Acta Medica Ruha

International Journal of Medicine and Health Sciences

Review **Volume: 1 Issue: 3 Year: 2023** **DOI: 10.5281/zenodo.8340188**

The 24 Big Challenges of Artificial Inteligence Adoption in Healthcare

Sağlıkta Yapay Zeka Uygulamasının 24 Büyük Zorluğu

Rebeca Tenajas[1], David Miraut[2]

1 Departamento de Medicina de Familia, Centro de Salud de Arroyomolinos, 28939 Arroyomolinos. España
https://orcid.org/0000-0001-8815-7341

2 Advanced Healthcare Technologies Department. GMV Innovating Solutions. Calle Grisolia 4. 28760 Tres Cantos. España
https://orcid.org/0000-0003-1648-5308

Abstract

Introduction: The integration of Artificial Intelligence (AI) into medical disciplines has shown significant potential. However, while an abundance of literature is enthusiastic about the potential and promises of AI, particularly in medical diagnostics, there is a distinct of discussion concerning the multitude of challenges associated with its widespread adoption in practical medical settings.
Objective: This study aims to thoroughly analyze the challenges associated with adopting artificial intelligence technologies in medical practice. It provides a realistic perspective on the progress of this technology, countering the often overly idealized viewpoints that primarily showcase advancements in prototypes and technological demonstrators within controlled laboratory conditions.
Method: The research design of this study is grounded in the method of document analysis/review. In this context, numerous scientific works were explored through platforms such as Google Scholar, PubMed, BioMed Central, Cochrane, and various scietific databases. Access to articles was obtained, followed by meticulous data analysis and assessments. Search criteria were adjusted based for each of the challenges under examination.
Results:A total of 24 significant challenges have been identified, intricately interconnected, and dissected using examples that illustrate the maturity level of AI-based developments within the medical domain. These challenges have been categorized into three main categories based on their nature. Each section has been wriiten in a way that can be independently comprehended. The future holds great promise, as underscored by numerous articles showcasing the remarkable advancements arising from the synergy between medicine and artificial intelligence. Hence, there is a need to develop critical thinking to discern the benefits, current limitations, and new paths to overcome them.
Conclusion: None of the challenges holds greater importance than the others. The evolution of artificial intelligence in medicine entails collectively overcoming these challenges, using strategies to maximize benefits for both patients and medical experts.

Keywords: Artificial Intelligence, Radiology, Learning Curve, Deep Learning, Health Centers.

Özet

Giriş: Yapay Zeka'nın (YZ) tıbbi disiplinlere entegrasyonu önemli bir potansiyel göstermiştir. Ancak, AI'ın potansiyeli ve vaatleri hakkında coşkulu bir literatür bolluğu olmasına rağmen, özellikle tıbbi teşhislerde, pratik tıbbi ortamlarda yaygın olarak benimsenmesiyle ilişkilendirilen birçok zorluk hakkında tartışma eksikliği vardır.
Amaç: Bu çalışma, tıbbi uygulamada yapay zeka teknolojilerini benimseme ile ilişkilendirilen zorlukları ayrıntılı olarak analiz etmeyi amaçlamaktadır. Bu teknolojinin ilerlemesi hakkında gerçekçi bir perspektif sunar, kontrol altındaki laboratuvar koşullarında prototiplerin ve teknolojik göstericilerin ilerlemesini öne çıkaran aşırı idealize edilmiş görüşlere karşı koyar.
Yöntem: Bu çalışmanın araştırma tasarımı, belge analizi/inceleme yöntemine dayanmaktadır. Bu bağlamda, Google Scholar, PubMed, BioMed Central, Cochrane ve çeşitli bilimsel veritabanları gibi platformlar aracılığıyla

Corresponding Author: David Miraut, e-mail: dmiraut@gmv.com
Received: 13.08.2023, Accepted: 03.09.2023, Published Online: 30.09.2023
Cite: Tenajas R, et al. The 24 Big Challenges of Artificial Inteligence Adoption in Healthcare. Acta Medica Ruha. 2023;1(3):432-467. https://doi.org/10.5281/zenodo.8340188

birçok bilimsel çalışma keşfedildi. Makalelere erişim sağlandı, ardından dikkatli veri analizi ve değerlendirmeleri yapıldı. İncelenen her zorluk için arama kriterleri ayarlandı.
Bulgular: Toplamda 24 önemli zorluk tespit edildi, bunlar iç içe geçmiş ve tıbbi alandaki AI tabanlı gelişmelerin olgunluk seviyesini örneklendiren örneklerle ayrıntılı olarak incelendi. Bu zorluklar doğalarına göre üç ana kategoriye ayrılmıştır. Her bölüm, bağımsız olarak anlaşılabilir bir şekilde yazılmıştır. Gelecek, tıp ve yapay zeka arasındaki sinerjiden kaynaklanan dikkat çekici ilerlemeleri vurgulayan sayısız makale ile büyük vaatlerde bulunmaktadır. Bu nedenle, faydaları, mevcut sınırlamaları ve bunların üstesinden gelmek için yeni yolları ayırt edebilmek için eleştirel düşünmeyi geliştirmek gerekmektedir.
Sonuç: Zorlukların hiçbiri diğerlerinden daha önemli değildir. Tıpta yapay zekanın evrimi, hem hastalar hem de tıbbi uzmanlar için faydaları en üst düzeye çıkarmak için stratejiler kullanarak bu zorlukların kolektif olarak üstesinden gelinmesini gerektirir.

Anahtar Kelimeler: Yapay Zeka, Radyoloji, Öğrenme Eğrisi, Derin Öğrenme, Sağlık Merkezleri.

INTRODUCTION

The integration of Artificial Intelligence (AI) into medical disciplines has shown significant potential. However, while an abundance of literature is enthusiastic about the potential and promises of AI, particularly in medical diagnostics, there is a distinct of discussion concerning the multitude of challenges associated with its widespread adoption in practical medical settings (1).

Research prototypes and technological demonstrators –usually shown by researchers from academic institutions and startup enterprises– are considerably different from full-fledged, operational systems in active clinical environments. Medical practice requires that any technological adjunct to it, AI-based or otherwise, adheres to strict standards of design, planning, and validation. Such observance ensures that these systems not only augment the abilities of medical professionals but do so in a manner as reliable as any certified medical equipment (2).

This paper provides a comprehensive exploration of 24 significant challenges that are intrinsic to the adoption of AI technologies in healthcare. These challenges are catalogued into three main themes: eight associated with the development of software AI-based solutions; another eight linked to their adaptation, integration, and operation within existing healthcare facilities; and the final eight focusing on the ethical, societal, and equality implications in healthcare access. We must that each challenge is complex enough to be the subject of an independent research article. This review, however, aims to offer a balanced andcomprehensiveexploration of the real-world intricacies facing the deployment of AI solutions in clinical practices. While the merits of AI have been highlighted in a prior research work (3) published in this journal, the rate at which these advantages mature into accessible solutions for clinicians appears more restrained.

It is not surprising that Mr. Sayar, in his recent article titled "Use of Artificial Intelligence in Medicine" (3), has focused his attention on the technological advancements impacting the field of Radiology. The recent advances in computational vision and the evolution of Deep Learning have synergistically elevated Radiology's potential for enhanced diagnostics and treatment planning (4). The discipline's intrinsic dependence on imaging has always encouraged a close relationship with evolving computational technologies (5). Nevertheless, the journey from a proof-of-concept algorithm in a controlled environment to an operationally reliable tool in diverse clinical settings is riddlewith a range of complex challenges. The realization of AI's potential in medicine is not simply a matter of technological advancement but also involves skilful manoeuvring through the complexities of software development, deep learning training, the peculiarities of healthcare operations, and ethical considerations.

The software development for AI in healthcare is not merely a replication of traditional software engineering. The vast and diverse data requirements, the need for interpretability, and the twists and turns of deploying machine learning models in production are unique and demand special attention (6). Furthermore, within the security of healthcare facility, the integration of AI solutions requires the seamless merger with legacy systems, the guarantee of uninterrupted operation, and provisions for regular updates without compromising patient care.

Moreover, the potential for AI to reshape healthcare is not solely technological; it is deeply societal. The questions of who benefits from AI, how it impacts the patient-doctor relationship, and how it might inadvertently exacerbate health disparities are of paramount importance (7). These considerations ensure that the technology remains a tool for equitable healthcare advancement and doesn't devolve into a source of fragmentation or disparity.

Artificial intelligence, therefore, has shown significant potential in shaping the future of healthcare, with impacts spanning across diagnostics and personalized medicine. However, it should be noted that these potential benefits come with a set of challenges and ethical considerations. Therefore, while the promise of artificial intelligence in healthcare is substantial, careful consideration and management of these challenges are necessary in order to fully realize its potential.

HEALTHCARE AI TECHNOLOGICAL DEVELOPMENT CHALLENGES

The first section is focused on the 8 challenges related to the development of artificial intelligence applications in the field of medicine. While some of these challenges are common to all types of software built on Deep Learning algorithms, the majority are deep-rooted to the medical domain, given the sensitivity of patient data and their potential ailments.

Voracity of large and curated datasets for Deep Learning medical algorithms

Datasets are the raw material to build deep learning algorithms, their availability, size, and quality fundamentally shape the performance of these models. As remarked by Mr. Sayar (3), one salient example is the field of medical imaging, where large datasets of curated images are used to train machine learning models to detect and diagnose disease, improving both speed and accuracy over human examination alone (8). The impact of deep learning algorithms on medical diagnosis, prognosis, and treatment planning is currently being widely studied and acknowledged across the globe (9). However, the production and acquisition of large, high-quality and labelled datasets, particularly in healthcare, are fraught with challenges that slow the pace of development and the application of these powerful technologies.

Data is often scarce in healthcare settings due to factors including but not limited to data privacy regulations, the nature of rare diseases, and the gradual pace of public medical research data generation. Given that deep learning models require large amounts of data to accurately learn and predict outcomes, scarcity poses a significant challenge. Privacy regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe, restrict the use and sharing of medical data, further limiting its availability (10). The development and maintenance of large databases require extensive resources, making it particularly challenging for rare diseases where data is inherently limited (11) (see section 0).

Medical data is also typically unlabelled, which necessitates the laborious task of manually annotating the data. A team of professionals with medical expertise is required to ensure accurate annotation, which can be a lengthy and costly process. Unlike other fields, healthcare data cannot be crowd-sourced due to its specialized nature and privacy concerns. Thus, the labelling process becomes a bottleneck in the use of deep learning for healthcare applications (4).

The value of curated datasets, which have been carefully collected, cleaned, and labelled, cannot be emphasized enough. However, their creation involves a significant investment of time, effort, and expertise. The benefit is twofold: high-quality datasets improve the performance of individual algorithms, and they allow researcher to advance the field as a whole by providing benchmarks against which new algorithms can be compared. This standardization, achieved through the use of curated datasets, propels forward the development of deep learning applications in medicine (12).

Despite these challenges, the healthcare field has seen some notable successes in the acquisition of medical data. Institutions like the Stanford School of Medicine, the Massachusetts General Hospital, and the Mayo Clinic, among others, have been successful in building large databases. Public datasets such as the MIMIC-III (Medical Information Mart for Intensive Care III), which contains de-identified health data associated with over forty thousand patients, or EchoNet-LVH dataset, which includes 12,000 labeled echocardiogram videos and human expert annotations (measurements, tracings, and calculations) to provide a baseline to study cardiac chamber size and wall thickness, are serving as significant sources for deep learning research (13–16).

The procurement of large and curated datasets will be a persistent challenge in the application of deep learning algorithms in healthcare, but it is a hurdle that can and must be overcome. It will take concerted effort and cooperation among multiple stakeholders, including medical institutions, regulatory bodies, and technology companies. As this issue continues to be addressed, the healthcare field will be poised to reap the full benefits of deep learning technologies.

Data quality and integrity, a prerequisite for the development of reliable AI models

The development of trustworthy artificial intelligence models in healthcare heavily relies on data quality and integrity (17). While this factor might seem self-evident, its implementation in practice proves to be a challenge. Although most Deep Learning systems are slightly robust to noise, the well-known 'garbage in, garbage out' principle serves as a reminder of how crucial high-quality data is in influencing the performance of AI systems. This factor Is especially important in critical sectors, like Healthcare, as performance of AI systems is directly influenced by the quality of the input data (18). Thus, the task of ensuring the cleanliness, completeness, and representativeness of health data is both a significant challenge and an essential area for improvement.

Clinical data is the primary source of information for health-focused AI applications. This kind of data, derived from a variety of sources including electronic health records, clinical trials, or digital wearables, is often fragmented and laden with noise, biases, or errors. The fragmentation issue arises due to the disparate nature of health data sources, each carrying their own unique structure and format (19). Merging these disparate data sources into a cohesive whole that can be used by an AI model often requires substantial cleaning and reformatting efforts.

The presence of noise, biases, or errors introduces additional complexities. These inaccuracies can act as confounders, impeding the generalizability of AI models and potentially leading to flawed predictions. The noise in data may arise due to measurement errors, data entry mistakes, or even inherent biological variability (20). Biases can be introduced during the data collection process, often stemming from the population sample, the study design, or the collection method itself (21). Data errors, on the other hand, can arise from a variety of reasons, including data entry mistakes, transmission errors, or missing data (22).

For these reasons, several approaches are usually undertaken to enhance data quality and integrity. Methods to address the issue of fragmentation include the development of standard data formats and structures that can facilitate data integration (23). Noise reduction techniques and validation checks can help minimize errors and ensure data accuracy (24). As for biases, robust study design and sample selection techniques can help limit their introduction in the first place, and statistical methods can help adjust for them once present.

To illustrate, The Cancer Genome Atlas (TCGA) provides an example of a successful integration of diverse data types from various sources. By standardizing data structures and formats, TCGA was able to create a rich resource for cancer genomics research (25). Similarly, the Framingham Heart Study, a longitudinal study that has been ongoing since 1948, demonstrates the effectiveness of meticulous study design and robust statistical methods in minimizing biases and ensuring the quality and integrity of data collected over time (26).

Although the task of maintaining data quality and integrity presents a substantial challenge, it is not unconquerable. The prospect of AI in healthcare hinges on the capability to confront these challenges directly, employing robust strategies and methodologies to the cleanliness, completeness, and representativeness of health data.

Lack of interpretability and transparency in AI models

Since the dawn of artificial intelligence in medicine, the methodologies for extracting meaningful patterns from medical data have experienced remarkable evolutions. In the initial stages of AI-driven medical research, considerable emphasis was placed on hand-crafting features from the data, which then served as an input to machine learning algorithms (27). This manual feature engineering required lot of effort to be designed, given the constraints of computational power and the limited understanding of algorithmic designs that could capture intricate data patterns. However, these features were often dependent on the domain knowledge and intuition of the researchers, which introduced potential biases and might have omitted significant, non-intuitive characteristics (28).

Transitioning from these early strategies, the advent of deep learning and, more specifically, deep neural networks has brought about a paradigm shift in AI approaches within the medical domain (29). These models are often categorized as "end-to-end" since they are designed to learn features directly from raw data, bypassing the often labor-intensive process of manual feature design and extraction. While such models have been shown to outperform their predecessors in numerous medical applications, such as diagnostic imaging and predictive analytics, they come with a different set of challenges and considerations.

One significant aspect of these neural models is the interpretability, or more aptly, the lack thereof (30). The features learned by deep networks tend to be more abstract and non-linear, making them more difficult to interpret in a comprehensible manner by human researchers. This raises questions about transparency, trustworthiness, and the very nature of knowledge in AI-driven medical applications. While these models provide results with high accuracy, the inability to understand the reasoning behind these predictions could be a potential stumbling block in the broader acceptance and application of such models in clinical settings (31).

The concern surrounding the 'black box' nature of AI models, first delineated by Holzinger et al. in 2017, is a palpable barrier to their extensive adoption, particularly in the medical community (32) and end-to-end models.

A black box model is defined by its extreme mathematical complexity, which obscures the very nature of its internal decision-making mechanisms. In essence, a black box AI model takes an input, processes it through a complex web of computations, and outputs a result, all without offering clear insight into how that result was reached. This absence of transparency and interpretability can be disturbing for professionals in any field. Still, in medicine, where decisions directly impact human lives and can have profound consequences, the lack of understanding of how decisions are reached is especially concerning.

A physician's trust in a tool or a procedure often arises from a clear understanding of its mechanisms, and it is this clarity that AI models often lack (33). The integrity of the patient-doctor relationship and the implicit trust placed in a physician's decision-making are not to be taken lightly. These factors are fundamentally based on the understanding of the methods, procedures, and decisions made. When AI is introduced without proper interpretability, it jeopardizes this trust and is likely to be met with resistance.

Moreover, the opacity of AI decision-making raises ethical and legal questions. The moral responsibility of a decision in medicine usually lies with the clinician making the decision. However, if an AI tool is involved, and an error occurs, it may be difficult to attribute responsibility (34). Wachter et al. (35) pointed out the challenges in this regard, particularly because a clear comprehension of how the AI arrived at its decision is often difficult to trace.

Explainable AI emerges as a promising solution to these concerns. The creation of AI models that maintain high performance while also offering transparency in their workings is an hot field of research (36). These models strive to strike a balance between the accuracy of a deep learning model and the transparency of a simpler model.

In theory, the development of explainable AI seems straightforward, but in practice, it is a complex undertaking, fraught with challenges. One such challenge is maintaining the balance between model complexity and interpretability. Simpler models are typically more interpretable, but they often compromise on accuracy. On the other hand, more complex models, while highly accurate, sacrifice transparency and interpretability (37).

Research into explainable AI is a pressing subject, given the rising usage of AI in various fields, particularly in healthcare. Designing these models is not just a scientific challenge but also an issue of trust, ethics, and legal responsibility. From a scientific perspective, it represents one of the frontiers of AI research. From an ethical perspective, it poses questions about responsibility, transparency, and trust. From a legal standpoint, it begs the question of culpability when errors occur (35). From the government certification agencies for medical products, there is a growing requirement for a comprehensive examination of the interpretability of AI models used in medicine, despite it being a complex challenge that depends on each type and architecture of neural network (usually preserved as industrial secret).

The integration of AI into clinical practice must grapple with the ongoing issue of the black box problem. Physicians' trust, ethical responsibility, and legal accountability are vital considerations that must be addressed as we advance further into this technological age. The development of explainable AI offers a potential solution, but substantial challenges must be overcome before we can fully reap the benefits of AI in a responsible, transparent, and trustworthy manner.

The task of handling unstructured medical data

The digital revolution in the medical sector has seen an exponential increase in the generation of computerized medical data. Much of this data, encompassing clinical notes, radiology images, pathology reports, among others, exists in an unstructured format, posing a considerable challenge in leveraging this data for advanced computational applications such as artificial intelligence. As a result, one significant task faced by the medical community is the conversion of unstructured data into a form that can be feed into AI algorithms (19). It is worth noting that this process is not trivial; it involves labor-intensive and time-consuming operations that require sophisticated methodologies and tools.

The complexity of the process arises from the arbitrary and diverse nature of unstructured medical data, which is typically characterized by a lack of predefined format, making it difficult to classify, interpret, or utilize in AI models (38). It also contains a wealth of information critical to patient care and clinical decision making, including details about a patient's history, condition, treatment options, and prognosis. Consequently, the conversion process must not only be meticulous but also comprehensive, ensuring that the rich information content is preserved while being made accessible to AI models.

Nevertheless, the benefits of achieving this task are significant. The ability of AI models to handle and interpret unstructured data could significantly increase the utilization of medical data, thereby augmenting the quality of AI-assisted clinical decision-making. For instance, in a study by Rajkomar et al. (39), a deep learning model was developed that could make accurate predictions about various patient outcomes, such as length of hospital stay and readmission likelihood, by utilizing electronic health record data, including unstructured data like clinical notes. This example illustrates the potential value of unstructured data in enriching AI models, which in turn can enhance the quality of care delivered to patients.

However, the development of AI models capable of managing unstructured medical data remains a challenging task due to various reasons. For example, the unique characteristics of medical data, including its diversity, sensitivity, custom acronyms, and size, necessitate specialized and advanced techniques for processing and handling it (40). Furthermore, the absence of standardization in the representation of unstructured medical data adds to the complexity of the task. Moreover, ethical and legal considerations surrounding the use of medical data, particularly in the context of AI, present another layer of challenges that need to be addressed.

Overcoming these challenges necessitates a multidisciplinary approach that combines expertise in areas such as data science, machine learning, medical informatics, and healthcare law and ethics. It also requires collaboration between various stakeholders, including clinicians, data scientists, ethicists, and policymakers, to ensure the responsible and effective use of unstructured medical data in AI (41).

Handling unstructured and multimodal medical data is a challenging yet essential task in the field of medical AI. It demands sophisticated methodologies, collaborative efforts, and the addressing of ethical and legal considerations. However, achieving this task could significantly enhance the quality of AI-assisted clinical decision-making.

The development of AI tools that can handle multiple comorbidities

Initial efforts in AI predominantly targeted singular diseases (12). This approach aligns with traditional biomedical research, which generally prioritizes understanding individual diseases in isolation (42). While this strategy is valuable, it fails to account for the intricate relationship between multiple concurrent diseases in a patient. This issue is a stark contrast to real-world clinical practice, where patients often present with multiple concurrent diseases or comorbidities (43). In reality, patients often present with comorbid conditions that can modify disease expression, prognosis, and response to treatment (44). Recognizing and addressing this discrepancy requires the development of advanced AI models capable of managing this level of complexity.

Moreover, developing AI models capable of handling multiple comorbidities is far from trivial. Comorbidities are not merely a collection of individual diseases but a complex network of interconnected conditions (45). Consequently, AI models must account for the non trivial interconnections among different diseases. Moreover, these models must factor in how the presence of multiple comorbidities might affect a patient's health status, treatment response, and prognosis (46). Consequently, the development of such models involves sophisticated computational methodologies and data management, requiring extensive knowledge in both medical science, software engineering and AI technologies (47).

An important step in this development process is the availability of high-quality, comprehensive patient data (see section 0). The creation of these models depends on access to extensive datasets detailing patients' disease history, including all concurrent conditions and their progression over time. While such data is typically available in electronic health records (EHRs), several barriers exist to its utilization (48). These include issues relating to data privacy and security, data quality and consistency, and the lack of standardized methodologies for extracting meaningful information from EHRs (24).

Despite these challenges, significant progresses have been made in the development of AI models capable of handling multiple comorbidities. For example, a model proposed by Choi et al. utilizes graph-based deep learning to predict future comorbidities in patients (49). This model, known as the Medical RelationNet, uses EHR data to construct a patient-disease-time graph, which it then uses to predict the future onset of comorbid conditions. While this model represents a significant advancement in this field, more work is needed to develop AI models that can not only predict comorbidities but also provide personalized treatment recommendations (see section 0).

Developing AI algorithms that can effectively deal with rare diseases

The capacity of artificial intelligence and machine learning (ML) to manage and to diagnose rare diseases represents a domain still in its nascent stage. The fundamental challenge resides in the scarcity of data pertinent to these conditions. These diseases, while individually infrequent, collectively affect a significant portion of the global population (50).

In the domain of machine learning and Statistics, the "Long Tail" refers to the distribution of classes where a limited number of classes have abundant instances (the "head") and a large number of classes have few instances (the "tail"). In this distribution, rare diseases fall into the "tail" segment, thus posing unique difficulties to the ML algorithm development. Despite the considerable improvements in AI techniques, algorithms are still primarily reliant on large datasets for effective training (51).

The relative paucity of shared public data concerning rare diseases affects the performance of AI models in two primary ways. First, the limited availability of data impacts the model's ability to recognize the distinguishing patterns and features of rare diseases (52). Second, the lack of variability in the sparse data hinders the generalizability of the model. These two facets form the core obstacles in training AI models to effectively manage rare diseases.

One proposed solution to this problem involves the utilization of transfer learning. Transfer learning enables the adaptation of a pre-trained model on a new task with comparatively fewer data, exploiting the commonalities and differences between the tasks. This method has been successfully applied in several medical imaging contexts (53,54). Nonetheless, it requires a careful design to ensure the similarity between the source and target tasks, which might not always be feasible in the case of rare diseases.

In parallel, data augmentation techniques can be employed to artificially inflate the limited dataset size. Methods such as synthetic minority over-sampling technique (SMOTE) have been employed with encouraging results in other areas of medicine (55). Moreover, advances in synthetic data generation, offer promising directions for addressing data scarcity in rare diseases (56).

Collaboration and data sharing among healthcare institutions and researchers around the globe could markedly enhance the capacity to develop efficient AI models for rare diseases. In an exemplar case, an international collaborative project called 'Genomics England' has resulted in the sequencing of 100,000 genomes from around 85,000 NHS patients, including many with rare diseases, creating a valuable dataset that could be instrumental for the development of certain AI models (57).

Ensuring the continuous learning and adaptation of AI models while maintaining their safety and performance

The advancement and expansion of medical knowledge necessitate a periodic update of the AI models. Notwithstanding the perceived benefits of such updates, they carry potential alterations in model behaviour, thereby eliciting unforeseen consequences (6).

These AI models might play become 7/24h assistants in diagnostics, prognostics, treatment strategies, and managing public health emergencies, thereby emphasizing the need for ensuring safety and efficacy. The premise of AI models' success rests on their ability to learn, adapt, and improve performance continuously over time, which implies training on expansive and diverse datasets (1). The concept of frozen or static models falls short in the medical domain, given the inherent dynamic nature of medical science and the necessity of staying up-to-date with new research, best practices, and clinical guidelines (33).

Integrating new knowledge into AI models necessitates an understanding of the repercussions this might have on their behaviour. These transformations, although intended to enhance the performance and applicability of the model, might bear the potential to sway the model's behaviour, inducing unanticipated outcomes (58). Therefore, any update process needs to be structured and monitored to ensure a balance between the incorporation of new knowledge and the safety and efficacy of the model. The focus, thus, is to create secure environments that enables continuous learning and adaptation without compromising model safety or performance.

Regulating updates to AI models poses a set of formidable challenges (see section 0). The maintenance of an AI model's safety and efficacy is an iterative process, involving a cyclic methodology of update, testing, validation, and deployment. It is necessary to ensure that an update does not lead to "catastrophic forgetting," where the model may lose its ability to perform tasks it was previously trained for. One such method to mitigate this issue is 'elastic weight consolidation,' which protects the parameters important for previously learned tasks during the learning of new ones (59), although it is quite dependent on the underlying Deep Learning architecture. Another approach is the development of automated and continuous monitoring systems to detect and mitigate unexpected behaviour changes in deployed models (60).

However, not all model updates result in significantly altered behaviours. In fact, some updates, such as those involving bug fixes or minor adjustments, may be largely benign and have little to no impact on the safety or performance of the AI system. This highlights the need for a systematic and graded approach to updates, which carefully considers the potential impact of each change and prioritizes thorough testing and validation for more significant updates.

The challenge of balancing updates with safety and efficacy of AI models is another pressing issue. A framework is required that enables the incorporation of novel medical knowledge, supports the ability to adapt and improve, and ensures the safety and performance of the model are upheld. Such a framework will benefit from collaborative efforts from the spheres of medical science, artificial intelligence, and regulations.

Designing user interfaces for AI tools that are intuitive and user-friendly

The employment of artificial intelligence in healthcare is increasing rapidly and is reshaping many aspects of medical practice, from diagnosis and treatment planning to patient management and follow-up (2). This proliferation has engendered an imperative need for effective user interfaces (UI) that enhance rather than obstruct the communication between the human end-users, predominantly healthcare providers, and AI tools. The UI is essentially the point of interaction between the users and the AI tools, and its design determines the degree of utilization and adoption of such systems in the clinical practice (61).

Efficiency, intuitiveness, and user-friendly attributes are often designed carefully in UIs for AI tools (62). Efficiency sums up the need for a swift and precise response, necessitating the interface to possess capabilities for timely data retrieval and presentation, fast data entry, and immediate responses to user commands. An intuitive design allows users to predict the behaviour of the AI tools without requiring extensive training or having to resort to a user manual. User-friendly interfaces, on the other hand, prioritize the ease of use and understandability, making them accessible and comfortable for users with varying levels of technical proficiency.

A judicious integration of these principles into UI design of AI tools could improve the effectiveness and acceptance among healthcare providers. Research conducted by Patel et al. has demonstrated that the appropriate design of UIs for AI tools in healthcare could result in improved user satisfaction, increased productivity, and more precise medical decision-making (63).

Yet, suboptimal UI design may present considerable challenges, affecting user acceptance and utilization. A poorly designed UI can increase cognitive load, resulting in user frustration and decreased productivity (64). More concerning is the potential for patient harm. For instance, incorrect interpretation or misunderstanding caused by a baroque interface could lead to misdiagnoses, inappropriate treatments, or other detrimental clinical decisions.

As we delve further into the implications of UI design in the context of AI in healthcare, it's valuable to reflect on the example of IBM's Watson for Oncology. As reported in Strickland (65), despite the system's powerful underlying AI technology, physicians were often frustrated by the complex and unintuitive interface, leading to low adoption rates and criticisms of its applicability in real-world practice. This experience underscores the critical role of interface design in AI tool acceptance and effectiveness.

Therefore, the design of UI for AI tools in healthcare should be a meticulous and thoughtful process, requiring a multi-disciplinary approach that includes not only designers and AI specialists but also end-users, the healthcare providers themselves (66). This would ensure that the AI tools are indeed suited to the real-world, high-demand environment of healthcare provision, aligning with the unique requirements and workflows of the sector.

While the potential of AI in healthcare is vast, the success of AI tools will be dependent not just on the sophistication of the underlying algorithms, but also on the effectiveness and usability of the interfaces through which users interact with these systems. Future research should continue to explore the optimal design principles for AI tool interfaces in healthcare, ensuring their alignment with user needs and clinical workflows.

HEALTHCARE AI INTEGRATION AND OPERATIONALIZATION CHALLENGES

The current technological revolution is by no means the first experienced by the medical world. Medicine has been closely tied to disruptive technological changes since its inception (67), as a significant portion of human ingenuity is invested in improving the quality of life for others and, consequently, in combating illness with all the tools at our disposal.

This factor necessitates that the implementation of AI-based solutions must compete and adapt within a framework of previously deployed technologies, which have seldom been designed to be scalable or collaborate with systems of such advanced nature.

The subsequent sections delve into the challenges surrounding these matters, which must evolve concurrently with the adoption of these technologies that offer us so many advantages (3).

The integration of AI into clinical workflow

Artificial intelligence applications in healthcare hold a remarkable potential to transform clinical practice through streamlining workflows, improving diagnostic accuracy, and enhancing patient care (67). Thus, it is necessary to consider how these systems integrate into existing healthcare practices, how they interact with healthcare information technology (IT) systems, and how user-friendly they are to the clinicians who will be using them in real-time.

In the current landscape, many AI tools have been created as stand-alone systems, posing a significant challenge to their incorporation into the clinical workflow. This separation from established systems can result in additional work for healthcare providers, as these systems often require additional and customized interface and data entry process to interact with. Rather than augmenting the clinician's work, these stand-alone AI systems may disrupt the delivery of patient care by adding another layer of complexity to the provider's tasks. This issue is highlighted in the research by Blease et al., where the authors underline the necessity of integrating AI systems into the existing healthcare IT infrastructure to ensure seamless operation and avoid unnecessary burden on clinicians (68).

Addressing this challenge necessitates the development of AI systems that are not only interoperable with existing healthcare IT systems but also align well with the clinical workflow. This factor involves a full understanding of the needs, preferences, and routines of the healthcare providers who will be using these systems in their practice. The AI tools should be designed to augment the clinician's work rather than replace it, offering assistance in real-time without disrupting the clinical workflow. Integration can be facilitated by creating AI systems that are compatible with existing electronic health records (EHRs) and other healthcare IT systems. Such an approach was presented by Rajkomar et al. where the authors showed how machine learning models can be integrated into EHRs to predict medical events (39).

However, integration and interoperability are not sufficient to ensure the effective use of AI in clinical practice. Another key consideration is the training and preparation of end-users: the clinicians and their assistants (69). Despite the advancement of AI technologies, the proficiency and confidence of clinicians in using these tools remain limited. This limitation can lead to a lack of trust and adoption, inhibiting the full potential of AI in healthcare. As noted by Johnson et al., it is necessary to provide healthcare providers with adequate training to use AI tools effectively(69), including an understanding of their underlying principles, the potential benefits, and the limitations (70).

The technical difficulties of integrating AI with existing health IT systems

The implementation of AI into existing health IT systems poses a significant technical challenge, especially considering that many healthcare organizations continue to use older IT infrastructure, which are not easy to integrate to improve patient monitoring and administrative processes powered by AI-based software (71).

A fundamental issue to address in the integration of AI with current health IT systems is the heterogeneity and fragmentation of these systems. Many hospitals and health institutions operate with a multitude of disparate and often outdated IT systems, not designed for interoperability (1). These systems tend to be siloed and only compatible with same vendor tools, making data extraction, sharing, and utilization for AI applications a complex task (72).

For example, the adoption of the Epic System by Kaiser Permanente, a leading healthcare provider in the United States, highlights the struggle to streamline patient records across different facilities due to the significant variance in data structure and language (see section 0).

Moreover, data quality and consistency present notable hurdles. Deep Learning models rely on large, structured, and high-quality datasets to produce meaningful and accurate results (6). However, in healthcare settings, data are frequently unstructured, heterogeneous, and inconsistent due to variations in input methods and procedures, potentially leading to issues with accuracy and reliability in AI applications (73) (see section 0). Furthermore, the quality of data can be compromised due to errors, omissions, and biases in recording practices, which pose significant challenges to the training of AI models. For instance, a study in the UK demonstrated that inconsistent coding practices in electronic health records led to substantial discrepancies in stroke event recording, an issue that could significantly impact AI models' performance in predicting stroke events (74).

Additionally, the technological readiness and capacity of the existing health IT systems play a crucial role in the successful integration of AI. These older systems often lack the required computational power and capabilities for advanced AI applications (1). Consequently, the process may necessitate substantial upgrades or complete renovations of the existing infrastructure, which could entail significant financial investment and potential workflow disruptions (47).

Furthermore, the complexities related to data privacy, security, and regulatory compliance cannot be understated. Health data are sensitive, and their management involves strict regulatory frameworks to ensure patient privacy and data security (75). The integration of AI solutions into health IT systems necessitates robust mechanisms to ensure data are handled, processed, and stored in ways that comply with laws such as the Health Insurance Portability and Accountability Act in the United States or the General Data Protection Regulation in the European Union.

AI validation and regulation

The introduction of artificial intelligence in healthcare has opened new possibilities for diagnosis, prognosis, and management of disease conditions. Its effective and safe application in real-world clinical settings necessitates a thorough validation process that recognizes and addresses the innate complexities and variations in clinical practices, patient populations, and disease patterns. These elements have been identified to significantly impact the effectiveness of AI in actual healthcare environments, leading to its performance that is often not as promising as it is in controlled research scenarios (76).

The heterogeneity inherent in real-world patient populations and clinical practices poses a unique challenge to AI validation. Disease manifestation and patient response can vary significantly due to a myriad of factors such as age, gender, ethnicity, comorbidities, and socioeconomic status, among others (77). This considerable variation requires AI systems to be adaptable and generalizable to deliver precise and effective care. Similarly, the shifting patterns of diseases over time, influenced by factors such as lifestyle changes, genetic evolution, and environmental changes, necessitate AI's adaptability to continue being effective (78).

To achieve this level of adaptability and generalizability, rigorous large-scale, multicenter clinical trials are recommended as they expose AI systems to a wide range of clinical and demographic variations. These trials serve as a crucial validation process that evaluates not only AI's accuracy but also its safety, effectiveness, and ability to generalize its learning across diverse scenarios (79). However, conducting these trials comes with its own set of challenges. Logistical difficulties, such as data sharing issues, patient privacy concerns, and difficulty in standardizing trial protocols across different centers are some of the many hurdles. Furthermore, the financial burden associated with these large-scale trials can be substantial, often proving prohibitive, particularly for smaller companies and institutions (80).

The second layer of complexities in validating AI for healthcare arises from the current regulatory frameworks. AI, in its essence, is a continuously learning and adapting system. This characteristic, while being one of its strengths, also presents unique regulatory challenges (see section 0). The Food and Drug Administration (FDA) and similar regulatory bodies have long-established protocols for approving medical devices and interventions. However, these traditional protocols are ill-equipped to handle the dynamic nature of AI systems. The current regulations demand a static validation process that is incompatible with the continuous learning and adaptation of AI (81).

To address these challenges, there is an immediate necessity for innovative validation strategies and updated regulatory norms. The validation strategies need to take into consideration the heterogeneity of real-world healthcare settings and the dynamic nature of AI, designing comprehensive protocols that allow rigorous testing of these systems. Updated regulatory norms, on the other hand, should accommodate the evolving nature of AI systems while ensuring patient safety and efficacy of AI interventions (82).

However, the development and implementation of such strategies and norms will require concerted efforts from all stakeholders, including researchers, clinicians, regulators, and patients. Moreover, a balance needs to be struck between encouraging innovation and maintaining stringent safety and efficacy standards. It is, therefore, a subject of ongoing research and discussion.

Managing the expectations of healthcare providers and patients regarding AI

An overestimation of AI's capabilities or an underestimation of its limitations may lead to disillusionment among healthcare providers and patients, potentially inhibiting the acceptance and integration of AI in healthcare settings.

Understanding the transformative power of AI in healthcare requires a nuanced appreciation of its capabilities and limitations. AI tools, such as machine learning algorithms, can learn from and make predictions based on training data, thereby enhancing diagnostic accuracy (12). However, they are dependent on the quality and quantity of data they are trained on, and they require continual refinement to maintain their predictive accuracy (7). These tools do not replace healthcare providers; rather, they augment their ability to make informed decisions about patient care.

Contrary to the notion of AI as a universal solution to healthcare's problems, its application is often context-specific. For instance, AI tools that are effective in tertiary care settings may not be equally effective in primary care or community healthcare settings due to differences in the nature and volume of data, technological infrastructure, and the level of training of healthcare providers (83).

Moreover, the ethical implications of using AI in healthcare should not be overlooked. Issues related to data privacy, algorithmic bias, and the transparency of AI decisions are major concerns that need to be addressed (34). In addition, the integration of AI into healthcare systems may have implications for the patient-provider relationship, potentially disrupting traditional models of care.

Managing expectations of AI in healthcare, therefore, is a delicate balancing act. It necessitates an open, honest, and ongoing dialogue among stakeholders, including healthcare providers, patients, policymakers, AI developers, and researchers. This dialogue should foster an understanding of AI's capabilities and limitations, and the ways in which AI can be responsibly integrated into healthcare practice.

The role of professional bodies, such as the American Medical Association and the British Medical Association, is pivotal in shaping this dialogue. Through their policy statements and guidelines, they can help define the role of AI in healthcare, provide recommendations for its ethical use, and promote education and training for healthcare providers.

Patients, too, have an essential role to play in this dialogue. Patient advocacy groups can empower patients to understand and engage with AI technologies, making them active participants in their care.

The application of AI in healthcare is not an end in itself, but a means to an end. Its ultimate goal should be to improve healthcare outcomes and enhance patient well-being. Any hype surrounding AI should not distract from this goal. A sober appraisal of AI, based on rigorous scientific evidence and thoughtful ethical deliberation, is necessary to realize its potential and avoid pitfalls.

The need for robust cybersecurity measures for AI systems in healthcare

As AI technological advancements continue to permeate the healthcare industry, the risks associated with cybersecurity vulnerabilities concurrently escalate, underscoring the necessity for robust protective measures.

In recent years, the healthcare sector has been increasingly susceptible to a spectrum of cyber threats (84). Among the most prominent concerns are data breaches that could compromise sensitive patient information. The sheer volume of data that healthcare systems process, coupled with its intrinsic sensitivity, makes it an attractive target for cybercriminal activities.

For example, in 2015, Anthem, one of the largest health insurance companies in the United States, experienced a cyberattack that led to the exposure of nearly 78.8 million records containing personal patient information (85,86). Another example is the WannaCry ransom ware attack that occurred on May 12, 2017, impacted 230,000 systems across more than 150 countries, including the UK's National Health Service (NHS) (87). Consequently, from May 13 to 16, 2017, five NHS Trusts were compelled to redirect Accident and Emergency patients to unaffected Trusts, and several Trusts encountered difficulties with their CT and MRI imaging systems. This disruption led to the cancellation of nearly 20,000 appointments or surgeries and incurred a cost of nearly £92 million for the NHS (88). Not only did these breaches result in a significant financial impact, but it also highlighted the potential for severe harm to patient trust in healthcare systems.

Beyond data breaches, there lies an arguably more sinister threat in the form of tampering with AI algorithms. The complexity and opaque nature of some AI systems can render them susceptible to adversarial attacks. Adversarial machine learning, a field that investigates how AI systems can be fooled or manipulated, has documented a variety of potential attack vectors (89). One of the most worrisome is the "poisoning" of machine learning models, wherein attackers insert misleading data into the training phase, leading the model to produce incorrect outputs. In a healthcare setting, such manipulation could lead to incorrect diagnoses, flawed treatment recommendations, and even directly endanger patient safety (90).

Given the severity of these threats, the necessity of robust cybersecurity measures in healthcare AI systems cannot be overstated. It is not sufficient to treat cybersecurity as an ancillary concern; rather, it must be integrated as a core component of the design, implementation, and operation of healthcare AI systems. Such measures may encompass a variety of strategies, from hardened system security and encryption practices to intrusion detection and response systems (2).

Moreover, there is a compelling need for a collaborative approach to cybersecurity in healthcare AI. A multi-stakeholder model involving healthcare providers, AI developers, regulatory bodies, and even patients themselves can enhance shared understanding, establish best practices, and facilitate rapid response to new threats (91). Additionally, international cooperation is vital, given the global nature of cyber threats and the international reach of many healthcare and technology companies.

The importance of obtaining sufficient investment for AI research and implementation

It is necessary to underscore the exigency of securing adequate financial resources for the continuation and expansion of research and implementation in artificial intelligence within the healthcare sector. The current state of AI technology, though promising, remains in its infancy and demands significant funding for further research, rigorous validation, and seamless integration into healthcare systems (2).

There is an expanding body of literature that recognizes the enormous potential of AI to revolutionize healthcare, offering transformative solutions for patient care, disease prevention, diagnosis, and treatment (1). For instance, studies have shown that AI can assist clinicians in accurately diagnosing diseases, such as cancer, by analyzing medical images and detecting anomalies that might be missed by the human eye (12,92). Despite these breakthroughs, the advancement of AI in healthcare is presently stymied by multiple barriers, not least of which is the lack of sufficient funding.

Securing adequate funding for AI research and implementation in healthcare, however, is a complex endeavour. The investment managers have to deal with competing priorities vie for finite resources (93). The significance of this challenge cannot be overstated, given the myriad issues that persist in healthcare, such as chronic disease management, elderly care, and health inequality, all of which necessitate immediate and tangible solutions. In such a context, prioritizing funding for a still emerging technology can be a contentious issue.

A particular concern is the validation of AI technology. Unlike traditional pharmaceutical or biomedical research, AI applications require unique validation processes, often involving large and diverse datasets (94). Furthermore, AI applications need to be validated in a real-world clinical setting, a process that is both time-consuming and costly. The financial burden of validation hence poses a significant barrier to the deployment and adoption of AI in healthcare.

Additionally, the integration of AI into existing healthcare systems is another substantial financial endeavor. Even with promising AI applications, the transition from development to integration is fraught with obstacles. For instance, the deployment of AI necessitates changes in the existing workflow, training for healthcare professionals, and infrastructure upgrades, all of which are costly endeavors (95).

To ensure that the potential benefits of AI in healthcare are realized, adequate investment must be secured. However, this is not merely a question of allocating more funds. Rather, it necessitates a multi-faceted approach that includes policy changes, fostering collaborations between public and private entities, and developing novel funding models that can sustain the growth and adoption of AI in healthcare. The precise direction and shape these efforts will take remains to be seen, but it seems evident that securing the financial resources necessary for AI research and integration requires serious consideration.

Patient-Centric AI Design

The ascent of artificial intelligence into the healthcare domain has inspired fervent discussions regarding its potential advantages and challenges. Among the latter, a subtle yet vital issue arises: the tension between AI's innate drive to generalize and the clinical imperative to individualize. As the integration of AI tools becomes more prevalent in healthcare, striking the right balance between these contrasting tendencies becomes more important. Consequently, patient-centered AI design emerges as a focal area of interest, intertwining technology's capability with medicine's principles.

Traditionally, medicine has flourished on the philosophy of patient-centricity. When a clinician engages with a patient, they aren't just observing symptoms; they're reading a narrative. One shaped by the patient's genetics, environment, history, and lifestyle. This principle, rooted in the Hippocratic oath, is what has shaped medical decisions, where emphasis is not solely on objective signs but also on the patient's subjective experience (96). With AI's foray into this physician-patient space, the objective may overshadow the subjective, necessitating a design framework that equally respects both.

AI thrives on patterns discernible from large-scale datasets. In this context, these patterns allow for predictions, assessments, and interventions. Yet, human biology and experience, by their nature, do not always conform to discernible patterns. Herein lies the challenge: the granularity of individual experience may be lost in the vastness of data (34).

Moreover, there's the issue of representation. If AI is to be truly patient-centric, it must be trained on diverse datasets, ensuring that every patient demographic is adequately represented, thereby preventing potential biases and misrepresentations (77).

Building a bridge between AI's pattern-seeking essence and the particulars of individual patient stories requires deliberate action. Multi-modal data integration, which brings together diverse sources of patient data, stands as a promising avenue. By considering genetic, physiological, psychosocial, and other variables, AI can foster a richer, more nuanced understanding of individual health profiles.

Ethics remains at the core of this integration. Beyond the crucial matter of data privacy, there is an inherent duty to ensure AI tools are developed and utilized with transparency and respect for patient autonomy (82).

AI Governance and Oversight

While there is general consensus on AI's transformative potential, the medical community remains split on its trajectory, especially concerning the necessary governance and supervising structures.

One of the more remarkable concerns in the introduction of AI into healthcare revolves around the notion of trust. For healthcare professionals to adopt and, more importantly, rely on AI-driven systems, they need to trust the decisions and recommendations these tools make. But trust in AI isn't binary; it requires an understanding of the system's mechanisms, or at the very least, its logic. The non-transparent nature of certain advanced algorithms (see section 0), which make them almost inscrutable, poses a challenge in building this trust (97). Hence, it becomes vital for any governance framework to emphasize interpretability and transparency as core principles for AI tools in healthcare.

Another governance challenge emerges from the data upon which AI systems are trained. Medical data, inherently complex and multifaceted, can sometimes be unrepresentative or contain implicit biases. Systems trained on such data can inadvertently perpetuate or even exaggerate these biases, leading to imprecise or biased clinical recommendations (77). Thus, an integral part of AI governance must involve rigorous data validation and constant oversight to ensure data quality and representativeness.

The AI-governance nexus also intersects with regulatory standards that already exist in healthcare. How should regulatory bodies approach AI? Is there a need for new structures, or can AI be accommodated within the current regulatory frameworks? The dynamic and evolving nature of AI, wherein systems continuously learn and adapt, poses a unique challenge. Unlike static medical devices or interventions, AI-driven systems can evolve post-deployment, necessitating periodic reassessment and validation (98).

From an ethical vantage, there are additional layers of complexity. Issues like patient consent, data privacy, and the potential for misuse take center stage in discussions about AI's integration into healthcare (99). AI governance frameworks should be constructed with these ethical considerations at their core, perhaps even mandating ethics committees specifically for AI-related interventions in medical settings.

The marriage of AI and healthcare, while promising, is rife with challenges that demand well-thought-out governance and oversight. Constructing such frameworks requires a multidisciplinary approach, combining technological expertise with clinical, ethical, and regulatory insights. Only with such comprehensive governance can we ensure that the transformative potential of AI in healthcare is realized responsibly and safely.

HEALTHCARE AI ETHICAL, SOCIAL, AND EQUITY CHALLENGES

The adoption of artificial intelligence in healthcare marks a significant transition, affecting patient care strategies, disease prediction methods, and the direction of medical research (2). As AI technologies become more commonplace in healthcare settings, a range of ethical, social, and equity considerations emerges, warranting thoughtful analysis. While the advantages of AI—such as enhancing diagnostic accuracy, streamlining clinical processes, and tailoring patient care—are evident, there are concurrent concerns about patient data protection, potential biases in AI systems, and equitable access to these technologies.

The following sections delve into the challenges revolving around these matters. While they possess a more human dimension, they cannot be detached from the challenges of a more technological nature addressed in the preceding sections.

Data privacy and patients' data anonymization

Medical practices are about to experience profound shifts with the emergence of artificial intelligence, particularly in predictive modelling derived from patient health records.. These records, however, are typically replete with highly sensitive personal data, presenting an intricate quandary of data privacy and protection (100). The proliferation of data breaches in recent years has underscored the urgent necessity of safeguarding such data. In an age where privacy has come to the fore as a global concern, AI medical practitioners face the task of navigating a delicate equilibrium. On the one hand, they must advance AI-driven healthcare innovations. On the other, they must uphold patients' rights and confidentiality by ensuring data protection (101).

The challenge is further amplified by the introduction of robust data protection legislation like the European General Data Protection Regulation (GDPR), which imposes strict constraints on data usage (102). GDPR and analogous legislation enacted worldwide delineate a series of stringent conditions for the acquisition, storage, processing, and disclosure of personal data, thus establishing an additional regulatory obstacle for AI applications in healthcare. Navigating the legal landscape, AI experts must formulate a balanced strategy that respects data protection mandates while also fostering the progression of AI within healthcare.

The dichotomy between safeguarding data privacy and fostering AI's progress in healthcare can be traversed through strategies that employ data anonymization and de-identification. Anonymization, as defined by the GDPR, is the irreversible process of transforming personal data such that the data subject cannot be identified. This practice, when employed judiciously, can act as a potent instrument for achieving a balanced approach. Data anonymization techniques, such as k-anonymity, l-diversity, and t-closeness, aim to mitigate the risk of re-identification while maintaining the utility of the dataset for research and AI-based analysis (103).

The potential of these techniques is exhibited in a study by El Emam et al., wherein a de-identified dataset was successfully utilized for a drug safety surveillance project without breaching patient privacy (104). While the example demonstrates the feasibility of anonymization and de-identification, it also emphasizes the care that needs to be taken to ensure the techniques are applied rigorously and correctly. Inaccurate or inappropriate de-identification can result in residual risks of re-identification, thus compromising data security (105).

Although there's no one-size-fits-all solution to the complex puzzle of data privacy in AI-driven healthcare modeling, the prudent application of anonymization and de-identification methods holds promise. AI practitioners must strive to incorporate these methods into their data protection strategies to uphold patient privacy, abide by legal norms, and sustain the progress of AI in healthcare.

The potential widening of health disparities

Without a careful implementation, AI systems may inadvertently magnify health disparities.Health disparities can be thought of as the unequal distribution of health resources or outcomes across different groups, typically associated with socioeconomic status, race, ethnicity, or geographic location. It has long been documented in the literature that certain communities experience disproportionately negative health outcomes due to such disparities (106). If AI-based tools, which have the potential to be a great equalizer, are trained primarily on data from specific demographic or socio-economic groups, they may not function as effectively when applied to other populations.

The issue at hand is that medical AI systems, like all machine learning algorithms, are only as good as the data on which they are trained. The machine learning models used in AI systems do not generate knowledge or understanding independently; they learn patterns from the input data they are given (107). If the data used to train these models are not representative of the diversity of populations that will use the healthcare services, then the AI system may generate biased results, which can lead to suboptimal care for underrepresented groups.

A notable example of this issue can be seen in a study by Obermeyer et al. (7), where an AI system used to identify and prioritize patients for high-risk care management programs was found to be biased against African Americans. Despite having similar health needs as white patients, African Americans were less likely to be identified as high-risk by the AI system. The source of the bias was traced back to the algorithm's training data, which used healthcare costs as a proxy for health needs. Given that African Americans tend to spend less on healthcare due to existing disparities, the AI system was inadvertently perpetuating these disparities.

The potential exacerbation of health disparities by AI systems presents a significant challenge to healthcare providers, researchers, and policymakers. However, it also presents an opportunity to rethink how we design and implement AI systems in healthcare. Greater diversity in data collection and more inclusive algorithmic design can ensure that AI tools provide equitable health outcomes. By incorporating data from a wider range of demographic and socio-economic groups, we can create AI models that generalize better across populations. This requires a coordinated effort from all stakeholders, including healthcare providers, AI developers, patients, and regulatory bodies.

AI systems should be developed and deployed with clear, understandable explanations of how they operate and make decisions (see section 0). This will empower patients and healthcare providers to make informed decisions, potentially alleviating some of the distrust that may arise from opaque AI systems.

While AI has the potential to democratize healthcare, it also has the potential to widen existing health disparities if not implemented thoughtfully. The key to ensuring equity in AI applications lies in diverse data collection, inclusive algorithm design, and transparent implementation practices. These efforts can help mitigate potential bias in AI systems and contribute to more equitable health outcomes.

Maintaining a humanistic approach to healthcare in an AI-driven environment

The future ubiquity of AI in the healthcare sector has raised concerns about the maintenance of a humanistic approach to patient care. This approach is fundamental in the medical practice, it emphasizes the importance of empathy, compassion, and interpersonal relationships, qualities that are not readily reproducible by AI (95).

One aspect worth discussing is the preservation of empathy in an AI-mediated healthcare context. Empathy is a fundamental aspect of the physician-patient relationship, it allows healthcare providers to understand and resonate with patients' emotional states, leading to improved therapeutic outcomes (108). The reliance on AI tools, while expediting healthcare delivery, could potentially depersonalize patient care, creating an environment that might feel cold and devoid of human warmth and understanding. However, to reconcile AI integration with empathetic care, it is important to recognize the distinct roles of AI and healthcare professionals. AI excels in standardizing and streamlining tasks, analyzing vast amounts of data, and delivering evidence-based predictions, whereas healthcare professionals bring a nuanced understanding of human emotions, patient narratives, and holistic care (109).

In addition to empathy, compassion is an integral part of the therapeutic alliance between physicians and patients. The experience of compassion, not just as a sentiment but as a committed action to relieve suffering, underlies every clinical encounter (110). As AI tools become more embedded within clinical practice, it is essential to ensure that they are used as instruments that aid compassionate care, rather than as replacements for human providers. For instance, AI could be leveraged to handle administrative burdens, freeing clinicians to spend more quality time with patients, thereby fostering a deeper connection (111).

Moreover, the preservation of interpersonal relationships in an increasingly digital healthcare ecosystem cannot be forgotten. The art of medicine relies significantly on the physician-patient relationship, a delicate dynamic built upon mutual trust and respect. This relationship is decisive for eliciting patient histories, making accurate diagnoses, and implementing effective treatment strategies (112). In the midst of the rapid digitalization of healthcare, there is a risk of this relationship becoming more transactional and less personal. Therefore, the challenge lies in the strategic implementation of AI to supplement rather than supplant these human connections (113).

The emergence of AI as a driving force in healthcare necessitates a conscientious appraisal of the roles that empathy, compassion, and the physician-patient relationship play in healthcare delivery. As we integrate AI into the healthcare scene, it should be regarded as a tool to enhance, not replace, the humanistic qualities that are central to the practice of medicine(114-115). It is through this careful and mindful integration of AI that we can uphold the humanistic values that form the heart of healthcare, despite the changing landscape.

Achieving broad-based consensus on the ethical use of AI in healthcare

As artificial intelligence becomes an integral part of modern healthcare, it gives rise to ethical challenges related to data privacy, informed consent, accountability, and equitable access, necessitating broad-based consensus from all stakeholders (82).

The integration of artificial intelligence in healthcare necessitates stringent measures to safeguard patient data privacy. It is essential to consider that the use of large data sets, a key element in training AI models, could potentially compromise patient anonymity if not appropriately de-identified and secured (34). Despite the existence of data protection regulations like the General Data Protection Regulation (GDPR) in the EU and the Health Insurance Portability and Accountability Act (HIPAA) in the US, the question persists as to whether these are adequate to address the unique privacy concerns posed by AI in healthcare (35). Therefore, strategies to ensure the ethical use of patient data should be developed, taking into account not only the de-identification and secure storage of data but also the ethical considerations surrounding its use.

Concurrently, the matter of informed consent for AI-assisted care is matter of great importance. The integration of AI into healthcare necessitates a re-evaluation of the traditional model of informed consent. Ordinarily, informed consent involves explaining the risks and benefits of a proposed intervention to a patient, but AI's complexity and current inherent unpredictability could challenge this model (116). Given that AI algorithms often operate as 'black boxes' with decision-making processes that are opaque even to the developers, it can be challenging for patients to fully understand the risks and benefits of AI-assisted care (81). This issue underscores the necessity for redefining informed consent in the era of AI, ensuring it encompasses the unique challenges posed by these technologies.

Further, accountability in the case of AI errors needs to be established. An error made by an AI system can lead to misdiagnosis or mistreatment, posing serious risks to patients. It can be challenging to determine responsibility in such cases, given that AI's decision-making process can be complex and opaque. Current malpractice laws and regulations may not be sufficient to address this new paradigm, and therefore novel legal frameworks are necessary to identify the responsible parties when AI causes harm (117).

Lastly, the potential benefits of AI in healthcare should be equitably distributed. Access to AI technologies in healthcare is likely to be initially concentrated in affluent urban areas, possibly exacerbating existing health disparities (118). For instance, rural areas and low-income regions, both within and between countries, may not have the same access to AI technologies, potentially leaving these populations at a disadvantage (119,120). Hence, it is critical to ensure that the benefits of AI in healthcare are widely available and that mechanisms are in place to avoid exacerbating health disparities.

While AI presents numerous opportunities for advancing healthcare, it also raises significant ethical issues that need to be addressed. Broad-based consensus among all stakeholders is needed to ensure the ethical use of AI in healthcare, considering the unique challenges posed by data privacy, informed consent, accountability, and equitable access.

Managing the potential workforce implications of AI

Artificial Intelligence integration has raised concerns about potential workforce implications, necessitating a thorough understanding and careful management of these changes to uphold high-quality patient care.

One of the primary aspects to consider is job displacement or redundancy due to AI. AI has been proclaimed for its ability to automate routine and tedious tasks, which in a healthcare setting range from administrative responsibilities to patient triage (68). This enhanced efficiency raises the question of whether some roles currently fulfilled by humans may become unnecessary. Previous studies have suggested that while AI has the potential to automate certain tasks, the scope for complete replacement of healthcare professionals remains limited (107). The integration of AI, therefore, does not equate to total job displacement but a reconfiguration of roles, necessitating a nuanced understanding of its implications.

The advent of AI has generated a shift in the roles and responsibilities of healthcare professionals by making them more time-efficient. It is a shift towards tasks that necessitate human judgment, empathy, and complex decision-making, rather than tasks that are routinized or systematic in nature. The British Medical Journal's study on the integration of AI in healthcare (1) suggests that clinicians will increase their function as data interpreters, as

AI algorithms process and provide substantial medical data. This evolving role will require that healthcare professionals need to be competent in understanding, interpreting, and applying AI-generated data in their practice.

In managing this transition, it will be necessary to offer training and support to the existing workforce. A study conducted by the National Academy of Medicine (121) suggested the need for ongoing education and training programs that enable healthcare professionals to understand the potential and limitations of AI and to efficiently apply it in clinical practice. There is also an emerging need for interdisciplinary collaboration, particularly with data scientists and technologists, to provide suitable support for AI in healthcare.

Furthermore, AI's integration into healthcare prompts a reevaluation of ethical considerations, particularly regarding patient care and data privacy. An article published in the Lancet (82) proposes that healthcare institutions should uphold transparency in AI application, providing patients with a clear understanding of how their health data is utilized and the role of AI in their care. Ethical principles, such as autonomy, beneficence, and justice, must not be overshadowed by the technological advancement brought about by AI.

The integration of AI into the healthcare workforce poses both opportunities and challenges. A careful approach is required to manage these workforce changes, emphasizing training, interdisciplinary collaboration, and ethical considerations (122). Doing so will ensure that while the healthcare sector leverages the benefits of AI, it maintains the central principle of patient care.

Impact and long-term results assessment of AI medical applications

As artificial intelligence-based solutions become established in the services of various medical specialties, and as solutions that interoperate with each other begin to emerge, the comprehensive care of the patient will be enhanced in the long term.

Firstly, the introduction of AI into healthcare will led to an enhanced quality of care. Machine learning algorithms are being employed in disease diagnosis, where they process and analyze large volumes of data and identify patterns that might elude the human eye. For instance, an AI algorithm has been developed that can detect malignant melanomas with a precision equal to, if not surpassing, that of dermatologists (12). However, it is essential to consider the limitations and biases inherent in AI algorithms. The training data for these algorithms can often be skewed towards certain demographics, thereby affecting the generalizability of results (123).

Further, the adoption of AI has the potential to influence patient outcomes positively. A salient example is one of the Google's DeepMind projects, which developed an AI system that can predict acute kidney injury up to 48 hours before it occurs, potentially giving doctors a significant lead time to intervene (124). However, this presupposes seamless integration with existing healthcare systems and workflows, which is often not the case (see sections 0 and 0).

AI-driven systems, through machine learning algorithms and predictive analytics, offer potential to optimize inventory levels, reduce waste, and streamline procurement processes. A study by Rajkomar et al. highlighted the ability of deep learning algorithms to predict patient admissions, which can be extrapolated to foresee material and equipment usage patterns, thus aiding in more precise resource allocation (6). A study by Jha and Topol discussed how AI could automate routine tasks, thereby freeing up time for physicians to focus on complex tasks and direct patient care (125). Future artificial intelligence systems will have the capability to detect and alert users when instruments or other devices are not in optimal conditions for their use, thereby promoting proactive maintenance and extending their operational lifespan (126). Nonetheless, the risk of job displacement due to AI cannot be discounted and merits thorough evaluation. As the integration of AI into hospital resource management continues, stakeholders should critically assess the effectiveness and ethical considerations of such applications, ensuring that benefits do not compromise patient safety and care quality.

Biases in AI algorithms, issues in AI implementation, and concerns about job displacement warrant rigorous and ongoing scrutiny. Hence, the integration of AI into healthcare requires a delicate balance, wherein the benefits are maximized, and the drawbacks are mitigated.

Intellectual property and collaboration challenges in medical AI applications

As AI becomes increasingly central to healthcare, the strategies and legal frameworks surrounding intellectual property (IP) rights are key to mediating successful collaboration in this area.

One of the most universal issues regarding IP in medical AI applications pertains to the ownership of algorithms. AI models are generally built on vast datasets to predict health outcomes or suggest potential treatments, among other applications (7). In many cases, the development of these models is carried out by multiple stakeholders—ranging from academic institutions and hospitals to private corporations—which raises complex questions about the ownership of the final product (127).

The complex nature of IP rights in this scenario often leads to protracted disputes over the ownership of algorithms, impeding the sharing of knowledge and slowing down the development of beneficial AI solutions. However, institutions that attempt to bypass these issues by keeping their algorithms entirely private(as industrial secret) face a different set of challenges. Concealing AI models can impede their full validation and hinder the broader scientific community from verifying their accuracy and safety (128).

Data rights also play a significant role in the IP landscape. Many AI models require vast amounts of patient data, and while anonymized datasets can help protect patient privacy, they also introduce other complexities (129). For instance, there are ongoing debates about whether or not these anonymized datasets can be considered a form of IP. Additionally, the sensitive nature of this data necessitates stringent security measures and carefully negotiated data sharing agreements between institutions (130).

While the landscape is fraught with challenges, several strategies have been proposed to help mediate these IP and collaboration hurdles. A potential solution is to establish a more transparent and standardized framework for the ownership and sharing of AI models and datasets. This could involve measures such as mutually agreed contracts and licensing agreements that define clear parameters for ownership, access, and revenue sharing between collaborating institutions. This approach can help establish a more level playing field and encourage collaboration by reducing the risk of disputes.

Despite these potential solutions, it is important to acknowledge that the landscape of IP and collaboration in medical AI applications is complex and constantly evolving. Therefore, there is an ongoing need for discussion and adaptation to address these challenges and ensure the continued growth and success of AI in medicine.

Global applicability of AI

The global applicability of AI models presents considerable challenges, primarily due to differences in demographics, disease patterns, healthcare practices, and data standards across various countries and healthcare systems.

As indicated along the article, Deep Learning models, by their very nature, operate based on the premise of learning from large amounts of training data. These models make predictions or decisions based on patterns observed in the provided data sets. Consequently, if an AI model is trained on data from a particular demographic or healthcare system, its efficacy may diminish significantly when applied to a different demographic or healthcare system (77). This phenomenon has its roots in the inherent biases of the training data, which may not encompass the variability and diversity present in the global population (7).

Data standards also play a substantial role in the global applicability of AI. Various healthcare systems have different standards and protocols for recording and managing healthcare data (see section 0). The disparities in these standards can lead to inconsistencies in the performance of AI models when applied to different healthcare systems. For instance, the usage of electronic health records (EHR) varies greatly worldwide, affecting the quality and compatibility of data used in AI model training (131).

The challenge, therefore, lies in developing universally applicable AI models or in adapting existing models to different settings. However, achieving this goal is not a straightforward task. An attractive strategy to mitigate this problem is the use of federated learning techniques, where AI models are trained across multiple decentralized devices or servers holding local data samples. This allows the AI model to learn from a diverse array of data sources while addressing privacy concerns associated with data centralization (132,133). Additionally, it is important to ensure that the data used for model training represent a wide demographic range and disease patterns. The inclusion of diverse and representative data aids in overcoming biases and improves the robustness of the models.

However, even as we strive for universality, it must be recognized that adaptations are sometimes necessary, and one-size-fits-all solutions may not be applicable in all contexts. In these instances, transfer learning can be applied, a machine learning method that leverages the knowledge gained from one problem to solve a different, yet related, problem (134). It allows an AI model trained on a specific setting to be fine-tuned for a different population target, thereby ensuring the wider applicability of AI in global healthcare.

The governmental agencies of each region are making a special effort to ensure that AI-based medical products yield favourable outcomes across the population cohorts of their respective countries. Thus, this challenge is closely intertwined with many of the previously outlined challenges, as model explainability, data quality and traceability, and secure integration into hospital systems' infrastructure are essential in attaining this objective.

CONCLUSIONS

Throughout this extensive review article, we have examined the 24 primary challenges confronting the actual deployment of AI-based healthcare applications. We have strived to provide a rigorous perspective, accompanied by examples of published cases to illustrate each of these challenges.

None of the challenges holds greater importance than the others. In fact, all of them are intricately interconnected (which is why the descriptions of some allude to the others). Thus, a comprehensive approach is imperative for the future adoption of such technology in a domain as sensitive as healthcare.

There is no doubt that we are fortunate to live at the dawn of a bright future, and such applications will enhance physicians' capabilities to be more effective in many ways. The impending revolution will confront each of the current challenges and gradually overcome them. AI-based solutions in the field of medicine will not replace doctors; rather, it will be the physicians who adopt these advancements and tools that will replace those who do not exploit them.

Conflict of Interest Statement: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Financial Support: The authors received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

1. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature medicine. 2019;25: 44–56.

2. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2: 230–243. doi:10.1136/svn-2017-000101

3. Sayar B. Tıp Alanında Yapay Zekânın Kullanımı: Araştırma Makalesi. Acta Medica Ruha. 2023;1: 27–33. doi:10.5281/zenodo.7749209

4. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical Image Analysis. 2017;42: 60–88. doi:10.1016/j.media.2017.07.005

5. Liew C. The future of radiology augmented with Artificial Intelligence: A strategy for success. European Journal of Radiology. 2018;102: 152–156. doi:10.1016/j.ejrad.2018.03.019

6. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019;380: 1347–1358. doi:10.1056/NEJMra1814259

7. Obermeyer Z, Emanuel EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016;375: 1216–1219. doi:10.1056/NEJMp1606181

8. Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology. 2017;284: 574–582. doi:10.1148/radiol.2017162326

9. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLOS Medicine. 2018;15: e1002686. doi:10.1371/journal.pmed.1002686

10. Belle A, Thiagarajan R, Soroushmehr SMR, Navidi F, Beard DA, Najarian K. Big Data Analytics in Healthcare. BioMed Research International. 2015;2015: e370194. doi:10.1155/2015/370194

11. Baytas IM, Xiao C, Zhang X, Wang F, Jain AK, Zhou J. Patient Subtyping via Time-Aware LSTM Networks. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: Association for Computing Machinery; 2017. pp. 65–74. doi:10.1145/3097983.3097997

12. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. nature. 2017;542: 115–118.

13. Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data. 2016;3: 160035. doi:10.1038/sdata.2016.35

14. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature. 2020;580: 252–256.

15. Ghorbani A, Ouyang D, Abid A, He B, Chen JH, Harrington RA, et al. Deep learning interpretation of echocardiograms. NPJ digital medicine. 2020;3: 1–10.

16. He B, Kwan AC, Cho JH, Yuan N, Pollick C, Shiota T, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. Nature. 2023;616: 520–524. doi:10.1038/s41586-023-05947-3

17. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med. 2019;25: 24–29. doi:10.1038/s41591-018-0316-z

18. Wang RY, Strong DM. Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems. 1996;12: 5–33. doi:10.1080/07421222.1996.11518099

19. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet. 2012;13: 395–405. doi:10.1038/nrg3208

20. Zook M, Barocas S, Boyd D, Crawford K, Keller E, Gangadharan SP, et al. Ten simple rules for responsible big data research. PLOS Computational Biology. 2017;13: e1005399. doi:10.1371/journal.pcbi.1005399

21. Schriger DL. Annals of Emergency Medicine. Annals of Emergency Medicine. 2008;52: 480. doi:10.1016/j.annemergmed.2008.06.461

22. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The Lancet Digital Health. 2019;1: e271–e297. doi:10.1016/S2589-7500(19)30123-2

23. Ohno-Machado L. Realizing the full potential of electronic health records: the role of natural language processing. Journal of the American Medical Informatics Association. 2011;18: 539. doi:10.1136/amiajnl-2011-000501

24. Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, et al. Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. Journal of the American Medical Informatics Association. 2007;14: 1–9. doi:10.1197/jamia.M2273

25. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013;45: 1113–1120. doi:10.1038/ng.2764

26. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. The Lancet. 2014;383: 999–1008. doi:10.1016/S0140-6736(13)61752-3

27. Bishop CM, Nasrabadi NM. Pattern recognition and machine learning. Springer; 2006.

28. Goodfellow I, Bengio Y, Courville A. Deep learning. MIT press; 2016.

29. LeCun Y, Bengio Y, Hinton G. Deep learning. nature. 2015;521: 436–444.

30. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature machine intelligence. 2019;1: 206–215.

31. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface. 2018;15: 20170387.

32. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain? 2017 [cited 6 Aug 2023]. doi:10.48550/ARXIV.1712.09923

33. Coiera E. Guide to Health Informatics. CRC Press; 2015.

34. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care — Addressing Ethical Challenges. N Engl J Med. 2018;378: 981–983. doi:10.1056/NEJMp1714229

35. Wachter S, Mittelstadt B, Floridi L. Transparent, explainable, and accountable AI for robotics. Science Robotics. 2017;2: eaan6080. doi:10.1126/scirobotics.aan6080

36. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A Survey of Methods for Explaining Black Box Models. ACM Comput Surv. 2019;51: 1–42. doi:10.1145/3236009

37. Doshi-Velez F, Kim B. Towards A Rigorous Science of Interpretable Machine Learning. arXiv; 2017. Available: http://arxiv.org/abs/1702.08608

38. Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PRO, Bernstam EV, et al. Caveats for the Use of Operational Electronic Health Record Data in Comparative Effectiveness Research. Med Care. 2013;51: S30–S37. doi:10.1097/MLR.0b013e31829b1dbd

39. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. npj Digital Med. 2018;1: 1–10. doi:10.1038/s41746-018-0029-1

40. Murdoch TB, Detsky AS. The Inevitable Application of Big Data to Health Care. JAMA. 2013;309: 1351–1352. doi:10.1001/jama.2013.393

41. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. JAMA. 2018;319: 1317–1318. doi:10.1001/jama.2017.18391

42. Loscalzo J, Kohane I, Barabasi A-L. Human disease classification in the postgenomic era: A complex systems approach to human pathobiology. Molecular Systems Biology. 2007;3: 124. doi:10.1038/msb4100163

43. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012;380: 37–43. doi:10.1016/S0140-6736(12)60240-2

44. Tinetti ME, Fried TR, Boyd CM. Designing Health Care for the Most Common Chronic Condition—Multimorbidity. JAMA. 2012;307: 2493–2494. doi:10.1001/jama.2012.5265

45. Liu J, Ma J, Wang J, Zeng DD, Song H, Wang L, et al. Comorbidity Analysis According to Sex and Age in Hypertension Patients in China. International Journal of Medical Sciences. 2016;13: 99–107. doi:10.7150/ijms.13456

46. McGovern A, Hinton W, Correa A, Munro N, Whyte M, Lusignan S de. Real-world evidence studies into treatment adherence, thresholds for intervention and disparities in treatment in people with type 2 diabetes in the UK. BMJ Open. 2016;6: e012801. doi:10.1136/bmjopen-2016-012801

47. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst. 2014;2: 3. doi:10.1186/2047-2501-2-3

48. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. Journal of the American Medical Informatics Association. 2013;20: 144–151. doi:10.1136/amiajnl-2011-000681

49. Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. Proceedings of the 1st Machine Learning for Healthcare Conference. PMLR; 2016. pp. 301–318. Available: https://proceedings.mlr.press/v56/Choi16.html

50. Schieppati A, Henter J-I, Daina E, Aperia A. Why rare diseases are an important medical and social issue. The Lancet. 2008;371: 2039–2041. doi:10.1016/S0140-6736(08)60872-7

51. Domingos P. A few useful things to know about machine learning. Commun ACM. 2012;55: 78–87. doi:10.1145/2347736.2347755

52. Shortliffe EH, Sepúlveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. JAMA. 2018;320: 2199–2200. doi:10.1001/jama.2018.17163

53. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Transactions on Medical Imaging. 2016;35: 1299–1312. doi:10.1109/TMI.2016.2535302

54. Cheplygina V, de Bruijne M, Pluim JPW. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Medical Image Analysis. 2019;54: 280–296. doi:10.1016/j.media.2019.03.009

55. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research. 2002;16: 321–357. doi:10.1613/jair.953

56. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM. 2020;63: 139–144. doi:10.1145/3422622

57. Turnbull C, Scott RH, Thomas E, Jones L, Murugaesu N, Pretty FB, et al. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. BMJ. 2018;361: k1687. doi:10.1136/bmj.k1687

58. Amodei D, Olah C, Steinhardt J, Christiano P, Schulman J, Mané D. Concrete Problems in AI Safety. 2016 [cited 6 Aug 2023]. doi:10.48550/ARXIV.1606.06565

59. Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, Rusu AA, et al. Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences. 2017;114: 3521–3526. doi:10.1073/pnas.1611835114

60. Sculley D, Holt G, Golovin D, Davydov E, Phillips T, Ebner D, et al. Hidden Technical Debt in Machine Learning Systems. Advances in Neural Information Processing Systems. Curran Associates, Inc.; 2015. Available: https://papers.nips.cc/paper_files/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html

61. Miller K, Mosby D, Capan M, Kowalski R, Ratwani R, Noaiseh Y, et al. Interface, information, interaction: a narrative review of design and functional requirements for clinical decision support. Journal of the American Medical Informatics Association. 2018;25: 585–592. doi:10.1093/jamia/ocx118

62. Zhang J, Walji MF. TURF: Toward a unified framework of EHR usability. Journal of Biomedical Informatics. 2011;44: 1056–1067. doi:10.1016/j.jbi.2011.08.005

63. Patel VL, Kaufman DR. Cognitive Science and Biomedical Informatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. London: Springer; 2014. pp. 109–148. doi:10.1007/978-1-4471-4474-8_4

64. Horsky J, McColgan K, Pang JE, Melnikas AJ, Linder JA, Schnipper JL, et al. Complementary methods of system usability evaluation: Surveys and observations during software design and development cycles. Journal of Biomedical Informatics. 2010;43: 782–790. doi:10.1016/j.jbi.2010.05.010

65. Strickland E. IBM Watson, heal thyself: How IBM overpromised and underdelivered on AI health care. IEEE Spectrum. 2019;56: 24–31. doi:10.1109/MSPEC.2019.8678513

66. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: An interpretative review. International Journal of Medical Informatics. 2013;82: e73–e86. doi:10.1016/j.ijmedinf.2012.10.007

67. Tenajas R, Miraut D, Illana CI, Alonso-Gonzalez R, Arias-Valcayo F, Herraiz JL. Recent Advances in Artificial Intelligence-Assisted Ultrasound Scanning. Applied Sciences. 2023;13: 3693. doi:10.3390/app13063693

68. Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, DesRoches CM. Artificial Intelligence and the Future of Primary Care: Exploratory Qualitative Study of UK General Practitioners' Views. Journal of Medical Internet Research. 2019;21: e12802. doi:10.2196/12802

69. Tenajas R, Miraut D. Echoes in space: Online training and AI's potential in advancing ultrasound competency. WFUMB Ultrasound Open. 2023;1: 100015. doi:10.1016/j.wfumbo.2023.100015

70. Johnson KB, Wei W, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci. 2021;14: 86–93. doi:10.1111/cts.12884

71. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics. 2018;19: 1236–1246. doi:10.1093/bib/bbx044

72. Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. London: Springer; 2014. doi:10.1007/978-1-4471-4474-8

73. Hripcsak G, Albers DJ. Next-generation phenotyping of electronic health records. Journal of the American Medical Informatics Association. 2013;20: 117–121. doi:10.1136/amiajnl-2012-001145

74. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, Staa T van, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. BMJ. 2013;346: f2350. doi:10.1136/bmj.f2350

75. Gostin LO, Halabi SF, Wilson K. Health Data and Privacy in the Digital Era. JAMA. 2018;320: 233–234. doi:10.1001/jama.2018.8374

76. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019;28: 231–237. doi:10.1136/bmjqs-2018-008370

77. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. JAMA Intern Med. 2018;178: 1544–1547. doi:10.1001/jamainternmed.2018.3763

78. Naylor CD. On the Prospects for a (Deep) Learning Health Care System. JAMA. 2018;320: 1099–1100. doi:10.1001/jama.2018.11103

79. London AJ. Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability. Hastings Center Report. 2019;49: 15–21. doi:10.1002/hast.973

80. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine Learning for Medical Imaging. Radiographics. 2017;37: 505–515. doi:10.1148/rg.2017160130

81. Price WN II, Gerke S, Cohen IG. Potential Liability for Physicians Using Artificial Intelligence. JAMA. 2019;322: 1765–1766. doi:10.1001/jama.2019.15064

82. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. PLOS Medicine. 2018;15: e1002689. doi:10.1371/journal.pmed.1002689

83. Adkins DE. Machine Learning and Electronic Health Records: A Paradigm Shift. AJP. 2017;174: 93–94. doi:10.1176/appi.ajp.2016.16101169

84. Neprash HT, McGlave CC, Cross DA, Virnig BA, Puskarich MA, Huling JD, et al. Trends in Ransomware Attacks on US Hospitals, Clinics, and Other Health Care Delivery Organizations, 2016-2021. JAMA Health Forum. 2022;3: e224873. doi:10.1001/jamahealthforum.2022.4873

85. Yeo LH, Banfield J. Human Factors in Electronic Health Records Cybersecurity Breach: An Exploratory Analysis. Perspect Health Inf Manag. 2022;19: 1i.

86. Cartwright AJ. The elephant in the room: cybersecurity in healthcare. J Clin Monit Comput. 2023 [cited 13 Aug 2023]. doi:10.1007/s10877-023-01013-5

87. Mayor S. Sixty seconds on… the WannaCry cyberattack. BMJ. 2018;361: k1750. doi:10.1136/bmj.k1750

88. Martin G, Ghafur S, Kinross J, Hankin C, Darzi A. WannaCry—a year on. BMJ. 2018;361: k2381. doi:10.1136/bmj.k2381

89. Biggio B, Roli F. Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning. Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. New York, NY, USA: Association for Computing Machinery; 2018. pp. 2154–2156. doi:10.1145/3243734.3264418

90. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. Science. 2019;363: 1287–1289. doi:10.1126/science.aaw4399

91. Keesara S, Jonas A, Schulman K. Covid-19 and Health Care's Digital Revolution. New England Journal of Medicine. 2020;382: e82. doi:10.1056/NEJMp2005835

92. Lång K, Josefsson V, Larsson A-M, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. The Lancet Oncology. 2023;24: 936–944. doi:10.1016/S1470-2045(23)00298-X

93. Zuparov IB, Ibragimova MN, Norbutayeva MK, Otaxonov PE, Normamatov SF, Safarov UQ, et al. Modern directions and perspectives of using medical information systems. Modern directions and perspectives of using medical information systems. 2023;2: 1218–1233.

94. Stead WW. Clinical Implications and Challenges of Artificial Intelligence and Deep Learning. JAMA. 2018;320: 1107–1108. doi:10.1001/jama.2018.11029

95. Verghese A, Shah NH, Harrington RA. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. JAMA. 2018;319: 19–20. doi:10.1001/jama.2017.19198

96. Laine C, Davidoff F. Patient-Centered Medicine: A Professional Evolution. JAMA. 1996;275: 152–156. doi:10.1001/jama.1996.03530260066035

97. Castelvecchi D. Can we open the black box of AI? Nature News. 2016;538: 20. doi:10.1038/538020a

98. Gerke S, Yeung S, Cohen IG. Ethical and Legal Aspects of Ambient Intelligence in Hospitals. JAMA. 2020;323: 601–602. doi:10.1001/jama.2019.21699

99. Goodman B, Flaxman S. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation." AI Magazine. 2017;38: 50–57. doi:10.1609/aimag.v38i3.2741

100. Rakha NA. The Ethics of Data Mining: Lessons from the Cambridge Analytica Scandal. International Journal of Cyber Law. 2023;1. doi:10.59022/clr.24

101. Carmi L, Zohar M, Riva GM. The European General Data Protection Regulation (GDPR) in mHealth: Theoretical and practical aspects for practitioners' use. Med Sci Law. 2023;63: 61–68. doi:10.1177/00258024221118411

102. European Union. European Union. Regulation 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation, "GDPR"). Regulation (EU) 2016/679 of the European Parliament and of the Council. 2016 p. 2016.

103. Sweeney L. k-anonymity: a model for protecting privacy. Int J Unc Fuzz Knowl Based Syst. 2002;10: 557–570. doi:10.1142/S0218488502001648

104. Emam KE, Jonker E, Arbuckle L, Malin B. A Systematic Review of Re-Identification Attacks on Health Data. PLOS ONE. 2011;6: e28071. doi:10.1371/journal.pone.0028071

105. Islam SMR, Kwak D, Kabir MDH, Hossain M, Kwak K-S. The Internet of Things for Health Care: A Comprehensive Survey. IEEE Access. 2015;3: 678–708. doi:10.1109/ACCESS.2015.2437951

106. Braveman P. health disparities and health equity: Concepts and Measurement. Annual Review of Public Health. 2006;27: 167–194. doi:10.1146/annurev.publhealth.27.021405.102103

107. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6: 94–98. doi:10.7861/futurehosp.6-2-94

108. Derksen F, Bensing J, Lagro-Janssen A. Effectiveness of empathy in general practice: a systematic review. Br J Gen Pract. 2013;63: e76–e84. doi:10.3399/bjgp13X660814

109. Topol E. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. Hachette UK; 2019.

110. Sinclair S, Norris JM, McConnell SJ, Chochinov HM, Hack TF, Hagen NA, et al. Compassion: a scoping review of the healthcare literature. BMC Palliative Care. 2016;15: 6. doi:10.1186/s12904-016-0080-0

111. Meyer AND, Singh H. The Path to Diagnostic Excellence Includes Feedback to Calibrate How Clinicians Think. JAMA. 2019;321: 737–738. doi:10.1001/jama.2019.0113

112. Ridd M, Shaw A, Lewis G, Salisbury C. The patient–doctor relationship: a synthesis of the qualitative literature on patients' perspectives. Br J Gen Pract. 2009;59: e116–e133. doi:10.3399/bjgp09X420248

113. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare. 2023;11: 887. doi:10.3390/healthcare11060887

114. Yeo YH, Samaan JS, Ng WH, Ting P-S, Trivedi H, Vipani A, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. medRxiv; 2023. p. 2023.02.06.23285449. doi:10.1101/2023.02.06.23285449

115. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. JAMA Internal Medicine. 2023;183: 589–596. doi:10.1001/jamainternmed.2023.1838

116. O'Sullivan S, Nevejans N, Allen C, Blyth A, Leonard S, Pagallo U, et al. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. Int J Med Robot. 2019;15: e1968. doi:10.1002/rcs.1968

117. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020;46: 205–211. doi:10.1136/medethics-2019-105586

118. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth. 2017;3: 38. doi:10.21037/mhealth.2017.08.07

119. Syed ST, Gerber BS, Sharp LK. Traveling Towards Disease: Transportation Barriers to Health Care Access. J Community Health. 2013;38: 976–993. doi:10.1007/s10900-013-9681-1

120. Guo J, Li B. The Application of Medical Artificial Intelligence Technology in Rural Areas of Developing Countries. Health Equity. 2018;2: 174–181. doi:10.1089/heq.2018.0037

121. Matheny ME, Whicher D, Thadaney Israni S. Artificial Intelligence in Health Care: A Report From the National Academy of Medicine. JAMA. 2020;323: 509–510. doi:10.1001/jama.2019.21579

122. Nick Bostrom,, Eliezer Yudkowsky. The Ethics of Artificial Intelligence. Artificial Intelligence Safety and Security. Chapman and Hall/CRC; 2018.

123. Rezk E, Eltorki M, El-Dakhakhni W. Leveraging Artificial Intelligence to Improve the Diversity of Dermatological Skin Color Pathology: Protocol for an Algorithm Development and Validation Study. JMIR Research Protocols. 2022;11: e34896. doi:10.2196/34896

124. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med. 2018;24: 1716–1720. doi:10.1038/s41591-018-0213-5

125. Jha S, Topol EJ. Adapting to Artificial Intelligence: Radiologists and Pathologists as Information Specialists. JAMA. 2016;316: 2353–2354. doi:10.1001/jama.2016.17438

126. Tenajas R, Miraut D. Rethinking ultrasound probe maintenance in the era of AI. WFUMB Ultrasound Open. 2023;1: 100014. doi:10.1016/j.wfumbo.2023.100014

127. Bierer BE, Crosas M, Pierce HH. Data Authorship as an Incentive to Data Sharing. N Engl J Med. 2017;376: 1684–1687. doi:10.1056/NEJMsb1616595

128. Cohen IG, Amarasingham R, Shah A, Xie B, Lo B. The legal and ethical concerns that arise from using complex predictive analytics in health care. Health Aff (Millwood). 2014;33: 1139–1147. doi:10.1377/hlthaff.2014.0048

129. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying personal genomes by surname inference. Science. 2013;339: 321–324. doi:10.1126/science.1229566

130. Sweeney L. Simple Demographics Often Identify People Uniquely. Carnegie Mellon University Pittsburgh.

131. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. J Am Med Inform Assoc. 2017;24: 1142–1148. doi:10.1093/jamia/ocx080

132. Ramírez DH, Díaz LP, Pérez ÁC, Peña BI, Buenache AB, García JMA, et al. uTile PET: a Privacy Preserving Solution for Collaborative Data-Driven Projects. 2021 [cited 13 Aug 2023]. Available: https://easychair.org/publications/preprint/bcM6

133. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with Federated Learning. npj Digit Med. 2020;3: 1–7. doi:10.1038/s41746-020-00323-1

134. Pan SJ, Yang Q. A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering. 2010;22: 1345–1359. doi:10.1109/TKDE.2009.191