In addition, generative AI by identifying specific biological processes that play a role in disease can help to pinpoint new targets for drug development, which can ultimately lead to the development of more effective treatments.

Medical diagnosis

Generative models can be trained on vast datasets of medical records and imagery (like MRIs and CT scans) to identify patterns related to diseases. For instance, GANs have been used for image reconstruction, synthesis, segmentation, registration and classification [5, 9, 37, 39]. Moreover, GANs can be used to generate synthetic medical images that can be used to train machine learning models for image-based diagnosis or augment medical datasets. LLMs can enhance the output of multiple CAD networks, such as diagnosis networks, lesion segmentation networks and report generation networks, by summarising and reorganizing the information presented in natural language text format. This can create a more user-friendly and understandable system for patients compared to conventional CAD systems.

EHRs and other patient records are rich repositories of data, and LLMs can be used to analyse these records in a sophisticated manner [40]. They can process and understand the information and terminology used in these records, which allows them to extract and interpret complex medical information. This capability extends beyond simple keyword matching, as LLMs can infer meaning from incomplete information, and even draw on a vast medical corpus to make sense of the data. Moreover, LLMs can integrate and analyse information from multiple sources within the EHR. They can correlate data from lab results, physician’s notes and medical imaging reports to generate a more holistic view of the patient’s health [10]. This can be particularly useful in complex cases where the patient has multiple conditions or...
symptoms that may be related. LLMs, like GPT-4, have shown medical knowledge despite lacking medicine-specific training [10, 29]. One of the most impressive aspects of these models is their ability to apply this knowledge in decision-making tasks [10]. For example, when presented with a hypothetical patient scenario, an LLM can generate a list of potential diagnoses based on the symptoms described, suggest appropriate tests to confirm the diagnosis and even propose a treatment plan. In some studies, these models have shown near-passing performance on medical exams, demonstrating a level of understanding comparable to that of a medical student [29]. However, limits exist, and the models' outputs may carry certain risks and cannot fully substitute outpatient physicians' clinical judgement and decision-making abilities [14].

Clinical documentation and healthcare administration
LLMs such as GPT-4 and PALM-2 can be used to generate summaries of patient data [41]. This could be particularly useful in healthcare settings where large amounts of data are collected and need to be interpreted quickly and accurately. For instance, an EHR may contain patient data such as medical history, medications, allergies and laboratory results. A generative AI model could be trained to read through this data, understand the key points and generate a concise summary. This summary could highlight critical information such as diagnosis, prescribed medications and recommended treatments. It could also identify trends in the patient's health over time. By automating this process, healthcare providers could save time and ensure that nothing important is overlooked. Furthermore, these summaries could be used to improve communication between different healthcare providers and between providers and patients, as they provide a clear and concise overview of the patient's health status. The ability of LLMs to automate such processes can alleviate the current documentation burden and the consequent burnout many physicians across the world face [41]. Currently, many clinicians, due to organisational policies or health insurance requirements, are required to fill in lengthy documentation beyond what is required for routine clinical care. Studies have shown that many physicians spend over 1 h of time on electronic health record tasks for every hour of direct clinical face time [42]. Additionally, the cognitive load and frustration associated with documentation can reduce work satisfaction, contributing to their burnout [43]. Implementation of natural language processing tools to automate documentation could lessen this burden. An LLM embedded in the relevant information platform can undertake the documentation and provide draft versions for the clinician to approve [40, 41]. For example, hospitals can use LLMs to generate routine progress notes and discharge summaries [44].

Further to this, there is potential for these LLM-based applications to reduce medical errors and capturing missed information by providing a layer of scrutiny when embedded in EHRs [45]. In addition to automating documentation, LLMs integrated into EHRs could help reduce medical errors and ensure important information is not
missed. Studies have found that many hospital patients will experience a preventable medical error, often due to issues like misdiagnosis, prescription mistakes or examination findings that are not followed up correctly [46]. Also, LLMs have the potential to serve as a decision support tool by analysing patient charts and flagging discrepancies or gaps in care [45]. For example, an LLM could cross-reference current symptoms and diagnostics against past medical history to prompt physicians about conditions that require further investigation. Additionally, they could scan medication lists and warn of potential adverse interactions or contraindications.

Generative AI can also be used to automate routine tasks in healthcare, such as scheduling appointments, processing claims and managing patient records [47]. For example, AI models can be used to develop intelligent scheduling systems. These systems can interact with patients through chatbots or voice assistants to schedule, reschedule or cancel appointments. They can consider factors such as doctor's availability, patient's preferred time and urgency of the appointment to optimize the scheduling process. Generative AI can also automate the process of insurance claims. It can read and understand the claim documents, verify the information, check for any discrepancies and process the claim. This can significantly reduce the time taken to process claims and minimise errors. By automating these tasks, healthcare providers can save time and resources and improve the patient experience as they get faster responses and more efficient service.

8. Personalized medicine

Generative AI can analyse a patient’s genetic makeup, lifestyle and medical history to predict how they might respond to different treatments [48]. This is achieved by training the AI on large datasets of patient information, allowing it to identify patterns and correlations that might not be immediately apparent to human doctors. For example, the AI might notice that patients with a certain genetic marker respond particularly well to a specific medication. This information can then be used to create a personalized treatment plan that is tailored to the individual patient's needs. This approach can lead to more effective treatment, as it considers the unique factors that might affect a patient's response to medication. It can also lead to improved patient outcomes, as treatments can be optimized based on the AI's predictions [48].

Generative AI can also be utilised in the field of mental health, particularly in the creation of interactive tools for cognitive behavioural therapy (CBT) [49, 50]. CBT is a type of psychotherapy that helps patients manage their conditions by changing the way they think and behave. Generative AI can be used to create personalized scenarios and responses that are tailored to the individual patient's needs. For example, the AI might generate a scenario that triggers a patient's anxiety, and then guide the patient through a series of responses to help them manage their reaction. This can provide patients with a safe and controlled environment in which to practice their coping strategies, potentially leading to improved mental health outcomes.
9. Medical education and training

In the context of medical education and training, this technology can be used to generate a wide variety of virtual patient cases. These cases can be based on a diverse range of medical conditions, patient demographics and clinical scenarios, providing a comprehensive learning platform for medical students and healthcare professionals [51, 52]. One of the primary benefits of using generative AI in medical education is the ability to create a safe and controlled learning environment. Medical students can interact with these virtual patients, make diagnoses and propose treatment plans without any risk to real patients. This allows students to make mistakes and learn from them in a low-stake setting. Generative AI can also create patient cases that are rare or complex, giving students the opportunity to gain experience and knowledge in areas they might not encounter frequently in their clinical practice. This can be particularly beneficial in preparing students for unexpected situations and enhancing their problem-solving skills. Furthermore, the use of AI in medical education can provide a more personalized learning experience. The AI can adapt to the learning pace and style of each individual, presenting cases that are more relevant to their learning needs. For example, if a student is struggling with a particular medical condition, the AI can generate more cases related to that condition for additional practice.

In addition to creating virtual patient cases, generative AI can also be used to simulate conversations between healthcare professionals and patients [51, 52]. This can help students improve their communication skills and learn how to deliver difficult news in a sensitive and empathetic manner. Moreover, the integration of AI in medical education can provide valuable data for educators. The AI can track the performance of students, identify areas of improvement and provide feedback, helping educators to refine their teaching strategies and curricula.

10. Conclusion

Healthcare systems worldwide face crises of affordability, access and inconsistent quality that now endanger public health [71]. Generative AI presents solutions to begin rectifying these systemic failures through responsible implementation guided by scientific best practices. Validated frameworks like the TAM and NASSS model provide actionable roadmaps for change management, stakeholder alignment and impact optimization [58, 59]. They allow anticipating adoption barriers related to perceived value, usability, risks and more while delineating interventions to drive acceptance. With meticulous planning grounded in evidence, generative AI can transform productivity, insight and care enhancement. Use cases like workflow and documentation automation, personalized predictive analytics, and patient education chatbots confirm vast potential [26, 41, 45], provided the technology supports rather than supplants human expertise. Structured integrations emphasizing clinician control safeguard quality while unlocking efficiency. Thoughtful translation is essential, but implementation science provides proven guidance.
The time for debate has passed. Patients worldwide stand to benefit, and responsible leaders must act urgently. Strategic pilots, iterative scaling and governance emphasizing ethics alongside innovation will realize long-overdue progress. Generative AI cannot single-handedly fix broken systems, but carefully facilitated adoption can catalyse reform while upholding healthcare’s humanitarian obligations. The approach, not just technology, defines success. Guided by wisdom and compassion, generative AI may help restore healthcare ideals so many now lack: quality, affordability and humane care for all.

References


