



Review article

Ethical and regulatory challenges of AI technologies in healthcare: A narrative review



Ciro Mennella, Umberto Maniscalco*, Giuseppe De Pietro, Massimo Esposito

Institute for High-Performance Computing and Networking (ICAR) - Research National Council of Italy (CNR), Italy

ARTICLE INFO

Keywords:

Artificial intelligence
 Technologies
 Decision-making
 Healthcare
 Ethics
 Regulatory guidelines

ABSTRACT

Over the past decade, there has been a notable surge in AI-driven research, specifically geared toward enhancing crucial clinical processes and outcomes. The potential of AI-powered decision support systems to streamline clinical workflows, assist in diagnostics, and enable personalized treatment is increasingly evident. Nevertheless, the introduction of these cutting-edge solutions poses substantial challenges in clinical and care environments, necessitating a thorough exploration of ethical, legal, and regulatory considerations.

A robust governance framework is imperative to foster the acceptance and successful implementation of AI in healthcare. This article delves deep into the critical ethical and regulatory concerns entangled with the deployment of AI systems in clinical practice. It not only provides a comprehensive overview of the role of AI technologies but also offers an insightful perspective on the ethical and regulatory challenges, making a pioneering contribution to the field.

This research aims to address the current challenges in digital healthcare by presenting valuable recommendations for all stakeholders eager to advance the development and implementation of innovative AI systems.

1. Introduction

Reforming and improving long-term care programs and healthcare outcomes face significant challenges under current healthcare policies. Recent technological advancements have generated interest in cutting-edge healthcare applications, turning this field into a growing area of research. Monitoring and decision support systems based on artificial intelligence (AI) show promise in extending individualized care programs to remote and home settings.

The World Health Organization (WHO) has recognized the crucial role of technology in seamlessly integrating long-term healthcare services into patients' daily lives [1]. Technology is considered vital in achieving universal health coverage for all age groups by promoting cost-effective integration and service delivery. Technological advancements, particularly in AI and robotics, are now being harnessed to support healthcare research and practice.

Over the past decade, research into utilizing AI to enhance critical clinical processes and outcomes has steadily expanded. AI-based decision support systems, in particular, have the potential to optimize clinical workflows, improve patient safety, aid in diagnosis, and enable personalized treatment [2]. Various digital technologies, including smart mobile and wearable devices, have been developed to gather data and process information for assessing health and tracking individualized therapy progress. Additionally, social and

* Corresponding author.

E-mail address: umberto.maniscalco@icar.cnr.it (U. Maniscalco).

<https://doi.org/10.1016/j.heliyon.2024.e26297>

Received 27 December 2023; Accepted 9 February 2024

Available online 15 February 2024

2405-8440/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

physical assistance systems are being deployed to aid individuals in their recovery from illness, injury, or sensory, motor, and cognitive impairments [3].

In the context of decision support systems, AI-based tools have demonstrated the potential to optimize clinical workflows, enhance patient safety, aid in diagnosis, and facilitate personalized treatment [4]. Cutting-edge AI technologies are transforming healthcare, achieving notable successes in different medical areas. For instance, in skin disease identification, a smart learning system displayed remarkable accuracy, outperforming other methods and offering quick assistance through a mobile app [5]. Similarly, a model predicting breast cancer spread achieved high accuracy, aiding doctors in precise analysis and potentially preventing complications [6]. Different unique models for timely COVID-19 identification offered a high-accuracy solution, reducing reliance on time-consuming tests [7,8]. In cancer imaging, technology improved accuracy, providing valuable insights for doctors [9]. Moreover, the exponential growth of medical data is reshaping the landscape of security and privacy in healthcare. Robust security upgrades are crucial to guarantee the safe handling of vast datasets, and advancements in technology play a pivotal role in enhancing secure data storage, as discussed by recent studies [10]. These strides underscore the transformative influence of AI technologies and medical data, promising to significantly elevate healthcare efficiency and enhance patient outcomes.

While these AI-driven innovations promise to improve healthcare outcomes, ethical and regulatory challenges must be addressed. The rapid evolution of AI in healthcare has led to the emergence of tools and applications that often lack regulatory approvals, posing ethical and legal concerns. Therefore, it is crucial to comprehensively explore and understand the ethical and regulatory challenges associated with AI technologies in healthcare to ensure responsible development and practical implementation. This research aims to fill the existing gap in the literature by providing valuable insights into these challenges and contributing to the responsible integration of AI-driven healthcare applications.

The integration of AI in healthcare, while promising, brings about substantial challenges related to ethics, legality, and regulations. The need to ensure patient safety, privacy, and compliance with existing healthcare standards makes it imperative to address these challenges. The rapid evolution of AI, particularly in the domain of healthcare, has led to the emergence of numerous tools and applications that often lack regulatory approvals. Despite being promising, these innovations pose ethical and regulatory challenges, forming the basis for research in this area.

AI is reshaping the landscape of decision-support systems in healthcare by making them more effective, efficient, and patient-centered. These AI-driven systems are not only enhancing the way healthcare professionals interact with data but are also playing a crucial role in monitoring and assisting patients, ultimately improving the overall quality of healthcare delivery.

In this context, AI healthcare applications driven by advanced computational algorithms hold the potential to implement individual and efficient integrated programs for patients, particularly in remote and home care settings. However, it is important to acknowledge that these significant innovations pose substantial future challenges in clinical and care settings, encompassing ethical, legal, and regulatory considerations.

The practical implications of AI in improving healthcare outcomes make it a crucial area of investigation. However, while there is research on the broader applications of AI in healthcare, the specific considerations for related challenges are under-explored. This article seeks to shed light on the key ethical and regulatory issues, rules, and principles associated with the introduction of AI technologies in healthcare services.

The research aims to provide valuable insights into the ethical and regulatory challenges posed by AI technologies in healthcare. By addressing these challenges, the research contributes to the responsible development and practical implementation of AI-driven healthcare applications. The outcomes of this research can inform policymakers, healthcare professionals, and developers, ensuring that AI benefits are realized without compromising ethical standards or regulatory compliance.

2. Materials & methods

The primary objective of this study was to provide a comprehensive overview of the existing evidence concerning AI technologies, examining both technical aspects and regulatory considerations. The focal point was the identification of categories within clinical practice, aiming to deepen understanding of the diverse technologies involved and the overarching framework of ethics and legal guidelines governing the application of AI in healthcare.

To achieve this, a rigorous narrative review was undertaken, scrutinizing literature from a wide array of sources, including books, published reports, newsletters, media reports, and electronic or paper-based journal articles [11,12].

In the exploration of literature, reports, and studies related to the ethical and regulatory framework, thorough searches were conducted in relevant databases. International organizations' websites were consulted, and comprehensive searches were performed on Scopus, ACM Digital Library, and PubMed using the following query: 'Artificial intelligence' AND 'clinical decision' AND (ethics OR law OR regulation).

The information gleaned from the identified references underwent systematic assessment for relevance and was judiciously utilized in constructing the narrative review.

This review aims to offer a comprehensive analysis of AI categories within the healthcare field, shedding light on the diverse technologies that have evolved in this dynamic and rapidly advancing field. Consequently, it strives to provide a nuanced overview of the ethical and regulatory landscape associated with the integration of AI in healthcare, addressing the main challenges linked to incorporating such technologies into clinical practice.

The structure of the review is outlined as follows.

Section 3 presents a comprehensive overview of emerging trends in AI technologies employed as decision-support and assistive systems in healthcare. This section delves into the specific role of AI in bolstering digital healthcare.

Table 1
Description of AI system autonomy levels.

Level	Definition	Example	Final decision
0	No presence of AI	Standard care.	Human
1	AI suggests decision to human	Clinicians consider AI recommendations but ultimately make the final decision on treatment and therapy.	Human
2	AI makes decisions, with permanent human supervision	AI makes clinical decisions on treatment and therapy, with human doctors providing ongoing supervision.	Human
3	AI makes decisions, with no continuous human supervision but human backup available	AI autonomously makes clinical decisions regarding treatment and therapy, but it can alert human users in case of uncertainty, minimizing the need for constant supervision.	AI
4	AI makes decisions, with no human backup available	AI autonomously govern the clinical decision with no human backup.	AI

Table 2
Description of existing algorithm categories.

Category	Definition	Example
Rule-based algorithm [20]	The system follows a set of rules predefined by experts.	The expert defines the knowledge representation of a phenomenon and integrates this model into the system.
ML-based algorithm [21]	The AI system facilitates the incorporation of intricate knowledge representations using statistics and probability theory.	The focus is not on defining a prior knowledge model, but rather on the collection of data and their integration into a training set. This approach allows for the continuous development of knowledge in a specific application domain through the ongoing utilization of data.

Section 4 succinctly summarizes prevailing ethical principles and regulatory aspects intended to guide the development and deployment of AI technologies in the healthcare domain.

Section 5 discusses the primary challenges associated with the utilization of AI technologies in clinical practice.

Section 6 concludes the work by emphasizing key issues that require attention to steer the future development of AI technologies within the realm of digital healthcare services.

3. The impact of AI in healthcare

3.1. Foundations

The concept of AI was first discussed in 1956 [13], referring to technology used to mimic human behavior. Since then, the field has made remarkable strides in development. As a subfield of AI, Machine Learning (ML) was conceptualized by Arthur Samuel in 1959 [14]. He emphasized the importance of systems that automatically learn from experience instead of being programmed. In the 1980s, ML demonstrated great potential in computer foresight and predictive analytics, including clinical practice and machine translation [15]. Deep Learning (DL), a subfield of ML, has ushered in new breakthroughs in information technology. DL may study underlying features in data from multiple processing layers using neural networks, similar to the human brain [16]. Since the 2010s, DL has garnered immense attention in many fields, especially in image recognition and speech recognition [17].

The term AI generally refers to the performance of tasks that are commonly associated with intelligent beings by software and/or devices. A specific definition of AI in a recommendation by the OECD Council on Artificial Intelligence states: ‘An artificial intelligence system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions that affect real or virtual environments.

The basis of AI technologies is algorithms and they may operate with different levels of autonomy (see Table 1) [18].

Two main approaches are actually proposed to develop AI algorithms: the rule-based approach and the ML based approach whose definitions are described in Table 2 [19]. These algorithms are translated into computer code that contains instructions for the rapid analysis and transformation of data into conclusions, information, or other results.

The huge amounts of data and the ability to analyze data rapidly feed AI to perform increasingly complex tasks [22]. For this reason, AI in medicine raises the idea that AI replaces doctors and human decision-making. However, applications of AI are still relatively new and AI is not yet routinely used in clinical decision-making. Few of these systems have been evaluated in clinical studies [23].

The use of AI in clinical care is expected to bring about the major changes schematically reported in Fig. 1. There are four main trends recognized by the WHO: i) the evolution of the role of the patient in the clinical care process; ii) the shift from hospital to

community-based care; iii) the use of AI to provide clinical care outside the formal healthcare system; iv) the use of AI for resource allocation and prioritization [24]. Each of these trends has ethical implications, which will be discussed later.

Patients already take significant responsibility for their care, including taking medication, improving their diet and nutrition, engaging in physical activity, treating wounds, or administering injections. However, AI could further change the way patients independently manage their medical conditions, particularly in chronic disease conditions. AI could help in self-care, for instance through conversational agents such as 'chatbots', health monitoring tools and risk prediction technologies for prevention programs [25]. While the move to patient-based care may be seen as empowering and beneficial for some patients, others may find the added responsibility stressful and may limit an individual's access to formal healthcare services.

Telemedicine is part of a broad revolution marking the transition from hospital to home care, and the use of AI technologies is helping to accelerate this journey. The shift to home care has been partly facilitated by the increase in the use of search engines (which rely on algorithms) for medical information, as well as the growth in the number of text or voice chatbots for healthcare [26]. In addition, AI technologies can play a more active role in managing patients' health outside clinical settings, such as in 'just-in-time adaptive interventions'. These rely on sensors to provide patients with specific interventions according to previously collected data [27]. The growth and use of wearable sensors and devices may improve the effectiveness of 'just-in-time adaptive interventions', but also raise concerns in light of the amount of data these technologies are collecting, how they are being used, and the burden these technologies may shift to patients.

The increasing use of digital self-management applications and technologies also raises broader questions about whether these technologies should be regulated as clinical applications, thus requiring more regulatory control, or as 'wellness applications', requiring less regulatory control. Many digital self-management technologies are likely to fall in a 'grey area' between these two categories and may present a risk if they are used by patients for their own disease management or clinical care, but remain largely unregulated or could be used without prior medical advice. These concerns are exacerbated by the distribution of such applications by entities that are not part of the formal healthcare system. Indeed, AI applications in healthcare are no longer used exclusively in healthcare (or home care) systems, as AI technologies for health can easily be acquired and used by entities in the non-health system. Emerging is therefore the issue concerning the use of AI to extend 'clinical' care beyond the formal healthcare system.

Finally, with the trend towards self-management, the use of mobile and wearable devices driven by software capable of acquiring and processing data through sophisticated algorithms has increased [28]. Self-management systems empower individuals to take an active role in managing their health. They provide the necessary resources and tools for patients to track their health, adhere to treatment plans, and make informed decisions about their well-being. This approach fosters a proactive approach to healthcare, empowering patients to become partners in their health management. Wearable technologies include those placed in the body (artificial limbs, smart implants), on the body (insulin pump patches, electroencephalogram devices), or near the body (activity trackers, smart watches, and smart glasses). Wearable devices will create more opportunities to monitor a person's health and capture more data to predict health risks, often more efficiently and in a more timely manner. Although such monitoring of 'healthy' individuals may generate data to predict or detect health risks or improve a person's treatment when necessary, it raises concerns as it allows for near-constant surveillance and the collection of excessive data that would otherwise have to remain unknown or uncollected. Such data collection also contributes to the growing practice of 'bio-surveillance', a form of surveillance of health and other biometric data, such as facial features, fingerprints, temperature and pulse [29]. The growth of biosurveillance raises significant ethical and legal concerns, including the use of such data for medical and non-medical purposes for which explicit consent may not have been obtained or the re-use of such data for non-health purposes by a government or company, such as within the criminal justice or immigration systems. Therefore, such data should be subject to the same levels of protection and security as data collected on an individual in a formal clinical care setting.

3.2. Clinical applications

Digital technologies are at the forefront of transforming healthcare practices. Recent innovations hold the promise of improving preventive measures, facilitating early detection of severe illnesses, and enabling remote management of chronic conditions beyond the confines of traditional healthcare settings. These developments open up new possibilities for delivering healthcare services at any time and in any place, aligning with the era of disruptive and minimally invasive medicine.

Within the realm of digital health technologies, a diverse array of innovative healthcare tools has emerged, such as health information technologies, telemedicine applications, robotic platforms, mobile and wearable devices, and Internet of Things (IoT) networks. While these technologies may differ significantly from a technical perspective, they all share a common objective: to provide decision support in the context of healthcare practice. They accomplish this by gathering data during their use, consequently enriching the informativeness and effectiveness of medical practice through the analysis of the recorded information.

A decision support system in healthcare can be considered a computerized tool or software that assists healthcare professionals, including doctors, nurses, and administrators, in making informed and evidence-based decisions related to patient care, treatment options, and healthcare management. It integrates patient data, medical knowledge, and analytical tools to provide real-time information and recommendations, aiding healthcare providers in diagnosing conditions, creating treatment plans, and optimizing healthcare operations. Decision support systems have the potential to enhance the quality of care, reduce errors, and improve efficiency by offering valuable insights and suggestions based on the latest medical research and patient data.

In detail, AI is playing a pivotal role in revolutionizing decision support systems in healthcare, fundamentally transforming the way data is collected, analyzed, and utilized. The integration of AI introduces several significant advancements. AI algorithms can process vast amounts of patient data rapidly and with a high degree of accuracy. This allows for more nuanced and precise

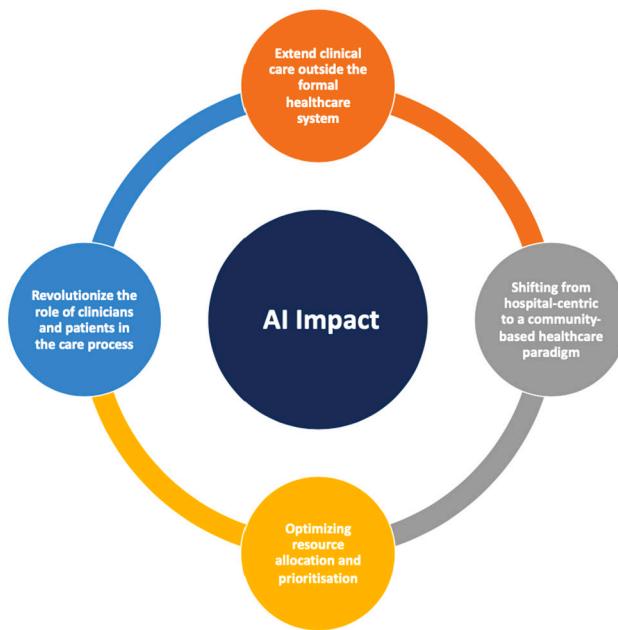


Fig. 1. The AI impact on future healthcare system.

diagnostic and treatment recommendations. AI can identify patterns and anomalies in patient data that might be challenging for human professionals to detect. AI enables the tailoring of treatment plans to individual patients. By analyzing a patient's unique health data, genetic information, and treatment history, AI can suggest personalized therapies that are more effective and have fewer side effects. This personalized approach can lead to better outcomes and a higher quality of care. AI-driven decision support systems can use predictive data analysis to anticipate potential health issues. By continuously monitoring and analyzing patient data, AI can alert healthcare providers to early signs of disease or complications, enabling proactive intervention and preventive measures. AI can monitor patients in real-time, whether they are in a healthcare facility or at home. Wearable devices and sensors connected to AI systems can provide continuous updates on a patient's health status. This real-time monitoring is particularly valuable for chronic disease management and remote patient care. AI technology allows also to extraction of valuable information from unstructured data sources such as medical notes and reports. This capability simplifies the retrieval of relevant patient information, aiding in diagnosis and treatment planning. AI-powered care assistive systems enhance the patient experience during remote consultations. They can assist in collecting and interpreting patient data during telehealth visits, ensuring that healthcare providers have access to the information they need for informed decision-making.

To delve deeper into the discussion on AI technologies, these systems can be methodically categorized into two primary groups: diagnosis support systems, and care assistive systems.

The subsequent paragraphs offer an in-depth examination of each system category, elucidating their distinct healthcare functions as shown in Fig. 2. A comprehensive overview of the AI technologies adopted for each category and current trends as evidenced in existing literature are reported.

3.2.1. Diagnosis support systems

Diagnosis support systems are designed to assist healthcare professionals in accurately diagnosing medical conditions, often by analyzing patient data, medical records, and clinical information. They aid in the formulation of precise diagnoses and the selection of appropriate treatment plans.

The use of AI in disease diagnosis and treatment has been a focus of research since the 1970s when MYCIN, developed at Stanford, was used for diagnosing blood-borne bacterial infections [30].

In various medical fields, researchers have harnessed a range of AI-based techniques to detect diseases that require early diagnosis. Fig. 3 offers a comprehensive summary of the distribution of medical areas of interest, considering the results of a recent systematic review [31].

AI is proving to be a valuable tool for image analysis and is increasingly employed by professionals in radiology for early disease diagnosis and the reduction of diagnostic errors in preventive medicine. AI also aids in analyzing images and signals from various diagnostic tools to support decision-making. For instance, the Ultromics platform, implemented in Oxford, utilizes AI to analyze echocardiography scans, detecting heartbeat patterns and ischemic heart disease [32]. AI has shown promising results in the early detection of diseases such as breast and skin cancer, eye diseases, and pneumonia using various imaging techniques [33–35]. More recently, a system integrated with the advanced DL model enhanced precision in identifying ductal carcinoma in breast cancer imaging, providing valuable insights for practitioners [9]. Furthermore, AI is becoming an integral part of clinical practice, aiding

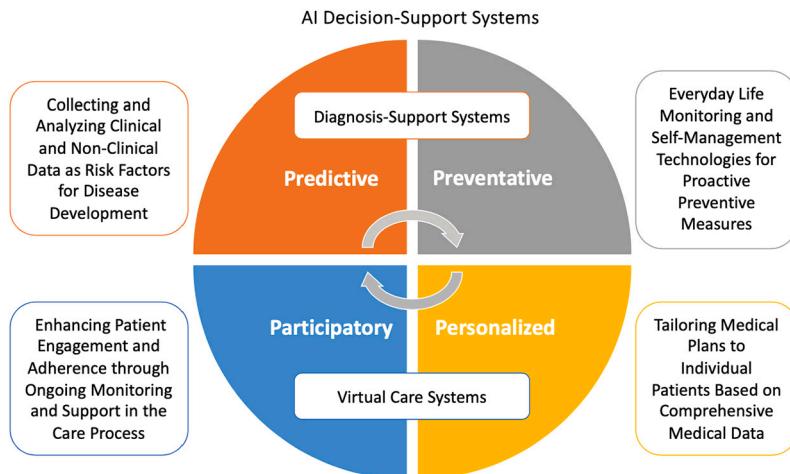


Fig. 2. Category and characteristics of AI decision-support systems.

in diagnostic and therapeutic imaging analysis within the context of the prostate cancer pathway [36]. This integration facilitates enhanced risk stratification and enables more precisely targeted subsequent management.

Further, surprising speech pattern analysis with AI has been shown to predict psychotic occurrences and identify features of neurological diseases like Parkinson's disease [37,38]. In a recent study, AI-based models predicted the onset of diabetes [39]. Additionally, AI has been instrumental in assisting the public in the battle against the virus, as it has played a crucial role in the diagnosis of COVID-19 using various imaging techniques, such as computed tomography (CT), X-rays, magnetic resonance imaging (MRI), and ultrasound (US) [40–42].

AI is significantly impacting clinical decision-making and disease diagnosis. It can process, analyze, and report large volumes of data from different sources, aiding in disease diagnosis and clinical decision-making. AI has the potential to assist physicians in making more informed clinical decisions and, in some cases, may even replace human decisions in therapeutic domains [43]. Moreover, investigations employing computer-aided diagnostics have shown remarkable sensitivity, accuracy, and specificity in detecting subtle radiographic abnormalities, contributing to advancements in public health. However, it's worth noting that the assessment of AI outcomes in imaging studies often focuses on lesion detection, potentially overlooking the biological severity and type of a lesion. This approach may lead to a skewed interpretation of AI output. Additionally, the use of non-patient-related radiological and pathological endpoints may increase sensitivity but at the cost of higher false positives and potentially overestimate diagnosis by detecting minor abnormalities that could mimic subclinical disease [44]. Furthermore, AI has found application in motion analysis, with machine learning-driven video analysis showcasing the capacity of computers to automate the identification of gait abnormalities and associated pathologies in individuals afflicted with orthopedic and neurological disorders [45,46].

Despite significant advancements in recent years, the field of precise clinical diagnostics still faces several challenges that demand continuous improvement to effectively combat emerging illnesses and diseases. Even healthcare professionals acknowledge the obstacles that need to be addressed before illnesses can be accurately detected in collaboration with AI. Currently, doctors are not entirely reliant on AI-based approaches because they are uncertain about their ability to predict diseases and associated symptoms. Therefore, substantial efforts are needed to train AI-based systems, enhancing their accuracy in disease diagnosis. Consequently, future AI-based research should take into account the aforementioned limitations to establish a mutually beneficial relationship between AI and clinicians. Utilizing a unified model for disease diagnosis across different institutions can significantly improve accuracy, thereby aiding in the early diagnosis of diseases.

3.2.2. Care assistive systems

Care assistive systems encompass a diverse range of versatile systems that offer comprehensive monitoring and support capabilities across various healthcare settings. They enable real-time monitoring and provide assistance, fostering patient engagement and compliance in care. Whether the patient is physically present with the healthcare provider or in a remote location, these systems ensure accessibility and convenience, thereby improving the overall patient experience.

Due to remarkable technological advancements, AI has ushered in innovative applications within the realm of care assistive [23]. Significant progress was made in wearable devices capable of measuring physiological changes and facilitating real-time patient monitoring [47]. Remote patient monitoring, a subset of telehealth, enables healthcare providers to remotely monitor and assess patient conditions, reducing the reliance on traditional in-person visits. This approach leverages sensors and communication technologies, simplifying the process of remotely collecting and evaluating health data, and empowering patients to take control of their health [48,49].

Traditionally, patient monitoring systems relied heavily on clinicians' time management and invasive methods requiring skin contact. However, patient remote monitoring in healthcare now incorporates innovative IoT techniques, including contact-based sensors, wearable devices, and telehealth applications. These technologies enable the examination of vital signs and physiological

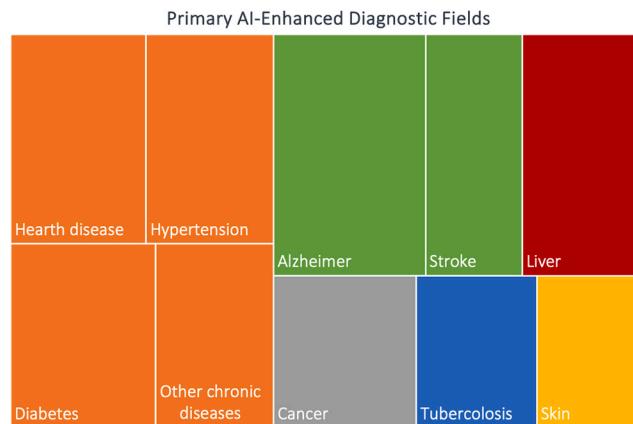


Fig. 3. Literature evidences of diseases diagnosis using AI.

variables, such as motion recognition, which supports medical decision-making and therapeutic strategies for various conditions, including psychological illnesses and movement disorders [50]. Healthcare providers have also harnessed remote patient monitoring platforms to ensure the continuity of patient care during the COVID-19 pandemic.

Conventional AI techniques are commonly employed in virtual applications to detect early signs of patient deterioration, understand patient behavior patterns through reinforcement learning, and tailor the monitoring of patient health parameters via federated learning. AI plays a pivotal role in managing chronic diseases, including diabetes mellitus, hypertension, sleep apnea, and chronic bronchial asthma, through non-invasive, wearable sensors [51]. Smart homes equipped with sensors that monitor physiological variables such as respiratory rate, pulse rate, breathing waveform, blood pressure, and ECG can aid residents in their daily activities and alert caregivers when assistance is required. Additionally, smart mobile and wearable devices allow users to collect data and monitor progress toward personalized therapy goals [52,53]. Inertial sensors in wearable technology can assess individuals' adherence to exercise regimens, particularly in rehabilitation programs [54,55]. Recent reviews highlight the potential of AI integration into wearable technologies while addressing the issue of user retention, and underscore the need for patient education to improve AI acceptance [56–58]. AI-driven data processing from sensors has the potential to monitor patterns in physiological measurements, positional data, and kinematic data, offering insights into improving athletic performance. AI can enhance injury prediction models, increase the diagnostic accuracy of risk stratification systems, enable continuous patient health monitoring, and enhance the patient experience. Despite these benefits, the adoption of AI in wearable devices faces challenges such as missing data, socioeconomic bias, data security, outliers, signal noise, and the acquisition of high-quality data using wearable technology [59,60]. Patient acceptance also presents a critical hurdle to the widespread adoption of these technologies. However, AI's transformative potential in remote patient monitoring is accompanied by challenges related to privacy, signal processing, data volume, uncertainty, imbalanced datasets, feature extraction, and explainability [50].

The integration of AI language models can further revolutionize patient care, with the potential for digital applications to assist patients in managing their treatment regimens. These virtual assistants, similar to personal healthcare advisors, can provide reminders for medication adherence and offer health status updates. The rise of virtual patient assistants using the AI natural language model exemplifies the application of AI in healthcare. These virtual assistants can guide patients in managing chronic conditions like diabetes, prescribe over-the-counter medications, and offer guidance on remote therapy sessions. Physically and socially supportive robots have proven invaluable in aiding individuals in their recovery from injuries or illnesses, bridging cognitive, motor, or sensory deficits. These technologies contribute significantly to enhancing functional abilities, independence, and overall well-being [53]. ML methods have also found application in the evaluation of patient data, clinical decision support, and diagnostic imaging. In therapy, artificial cognitive applications can assess rehabilitation sessions based on machine-generated signals [52]. The incorporation of AI-driven tools, such as the AI natural language model, into rehabilitation sessions can complement traditional therapy, offering personalized guidance and monitoring for patients during their recovery journey [61]. Furthermore, the AI language model can assist individuals in practicing speech and language skills, making it accessible both at home and outside. Studies have demonstrated ChatGPT's ability to redraft text empathetically, enhancing peer-to-peer mental health support and various community-based self-managed therapy tasks, including cognitive behavioral therapy [62]. Furthermore, the utilization of apps and online portals for patient-physician communication has significantly improved patient engagement rates. Healthcare apps securely store and distribute patient data in the cloud, providing patients with easy access to their health information and facilitating better health outcomes. AI-based medical consultation apps allow patients to obtain information (non-emergency) and even offer medication reminders. AI language model's integration into healthcare apps streamlines time-consuming tasks like summarization, note-taking, and report generation, making healthcare more efficient. These apps also assist patients in symptom checking, appointment scheduling, medication management, patient education, and the self-management of chronic diseases [63]. Various digital platforms, including mobile apps, voice assistants, and websites, can facilitate access to care assistive. However, the utilization of virtual assistants in healthcare is not without its challenges, encompassing ethical concerns, data interpretation, privacy, security, consent, and liability issues [63].

AI is sparking a transformative wave in patient care, touching upon a multitude of healthcare domains, ranging from remote monitoring to injury prevention and virtual assistance. Through ongoing innovation and comprehensive education, AI stands poised to elevate the quality and global reach of healthcare services. Yet, it simultaneously raises pertinent questions regarding ethical considerations and the evolving regulatory landscape.

4. The regulatory framework

4.1. Ethical principles and guidelines

The ethical foundation in medicine rests upon a set of fundamental principles that guide healthcare professionals in delivering compassionate and patient-centered care. Central to these principles is respect for patient autonomy, acknowledging individuals' rights to make informed decisions about their healthcare. Concurrently, the principle of beneficence underscores the healthcare provider's duty to act in the best interests of patients, promoting well-being and striving to maximize positive outcomes while minimizing harm, in adherence to the principle of non-maleficence [64].

Justice in medical ethics emphasizes the equitable distribution of healthcare resources, treatments, and opportunities, addressing disparities and ensuring universal access. Veracity and confidentiality highlight the importance of honest and transparent communication while safeguarding patient privacy.

Fidelity, or professional faithfulness, underscores the commitment of healthcare professionals to fulfill their duties and obligations, maintaining trust within the physician-patient relationship and the broader healthcare system. These principles collectively form the ethical compass that guides decision-making in medicine, ensuring a balance between individual rights, societal equity, and the integrity of the medical profession. As the healthcare landscape evolves, these principles remain essential, fostering ethical practices that prioritize patient welfare and uphold the core values of the medical profession. Continuous reflection on these ethical considerations ensures that healthcare professionals navigate the complexities of medical practice with integrity, compassion, and an unwavering commitment to ethical standards.

When developing digital technologies for healthcare, it is essential to take into account the requirements for monitoring patient safety, privacy, traceability, accountability, and security. Furthermore, plans should be established to address any breaches that may occur.

The WHO initiated in 2019 the development of a framework to facilitate the integration of digital innovations and technology into healthcare. The WHO's guidelines for digital interventions in healthcare emphasize the importance of evaluating these technologies based on factors such as 'benefits, potential drawbacks, acceptability, feasibility, resource utilization, and considerations of equity.' These recommendations underscore that these digital tools should be perceived as essential aids in the quest for universal health coverage and long-term sustainability [65].

The ethical principles for applying AI in the field of healthcare are designed to provide guidance to developers, users, and regulators to enhance the design and utilization of these technologies while ensuring proper oversight.

At the heart of all ethical principles lie human dignity and the intrinsic worth of every individual. These foundational values underpin the ethical guidelines that outline duties and responsibilities within the sphere of developing, implementing, and continually evaluating AI technologies for healthcare. The European Regulation enacted on April 21, 2021, categorizes AI products with precision based on their potential risk to fundamental rights such as health and safety, dignity, freedom, equality, democracy, the right to be free from discrimination, and data protection.

Given this classification, ethical principles play a pivotal role for all stakeholders engaged in the responsible advancement, deployment, and assessment of AI technologies for healthcare. This inclusive group encompasses physicians, system developers, healthcare system administrators, health authority policymakers, as well as local and national governments. Ethical principles should serve as catalysts, encouraging and aiding governments and public sector agencies in adapting to the rapid evolution of AI technologies through legislation and regulation. Moreover, these principles should empower medical professionals to judiciously employ AI technologies in their practice.

Within a general framework of the use of AI techniques in the service of society, six fundamental principles in favor of the ethical development of such technologies have been identified in the literature. Some of them are fundamental principles commonly used in bioethics: beneficence and non-maleficence (i.e. to do no harm and minimize the benefit/risk trade-off), autonomy (respecting the individual's interest in making decisions), justice (ensuring fairness and that no person or group is subjected to discrimination, neglect, manipulation, domination or abuse). Other principles, on the other hand, principles draw from moral and legal standards. These principles emphasize maintaining the epistemological aspect of intelligibility, encompassing both the need for clear explanations of technology's operations and the responsibility to trace cause-and-effect relationships resulting from technology's actions. Furthermore, they underscore the importance of safeguarding and upholding individual privacy to empower individuals to retain control over sensitive information concerning themselves, thereby preserving their capacity for self-determination and, in turn, respecting their autonomy [66].

In a recent WHO initiative conducted in June 2021, a comprehensive set of indications, recommendations, and guidelines about the development, application, and utilization of AI technologies in medicine was established [24]. The WHO work offered a detailed exploration of fundamental ethical principles designed to guide the development and implementation of AI technologies. Strongly endorsing the adoption of this updated document within the AI domain, the subsequent paragraphs provide a comprehensive review of the key ethical guidelines delineated by WHO for a more thorough examination of the topic. Fig. 4 provides a schematic summary of recommended ethical principles and guidelines for AI in healthcare.

AI in Healthcare: Principles and Guidelines



Fig. 4. Ethical principles and guidelines.

Protection of autonomy The integration of AI may lead to scenarios where decision-making is either transferred to or shared with machines. Upholding the principle of autonomy necessitates that any expansion of machine autonomy should not compromise human autonomy. In the context of healthcare, this implies that individuals should retain complete control over healthcare systems and medical choices. AI systems should be meticulously and consistently designed to align with established principles and human rights, specifically focusing on their role in aiding individuals, whether they are healthcare professionals or patients, in making well-informed decisions. Respecting autonomy also encompasses the accompanying responsibilities of safeguarding privacy, maintaining confidentiality, and ensuring informed and valid consent through the implementation of appropriate legal frameworks for data protection.

Promoting welfare, safety and public interest AI technologies must adhere to strict regulatory standards concerning safety, accuracy, and effectiveness before they are made available to the public. These requirements are essential to ensure not only initial quality control but also to promote continuous improvement. As a result, individuals and entities involved in funding, development, and utilization of AI technologies carry an ongoing duty to evaluate and oversee the performance of AI algorithms to confirm their intended functionality.

Ensure transparency, explainability and intelligibility AI should be comprehensible to developers, users, and regulators. Achieving transparency necessitates the provision of adequate information, which should be documented or made available before an AI technology is designed and implemented. This commitment to transparency not only enhances the overall system quality but also serves as a protective measure for patient safety and public health. For instance, system evaluators rely on transparency to identify and rectify errors, and government regulators depend on it to carry out effective oversight.

AI technologies should be as explainable as possible, and tailored to the comprehension levels of their intended audience. Striking a balance between full algorithmic explainability (even at the cost of some accuracy) and enhanced accuracy (possibly at the expense of explainability) is a critical consideration.

Promoting accountability and responsibility Accountability can be established through the application of ‘human assurance,’ which entails the evaluation of AI technologies by both patients and physicians during their development and implementation. In the context of human assurance, regulatory principles are employed both upstream and downstream of the algorithm, creating human supervision points. The primary aim is to ensure that the algorithm maintains its medical effectiveness, remains open to evaluation, and upholds ethical accountability. Consequently, the utilization of AI technologies in the field of medicine necessitates accountability within intricate systems where responsibility is distributed among various stakeholders.

In cases where AI technologies make medical decisions that result in harm to individuals, accountability processes should unequivocally identify the respective roles of producers and clinical users in the harm incurred. To avert the diffusion of responsibility, where ‘everyone’s problem becomes nobody’s responsibility,’ a robust accountability model, often referred to as ‘collective responsibility,’ holds all parties involved in the development and deployment of AI technologies accountable. This approach encourages all stakeholders to act with integrity and minimize harm. It’s worth noting that this remains a continually evolving challenge and is yet to be fully addressed in the laws of many countries.

To ensure appropriate redress for individuals and groups adversely affected by decisions made by algorithm-based systems, mechanisms for compensation must be put in place. This should encompass access to prompt and effective remedies and redress facilitated by both government bodies and companies employing AI technologies in the healthcare sector.

Ensuring inclusivity and equity Inclusivity dictates that AI employed in healthcare should be intentionally designed to promote the broadest possible and equitable utilization and access, irrespective of factors such as age, gender, income, ability, or other distinguishing characteristics. AI technologies should not solely cater to the requirements and usage patterns prevalent in high-income settings; they must also be adaptable to accommodate various devices, telecommunications infrastructures, and data transfer capabilities, particularly in less economically privileged environments. Both industry and governments bear the responsibility of bridging the ‘digital divide’ within and between countries to ensure that new AI technologies are accessible on an equitable basis.

AI developers must ensure that AI data, especially training data, are free from sampling bias and characterized by accuracy, comprehensiveness, and diversity. Special provisions should be in place to safeguard the rights and well-being of vulnerable populations, coupled with mechanisms for redress in cases where biases and discrimination emerge or are alleged.

Promoting responsiveness and sustainability AI responsiveness necessitates that designers, developers, and users engage in a continuous, systematic, and transparent assessment of AI technology to ensure that it functions effectively, and appropriately, in accordance with the expectations and requirements of the specific context in which it is deployed. When an AI technology proves to be ineffective or causes dissatisfaction, the obligation to be responsive involves instituting a structured process to resolve the issue, which may include discontinuing the technology’s use. Therefore, AI technologies should only be introduced if they can be seamlessly integrated into the healthcare system and receive adequate support. Regrettably, in under-resourced healthcare systems, new technologies are frequently underutilized, left unrepaired, or not upgraded, which squanders precious resources that could have been invested in other beneficial interventions.

Sustainability also hinges on the proactive response of governments and companies to anticipate workplace disruptions. This includes providing training for healthcare professionals to adapt to the integration of AI and addressing potential job losses due to the adoption of automated systems for routine health functions and administrative tasks.

4.2. Legislative measures

It is imperative that AI models maintain simplicity in their properties and functions to ensure ease of operation by healthcare providers [67]. Nevertheless, several hurdles hinder the widespread adoption of AI in healthcare. These challenges encompass capacity limitations in developing and maintaining infrastructure to support AI processes, elevated costs associated with data storage and backup for research purposes, and the substantial expenses required to enhance data reliability [68]. AI algorithms, while powerful, are not without their limitations, including limited applicability beyond the training domain, and susceptibility to bias [69]. To address these challenges, healthcare stakeholders should develop and execute a carefully planned strategy for AI implementation in healthcare to handle the cost, technological infrastructure, and AI system integration into clinical workflows. Furthermore, clinicians frequently experience mistrust and lack of understanding concerning AI-based clinical decision support systems, primarily due to undisclosed risks and the reliability of such systems [70]. This skepticism acts as a substantial obstacle to broad adoption. To address this, there is a growing focus on implementing explainable AI solutions to boost user trust and navigate these issues [71].

All these pieces of evidence emphasize that the increasing integration of AI technologies into healthcare necessitates effective governance to address regulatory, ethical, and trust-related concerns [72,25]. A recent study also underscores the critical role of governing AI technologies at the healthcare system level in ensuring patient safety, healthcare system accountability, bolstering clinician confidence, enhancing acceptance, and delivering substantial healthcare benefits [73].

Maintaining control over regulated domains, especially healthcare, underscores the urgency of implementing national and international regulations. These regulations are essential for the responsible integration of AI-driven applications in healthcare while upholding the tenets of medical ethics [74]. This section concentrates on the proposal and refinement of five essential European acts, which together can form a unified regulatory framework for governing AI in healthcare (see Fig. 5).

In this endeavor, the European Union (EU) has already initiated actions by enacting the General Data Protection Regulation (GDPR) in 2018. GDPR is specifically crafted to protect personal data handled by data processors or controllers operating within the EU, setting a precedent for substantial regulatory changes in the United States and Canada [75].

Moreover, the European Commission has introduced the recent Artificial Intelligence Act (AIA), which is designed to address a range of risks linked to the extensive implementation of AI [76,77]. This set of regulations advocates for the responsible deployment of AI and endeavors to prevent or alleviate potential harms arising from specific technology applications. According to this proposed act, high-risk AI systems are obligated to undergo pre-deployment compliance assessments and post-market monitoring to ensure their adherence to all the requirements outlined by the AIA [78].

Furthermore, the European Union has recently implemented the Medical Device Regulation (MDR) 2017/745/EU [79], which replaces both the Medical Device Directive (MDD) 93/42/EC [80], and the Active Implantable Medical Device Directive (AIMDD) 90/385/EC [81] which aims to enhance the regulation of the medical device market.

In general, the implementation of the MDR does not remove any requirements from the replaced regulatory acts. Instead, it introduces new ones, emphasizing a life-cycle approach to device safety. To summarize, the primary changes brought about by the MDR’s entry into force and application for medical devices are as follows:

- i) the MDR enhances controls to ensure the safety and effectiveness of devices.

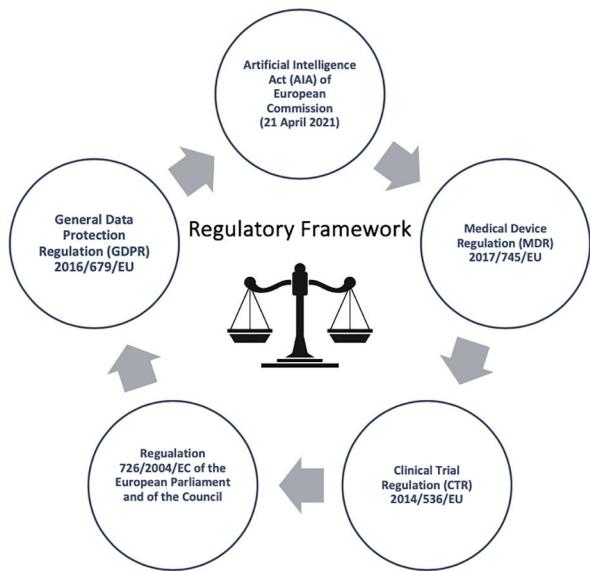


Fig. 5. Evidencing key European acts to govern AI in healthcare.

ii) the mechanism allowing for the acceleration of placing devices on the market or putting them into service through equivalence to existing devices is no longer applicable for all medical devices.

iii) post-marketing clinical follow-up is extended to all medical devices, leading to an increased importance and number of clinical evaluations and investigations.

According to Article 51(1) of the MDR, medical devices are classified into four main classes: I, IIa, IIb, and III. This classification depends on their intended purpose and inherent risks, based on the criteria specified in Annex VIII. In general, class I includes most non-invasive and non-active devices, representing the lowest risk. Class II(a) devices are of medium risk, while class II(b) devices are medium to high risk. Class III is reserved for high-risk devices. Regarding software, it is classified as class II(a) when it is intended to provide information for diagnostic or therapeutic decisions, unless these decisions could result in death or irreversible deterioration of a patient's health (in which case it is classified as class III), or if they lead to serious health deterioration or surgery (in which case it is classified as class II(b)). Furthermore, if the software is designed for monitoring physiological processes, it is classified as Class II(a), unless it is meant for monitoring vital physiological parameters, where changes in these parameters could pose an immediate danger to the patient (in which case it is classified as Class IIb).

All other software falls under class I (as per rule 11 of Annex VIII). However, it is essential to note that, concerning AI-based medical devices, class I risk classification is typically not applicable. From this point of view, however, it should be considered that, as far as AI-based MD is concerned, it is not plausible that risk class I could be applicable [82]. Nonetheless, AI technologies are medical devices in the light of their definition under the MDR. However, this conclusion requires further clarification. Specifically, the choice of applicable law hinges on whether a medical product incorporated into an AI device is considered 'ancillary', leading to the application of the MDR, or 'non-accessory', which triggers the application of laws related to medicines for human use [83]. In this context, one could argue that, according to Article 1 of the MDR, if the medicinal product incorporated into the device has an ancillary role concerning the device, it falls under the evaluation and authorization process defined by the MDR. However, if it serves a principal (i.e., non-accessory) function, the comprehensive product will be regulated by Directive 2001/83/EC of the European Parliament and of the Council, which pertains to medicinal products for human use, or by Regulation 726/2004/EC, which governs Community procedures for authorizing and overseeing medicinal products for human and veterinary use, including the establishment of the European Medicine Agency (EMA) when applicable.

Furthermore, when discussing accountability and transparency in AI, the Clinical Trial Regulation (CTR) offers relevant suggestions for regulation.

Of particular significance are the regulations surrounding two key aspects of the product lifecycle, particularly during the pre-market phase: the responsibilities of sponsors and investigators. A sponsor of a clinical study or experiment can be an individual, a company, an institution, or an organization that assumes the responsibility for initiating, managing, and financing the said investigation or experiment. On the other hand, an investigator in a clinical study or experiment is an individual responsible for conducting it at a clinical investigation site. Under the MDR, the sponsor, whether acting alone or in conjunction with the investigator, is obliged to adhere to various obligations concerning the execution of clinical investigations. These obligations, outlined in Article 11 of the MDR and Article 25 of the AIA, encompass the following tasks:

i) maintaining a copy of the EU Declaration of Conformity and technical documentation available for review by the national competent authority in the context of market surveillance;

ii) furnishing the national competent authority, upon request, with all the necessary information and documentation to demonstrate the conformity of a high-risk AI system;

iii) collaborating with the national competent authority, upon request, on any actions about the AI system.

Finally, the accountability process for regulating AI products may be extended to users, which encompasses health professionals, laypersons, legal entities, public authorities, agencies, and other organizations that utilize medical devices under their jurisdiction.

Article 29 AIA lays down a series of responsibilities for users employing high-risk AI systems. These responsibilities include adherence to the provided instructions for use, ensuring the data inputted aligns with the system's intended purpose, diligent monitoring of the system's operation as per the instructions, immediate notification of the developer or distributor, and discontinuation of the AI medical device use if there are concerns about its application not complying with the instructions for use. Users are also expected to report identified serious risks to the sponsor and investigator, utilize relevant information for fulfilling their obligations, and comply with data protection regulations specified in Article 35 of Regulation (EU) 2016/679 or Article 27 of Directive (EU) 2016/680, where applicable.

5. Open challenges

A plethora of challenges emerges in the healthcare landscape with the growing integration of AI. These challenges span technological advancements, regulatory considerations, and ethical dimensions.

This section delves into the nuanced implications of AI technologies in clinical practice, focusing the discussion on two pivotal aspects: the doctor-patient relationship and AI-driven clinical decision-making. Additionally, challenges related to health record data are addressed, offering insights into current practices regarding the utilization process of medical data.

5.1. The doctor-patient relationship

The patient-doctor relationship is a critical aspect of healthcare, characterized by mutual trust, effective communication, and collaboration. Establishing trust between patients and healthcare providers is foundational for successful medical care. It involves open and transparent communication, empathy, and a shared understanding of the patient's concerns, values, and treatment preferences [84].

Issues related to trust within the patient-doctor relationship are multifaceted. Patients place trust in the expertise and knowledge of their healthcare providers, relying on their guidance for accurate diagnoses and effective treatments. At the same time, healthcare providers trust in the information shared by patients to make informed decisions about their care. Exploring these trust dynamics is essential for a comprehensive understanding of the patient-doctor relationship, especially when integrating AI into clinical decision support systems.

The incorporation of AI prompts a reassessment of the conventional doctor-patient relationship.

Some argue that the traditional model of a nurturing relationship has become outdated. It appears that the concept of patients seeking a doctor's expert advice and placing themselves under their care is evolving into a model where patients actively participate in generating health knowledge and acquiring expertise to manage their illnesses [85].

The healing relationship must be understood as an idealistic picture of the relationship between 'expert doctors' and 'vulnerable patients'. This concept encompasses the factors that drive patients to seek professional medical assistance or leverage knowledge and technology for self-care. Whether one opts for professional medical services or self-directed care, the fiduciary responsibilities arising from this vulnerability remain consistent, regardless of the sources of expertise. These sources can include medical professionals, repositories of medical information and guidance, or other technologies and systems that support self-care, like telemedicine or easily accessible medical information on the internet.

In this context, the importance of redefining how fiduciary obligations of medicine are met takes on renewed significance, especially with the future deployment of AI in medicine. Questions of relevance have been raised, for instance, about the validity and effectiveness of medical knowledge available through Internet portals. Furthermore, as medical information becomes increasingly accessible through various means, the role of expertise as an indicator of trustworthiness does not undergo any change.

Built upon this foundation, the healing relationship model can be construed as an illustration of the ethical attributes and responsibilities intrinsic to medical practice. Traditionally, these principles were embodied by healthcare professionals, but they are now extending across diverse platforms and individuals, encompassing web portals, consumer device creators, wellness service providers, and others. While contemporary medicine has progressed beyond the conventional doctor-patient model delineated in the therapeutic relationship, the responsibilities associated with this connection have not evaporated. Instead, the diffusion and transfer of these responsibilities to new technological participants in medicine raise concerns regarding the oversight of AI integration in healthcare.

When assessing the impact of AI and algorithmic technologies on the doctor-patient relationship, the selection of metrics becomes paramount. If we solely gauge it in terms of cost-effectiveness or utility, the rationale for incorporating AI into healthcare and expanding its role is evident. Nonetheless, while algorithmic technologies might enhance efficiency and reduce the cost of treating more patients, they could potentially erode the non-mechanical aspects of care. One can distinguish between the effects of algorithmic systems (and their utility components) that contribute to the welfare of the patient or the practice of medicine, governed by established ethical standards, and those that benefit medical institutions and healthcare services.

The moral engagement intrinsic to the doctor-patient relationship, where treatment ideally stems from the practitioner's contextual and historically informed evaluation of the patient's condition, is challenging to replicate in interactions with AI systems. The role of the patient, the factors prompting individuals to seek medical assistance, and the vulnerability of patients remain unaltered

with the introduction of AI as a mediator or enhancer of medical care. What does transform is how care is administered, the possibilities for its delivery, and who delivers it. The delegation of care skills and responsibilities to AI systems can be disruptive in multiple respects.

The deployment of AI machines and robots in medicine, particularly when they demonstrate enhanced efficiency, precision, speed, and cost-effectiveness, appears appealing when considering their substitution for humans in tasks that are repetitive, tedious, perilous, demeaning, or exhausting. If used judiciously, AI has the potential to reduce the time healthcare professionals allocate to bureaucratic, routine tasks, or those that expose them to unnecessary risks, thereby offering a lower-risk, patient-focused environment.

Traditionally, clinical care and the doctor-patient relationship are ideally rooted in the physician's contextual and historically informed evaluation of the patient's condition. This type of care is challenging to replicate in technologically mediated healthcare. Patient data representations inherently confine the doctor's comprehension of the patient's case to quantifiable attributes. This can pose challenges as clinical assessments increasingly rely on data representations derived from sources like remote monitoring technologies or data collected without in-person interactions.

Patient data representations can be perceived as an 'objective' gauge of an individual's health and overall well-being, potentially diminishing the significance of contextual health elements or the perspective of the patient as a socially situated individual. These data representations can generate an 'illusion of certainty,' where 'objective' monitoring data are regarded as an accurate portrayal of the patient's condition, often overlooking the patient's interpersonal context and other unspoken knowledge [86].

Healthcare providers encounter this challenge when integrating AI systems into patient care protocols. The volume and intricacy of data and suggestions derived from technology can complicate the identification of crucial missing contextual details regarding a patient's condition. Depending solely on data obtained from health apps or monitoring technologies (e.g., smartwatches) as the primary information source about a patient's health may lead to the oversight of facets of a patient's well-being that aren't readily measurable. This encompasses vital aspects of mental health and overall well-being, such as the patient's emotional, mental, and social status. Consequently, a 'decontextualization' of the patient's state may transpire, wherein the patient relinquishes some influence over how their condition is conveyed and comprehended by healthcare professionals and caregivers. These scenarios imply that the interactions essential for cultivating the fundamental trust inherent in a traditional doctor-patient relationship may be impeded by technological intervention. Technologies that obstruct the conveyance of 'psychological signals and emotions' may impede the physician's comprehension of the patient's condition, jeopardizing 'the establishment of a physician-patient relationship grounded in trust and the pursuit of healing' [85].

Serving as an intermediary position between the doctor and the patient, AI systems modify the dynamics between healthcare providers and patients, as they delegate a portion of the patient's continuous care to a technological platform. This shift could lead to a growing divide between healthcare professionals and patients, potentially implying a missed chance to cultivate an intuitive grasp of the patient's health and overall well-being [86].

Relying on AI systems for clinical care or expert diagnostic capabilities may hinder the cultivation of expertise, professional networks, and the establishment of 'best practice' standards in the field of medicine. This phenomenon is commonly known as 'de-skilling' and contradicts the principles of 'human-centered AI' advocated by the WHO. Human-centered AI aims to bolster and enrich human skills and competency development rather than diminishing or substituting them [87].

As care becomes increasingly technologically mediated and involves non-professional entities, the development, maintenance, and enforcement of internal standards aimed at fulfilling moral obligations to patients may face potential compromises. There is a looming prospect that algorithmic systems could supplant the traditional roles of healthcare professionals, with a primary focus on efficiency and cost-effectiveness.

To safeguard against the erosion of comprehensive, patient-centered care, new care providers and entities outside the conventional medical community must place significant emphasis on upholding these moral obligations to benefit and respect patients. In doing so, they can ensure that healthcare remains not only technically 'efficient' but also holistic and genuinely beneficial.

A key role of human clinical expertise is to safeguard the well-being and safety of patients. When this human expertise is compromised through skill degradation or displaced by automation bias, it becomes essential for tests and clinical effectiveness trials to step in and bridge the gap, thus ensuring patient safety [88]. This trade-off is mirrored in the context of transparency and precision. Some scholars contend that medical AI systems may not require comprehensible explanations if their clinical accuracy and effectiveness can be consistently verified [89].

Automation in data acquisition, interpretation, diagnosis processing, and therapy identification cannot operate in complete isolation from human involvement. It necessitates continuous validation and, therefore, does not diminish the significance of the doctor-patient relationship's uniqueness. Each individual's ailment is inherently distinctive, and personal interaction remains the fundamental component of every diagnosis and treatment. In this context, machines cannot replace humans in a relationship founded on the interplay of complementary realms of autonomy, competence, and responsibility.

AI should be regarded solely as a tool to aid physicians in their decision-making processes, which must always be under human control and supervision. Ultimately, the responsibility for making the final decision rests with the doctor, as the machine's role is exclusively supportive, involving the gathering and data analysis, serving in an advisory capacity. It is important to underscore that an 'automated cognitive assistance system' in diagnostic and therapeutic procedures does not equate to an 'autonomous decision-making system' [88]. It assists by collating clinical and documentary data, comparing them with statistics on similar patients, and expediting the physician's analytical process.

An important concern arises: what happens when AI outperforms the capabilities of a doctor? In specific scenarios, this is indeed a technical possibility that must be considered. It is within this particular realm that the feared 'replacement' of machines for humans might potentially occur in the future.

However, a more immediate consequence might involve the delegation of decision-making to technology. Entrusting complex tasks to AI systems can result in the erosion of human and professional qualities. To maintain the doctor-patient relationship as one grounded in trust, in addition to care, it is crucial to preserve the essential role of the 'human doctor.' Only the human doctor possesses the unique capacities of empathy and genuine understanding that cannot be replicated by AI. While predetermined standards of behavior and codes of conduct, such as protocols and guidelines, provide support based on knowledge and experience in professional practice, the demands of diagnosis and treatment often necessitate deviation from these predetermined models.

It would be of great concern if the space seemingly left to the presumed neutrality of machines led to the 'neutralization' of the patient. The vast potential offered by AI should be viewed as a valuable opportunity through which technology can expand ethical horizons, enhancing the patient's opportunity to be heard and fostering a deeper connection with the progression of their illness. In this context, AI serves as a valuable tool that saves the physician time on routine tasks, allowing more time to be dedicated to the doctor-patient relationship.

5.2. *AI-powered clinical decision-making*

A substantial portion of the momentum in AI stems from the belief that applying these technologies in diagnosis, care, or health-care systems could enhance clinical and institutional decision-making.

Physicians and healthcare professionals are susceptible to various cognitive biases, leading to diagnostic errors. The National Academy of Sciences (NAS) has reported that approximately 5% of US adults seeking healthcare guidance are subject to misdiagnosis, with such errors contributing to 10% of all patient fatalities [90].

AI has the potential to mitigate inefficiencies and errors, leading to a more judicious allocation of resources, provided that the foundational data is both accurate and representative. Accountability plays a pivotal role in holding individuals and entities responsible for any adverse consequences of their actions. It is an indispensable element for upholding trust and safeguarding human rights.

Nevertheless, certain attributes of AI technologies introduce complexities to notions of accountability. These attributes include their opacity, reliance on human input, interactions, discretionary functions, scalability, capacity to uncover concealed insights and software intricacies. One particular challenge in ascribing responsibility arises from the 'control problem' associated with AI. In this scenario, AI developers and designers may not be held accountable, as AI-driven systems can operate independently and evolve in ways that the developer may assert are unpredictable [91]. Assigning responsibility to the developer could provide an incentive to take all possible measures to minimize harm to the patient. Such expectations are already well established for manufacturers of other commonly used medical technologies, including drug and vaccine manufacturers, medical device companies, and medical equipment manufacturers.

Another challenge is the 'traceability of harm,' which is a persistent issue in complex decision-making systems within healthcare and other domains, even in the absence of AI. Due to the involvement of numerous agents in AI development, ascribing responsibility is a complex task, entailing both legal and moral dimensions. The diffusion of responsibility can have adverse consequences, including the lack of compensation for individuals who have suffered harm, incomplete identification of the harm and its root causes, unaddressed harm, and potential erosion of societal trust in such technologies when it appears that neither developers nor users can be held accountable [92].

Physicians routinely employ various non-AI technologies in the diagnosis and treatment process, ranging from X-rays to computer software. When a medical professional commits an error in the utilization of such technology, they can be held accountable, especially if they have received training in its application [93]. However, in cases where an error arises from the algorithm or data used to train an AI technology, it may be more appropriate to assign responsibility to those involved in the development or testing of the AI system. This approach avoids placing the onus on the doctor to assess the AI technology's effectiveness and usefulness, as they might not possess the expertise to evaluate complex AI systems [94].

Numerous factors argue against placing exclusive responsibility on physicians for decisions made by AI technologies, many of which apply to assigning responsibility for the use of healthcare technologies beyond AI. To begin with, physicians do not wield control over an AI-driven technology or the recommendations it provides [95]. Nonetheless, physicians should not be entirely absolved of liability for inaccuracies in the content, as this is necessary to prevent 'automation bias' and encourage critical evaluation of whether the technology aligns with their requirements and those of the patients [92]. Automation bias occurs when a physician disregards errors that should have been identified through human-guided decision-making. While it's crucial for doctors to have trust in an algorithm, they should not set aside their own experience and judgment to blindly endorse a machine's recommendation. Automation bias occurs when a physician disregards errors that should have been identified through human-guided decision-making. While it's crucial for doctors to have trust in an algorithm, they should not set aside their own experience and judgment to blindly endorse a machine's recommendation [94].

Certain AI technologies may present not just a single decision but a range of options from which a physician must make a selection. When a physician makes an incorrect choice, determining the criteria for holding them accountable becomes a multifaceted challenge.

The complexity of attributing liability is magnified when AI technology is integrated into a healthcare system. In such cases, the developer, the institution, and the physician may all have contributed to a medical error, making it difficult to pinpoint full responsibility [90]. Consequently, accountability might not rest solely with the provider or developer of the technology; instead, it may lie with the government agency or institution responsible for selecting, validating, and implementing the technology.

Fortunately, as of today, the shift of decision-making in healthcare from humans to machines has not reached its culmination. While AI today is primarily proposed to augment human decision-making in the practice of public health and medicine, epistemic

authority has begun to debate the issue of why, in some circumstances, AI systems (as with the use of computer simulations) can move humans from the center of knowledge production [96,97]. The debates evidence a possible complete transfer of routine medical tasks to AI, prompting questions about the legality of such full delegation. Modern laws increasingly acknowledge an individual's right not to be solely subjected to automated decisions when these decisions hold substantial consequences. Additional concerns may emerge if human judgment is progressively supplanted by machine-driven assessment, giving rise to broader ethical considerations associated with the loss of human oversight, particularly if predictive healthcare becomes standard practice.

Nonetheless, it is improbable that AI in the field of medicine will attain complete autonomy; instead, it may reach a stage of conditional automation or necessitate ongoing human support [98].

Substituting human judgment with AI and relinquishing control over certain aspects of clinical care offers distinct advantages. Humans are capable of making decisions that may be less equitable and more biased compared to machines (the concern about bias in AI usage is further elaborated below).

The utilization of AI systems for specific, well-defined decisions can be entirely justified if there is compelling clinical evidence indicating that the system outperforms a human in that particular task. However, the transition to the application of AI technologies for more intricate aspects of clinical care presents a set of challenges.

One such challenge is the potential emergence of a 'disagreement between equals' between two competent experts: an AI machine and a doctor. In such scenarios, there is no feasible method for harmonizing decisions or engaging in a rationale with the algorithm since accessing it or committing to a change of mind is not possible. Furthermore, there are no clear-cut rules for determining which entity is correct, and whether a patient should place trust in technology or a doctor. The choice may hinge on factors that are not rooted in the 'competence' of either the machine or the doctor. Some argue that the algorithm's recommendation should be favored, as it incorporates the expertise of multiple individuals and a wealth of data points. However, opting for one over the other may result in an undesirable outcome. If the physician disregards the machine's recommendation, the AI's value addition can be limited. If the physician concurs with the machine's decision, it might erode their authority and diminish their responsibility [90].

The delegation of decision-making to AI-driven technologies and the subsequent loss of human control could impact various aspects of clinical care and the healthcare system.

While offering individuals more opportunities to share data and access autonomous health advice might enhance their sense of empowerment and self-care, it could also potentially trigger anxiety and fatigue [93]. As these technologies accumulate more personal data and incorporate it into the decision-making process of physicians, patients might find themselves gradually marginalized in shared decision-making, potentially compromising their ability to exercise free will or autonomy in health-related decisions [90].

The utilization of AI in medicine, especially when its use is not disclosed, poses challenges to the fundamental principles of informed consent and broader public trust in healthcare. While some may view the reduction of physician control over patients as a way to promote patient autonomy, there is an equally significant risk of relinquishing decision-making to AI technology. This risk is heightened if the technology is presented to the patient as offering a superior understanding of their health status and prognosis compared to a physician [94].

5.3. Health data records

The integration of AI technology in the healthcare sector faces notable limitations, particularly in the realm of managing health record data. Health data encompasses a patient's complete medical history, including details from physical examinations, investigations, and treatment records—all stored in electronic form.

The 2018 WHO-UNICEF Global Primary Health Care (PHC) Conference underscored the importance of harnessing AI techniques to augment existing systems and leverage data for electronic decision support and analytics [99]. The overarching objective is to attain Universal Health Coverage and advance healthcare equity for both individuals and populations. However, ethical challenges arise in the context of health records. The use of electronic health records introduces ethical issues related to beneficence, autonomy, fidelity, and justice. Autonomy is compromised when patients' health data are shared or linked without their knowledge. Fidelity is at risk when health organizations fail to take adequate precautions to safeguard identifiable health data. Justice is undermined when disparities exist in access to health information and services based on income, language, age, geography, literacy, and disability [100].

A comprehensive examination of health records and their implications necessitates the involvement of various stakeholders, including health personnel, leaders, policymakers, ethicists, and engineers.

Despite significant strides in computerization, challenges persist throughout the data life cycle—from collection and documentation to storage, management, sharing, and utilization [101,102]. Efforts to address these challenges are essential to fully realize the potential benefits of AI in healthcare.

Legislation for AI regulation in the EU emphasizes trustworthy and ethical AI with explicit processes and human oversight for high-risk systems. Equitable AI is proposed as a solution, emphasizing fairness in capacity-building strategies. Integrating data quality assessment and management with information governance in the big data environment is essential for ethical health information ecosystems.

To effectively navigate the complexities of integrating AI technology into healthcare, a comprehensive approach that combines data usage, data quality management, governance, and ethics is essential. In this context, key recommendations have been identified to guide stakeholders in the enhancement of data utilization in healthcare: transparency, use limitation, access and correction, data quality, and security [103].

Healthcare professionals play significant roles as creators, collectors, managers, and users of observational health data. This raises the question of whether we need new consent models from clinicians, patients, and other healthcare providers within the referral and integrated care network at different points in the data life cycle. Determining the adequacy of consent, its informativeness, and relevance becomes a critical consideration, prompting the need for clear criteria and decision points throughout the life cycle.

In promoting good governance, it is essential to establish consent or employ other legal processes, such as opting out or utilization for public health purposes, to regulate access and sharing. Adhering to ethical principles, data quality categories, and information governance is crucial, ensuring their relevance at various stages in the data life cycle. Innovative technologies like blockchain are actively being utilized to expedite the consent process for clinical trials, potentially alleviating concerns related to personal data privacy associated with integrated service delivery [104].

Establishing and maintaining trust is paramount in implementing sustainable data creation and collection practices.

It's crucial to recognize that community-driven health data repositories may not provide the level of privacy citizens assume, particularly in the context of free online services with user agreements allowing the service owner to utilize collected data. The deployment of applications and systems facilitating unethical and unlawful exchange of digital information erodes trust. Building and maintaining trust in the health information exchange network requires reciprocity, transparency, and mutual trust among all actors, including data custodians and providers. The trustworthiness of data custodians hinges on their competence, commitment, and motives. Patients and service users generally have high levels of trust in the professionalism of clinical teams and health services. However, unnecessary tests and procedures resulting from unclear data or misinterpretations of unstructured data can compromise patient autonomy and hinder informed decision-making.

In the United States, legislation such as the Health Insurance Portability and Accountability Act (HIPAA) specifically excludes protections once data leaves a covered entity. Controversy surrounds the transmission of data to large companies for data mining, primarily due to the lack of explicit consent from patients, placing control in the hands of system designers and owners. In contrast, the EU's GDPR empowers individuals with control over their personal data.

Careful consideration is needed for all data processes, as they may harbor unrecognized risks.

Transferring data from information systems to data repositories demands a commitment to security, safety, and accuracy [105]. Privacy is a paramount concern, and while privacy-preserving linkage techniques exist for integrating observational data from information systems, their accuracy and security are not always guaranteed [106]. Beyond the threat of re-identification during data, there exists the potential for data loss and compromise of data integrity.

For instance, a recent study underscored inaccuracies in cohort identification when utilizing vocabulary mappings within a common data model during the data process [107]. These mappings, integral to data, may suffer from inaccuracies stemming from programming bugs and errors that escape detection during quality assurance stages.

It is imperative to thoroughly identify, assess, and establish contingency plans to mitigate all risks associated with the data process. Vigilance in recognizing potential pitfalls and implementing safeguards is essential for ensuring the reliability and security of the data transfer process [108].

Transparent data processing is imperative, along with the ethical and secure sharing of research data, methodologies, and algorithms. Achieving reproducibility and generalizability while safeguarding patient privacy, financial investments, and intellectual property is essential.

Explainable AI (XAI) algorithms allow doctors to understand and validate results, aligning with the 'learned intermediary' principle where clinicians play a central role in decision-making. However, discerning between biased or inadequately explained AI guidance and the clinician's interpretation poses challenges. Collaborative efforts among clinicians can potentially address these biases and enhance the design of XAI, ensuring a critical appraisal of AI guidance in primary care that augments rather than undermines the patient-physician relationship [109].

Conscious efforts to develop discrimination-aware algorithms aim to reduce bias in databases and associated applications, ensuring that data mining models do not lead to discriminatory decisions against vulnerable groups, even if the underlying dataset is inherently biased. Empirical studies evaluating AI practices are necessary to highlight ethical concerns and propose solutions. Trials comparing AI-based practices with conventional methods are vital for evidence-based policymaking in the face of poorly regulated AI practices. Reducing the opacity and complexity of AI methods, including deep machine learning and neural networks, is imperative.

The ethical governance of AI platforms necessitates modifications to traditional medical research ethics principles. This includes considerations for informed consent, protecting individual and group-level harms and benefits, ensuring patient empowerment, preserving the patient-doctor relationship, and upholding research subject rights in AI-supported projects. Data protection regulations for research and personal health services must be robust and effectively enforced.

In conclusion, adopting an integrated strategy encompassing data utilization, quality assurance, governance, and ethical considerations is pivotal. Key recommendations involve ensuring consent remains a focal point throughout the data life cycle, fostering sustainable data practices, meticulously examining data processes, seamlessly integrating data governance with quality management, navigating ethical challenges posed by AI, and applying a structured ethical framework aligned with the data life cycle. These proactive measures collectively strive to uphold trust in current systems and anticipated developments.

6. Conclusions

This study serves as a valuable addition to the existing literature by aggregating insights into AI's applications in healthcare. It also provides a comprehensive perspective on the ethical and governance challenges encountered by stakeholders interested in introducing AI in healthcare.

In the ongoing narrative of healthcare, AI assumes a pivotal role. As the driving force behind precision medicine, it fulfills a pressing need for enhanced care. While initial efforts to provide diagnoses and treatment recommendations have been met with many challenges, the trajectory suggests that AI will eventually master this domain.

The primary hurdle to widespread AI adoption in healthcare does not revolve around the capabilities of the technology but rather centers on ensuring seamless integration into daily clinical practice. For comprehensive adoption, AI systems must receive regulatory approval, attain standardized protocols that guarantee uniform functionality, train clinicians in their use, secure financial support from public and private payer organizations, and undergo iterative supervision updates. While these adoption challenges will ultimately be surmounted, they will necessitate more time than the maturation of the technologies themselves.

AI technologies are making remarkable strides across a spectrum of healthcare applications, serving as a vital catalyst for advancements in medical diagnosis, virtual patient care, treatment adherence, and administrative efficiency. Nevertheless, AI confronts a constellation of technical, ethical, and governance challenges on its path to healthcare integration. Issues about data security and privacy loom large due to the utilization of sensitive, legally bound health data. Additionally, the limitations of AI in mirroring distinctly human emotional and behavioral traits, like compassion, can pose constraints on its utility.

Crucially, it is becoming increasingly evident that AI systems are not poised to replace human clinicians on a grand scale. Instead, they are poised to augment the capabilities of human caregivers. As time progresses, healthcare professionals are likely to shift their focus toward roles that leverage uniquely human skills. Meanwhile, AI is poised to evolve into an essential tool, enhancing the collaborative efforts of clinicians as they strive to deliver superior patient care.

AI undeniably offers significant advantages, but it is crucial to emphasize that it can never supplant the human connections that underlie collaborative healthcare decisions. The intricate dynamics of multidisciplinary teamwork elude machines, as they cannot forge the profound bonds that humans establish through their experiential knowledge in personalizing patient care. To tackle these challenges, the future governance of AI must prioritize human expertise and experience in overseeing these technologies, ensuring that the ultimate decisions, even when contrary to AI recommendations, remain in the hands of healthcare professionals.

Additionally, the lack of AI accountability continues to pose a significant obstacle to its adoption. AI systems operate as black boxes, taking inputs and generating outputs without disclosing their underlying measurements or reasoning, a challenge known as the black-box problem. To tackle this issue and prevent healthcare practitioners from being wrongly held responsible for AI errors, the implementation of standardized policies and governmental measures is imperative.

A holistic approach involving all stakeholders, including providers, payers, and patients, is crucial for understanding clinical needs before advancing the adoption of AI in healthcare. To ensure the secure and proficient use of AI in the future, it is advisable to include AI training in the curricula of healthcare professionals. This will enhance their understanding of potential risks and their expectations regarding AI performance.

Concluding, this review may have far-reaching implications across various stakeholders. It may serve as a valuable resource offering policymakers insights into potential pitfalls and hurdles in integrating these systems into healthcare. This knowledge may facilitate the development of frameworks that can balance innovation with the protection of patient safety, privacy, and ethical considerations. Healthcare professionals may benefit from the guidance provided, enabling them to navigate challenges, make informed decisions, and incorporate AI technologies responsibly into clinical workflows. Moreover, the research may contribute to responsible development by aiding developers and technologists in designing systems that can prioritize patient safety, privacy, and compliance with healthcare standards. The identification of ethical and regulatory challenges may also serve as a proactive approach to risk mitigation, contributing to the overall success and acceptance of AI technologies. Patient trust, a critical factor in healthcare, may be enhanced by addressing ethical concerns and ensuring compliance with regulations, fostering transparency in the implementation of AI technologies. Additionally, the research may support education and training initiatives for healthcare professionals and stakeholders, ensuring they may be well-informed and equipped to navigate the complexities of AI technologies. Lastly, the study underscores the need for continuous improvement in ethical guidelines and regulatory frameworks to keep pace with the evolving landscape of AI technology in healthcare.

CRediT authorship contribution statement

Ciro Mennella: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Umberto Maniscalco:** Writing – review & editing, Supervision, Methodology. **Giuseppe De Pietro:** Writing – review & editing. **Massimo Esposito:** Writing – review & editing, Validation, Supervision.

Declaration of competing interest

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version. This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue. The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

Acknowledgement

Ciro Mennella is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma (Via Alvaro del Portillo 21, 00128 Roma, Italy).

References

- [1] W.H. Organization, et al., Integrated care for older people: guidelines on community-level interventions to manage declines in intrinsic capacity, 2017.
- [2] W.H. Organization, et al., Ageism in artificial intelligence for health: who policy brief, 2022.
- [3] F. Magrabi, E. Ammenwerth, J.B. McNair, N.F. De Keizer, H. Hyppönen, P. Nykänen, M. Rigby, P.J. Scott, T. Vehko, Z.S.-Y. Wong, et al., Artificial intelligence in clinical decision support: challenges for evaluating ai and practical implications, *Yearb. Med. Inform.* 28 (2019) 128–134.
- [4] Y. Kumar, A. Koul, R. Singla, M.F. Ijaz, Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda, *J. Ambient Intell. Humaniz. Comput.* (2022) 1–28.
- [5] P.N. Srinivasu, J.G. SivaSai, M.F. Ijaz, A.K. Bhoi, W. Kim, J.J. Kang, Classification of skin disease using deep learning neural networks with mobilenet v2 and lstm, *Sensors* 21 (2021) 2852.
- [6] A. Vulli, P.N. Srinivasu, M.S.K. Sashank, J. Shafi, J. Choi, M.F. Ijaz, Fine-tuned densenet-169 for breast cancer metastasis prediction using fastai and 1-cycle policy, *Sensors* 22 (2022) 2988.
- [7] G.E. Rao, B. Rajitha, P.N. Srinivasu, M.F. Ijaz, M. Woźniak, Hybrid framework for respiratory lung diseases detection based on classical cnn and quantum classifiers from chest x-rays, *Biomed. Signal Process. Control* 88 (2024) 105567.
- [8] I. Ahmad, A. Merla, F. Ali, B. Shah, A.A. AlZubi, M.A. AlZubi, A deep transfer learning approach for Covid-19 detection and exploring a sense of belonging with diabetes, *Front. Public Health* 11 (2023).
- [9] S.P. Praveen, P.N. Srinivasu, J. Shafi, M. Woźniak, M.F. Ijaz, Resnet-32 and fastai for diagnoses of ductal carcinoma from 2d tissue slides, *Sci. Rep.* 12 (2022) 20804.
- [10] M. Kumar, S. Verma, A. Kumar, M.F. Ijaz, D.B. Rawat, et al., Anaf-iomt: a novel architectural framework for iomt-enabled smart healthcare system by enhancing security based on recc-vc, *IEEE Trans. Ind. Inform.* 18 (2022) 8936–8943.
- [11] R.F. Baumeister, M.R. Leary, Writing narrative literature reviews, *Rev. Gen. Psychol.* 1 (1997) 311–320.
- [12] R.E. Slavin, Best evidence synthesis: an intelligent alternative to meta-analysis, *J. Clin. Epidemiol.* 48 (1995) 9–18.
- [13] J. McCarthy, M.L. Minsky, N. Rochester, C.E. Shannon, A proposal for the dartmouth summer research project on artificial intelligence, August 31, 1955, *AI Mag.* 27 (2006) 12–14.
- [14] A.L. Samuel, Some studies in machine learning using the game of checkers, *IBM J. Res. Dev.* 44 (2000) 206–226.
- [15] Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 1798–1828.
- [16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (2015) 436–444.
- [17] J. Schmidhuber, Deep learning in neural networks: an overview, *Neural Netw.* 61 (2015) 85–117.
- [18] F. La Rue, Report of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression, Human Rights Council, 2011, p. 16.
- [19] P. O. for Economic Cooperation Development, Recommendation of the council on artificial intelligence (Oecd Legal Instruments. OECD/LEGAL/O449) 2019.
- [20] F.C. La Vattiatia, Ai-based medical devices: the applicable law in the European Union, *BioLaw J.-Riv. BioDiritto* (2022) 412–437.
- [21] R. Liu, M. Wang, T. Zheng, R. Zhang, N. Li, Z. Chen, H. Yan, Q. Shi, An artificial intelligence-based risk prediction model of myocardial infarction, *BMC Bioinform.* 23 (2022) 1–17.
- [22] O. Lamberg, R. Lehner, K. Chua, S.K. Wee, D.K. Rajeswaran, C.W.K. Kuah, W.T. Ang, P. Liang, D. Campolo, A. Hussain, et al., Neurorehabilitation from a distance: can intelligent technology support decentralized access to quality therapy?, *Front. Robot. AI* 8 (2021) 612415.
- [23] C. Mennella, U. Maniscalco, G. De Pietro, M. Esposito, The role of artificial intelligence in future rehabilitation services: a systematic literature review, *IEEE Access* (2023).
- [24] S.J. Russell, *Artificial Intelligence a Modern Approach*, Pearson Education, Inc., 2010.
- [25] W.H. Organization, et al., Ethics and Governance of Artificial Intelligence for Health, Who Guidance, 2021.
- [26] L.N.H. Service, The topol review: preparing the healthcare workforce to deliver the digital future, <https://topol.hee.nhs.uk/>, 2019.
- [27] T. Nadarzynski, O. Miles, A. Cowie, D. Ridge, Acceptability of artificial intelligence (ai)-led chatbot services in healthcare: a mixed-methods study, *Digit. Health* 5 (2019) 2055207619871808.
- [28] J. Roski, B. Hamilton, W. Chapman, J. Heffner, R. Trivedi, G. Del Fiol, R. Kukafka, P. Bleicher, H. Estiri, J. Klann, et al., How Artificial Intelligence Is Changing Health and Health Care, 2019.
- [29] A. Gamble, Artificial intelligence and mobile apps for mental healthcare: a social informatics perspective, *Aslib J. Inf. Manag.* 72 (2020) 509–523.
- [30] J. Bush, How ai is taking the scut work out of health care, *Harv. Bus. Rev.* 5 (2018).
- [31] Y. Kumar, A. Koul, R. Singla, M.F. Ijaz, Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda, *J. Ambient Intell. Humaniz. Comput.* (2022) 1–28.
- [32] P. Ghosh, Ai early diagnosis could save heart and cancer patients, *Science Correspondent, BBC News* (2018).
- [33] D. Wang, A. Khosla, R. Gargyea, H. Irshad, A.H. Beck, Deep learning for identifying metastatic breast cancer, *arXiv preprint, arXiv:1606.05718*, 2016.
- [34] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, *Nature* 542 (2017) 115–118.
- [35] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al., Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning, *arXiv preprint, arXiv:1711.05225*, 2017.
- [36] M. Thomas, S. Murali, B.S.S. Simpson, A. Freeman, A. Kirkham, D. Kelly, H.C. Whitaker, Y. Zhao, M. Emberton, J.M. Norris, Use of artificial intelligence in the detection of primary prostate cancer in multiparametric mri with its clinical outcomes: a protocol for a systematic review and meta-analysis, *BMJ Open* 13 (2023) e074009.
- [37] G. Bedi, F. Carrillo, G.A. Cecchi, D.F. Slezak, M. Sigman, N.B. Mota, S. Ribeiro, D.C. Javitt, M. Copelli, C.M. Corcoran, Automated analysis of free speech predicts psychosis onset in high-risk youths, *npj Schizophr.* 1 (2015) 1–7.
- [38] G. Cecchi, With AI, our words will be a window into our mental health, 2017.
- [39] C.-Y. Chou, D.-Y. Hsu, C.-H. Chou, Predicting the onset of diabetes with machine learning methods, *J. Pers. Med.* 13 (2023) 406.
- [40] A. Gudigar, U. Raghavendra, S. Nayak, C.P. Ooi, W.Y. Chan, M.R. Gangavarapu, C. Dharmik, J. Samanth, N.A. Kadri, K. Hasikin, et al., Role of artificial intelligence in Covid-19 detection, *Sensors* 21 (2021) 8045.
- [41] V.V. Khanna, K. Chadaga, N. Sampathila, S. Prabhu, R. Chadaga, S. Umakanth, Diagnosing Covid-19 using artificial intelligence: a comprehensive review, *Netw. Model. Anal. Health Inform. Bioinform.* 11 (2022) 25.
- [42] K.S. Krishnan, K.S. Krishnan, Vision transformer based Covid-19 detection using chest x-rays, in: 2021 6th International Conference on Signal Processing, Computing and Control (ISPC), IEEE, 2021, pp. 644–648.
- [43] S. Secinaro, D. Calandra, A. Secinaro, V. Muthurangu, P. Biancone, The role of artificial intelligence in healthcare: a structured literature review, *BMC Med. Inform. Decis. Mak.* 21 (2021) 1–23.
- [44] O. Oren, B.J. Gersh, D.L. Bhatt, Artificial intelligence in medical imaging: switching from radiographic pathological data to clinically meaningful endpoints, *Lancet Digit. Health* 2 (2020) e486–e488.
- [45] R. Aggarwal, S.S. Ganvir, et al., Artificial intelligence in physiotherapy, *Physiotherapy* 15 (2021) 55.

- [46] O. Lambercy, R. Lehner, K. Chua, S.K. Wee, D. Rajeswaran, C. Kuah, W. Ang, P. Liang, D. Campolo, A. Hussain, G. Aguirre-Ollinger, C. Guan, C. Kanzler, N. Wenderoth, R. Gassert, Neurorehabilitation from a distance: can intelligent technology support decentralized access to quality therapy?, *Front. Robot. AI* 8 (2021) 612415.
- [47] A. Natarajan, H.-W. Su, C. Heneghan, Assessment of physiological signs associated with Covid-19 measured using wearable devices, *npj Digit. Med.* 3 (2020) 156.
- [48] A. Al Kuwaiti, K. Nazer, A. Al-Reedy, S. Al-Shehri, A. Al-Muhanna, A.V. Subbarayalu, D. Al Muhamna, F.A. Al-Muhanna, A review of the role of artificial intelligence in healthcare, *J. Pers. Med.* 13 (2023) 951.
- [49] M.M. Baig, H. GholamiHosseini, A.A. Moqeem, F. Mirza, M. Lindén, A systematic review of wearable patient monitoring systems—current challenges and opportunities for clinical adoption, *J. Med. Syst.* 41 (2017) 1–9.
- [50] T. Shaik, X. Tao, N. Higgins, L. Li, R. Gururajan, X. Zhou, U.R. Acharya, Remote patient monitoring using artificial intelligence: current state, applications, and challenges, *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 13 (2023) e1485.
- [51] J. Kim, A.S. Campbell, J. Wang, Wearable non-invasive epidermal glucose sensors: a review, *Talanta* 177 (2018) 163–170.
- [52] D. Anderson, Artificial intelligence and applications in PM&R, *Am. J. Phys. Med. Rehabil.* 98 (2019) e128–e129.
- [53] D.D. Luxton, L.D. Riek, Artificial intelligence and robotics in rehabilitation, 2019, <https://doi.org/10.1037/0000129-031>.
- [54] C.L. Goldzweig, G. Orshansky, N.M. Paige, A.A. Towfigh, D.A. Hagstrom, I. Miale-Lye, J.M. Beroes, P.G. Shekelle, Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review, *Ann. Intern. Med.* 159 (2013) 677–687.
- [55] C.A. Sinsky, R. Willard-Grace, A.M. Schutzbank, T.A. Sinsky, D. Margolius, T. Bodenheimer, In search of joy in practice: a report of 23 high-functioning primary care practices, *Ann. Fam. Med.* 11 (2013) 272–278.
- [56] V.-T. Tran, C. Riveros, P. Ravaud, Patients' views of wearable devices and ai in healthcare: findings from the compare e-cohort, *npj Digit. Med.* 2 (53) (2019).
- [57] D. Solino-Fernandez, A. Ding, E. Bayro-Kaiser, E.L. Ding, Willingness to adopt wearable devices with behavioral and economic incentives by health insurance wellness programs: results of a US cross-sectional survey with multiple consumer health vignettes, *BMC Public Health* 19 (2019) 1–8.
- [58] Y. Gao, H. Li, Y. Luo, An empirical study of wearable technology acceptance in healthcare, *Ind. Manag. Data Syst.* 115 (2015) 1704–1723.
- [59] S. Chidambaram, Y. Maheswaran, K. Patel, V. Sounderajah, D.A. Hashimoto, K.P. Seastedt, A.H. McGregor, S.R. Markar, A. Darzi, Using artificial intelligence-enhanced sensing and wearable technology in sports medicine and performance optimisation, *Sensors* 22 (2022) 6920.
- [60] J.W. Gichoya, L.G. McCoy, L.A. Celi, M. Ghassemi, Equity in essence: a call for operationalising fairness in machine learning for healthcare, *BMJ Health Care Inf.* 28 (2021).
- [61] C. Mennella, U. Maniscalco, G. De Pietro, M. Esposito, A deep learning system to monitor and assess rehabilitation exercises in home-based remote and unsupervised conditions, *Comput. Biol. Med.* (2023) 107485.
- [62] A. Sharma, I.W. Lin, A.S. Miner, D.C. Atkins, T. Althoff, Human–ai collaboration enables more empathic conversations in text-based peer-to-peer mental health support, *Nat. Mach. Intell.* 5 (2023) 46–57.
- [63] M. Javaid, A. Haleem, R.P. Singh, Chatgpt for healthcare services: an emerging stage for an innovative perspective, in: *BenchCouncil Trans. Benchmarks, Stand. Eval.*, vol. 3, 2023, 100105.
- [64] T.L. Beauchamp, J.F. Childress, *Principles of Biomedical Ethics*, Oxford University Press, USA, 2001.
- [65] W. Guideline, *Recommendations on Digital Interventions for Health System Strengthening*, World Health Organization, 2019, 2020-10.
- [66] What is 'biosurveillance'? The Covid-19 measures getting under our skin, <https://medium.com/digital-freedom-fund/what-is-biosurveillance-c8bffe70d16f-2020>.
- [67] R. Brown, Challenges to successful ai implementation in healthcare, *Data Science Central*, 2022.
- [68] K. Tachkov, A. Zemplenyi, M. Kamusheva, M. Dimitrova, P. Siirtola, J. Pontén, B. Nemeth, Z. Kalo, G. Petrova, Barriers to use artificial intelligence methodologies in health technology assessment in central and East European countries, *Front. Public Health* 10 (2022) 921226.
- [69] G. Marcus, Deep learning: a critical appraisal, *arXiv preprint*, arXiv:1801.00631, 2018.
- [70] A. Choudhury, O. Asan, Impact of accountability, training, and human factors on the use of artificial intelligence in healthcare: exploring the perceptions of healthcare practitioners in the US, *Hum. Factors Healthc.* 2 (2022) 100021.
- [71] F. Giuste, W. Shi, Y. Zhu, T. Naren, M. Isgut, Y. Sha, L. Tong, M. Gupte, M.D. Wang, Explainable artificial intelligence methods in combating pandemics: a systematic review, *IEEE Rev. Biomed. Eng.* (2022).
- [72] S. Reddy, S. Allan, S. Coghlan, P. Cooper, A governance model for the application of ai in health care, *J. Am. Med. Inform. Assoc.* 27 (2020) 491–497.
- [73] J.S. Marwaha, A.B. Landman, G.A. Brat, T. Dunn, W.J. Gordon, Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation, *npj Digit. Med.* 5 (2022) 13.
- [74] A. House of Lords Select Committee, et al., *AI in the UK: Ready, Willing and Able*, House of Lords, 2018, p. 36.
- [75] M.B. Forcier, H. Gallois, S. Mullan, Y. Joly, Integrating artificial intelligence into health care through data access: can the gdpr act as a beacon for policymakers?, *J. Law Biosci.* 6 (2019) 317–335.
- [76] European Commission, *Proposal for a Regulation Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, COM (2021) 206 final; European Commission, Brussels, Belgium, 2021, p. 206.
- [77] European Commission, *Commission Staff Working Document, Impact Assessment, Accompanying the Proposal for a Regulation Laying down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act)*, SWD (2021) 84 final, Part 1/2, 2021, p. 1.
- [78] M. Schaake, *The European Commission's Artificial Intelligence Act*, Standford University Human-Centered Artificial Intelligence (BHAI), Standford, Canada, 2021, 2021-06.
- [79] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on Medical Devices, Amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and Repealing Council Directives 90/385/EEC and 93/42/EEC.
- [80] Council Directives 90/385/EEC and 93/42/EEC.
- [81] I. Habli, T. Lawton, Z. Porter, Artificial intelligence in health care: accountability and safety, *Bull. World Health Organ.* 98 (2020) 251.
- [82] N. Martelli, D. Eskenazy, C. Déan, J. Pineau, P. Prognon, G. Chatellier, M. Sapoval, O. Pellerin, New European regulation for medical devices: what is changing?, *Cardiovasc. Interv. Radiol.* 42 (2019) 1272–1278.
- [83] A. Migliore, On the new regulation of medical devices in Europe, *Expert Rev. Med. Devices* 14 (2017) 921–923.
- [84] S.D. Goold, M. Lipkin Jr, The doctor–patient relationship: challenges, opportunities, and strategies, *J. Gen. Intern. Med.* 14 (1999) S26.
- [85] L. Floridi, J. Cowls, A unified framework of five principles for ai in society, in: *Machine Learning and the City: Applications in Architecture and Urban Design*, 2022, pp. 535–545.
- [86] D. Lupton, M-health and health promotion: the digital cyborg and surveillance society, *Soc. Theory Health* 10 (2012) 229–244.
- [87] M. Coeckelbergh, E-care as craftsmanship: virtuous work, skilled engagement, and information technology in health care, *Med. Health Care Philos.* 16 (2013) 807–816.
- [88] D.N. Kluttz, D.K. Mulligan, Automated decision support technologies and the legal profession, *Berkeley Technol. Law J.* 34 (2019) 853–890.
- [89] C. of Europe, <https://rm.coe.int/responsability-and-ai-en/168097d9c5>, 2019.
- [90] E.J. Topol, High-performance medicine: the convergence of human and artificial intelligence, *Nat. Med.* 25 (2019) 44–56.
- [91] T. S. N. C. on Medical Ethics, In brief – artificial intelligence in healthcare, <https://smer.se/wp-content/uploads/2020/06/smer-2020-2-in-brief-artificial-intelligence-in-healthcare.pdf>, 2020.

- [92] K. Yeung, A study of the implications of advanced digital technologies (including ai systems) for the concept of responsibility within a human rights framework, *MSI-AUT*(2018)05, 2018.
- [93] T. Grote, P. Berens, On the ethics of algorithmic decision-making in healthcare, *J. Med. Ethics* 46 (2020) 205–211.
- [94] J. Ross, C. Webb, F. Rahman, *Artificial Intelligence in Healthcare*, Academy of Medical Royal Colleges, London, 2019.
- [95] M. Braun, P. Hummel, S. Beck, P. Dabrock, Primer on an ethics of ai-based decision support systems in the clinic, *J. Med. Ethics* 47 (2021) e3–3.
- [96] B. Babic, S. Gerke, T. Evgeniou, I.G. Cohen, Beware explanations from ai in health care, *Science* 373 (2021) 284–286.
- [97] J.M. Durán, Computer simulations in science and engineering, 2018.
- [98] P. Humphreys, The philosophical novelty of computer simulation methods, *Synthese* 169 (2009) 615–626.
- [99] W.H. Organization, et al., Framework on Integrated, People-Centred Health Services, World Health Organization, Geneva, 2016, p. 2019.
- [100] E.J. Layman, Ethical issues and the electronic health record, *Heal. Care Manag.* 27 (2008) 165–176.
- [101] S. De Lusignan, S.-T. Liaw, P. Krause, V. Curcin, M.T. Vicente, G. Michalakidis, L. Agreus, P. Leysen, N. Shaw, K. Mendis, Key concepts to assess the readiness of data for international research: data quality, lineage and provenance, extraction and processing errors, traceability, and curation, *Yearb. Med. Inform.* 20 (2011) 112–120.
- [102] S.-T. Liaw, G. Powell-Davies, C. Pearce, H. Britt, L. McGlynn, M.F. Harris, Optimising the use of observational electronic health record data: current issues, evolving opportunities, strategies and scope for collaboration, *Aust. Fam. Phys.* 45 (2016) 153–156.
- [103] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L.B. da Silva Santos, P.E. Bourne, et al., The fair guiding principles for scientific data management and stewardship, *Sci. Data* 3 (2016) 1–9.
- [104] M. Benchoufi, P. Ravaud, Blockchain technology for improving clinical research quality, *Trials* 18 (2017) 1–5.
- [105] F. Prasser, H. Spengler, R. Bild, J. Eicher, K.A. Kuhn, Privacy-enhancing etl-processes for biomedical data, *Int. J. Med. Inform.* 126 (2019) 72–81.
- [106] C. Culnane, B.I. Rubinstein, V. Teague, Health data in an open world, arXiv preprint, arXiv:1712.05627, 2017.
- [107] G.N. Guo, J. Jonnagaddala, S. Farshid, V. Huser, C. Reich, S.-T. Liaw, Comparison of the cohort selection performance of Australian medicines terminology to anatomical therapeutic chemical mappings, *J. Am. Med. Inform. Assoc.* 26 (2019) 1237–1246.
- [108] S.-T. Liaw, H. Liyanage, C. Kuziemsky, A.L. Terry, R. Schreiber, J. Jonnagaddala, S. de Lusignan, Ethical use of electronic health record data and artificial intelligence: recommendations of the primary care informatics working group of the international medical informatics association, *Yearb. Med. Inform.* 29 (2020) 051.
- [109] S.Y. Lin, M.R. Mahoney, C.A. Sinsky, Ten ways artificial intelligence will transform primary care, *J. Gen. Intern. Med.* 34 (2019) 1626–1630.