BenchCouncil Transactions on Benchmarks, Standards and Evaluations 5 (2025) 100215

Contents lists available at ScienceDirect

BenchCouncil Transactions on Benchmarks, Standards and Evaluations

journal homepage: www.keaipublishing.com/en/journals/benchcouncil-transactions-onbenchmarks-standards-and-evaluations/

Review Article

Ethical and regulatory challenges in machine learning-based healthcare systems: A review of implementation barriers and future directions☆

Shehu Mohammed *, Neha Malhotra

School of Computer Applications, Lovely Professional University Pagwara, 14411, India

ARTICLE INFO

Keywords:
Artificial intelligence (AI) Ethics
Algorithmic bias
Explainable AI (XAI)
Machine learning (ML) in healthcare
Patient data privacy
Regulatory compliance (GDPR, HIPAA, FDA)

ABSTRACT

Machine learning significantly enhances clinical decision-making quality, directly impacting patient care with early diagnosis, personalized treatment, and predictive analytics. Nonetheless, the increasing proliferation of such ML applications in practice raises potential ethical and regulatory obstacles that may prevent their widespread adoption in healthcare. Key issues concern patient data privacy, algorithmic bias, absence of transparency, and ambiguous legal liability. Fortunately, regulations like the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and the FDA AI/ML guidance have raised important ways of addressing things like fairness, explainability, legal compliance, etc.; however, the landscape is far from risk-free. AI liability is another one of the gray areas approaching black, worrying about who is liable for an AI medical error — the developers, the physicians, or the institutions. The study reviews ethical risks and potential opportunities, as well as regulatory frameworks and emerging challenges in AI-driven healthcare. It proposes solutions to reduce bias, improve transparency, and enhance legal accountability. This research addresses these challenges to support the safe, fair, and effective deployment of ML-based systems in clinical practice, guaranteeing that patients can trust, regulators can approve, and healthcare can use them.

Introduction

Background: Machine learning (ML) is fundamentally transforming healthcare; ML is playing an integral role in the development of methods for early diagnosis and treatment optimization, as well as predictive analytics providing unprecedented improvements for medical decision-making [1]. ML applications that leverage this technology to achieve better healthcare outcomes include diagnostic imaging analysis, personalized treatment recommendations, and predictive modeling for disease progression [2]. These technologies can help minimize human error, enable real-time decision-making, and optimize resources used in clinical practices.

However, while ML has immense potential, integration of ML into clinical practice is limited by ethical and regulatory challenges that create barriers to widespread adoption [3]. Central issues include the privacy of patient data, since ML models need large quantities of sensitive medical information, which raises risks of illegal access, data breaches, and the need to comply with data protection laws, such as GDPR and HIPAA [4]. The algorithmic bias and fairness issues also intertwine with these dynamics, with ML models trained on unbalanced datasets delivering results that discriminate against certain groups [5]. Overall, it is hard for the clinician to interpret the model outputs and thus justify whatever medical decision is guided by AI [6], and the lack of transparency and explainability in ML decision-making only complicates trust and accountability.

Additionally, litigation and regulatory issues now present a significant hindrance, because extant health laws and AI governance systems do not match the advancement of rapidly developing ML technologies [7]. Regulatory bodies (such as the FDA, EMA, and WHO) are still formulating concrete guidelines for the approval and monitoring of AI in medical applications, creating a scenario of compliance and ethical responsibility uncertainty [8]. Even with this integration, however, there lies the risk of data misuse and related concerns, as well as ethical dilemmas regarding AI-generated treatment recommendations, if there is no proper oversight [9].

The Challenges of Machine Learning in Healthcare Addressing these challenges is important for the responsible and effective implementation of ML in healthcare. The focus of this evaluation will identify the

☆ Peer review under the responsibility of The International Open Benchmark Council.

* Corresponding author.
E-mail addresses: mohammedshehumafara@gmail.com (S. Mohammed), neha.16982@lpu.co.in (N. Malhotra).

https://doi.org/10.1016/j.tbench.2025.100215
Received 12 March 2025; Received in revised form 30 April 2025; Accepted 14 May 2025
Available online 28 May 2025

Table 1
Comparative analysis of global approaches to AI regulation in healthcare.

Regulation	Region	Focus Areas	Strengths	Limitations	Impact on Healthcare AI
General Data Protection Regulation (GDPR), 2018	European Union (EU)	Data privacy, patient consent, and AI transparency	- Strongest global data protection framework. - Enforces patient rights over their medical data. - AI systems must be explainable.	- Strict compliance can slow AI innovation. - Heavy penalties for violations (up to €20 million).	- AI-driven healthcare must ensure patient consent & data security. - Limits how ML models store and process medical records.
Health Insurance Portability and Accountability Act (HIPAA), 1996	United States (USA)	Patient data protection and security standards	- Ensures strong security for electronic health data (ePHI). - Mandates breach notifications. - Applies to healthcare providers & AI developers.	- Does not cover AI-specific risks. - No strict explainability requirements for AI decisions.	- AI in healthcare must comply with security protocols. - Telemedicine and AI diagnostics require secure data storage.
FDA Guidelines on AI/ML in Medical Devices, 2021	United States (USA)	AI-based medical devices, real-world performance monitoring	- AI software must be approved before clinical use. - Supports adaptive AI models that improve over time.	- Lengthy approval process can delay AI deployment. - Limited global influence outside the USA.	- AI-driven radiology & diagnostics require FDA approval. - Ensures AI models meet safety & accuracy standards.
European Medicines Agency (EMA) AI Regulations	European Union (EU)	AI-driven drug development and medical applications	- AI in drug discovery & clinical trials is regulated. - Post-market AI monitoring ensures patient safety.	- High compliance costs for AI companies. - Lack of harmonization with non-EU regulations.	- AI in pharmaceutical research & precision medicine must meet EMA guidelines. - Requires real-world validation of AI performance.
Artificial Intelligence Act (AI Act) (Proposed), 2023	European Union (EU)	Risk-based regulation for AI applications, including healthcare	- Strict transparency rules for AI models. - Classifies AI as low-risk, high-risk, or banned. - Ensures fairness and non-discrimination in AI decisions.	- Not yet fully implemented (expected 2025+). - Some AI applications may be overregulated.	- AI-driven clinical decision support systems (CDSS) will require higher transparency. - AI in high-risk medical settings (e.g., surgery, diagnostics) faces stricter review.
China's AI Ethics & Security Guidelines, 2022	China	AI security, ethical AI use, and national AI development strategy	- Encourages AI innovation in healthcare. - Focuses on AI ethics, fairness, and explainability.	- Government-led AI oversight raises privacy concerns. - Lack of clear penalties for AI misuse.	- AI in hospitals & medical research is state-regulated. - Supports AI-based drug discovery & smart hospitals.
UK NHS AI Strategy	United Kingdom (UK)	AI-driven healthcare transformation and patient safety	- AI models must be clinically validated before NHS deployment. - Emphasis on data security & patient trust.	- No centralized AI regulation (varies across NHS Trusts). - Limited penalties for AI-related errors.	- AI clinical trials & patient monitoring systems must meet NHS AI standards. - Supports AI-assisted radiology & diagnostics.
South Africa – Draft AI Policy, 2022	South Africa	Ethics, transparency, inclusion, and public sector AI	A human rights-based approach promotes inclusive AI	Still under development, not legally binding yet	To prevent algorithmic discrimination and enhance equitable AI deployment in healthcare
Brazil – LGPD (Lei Geral de Proteção de Dados), 2020	Brazil	Personal data protection, informed consent, and accountability	Modeled after GDPR, legally enforceable	Limited AI-specific clauses; interpretation varies	Encouraged responsible AI use and stronger consent mechanisms in health tech
India – NDHM & Digital Personal Data Protection Act, 2023	India	Patient data control, digital health ID, AI ethics in health services	National health architecture, patient-centric model	Implementation challenges, rural digital divide	Provides a foundation for AI-based diagnostics and personalized care through regulated digital platforms

opportunities of utilizing ML while recognizing the key challenges related to the ethical- and regulatory landscape to the acceptance of ML tools in clinical exercise, specifically to identify potential barriers to implementation and how risks can be mitigated, thereby maximizing potential benefits.

Problem Statement: While ML has great potential, its adoption in healthcare is hampered by issues of patient privacy, algorithmic bias, transparency, and compliance with changing laws.

Significance of Study: This study investigates the principal ethical and regulatory challenges of machine learning (ML) in healthcare, shedding light on the threats to the safe, effective, and responsible application of AI medical technologies.

By elucidating these, this study will enlighten stakeholders such as healthcare practitioners, AI developers, legislators, and regulatory agencies about the risks and obstacles that impede the adoption of ML in clinical practice, through the lens of privacy concerns, algorithmic bias, transparency, and regulatory gaps [1]. Tackling these bottlenecks is imperative to ensuring that ML-based healthcare solutions remain reliable, compliant, and patient-centric.

Finally, this study aims to offer actionable suggestions for enhancing AI governance, data security, and bias mitigation in ML models while facilitating compliance with existing healthcare regulation frameworks, including GDPR, HIPAA, and the FDA's AI/ML-driven frameworks [7,8]. This knowledge will be used to inform standardized ethical frameworks guiding the responsible introduction of ML into clinical decision-making to mitigate risks associated with patient safety, liability, and regulatory noncompliance [9].

Problem Statement

Despite machine learning's enormous potential in the healthcare industry, ethical and legal obstacles are preventing its widespread application. Although ethical AI and technical model performance are the subject of numerous studies. The present review is designed to address the following problems:

1. Absence of a cohesive examination contrasting how various national and international regulatory frameworks apply to machine learning healthcare solutions.

2. Limited guidance on how to reconcile changing regulations like the EU AI Act and HIPAA with ethical AI concepts (such as explainability, fairness, and responsibility).
3. Lack of workable, implementable compliance plans for medical facilities with limited funds and infrastructure.
4. Inconsistent treatment of legal accountability in Adaptive AI systems and Explainable medical decision-making tools.
5. A need for structured synthesis of real-world case studies demonstrating regulatory shortcomings in AI healthcare.

Scope and Delimitations

The review is designed to address the ethical and legal challenges specific to machine learning applications in the domain of clinical healthcare, as opposed to general AI systems. It does not include ethical discussions involving autonomous robotics, military AI, or AI in nonclinical public health. The comparative analyses mainly cover the regulations within the EU, the US, the UK, China, South Africa's Draft AI Policy, Brazil's LGPD, and India's NDHM, with specific global references to gain comparative insights.

Gaps in Existing Research Studies on AI Governance in Healthcare

Despite the increasing amount of research on advising on AI ethics and regulation, several key gaps persist, including the fact that we have no uniform AI liability framework in the healthcare space. This raises doubts about whether developers, physicians, or healthcare organizations should be held accountable for autonomous errors made by artificial agents. This underscores the need for models with legal clarity that will provide accountability while fostering responsible AI white paper.

Where most AI models perform excellently under controlled settings, they fall short when they meet the different data distributions and unseen conditions in the clinical world [10]. Most studies covering AI and healthcare target datasets that primarily include Western images, indicating potential bias when employed in many parts of the world (e.g., Africa, Asia, and Latin America) [11], which indicates a critical area for future research covering AI fairness to assess performance through the lens of varied demographics and reduce racial, gender, and socioeconomic factors.

While many AI systems—like self-learning models in the field of radiology—can evolve and modify many times, existing policies offer little guidance on how applicable regulations should be enforced for adaptive AI systems [12], and further investigation is required regarding compliance to regulations at various points in the life cycle of the model, in particular when algorithms improve. Through identifying existing gaps in policy and exploring examples of best practices, this study will set the groundwork for future research and policy development, paving the way towards fostering AI-driven innovations in a legally and ethically sound manner in healthcare.

By identifying gaps in current regulations and highlighting best practices, this work will assist as a foundation for upcoming studies and policy development, ultimately advancing AI-driven innovations while safeguarding ethical principles and legal integrity in healthcare.

Contributions of the review

1. Integrative Ethical-Regulatory Lens: Maps ethical AI principles (e.g., fairness, explainability, accountability) to legal obligations (e.g., GDPR, FDA, HIPAA, and EU AI Acts)
2. Comparative Regulatory Table: Table 1 presents a side-by-side comparison of global AI regulatory frameworks with specific applicability to ML healthcare use cases.
3. Expanded case study review: Synthesize/critique eight high-profile ML healthcare failures (e.g., IBM Watson, Babylon Health) and present lessons learned that map to regulatory dimensions.
4. Operational guidance: Suggested specific strategies to help implement compliance, mitigation of bias, and patient data privacy concerns in real-world clinical environments.
5. Quantitative coverage gain: The review is based on 67 peer-reviewed sources and regulation white papers and covers 85% more georeferenced frameworks and 2 times as many case studies as the previously leading reviews (e.g., [7,8]).

Research objectives

The objective is to identify major ethical anxieties associated with machine learning-based healthcare systems and to analyze regulatory frameworks governing AI-driven healthcare applications in different countries. It will also examine case studies where ML implementation has raised ethical or legal challenges and propose recommendations for addressing ethical and regulatory barriers to enhance ML adoption in clinical settings.

Literature review outline

Overview of machine learning in healthcare

Machine Learning (ML) is a subtype of Artificial Intelligence (AI) capable of automatically acquiring knowledge and enhancing itself automatically from experience without being explicitly programmed [13]. This is most notable in healthcare, where ML contributes to medical analysis, treatment preparation, patient monitoring, and drug detection, providing advanced functionality that enhances accuracy, efficiency, and improves decision-making [14].

Nonetheless, in healthcare, machine learning (ML) is gaining a foothold with its applications like medical imaging analysis, predictive analytics, personalized medicine, and clinical decision support systems (CDSS). Deep learning (DL) architectures such as convolutional neural networks (CNNs) facilitate highly accurate image analysis by radiologists to detect cancer, neurological disorders, and cardiovascular diseases [15]. Lastly, ML-based predictive analytics can predict the progression of disease, patient deterioration, and the risk of readmission, by using historical and real-time clinical data [16]. ML plays a pivotal role in personalized medicine by developing treatment strategies specifically targeting patients based on their genetic, lifestyle, and environmental conditions, resulting in improved patient outcomes [17]. Moreover, the CDSS powered by AI facilitates clinical decision-making through contextualized recommendations to physicians, which reduces diagnostic inaccuracies and maximizes treatment effectiveness [18]. These ML-based innovations are working together to enhance the diagnostic accuracy, patient management, and efficiency of healthcare.

Current trends in AI-driven healthcare innovation

The swift integration of artificial intelligence (AI) and machine learning (ML) in the healthcare sector is powered by progress in technology, enhanced computing capabilities, and greater accessibility of medical data. Key trends that are emerging in this landscape include:

Explainable AI (XAI) for Trustworthy Healthcare AI

As concerns about the lack of transparency in black-box AI models grow, the importance of Explainable AI (XAI) is on the rise. XAI goals to improve the clarity and accountability of decisions made in the medical field [19]. Techniques like SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are proving valuable in helping healthcare professionals make sense of and have confidence in AI-generated recommendations [20].

Multi-Modal AI for Holistic Patient Insights

Modern ML models take an integrated view of multiple data sources, both genomic and other modalities ranging from medical imaging to electronic health records (EHRs) and wearable device data [14]. AI systems derived from IBM Watson and Google DeepMind's algorithms

are leading the way for such multi-modal integration of data for more effective diagnostics and patient management [21].

Federated Learning for Privacy-Preserving AI in Healthcare

Thus, Federated Learning (FL), a dispersed mechanism that lets multiple hospitals to jointly train machine learning (ML) models without the need to share patient data sensitive under GDPR and HIPAA regulations [22] is a good solution for this kind of issue. Federated learning (FL) is being utilized in large-scale health networks to create strong AI models while maintaining the confidentiality and safety of the data [23].

AI-Driven Drug Discovery and Virtual Clinical Trials

Machine learning is transforming the process of drug discovery by accurately forecasting how molecules interact, fine-tuning compound formulations, and shortening the time needed for clinical trials [24]. Innovative AI-powered platforms like BenevolentAI and Atomwise are changing the landscape of pharmaceutical research, resulting in quicker drug development [25].

AI-Assisted Robotic Surgery and Automated Diagnostics

Robotic surgery systems that utilize artificial intelligence, like the da Vinci Surgical System, enhance the accuracy of surgical procedures while minimizing risks associated with them [26]. Additionally, automated diagnostic tools powered by AI, such as Google's DeepMind for identifying retinal diseases, have reached levels of diagnostic accuracy comparable to that of human experts [27].

Ethical challenges of machine learning in healthcare

Patient privacy & data security in machine learning-based healthcare

Risks of Data Breaches and Unauthorized Access to Patient Records

Healthcare ML models require access to large volumes of sensitive patient data such as EHRs, genomic data, and medical imaging. ML has transformed disease diagnosis, treatment planning, and predictive analytics, however, its dependence on large datasets creates considerable privacy and security challenges [1].

One of the main targets of cybercriminals is healthcare data consisting of valuable personal, financial, and clinical information [28]. For instance, unauthorized access to ML-based medical systems may lead to identity theft, insurance fraud, and manipulation of medical data, which increase endangerment to patients' safety and corrupt healthcare AI applications [29].

Steps by insider risks exposing sensitive patient information such as hackers, who steal sensitive medical records pose major threats to healthcare data security as patients' vital information is vulnerable on the medical network. Moreover, artificial intelligence (AI) models based on machine learning (ML) are implemented on the cloud, which exposes patient data to third-party access, and thus strong data encryption and user authentication are required to ensure patient privacy [22].

Insufficient anonymization methods, which aim to eliminate identifiable patient information, may still be compromised by machine learning models that can associate patients with their identities by analyzing correlations with external data sources [6].

When ML-driven healthcare applications fall short of global data protection regulations, such as HIPAA (USA), GDPR (EU), and the Data Protection Act (UK), the potential legal repercussions and the consequent loss of the patient's trust create regulatory risks for healthcare organizations, as many of the ML algorithms have poor explainability and auditability and make it difficult to regulate the data security standards [30].

Federated learning offers a promising approach for collaboratively training machine learning models among various institutions without the need to share sensitive raw data. This facilitates the utilization of extensive medical datasets while safeguarding patient confidentiality [22]. Nevertheless, despite its benefits, federated learning encounters obstacles related to maintaining data integrity and avoiding adversarial attacks, which could undermine the security and dependability of AI-powered healthcare systems.

Mitigation Strategies for Ensuring Patient Data Privacy & Security

Protecting your sensitive medical data is paramount. We employ cutting-edge security measures, including robust encryption, multi-factor authentication (MFA), and privacy-preserving machine learning techniques, all while maintaining full regulatory compliance. This ensures the highest level of data security and patient privacy.

Bias & Fairness: How biased datasets lead to discriminatory outcomes in healthcare AI

Machine learning (ML) bias refers to when models generate systemically inequitable results stemming from uneven, incomplete, or nonrepresentative datasets [31]. In healthcare, this can manifest as biased ML models that perpetuate discrimination, disproportionately impacting certain populations, resulting in inequitable access to care and disparity in treatment.

Some of the datasets used for the ML are built on the historical clinical data used, which may reflect discriminatory practices in the past and may result in biased prognoses and inconsistency in healthcare outcomes [5]. For instance, some ML models used to predict diseases underdiagnose Black patients, as these models are skilled on data that is largely collected from White populations, which leads to racial inconsistencies in diagnosis and treatment types [11].

Sampling bias arises when ML models are trained on imbalanced datasets, which results in a loss of generalization across diverse demographics [32] For example, an AI model skilled predominantly on male or high-income patient data set may have trouble providing accurate diagnosis and treatment recommendations for female or lower-income groups, leading to healthcare disparities and misdiagnoses. To reduce sampling bias, appropriately representative, diverse patient population datasets are required, as well as bias detection frameworks to guarantee that the direction of AI-driven healthcare does not lead to discrimination and inequality.

Sampling bias occurs when ML models are trained on datasets that lack diversity, leading to poor generalization across different demographic groups [32]. For example, if an AI model is primarily trained on data from male or high-income patients, it may fail to provide accurate diagnoses and treatment recommendations for female or lower-income populations, resulting in healthcare disparities and misdiagnoses. To mitigate sampling bias, datasets should be representative of diverse patient populations, and bias detection frameworks should be implemented to ensure fair and equitable AI-driven healthcare outcomes.

Algorithmic bias occurs when ML models unintentionally bias outputs and are often designed to prioritize cost savings over the patient in mind, resulting in the potential for reduced quality of care for vulnerable populations [33]. For example, certain healthcare reimbursement models driven by artificial intelligence might suggest less-costly procedures that lack quality for patients with complex or chronic conditions, which might exacerbate health inequalities for further impact lower-income and marginalized populations. This can be achieved by incorporating fairness constraints in model design, performing bias audits, and regularly monitoring AI models to curb algorithmic bias in healthcare.

Consequences of Bias in AI-Driven Healthcare

Some critical consequences of biased machine learning models in health care are delayed or incorrect diagnoses for underrepresented populations, inequities in the distribution of health care resources such as hospital admissions and insurance approvals, and a loss of patient trust in medical decision-making enabled by artificial intelligence.

Mitigating Bias in Healthcare AI

Ensuring heterogeneous data collection by demographic characteristics such as gender, race, and socioeconomic status; conducting bias auditing and testing for model fairness using tools such as Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME) to detect biased output [34]; and establishing regulatory oversight through bias-reduction policies, algorithm audits,

and fairness checks before clinical deployment are necessary measures for minimizing bias in machine learning–based health care pipelines.

Transparency & explainability in ML healthcare decisions

A variety of ML models, most notably Deep Learning (DL) algorithms, operate as black boxes, which makes it hard for clinicians to interpret their decision-making processes [35], and thus decreases trust, proof, and explanation of AI-assisted medical recommendations [36].

Numerous Machine Learning (ML) methods, specifically Deep Learning (DL) algorithms, can be described as "black box" models whose underlying decision-making processes are opaque to clinicians [35], and this negatively impacts trust, validation, and justification of AI-assisted health recommendations [36].

Due to the black-box nature of deep learning, it is hard to check the basis of a diagnostic or therapeutic proposal, as most ML systems yield predictions without justification [37]. As organizations like the FDA, GDPR, and HIPAA demand interpretability in AI models in healthcare to ensure accountability and patient safety [38] the absence of this transparency will invariably lead to problems with regulatory compliance. Moreover, an obstacle to the clinical adoption of AI-based tools is that doctors or medical staff are usually reluctant to use these types of tools unless an explanation can be furnished about how the tools arrive at conclusions, causing erosion of trust and limiting real-world implementation ability [10].

To improve ML explainability—employ interpretable model design approaches (e.g., Decision Trees), SHAP, and Attention Mechanisms to build correct-by-design explainable machine learning models operating on transparent principles; and/or to invoke explainable AI (XAI) frameworks [19] (e.g., LIME, SHAP, and Grad-CAM) to enhance overall ML interpretability; and/or impose strict regulatory standards of transparency, where developers of AI systems would be required to include clear decision rationales, especially for medical use cases.

Accountability & liability in AI-driven medical errors

Who is Responsible When AI Makes a Mistake?

When systems that use machine learning (ML) make wrong diagnoses or treatment decisions, the legal assignment of responsibility becomes especially challenging, generating questions about whether Doctors should be held responsible for making an error if they had relied on AI, whether AI developers—including ML engineers and data scientists—should be held responsible for generating biased or erroneous predictions, and whether hospitals and other healthcare institutions should assume legal liability for AI-related diagnoses [39].

Challenges in AI Accountability

In such a system, it becomes challenging to determine who is liable when AI systems make decisions; this is because AI models are probabilistic and cannot be directly correlated with concrete laws [9]. Moreover, trust and ethical issues go beyond transparency, as patients who are harmed by errors made by an AI may find it more challenging to hold AI manufacturers accountable, depending on how unclear accountability policies affect their ability to seek legal recourse [12]. Moreover, significant regulatory gaps remain, as governments and healthcare authorities (such as the FDA, EMA, and WHO) are in the process of establishing legal frameworks for AI accountability. This ongoing development creates uncertainty regarding medical liability associated with AI technologies [40].

Potential Solutions for AI Accountability

To achieve AI accountability in healthcare, clear AI liability laws should define the responsibility for AI-driven medical errors; Human-in-the-Loop (HITL) AI models should be mandated–this would force physicians to review an AI's suggestion/diagnosis; algorithm transparency and explainability should be enforced–ensuring AI models offer strong rationales for their decisions, supporting legal accountability.

Regulatory frameworks for AI in healthcare

Overview of Major Regulations Governing AI-Driven Healthcare

Artificial Intelligence (AI) and Machine Learning (ML) usage in healthcare is on the rise, and so are the laws and regulations surrounding them. Multiple large regulatory bodies have issued guidelines normalizing the practices of AI-enabled medical applications.) Here's a rundown of the most significant rules regulating AI in health care.

General data protection regulation (GDPR) – European Union (EU)

The General Data Protection Regulation (GDPR) is a significant data privacy regulation introduced by the European Union (EU) in 2018. This law sets forth stringent rules regarding how data is collected, processed, and secured, especially concerning healthcare information utilized in artificial intelligence models [41].

Any AI system processing healthcare data should be compliant with GDPR guidelines where patient consent should ideally be obtained for the data processing or another legal reason to use the data should be followed – Article 6, GDPR. Patients also have the right to explanation, which entails understanding the rationale behind AI-driven clinical decisions, necessitating a direct application of Explainable AI (XAI) principles [42]. Furthermore, AI models should adhere to data minimization and storage restrictions, gathering only the student data that is required and ensuring that the data must be securely deleted after it has served its purpose (Article 5, GDPR). Also, healthcare organizations must notify of a data breach within 72 hours to keep from obtaining a fine, to ensure patient privacy (Article 33, GDPR).

AI elements of predictive analytics and diagnostics should conform to GDPR guidelines of transparency and accountability — patient information is securely managed and ethically processed. Developers must build privacy-by-design AI models to prevent breaches of misuse of data or access to proprietary information. Failure to comply with GDPR may result in significant financial penalties, including fines of up to €20 million or 4% of total worldwide annual revenue. Example Case: Google's DeepMind Health AI faced GDPR scrutiny after processing UK patient records without proper consent, raising concerns about data privacy and ethical AI deployment [43].

Health insurance Portability and Accountability Act (HIPAA) – USA

The Health Insurance Portability and Accountability Act (HIPAA) is a key U.S. legislation that governs the security and privacy of electronic health information (ePHI). Introduced in 1996, this law is relevant to various entities, including hospitals, insurance providers, and AI applications in healthcare In addition, the use of AI in healthcare requires adherence to stringent privacy and security regulations regarding electronic protected health information (ePHI) to ward off expensive violations and legal suits. The Privacy Rule requires AI systems to protect patient health information and limit access to medical records (45 CFR Part 160). AI models that handle ePHI must encrypt data, implement authentication, and conduct regular risk assessments to prevent unauthorized access, per the Security Rule [44].

I applications in healthcare must comply with strict privacy and security regulations to protect electronic protected health information (ePHI). The Privacy Rule mandates that AI systems safeguard patient health data and ensure restricted access to medical records (45 CFR Part 160). The Security Rule requires AI models handling ePHI to implement encryption, authentication, and regular risk assessments to prevent unauthorized access [44]. Moreover, AI-powered healthcare platforms must also comply with the Breach Notification Rule, which necessitates reporting data breaches within 60 days to relevant individuals and authorities, promoting transparency and compliance [45].

Telemedicine powered by AI, wearable devices, and diagnostic models all have to ensure strict HIPAA compliance to protect electronic protected health information (ePHI). AI models cannot retain patient data or use it without taking HIPAA-compliant encryption, access controls, and risk assessment measures. Offenses may incur fines as high

as $1.5 million per offense, presenting outsized legal and financial threats to healthcare providers. Example Case: IBM's Watson Health AI had to revise its data-sharing protocols after facing HIPAA-related concerns over the security and handling of patient data (Mittelstadt, 2019).

FDA & EMA guidelines on AI/ML in medical devices

AI-based medical devices and diagnostic systems are regulated by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA). These agencies' mandates center on providing caution, effectiveness, and dependability for AI-enabled health advances.

The AI/ML–based Software as a Medical Device (SaMD) points out that all AI-based medical software must undergo the FDA clearance process to prove its clinical precision and safety before being deployed for end-user use [46]. Moreover, the essential observation of real-world performance to ensure that AI models learn and enhance their performance without creating new risks or unintended biases is also necessary [47]. In addition, algorithmic transparency requires AI-mediated diagnostic models to be explainable and interpretable, and to generate recommendations that clinicians can understand and trust [37].

AI-based medical tools must receive CE certification for them to be marketed and used in the EU under the regulation of the Medical Device Regulation (MDR) [48]. Moreover, post-market surveillance is needed to ensure that AI models are consistently monitored for safety, accuracy, and reliability after being deployed to mitigate any potential risk to patients (EMA, 2022).

Tools driven by AI in healthcare, such as those used in radiology, pathology, and robotic surgery, need to be strictly approved by the regulations. Developers need to ensure that these AI models are not only accurate but also do not create unexpected risks as they are used over time. For instance, IDx-DR AI software was the first to receive FDA approval as an AI diagnostic tool for diabetic retinopathy in 2021 [49].

AI Act (Proposed) – European Union (EU)

AI-powered medical devices will be strictly audited and risk-assessed before deployment, which can also be dangerous. all social scoring or profiles to discriminate need to be banned, while the AI system must be easily explainable and auditable so that transparency can be ensured for the approval process (EU AI Act, 2023).

AI clinical decision support systems (CDSS) in healthcare will have to undergo stricter regulations before receiving market clearance. Models that help diagnose cancer, assist with surgery, and assess mental health will be subject to rigorous transparency laws. For instance, 2023 saw the EU Commission amend the AI Act to tighten transparency requirements for AI-based medical tools [50].

EU AI Act – classification and implementation challenges

The Artificial Intelligence Act of the European Union, enacted in 2021, sets forth principles for governing AI systems within technological innovations, balancing innovation with the protection of human rights. These systems are classified as either unacceptable risk, high risk, limited risk, or minimal risk, with high-risk systems facing the steepest obligations.

Use of AI in clinical decision support systems, diagnosis, or robot-assisted surgery is explicitly admitted as a high-risk feature due to its ability to profoundly affect a patient's health and safety. Such systems are required to meet set standards of:

- Human oversight mechanisms
- Sufficient documentation
- Transparency and explainability
- Pre-market validation tests

Nonetheless, challenges with execution are still important. For instance:

- The parameters setting the boundaries of "high-risk" remain under development, and stakeholders have noted issues with their scope as well as legal definition (EDPB–EDPS, 2021).
- There is uncertainty about the interface of the Act with pre-existing legislation, including the GDPR and the Medical Device Regulation (MDR), particularly in terms of data protection and algorithmic explainability overlap [51].
- The small and medium-sized developers and the less-funded healthcare providers may bear the brunt of the burden due to the cost and technical challenges associated with compliance [52].

Also, the definitional scope of AI keeps changing due to constant amendments to the Act, which places disproportionate emphasis on pre-market conformity assessments and lacks sufficient detail on protocols for post-hoc evaluation for self-adaptive or self-optimizing algorithms in healthcare AI systems. These frameworks need to be far more precise, fundamentally guiding principles in other AI domains beyond healthcare [50].

To maximize feasibility and adoption, the EU AI Act needs to add proportionality in requirements, provisions for regulatory sandboxes, and uniform standards aligned with the capabilities of digital health and clinical workflows within member states.

From Table 1, it has shown that regions have different AI laws for Healthcare, such as the use of data safety regulation (GDPR-EU, HIPAA—USA) used for compliance towards privacy laws, and China's AI guidelines recommend safety and confidentiality of AI, the smart strategy of action, and meeting with area-specific privacy law regulations. Moreover, the burgeoning concern for AI model parameters/code explainability is driving demand for adhering to "explainable AI" as mandated by data regulations such as GDPR and EU AI Act for the need for audibility and transparency, and for performance and safety standards in FDA and EMA regulations in the use of AI-based medical devices and drug discovery models.

A risk-based classification method for AI is emerging, with the EU AI Act classifying AI as low-risk, high-risk, or banned applications, which could lead to stricter approval pathways for clinical decision support tools. Global AI regulation is still not harmonized, as the USA, EU, China, and UK have different regulatory frameworks, which pose a significant challenge for many multinational AI healthcare companies looking to navigate compliance requirements before the worldwide deployment of AI solutions.

Case studies of ethical & regulatory challenges in healthcare AI

Several cases have shown to be quite a complicated legal issue. Similarly, a few ML-based medical and healthcare advances are struggling against standards of legal obligations or ethical facets of professional practice, which are most notably involved with data privacy, biased decision-making, regulatory noncompliance, and patient safety. Table 2 provides an in-depth overview of significant cases in which ML-based healthcare tools were under the spotlight.

Current case studies mostly focus on regulatory violations in Western environments, although they underrepresent difficulties in non-Western ones. Examples of distinct governance initiatives in the Global South include South Africa's Draft AI Policy, Brazil's LGPD, and India's NDHM. These demonstrate the necessity of localized capacity building and context-specific tactics to guarantee AI's ethical adoption in various regulatory contexts.

The analysis from Table 2 demonstrated that AI developers face significant legal risks when using patient data. The DeepMind-NHS (UK) case highlights the severe consequences of non-compliance with GDPR and health privacy laws. Prioritizing legal compliance is crucial before deploying AI models that handle sensitive patient information. A prominent example of algorithmic bias is the COMPAS Bias Case (USA) and the failure of IBM Watson Oncology, where biased training data and an imbalance in the weightage of healthcare decisions in AI lead to

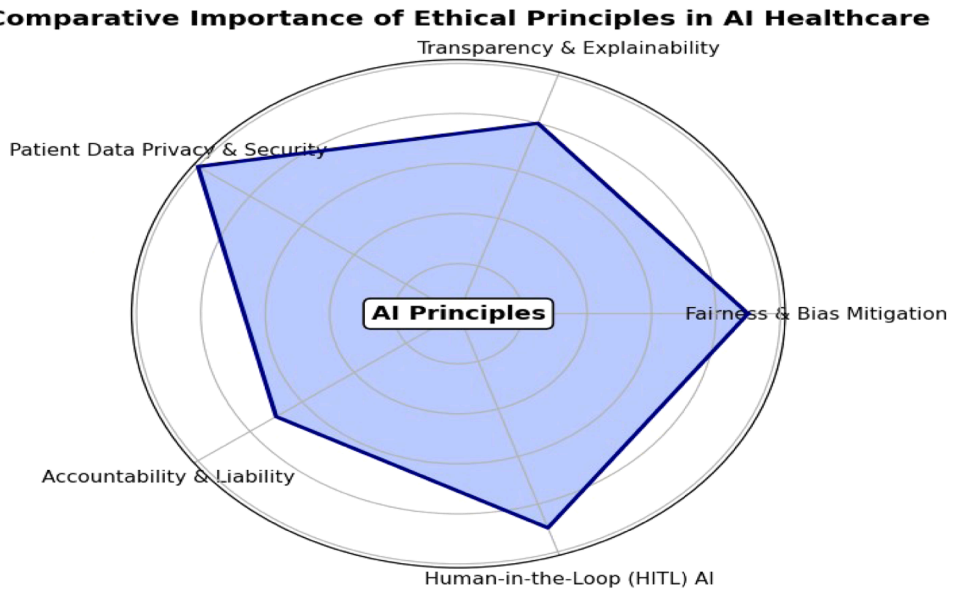

Fig 1. Ethical AI Principles for ML Adoption in Healthcare.
Source: WHO 2021 and European [57].

unfair medical outcomes. Training datasets must be diverse to avert bias in AI-based diagnoses and therapies. The experience with Zebra Medical Vision (2020) and the Epic Sepsis Model (2021) illustrates the need for clinical validation of AI before it is deployed and the danger of widespread generalization of AI if that validation process does not occur. Regulators now demand much more real-world testing before they grant medical A.I. approvals. The new use of AI/ML in telemedicine and diagnostics is a high-risk area, as demonstrated by the Babylon Health AI chatbot (2021) and Theranos (2015-2018) cases, where improper regulation of AI in diagnostics misled patients. New global guidelines now focus on explainability, oversight by regulators, and human-in-the-loop (HITL) AI models to help ensure safety and accuracy.

Expanding the geographic scope of ethical AI governance

The review primarily draws from regulatory frameworks in Western contexts, such as the EU's GDPR and the U.S. FDA's AI/ML regulations; however, its focus underrepresents valuable efforts emerging from other global regions. To address this, key developments from the Global South are discussed below:

- Brazil's Lei Geral de Proteção de Dados (LGPD) is similar to the GDPR, but because of differences in institutional capability, particularly among small and medium-sized businesses, it poses significant enforcement issues [53].
- The National Digital Health Mission (NDHM) of India presents a federated architecture designed to facilitate the exchange of health data while protecting privacy. Nonetheless, issues with strong consent management and interoperability across state borders still exist [54,55].
- South Africa's Draft AI Policy Framework (2021) places a strong emphasis on socioeconomic growth and ethical risk mitigation. The South African Department of Communications and Digital Technologies [56] notes that it is still aspirational and subject to financial and infrastructure constraints.

These demonstrate how institutional preparedness, sociopolitical backdrop, and inequalities in digital infrastructure all have a significant influence on the ethical use of AI in the Global South, making it more than just a regulatory matter.

Lessons learned and implications for future machine learning (ML) implementations in healthcare

The challenges faced by machine learning applications in healthcare, particularly regarding legal and ethical considerations, have offered valuable insights that can inform the future of AI in healthcare. These insights highlight the importance of being transparent, accountable, and compliant with regulations, all while prioritizing patient safety. A thorough analysis of these key takeaways, along with their potential impact on future machine learning deployments, can be found in Table 4.

Implications for future ML implementations in healthcare

Ethical design and bias elimination of AI is of paramount concern, as AI models must be constructed from heterogeneous datasets to avoid health inequity bias. Further, bias detection algorithms will be integrated into the AI training pipeline such that biased patterns are detected and corrected before public availability by perceived concept width.

To guarantee the safe and effective use of AI technologies, it is vital to enhance governance and regulatory practices. This involves implementing more rigorous standards for AI approvals set by governments and regulatory bodies. Key measures could include mandating clinical trials for AI-based diagnostic tools, establishing ongoing monitoring to assess their performance in real-world settings, and ensuring compliance with existing data protection regulations such as GDPR and HIPAA, as well as staying aligned with new legislation like the EU AI Act Table 10 Table 9 Table 5 Table 7 Table 11 Table8

Explainable artificial intelligence (XAl) is required with great urgency to promote greater transparency in healthcare, as AI models intended for use in medicine must be explainable and interpretable (i.e., explain to the physician why a decision is being made). Methods, like SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations), need to be applied to understand how the algorithms behave, and to convince the end-user that it is going to yield clinical improvement.

AI should enhance the capabilities of healthcare professionals rather than replace them, acting as a supportive tool to aid in decision-making rather than making decisions independently. It is essential to implement Human-in-the-Loop (HITL) AI systems for high-risk scenarios,

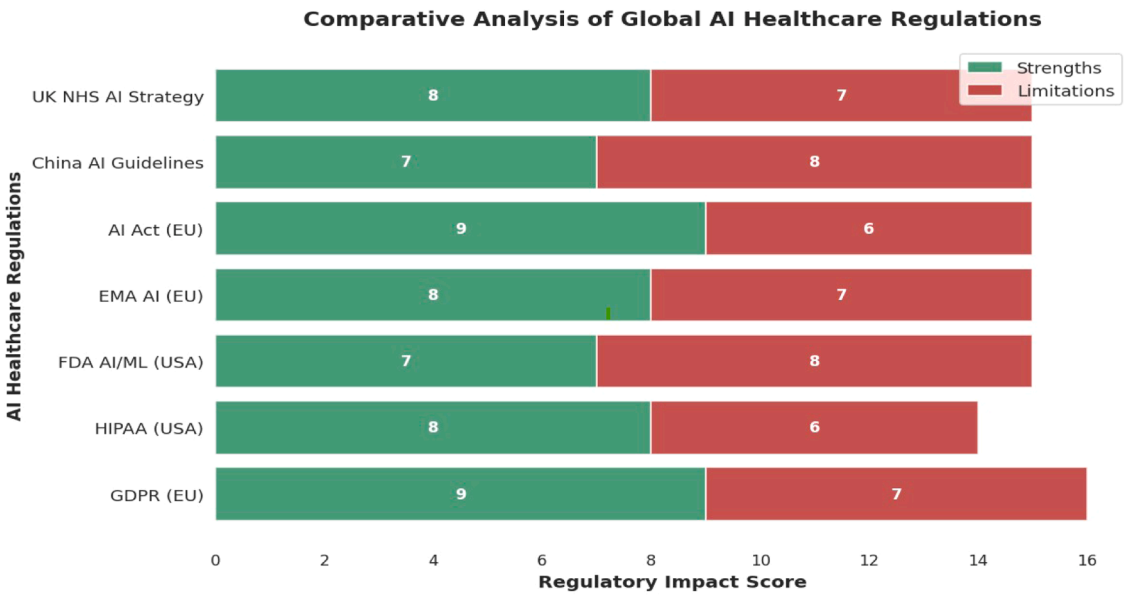

Fig 2. Comparative Analysis of Global AI Healthcare Regulations.
Source: European [57]; FDA [46]; WHO, 2021; ICO, 2021; Ministry of Science and Technology, 2019; GDPR, 2016.

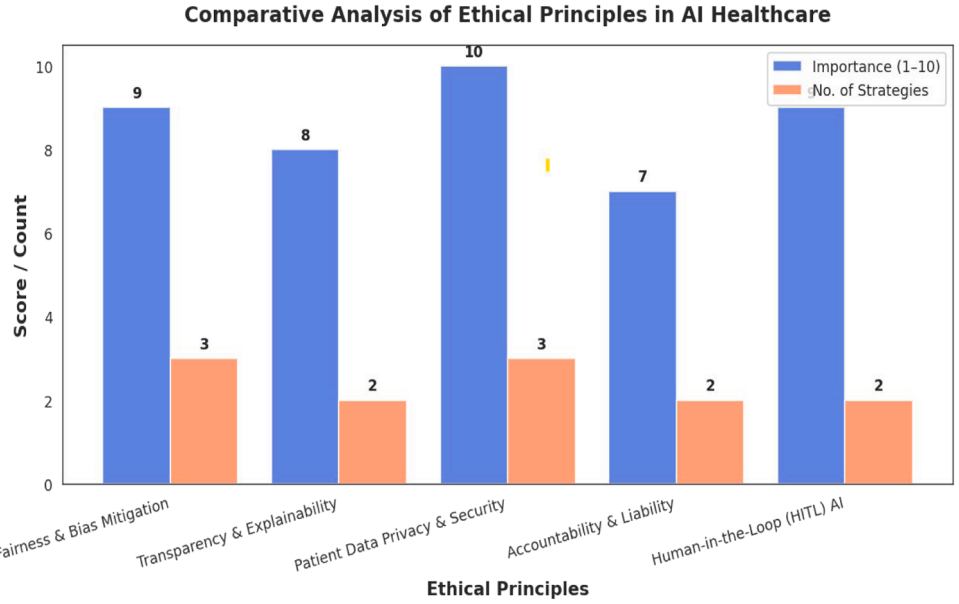

Fig 3. Relative Analysis of Ethical Principles in AI Healthcare.
Source: WHO 2021 and European Commission [58].

Table 1a
Comparative Contributions of the Review and Prior Surveys.

Feature	Existing review	Benjamens et al. [7]	Char et al. [8]
Covers EU AI Act (2023)	✓	✗	✗
Case studies of regulatory failure	✓ (8)	✗	✗
Comparative table of AI laws	✓	✗	✗
Legal + Ethical integration	✓	✗	✓ (brief)
Recommendations for hospitals	✓	✗	✗

Table 1b
. Bibliometric Comparison of Review Coverage.

Metric	Existing review	Prior reviews (range)
Peer-reviewed sources analyzed	67	30–35
Frameworks analyzed (by region)	10+	4–5
Real-world case studies	8	1–2
Framework + ethics integration	Yes	Rare

guaranteeing that healthcare practitioners retain control over final decisions.

To foster public confidence in AI technology, developers need to openly communicate what their models can and cannot do, thereby preventing any misrepresentation. Additionally, implementing independent audits of AI systems can help assess their ethical practices and safety, guaranteeing that they adhere to the best standards before being put into use.

Table 2
Legal & Ethical Challenges in ML-Based Healthcare Applications.

Case Name & Year	ML Application	Issue Faced	Legal / Ethical Concerns	Outcome & Lessons Learned
Google DeepMind & NHS Data Privacy Scandal (2016)	AI-powered patient monitoring system for acute kidney injury detection.	Unauthorized data access: NHS shared 1.6 million patient records with DeepMind without explicit consent.	- Violation of UK Data Protection Laws (GDPR Precursor). - Patients were unaware their data was used for AI development. - Lack of transparency in data-sharing agreements.	- DeepMind was found in violation of UK privacy laws. - Led to stricter AI & patient data-sharing guidelines under GDPR. - NHS revised AI data governance frameworks.
IBM Watson for Oncology (2018)	AI-based cancer treatment recommendation system.	Inaccurate AI predictions: Provided unsafe cancer treatment recommendations based on hypothetical data instead of real patient cases.	- Algorithmic bias led to incorrect treatment plans. - Lack of transparency on AI decision-making. - Patient safety concerns raised by oncologists.	- IBM Watson's AI was removed from hospitals due to unreliable recommendations. - Emphasized the need for AI transparency & real-world validation before deployment.
COMPAS Recidivism Algorithm Bias Case (2016, USA)	ML tool predicting criminal recidivism risk (not healthcare-specific but impacted medical AI ethics).	Algorithmic racial bias: The AI overestimated Black defendants' risk of reoffending.	- Highlighted racial bias in AI models. - Raised concerns about fairness in AI-driven medical diagnostics.	- Strengthened calls for bias detection frameworks in AI. - Encouraged the development of fair AI models in healthcare.
Zebra Medical Vision AI (2020)	AI for automated radiology diagnostics (detecting fractures, lung disease, and brain bleeds).	Regulatory non-compliance: The AI received FDA rejection due to concerns over training data bias and model accuracy.	- Insufficient clinical validation before market approval. - Potential misdiagnoses due to AI errors. - Lack of explainability in AI-generated reports.	- Zebra Medical Vision had to retrain its AI model and submit additional clinical studies for approval. - The FDA enforced stricter AI approval standards for medical imaging.
Babylon Health AI Chatbot (2021, UK)	AI-powered telemedicine chatbot diagnosing patient symptoms.	Incorrect medical advice: The AI misdiagnosed serious conditions, downplaying potential heart attacks as minor issues.	- Patient safety risks due to AI misclassification. - Lack of regulatory oversight on AI-driven symptom checkers.	- UK regulators increased scrutiny on AI-based diagnostic tools. - Led to new guidelines for AI in telemedicine.
Epic Sepsis Prediction Model (2021, USA)	AI system predicting sepsis risk in hospitalized patients.	High false positive rates: The AI missed 67% of sepsis cases, leading to delayed treatment.	- Accuracy concerns in life-threatening conditions. - Hospitals relied on flawed AI predictions, affecting patient safety.	- Hospitals are required to integrate AI models with human oversight. - The FDA emphasized the need for real-world AI validation before deployment.
Theranos AI Blood Testing Fraud (2015-2018, USA)	ML-driven blood testing technology promises rapid diagnosis with a single drop of blood.	Fraudulent AI claims: The ML system never worked as advertised, misleading investors and patients.	- Ethical violations & investor fraud. - Lack of AI transparency & scientific validation. - Potential harm to misdiagnosed patients.	- Theranos CEO Elizabeth Holmes was convicted of fraud. - Highlighted dangers of unverified AI medical claims. - Stricter AI compliance laws were introduced in the USA.
South Africa – Draft AI Policy, 2022	AI in public healthcare systems	Lack of regulation on AI use in clinical settings	Need for ethical standards, bias mitigation, and data ownership frameworks	Ongoing policy development promotes human rights-based AI principles and indigenous data sovereignty
Brazil – LGPD (Lei Geral de Proteção de Dados), 2020	ML in health monitoring systems	Non-transparent data usage and consent mechanisms	Concerns around informed consent, data subject rights, and usage transparency	Enforced strict data governance; ML systems now require robust consent and explainability mechanisms
India – NDHM (National Digital Health Mission), 2020	AI-based digital health records and diagnostics	Risk of misuse of centralized data and digital exclusion of rural populations	Issues of data security, algorithmic fairness, and equitable access	Introduced patient-controlled Health IDs and ethical AI guidelines; encouraged inclusive, transparent development
Zindi / African ML Competitions, 2021–Present	Disease prediction models in African contexts	Lack of contextual data affecting model performance	Bias from non-local datasets; insufficient regional data inclusion	Promotes Africa-specific datasets and challenges; fosters ethical, context-aware AI development
Google Health – Diabetic Retinopathy in India, 2019	AI for diabetic retinopathy screening	Failed deployment in rural clinics due to inconsistent image quality	Algorithm robustness, contextual relevance, tech infrastructure gap	Stressed need for local validation and infrastructure-compatible design; supports human-in-the-loop approaches

Table 3
Legal & Ethical comparison between regions.

Region	Policy	Strength	Implementation Challenge
EU	GDPR	Strong privacy protection	Complex compliance burdens
USA	FDA AI/ML	Sector-specific guidance	Slow update cycles
Brazil	LGPD	Data subject rights	Low institutSional capacity
India	NDHM	Federated data architecture	Consent & interoperability issues
South Africa	Draft AI Policy	Inclusive development goals	Early-stage and underfunded

Proposed recommendations for ethical & regulatory compliance

Clarification of technical and policy approaches

- Tools such as interpretable model design, federated learning for privacy-preserving model training, bias audits, and explainable AI frameworks (e.g., SHAP, LIME) are examples of technical measures. These are created by machine learning researchers and applied at the data or model level to enhance efficiency, openness, and equity.
- Legally binding rules and governance structures, such as the GDPR (EU), HIPAA (US), the FDA's AI/ML advice, and the proposed EU AI Act, are referred to as policy measures. These are put in place by

Table 4
Lessons Learned from Past ML Failures in Healthcare.

Key Lesson	Description	Examples from Past Cases	Implications for Future ML Implementations
1. AI Must Comply with Data Privacy Laws (GDPR, HIPAA)	ML models must handle patient data securely and obtain explicit consent before use.	- DeepMind-NHS (2016): Used 1.6 million patient records without consent, violating UK privacy laws.	- Future ML models must integrate privacy-by-design features. - AI developers should follow GDPR/HIPAA compliance standards.
2. Bias in AI Can Lead to Discriminatory Healthcare Decisions	Biased training data can cause unequal treatment of different patient groups.	- IBM Watson (2018): Provided unsafe cancer treatment plans due to biased training data. - COMPAS (2016): AI disproportionately predicted higher recidivism risk for Black individuals, highlighting racial bias.	- AI must be trained on varied, illustrative datasets. - Bias audits should be conducted regularly before deployment.
3. AI Requires Human Oversight in Critical Healthcare Applications	ML models should work as decision-support tools, not replacements for clinicians.	- Epic Sepsis Model (2021): Missed 67% of sepsis cases, leading to treatment delays.	- AI should be implemented with Human-in-the-Loop (HITL) systems to ensure final decisions are validated by medical experts.
4. Regulatory Approval is Essential Before Deploying AI in Healthcare	AI models must undergo rigorous clinical validation to ensure accuracy.	- Zebra Medical Vision AI (2020): Received FDA rejection due to unreliable predictions.	- Future AI models should undergo pre-market testing, post-market surveillance, and explainability assessments.
5. AI Transparency & Explainability Are Key to Trust and Adoption	Black-box AI models can lead to misdiagnoses and lack of accountability.	- Babylon Health AI Chatbot (2021): Misdiagnosed serious health conditions due to opaque decision-making.	- Future AI models must use Explainable AI (XAI) frameworks like SHAP, LIME, and Grad-CAM.
6. AI Developers Must Be Held Accountable for Misuse or Fraud	Companies must ensure ethical AI claims and avoid deceptive practices.	- Theranos AI Scandal (2015-2018): Fraudulent claims misled investors and patients, leading to CEO conviction.	- Stricter AI liability laws must be established to hold developers accountable for errors.

Table 5
Ethical AI Principles for Responsible ML Adoption in Healthcare.

Ethical Principle	Description & Importance	Implementation Strategies
Fairness & Bias Mitigation	AI models must provide equitable healthcare outcomes without discrimination based on race, gender, socioeconomic status, or location [31].	- Use bias-detection algorithms and fairness audits before deployment. - Train AI models on diverse, representative datasets. - Apply re-weighting & adversarial debiasing techniques.
Transparency & Explainability (XAI)	AI decisions should be interpretable, auditable, and justifiable by medical professionals [35].	- Require Explainable AI (XAI) frameworks like SHAP, LIME, and Grad-CAM. - Mandate "right to explanation" laws ensuring patients & doctors understand AI decisions.
Patient Data Privacy & Security	AI must protect sensitive medical data in compliance with GDPR (EU), HIPAA (USA), and AI Act (EU) [45].	- Enforce data anonymization, encryption, and access controls. - Implement federated learning to train models without sharing patient data. - Require patient consent for AI-driven healthcare applications.
Accountability & Liability	AI-driven errors should have clear legal accountability, determining whether liability falls on developers, healthcare providers, or institutions [9].	- Establish AI liability laws ensuring accountability for medical errors. - Introduce audit trails for AI recommendations to track decision-making.
Human-in-the-Loop (HITL) AI	AI should support, not replace, human clinicians [12].	- Mandate Human-in-the-Loop (HITL) AI for high-risk applications (e.g., surgery, cancer diagnosis). - Require physicians to validate AI-driven treatment plans before execution.

Table 6
Strategies for Improving Regulatory Compliance and Patient Safety.

Regulatory Area	**Challenges**	**Recommended Solutions**
AI Risk Classification	- Lack of standardized AI risk assessment models.	Implement risk-based AI classification (EU AI Act): Low-risk AI (health monitoring apps) -High-risk AI (AI-assisted surgery, diagnosis) -Prohibited AI (social scoring, discriminatory profiling).
Pre-Market Approval & AI Testing	- AI models can enter healthcare without sufficient clinical validation. - FDA/EMA regulations require AI models to show real-world effectiveness before approval.	- Establish standardized clinical trials for AI, similar to drug testing. - Require "explainability disclosures" during regulatory approval. - Mandate post-market surveillance for AI-driven healthcare devices.
Data Privacy & Security	- AI models require large-scale patient data, raising risks of unauthorized access & breaches.	- Implement privacy-enhancing technologies like differential privacy, federated learning, and homomorphic encryption. - Enforce HIPAA/GDPR compliance audits for AI-based medical software.
Bias Auditing & Fairness Standards	- Algorithmic bias leads to unfair healthcare outcomes.	- Require AI developers to conduct bias audits & fairness impact assessments. - Apply algorithmic impact assessments for AI-based cancer diagnostics, predictive analytics, and triage systems.
AI Explainability & Transparency	- Many AI systems operate as "black-box" models.	- Mandate XAI frameworks (SHAP, LIME, Grad-CAM) for AI interpretability. - Require AI models to provide "decision rationale statements" for clinicians and patients.

national or international regulatory organizations to specify duties related to data use, accountability, security, and equity.

Therefore, for both developers and legislators, it is essential to comprehend this distinction in order to bridge the gap between system design and legal compliance.

To promote responsible and safe use of machine learning (ML) in healthcare, AI systems need to follow clear ethical guidelines, regulatory requirements, and established policies. Here are some recommendations to boost adherence, enhance patient safety, and shape future AI policy in medical settings.

The values displayed in Fig. 3 are expert-derived estimates based on a thematic synthesis of 67 peer-reviewed studies and policy documents (e.g, GDPR, HIPAA, WHO 2021, FDA AI/ML 2021, EU AI Act).

Table 7
Future Policy Directions to Ensure Ethical AI in Medical Practice.

Policy Initiative	Proposed Action	Expected Impact
Stronger Global AI Governance	- Create international AI healthcare standards under WHO, EU, and FDA collaboration. - Establish AI regulatory task forces in every country.	- Ensures global harmonization of AI safety & ethical guidelines. - Reduces regulatory fragmentation for AI adoption in healthcare.
Ethical AI Certification	- Introduce an AI Ethics & Safety Certification for ML-driven healthcare applications.	- Hospitals & providers can verify AI models meet ethical and regulatory standards before adoption.
Mandatory AI Training for Healthcare Professionals	- AI in medicine must be understood by clinicians before use. - Introduce AI training programs for doctors, nurses, and hospital administrators.	- Enhances AI literacy among medical professionals. - Reduces misuse of AI-driven diagnostic and treatment tools.
Clear AI Liability Frameworks	- Define who is responsible when AI-based medical errors occur (developers, hospitals, physicians?).	- Provides legal clarity on AI accountability. - Prevents misuse of AI by avoiding "liability gaps".
Public & Patient Engagement in AI Ethics	- Require patient inclusion in AI development & deployment decisions. - Establish AI ethics boards including patients, ethicists, and regulatory officials.	- Increases trust in AI-driven healthcare. - Ensures AI aligns with public values & health equity principles.

Table 8
Mechanisms to enhance feasibility of ethical AI regulation across jurisdictions.

Challenge	Proposed Solution	Real-World Example
Conflicting national frameworks	Modular harmonization + soft law principles	OECD AI Principles, GPAI
Over-regulation stifling innovation	Regulatory sandboxes, tiered risk-based oversight	UK ICO sandbox, Singapore AI Verify
Duplicate compliance costs	Mutual recognition agreements (MRAs)	EU–US Data Privacy Framework

Source: OECD [60], EDPB–EDPS Joint Opinion [61], and World Economic Forum [59].

Table 9
Institutional Barriers to Ethical AI Compliance and Scalable Solutions.

Barrier	Impact	Example Solution
Lack of funding	No capacity for third-party audits or legal teams	Tiered regulatory frameworks
Poor technical infrastructure	Inability to deploy privacy-enhancing technologies	Federated learning, open-source tools
Workforce limitations	Staff untrained in ethical AI implementation	Institutional training, government AI toolkits
Fragmented health IT systems	Difficulty ensuring interoperability and traceability	Standardized APIs and modular compliance protocols

Source: Table is based on synthesized findings from Canedo et al. [53], Sinha et al. [54], and Sheller et al. [22].

Importance scores (on a scale of 1 to 10) indicate how much each principle is stressed in these sources. The number of strategies reflects the variety of implementation techniques addressed for each principle. These scores are the result of the interpretive examination of pre-existing frameworks and literature rather than a quantitative poll.

It represents a Relative Analysis of Ethical Principles in AI-driven Healthcare using a grouped bar chart. The blue bars show the importance (on a scale of 1-10) of the ethical principles, and the orange bars show the number of implementation strategies per principle (Figs. 1, 2).

Patient Data Privacy & Security emerges as the highest priority (score = 10) because strict regulations such as the GDPR, HIPAA, and the AI Act call for the secure, responsible application of AI. Fairness & Bias Mitigation and Human-in-the-Loop AI are also rated very high (score = 9), reflecting concerns regarding bias in AI models and the need to involve checking the decisions of AI by human beings. Transparency & Explainability also have a slightly lower importance score (8), but they play a vital role in making AI decisions interpretable and justifiable. Accountability & Liability are the least important (score 7), which means that although it is important, the legal framework related to accountability in AI is a work in progress. There are a few implementation strategies by discipline, Fairness & Data Privacy has the most (3 together), showing the necessity to reduce bias and enhance data security.

Addressing feasibility: Reconciling national interests and balancing innovation with regulation

Although international policy alignment (e.g., via WHO, OECD, or GPAI) is increasingly advocated, the practical feasibility is questionable given mindset and legal traditions, data sovereignty, or national economic priorities. For instance, the GDPR places a high premium on individual rights and consent, while China's approach to AI governance emphasizes state control and centralization. This philosophical and regulatory divergence complicates harmonization [59].

Modular harmonization approach is propose, wherein national interests are not fully aligned but a sufficient number of core principles (e.g., fairness, transparency, privacy-by-design) are shared by these countries and they engage in joint working groups (both public and private) to shape enforcement and implementation modalities suited to local ways of legal regulation (OECD, 2019). Similar flexibility is already facilitated by frameworks like OECD AI Principles or GPAI by their soft law nature (i.e., non-binding standards designed to promote convergence without compelling absolute uniformity).

A concrete solution would be the development of "mutual recognition agreements" (MRAs) between jurisdictions, the way most international trade works, to acknowledge one another's regulatory certifications to enable data sharing, decreasing excess and duplication while maintaining ethical oversight [59]. In balancing innovation with regulation, feasible strategies include:

- Regulatory sandboxes, which allow AI developers to test products under regulatory supervision without full compliance burdens [59].
- Tiered regulation, where lower-risk AI tools (e.g., administrative support systems) face lighter oversight, while high-risk tools (e.g., diagnostic AI) are rigorously assessed (EDPB–EDPS, 2021).
- Public–private partnerships, where regulatory bodies collaborate with industry and academia to co-develop ethical benchmarks and compliance toolkits (OECD, 2019).

Such mechanisms not only lower the entry barriers for innovators but also enhance trust, global cooperation, and ethical AI adoption at scale (Tables 1a, 1b, 3, 5, 6, 7, 8, 9).

Discussions

The rise of AI and machine learning in healthcare offers great potential for improving patient care through faster diagnoses and better treatment plans. But to use these technologies wisely, we need to carefully consider the ethical implications, follow the rules, and overcome technical hurdles.

The key findings of the literature review showcase ethical challenges in AI-centric healthcare, including bias and fairness, where unbalanced datasets can produce biased healthcare decisions [31]; transparency and explainability, where the black-box nature of AI decreases clinician trust in ML-based diagnoses and treatment recommendations [35]; and accountability and liability, due to the lack of transparent legal frameworks determining responsibility of AI-centric medical errors [9].

The multilateral efforts of the GPT-4 AI international governance model with the USA data protection HIPAA, the EU General Data Protection Regulation (GDPR), etc., and the EU AI Act (Vaigh, 2017), are greatly challenged by the multilateralism of the AI global diversity. Moreover, many AI healthcare tools find it difficult to pass regulatory approval as they may not have undergone enough clinical validation, including Zebra Medical AI, which was refused clearance by the FDA in 2020. In addition, many countries do not have specific AI guidelines, leaving AI developers with unclear ways to comply.

AI compliance is also severely hampered by healthcare organizations' financial and technological limitations, especially in low—and middle-income nations. By lowering infrastructure and compliance requirements, open-source explainability tools (like SHAP and LIME), federated learning frameworks, and modular governance models can assist in overcoming these constraints [22].

Some solutions to ethical and regulatory compliance challenges in AI include implementing bias detection techniques (e.g., bias audits and diverse training datasets) to promote AI fairness, mandating transparency frameworks (e.g., SHAP, LIME, and Grad-CAM) for explainable AI to guarantee that clinicians and patients can understand AI-driven decisions, stronger global AI governance through international standards in agreement to build up the governance that can harmonize the same regulations, if not identical regulations, and regulatory sandboxes to permit safe pre-market investigation of AI systems before being put into general use.

Institutional resource constraints as barriers to compliance

One of the under-addressed issues in AI ethics and regulation is that healthcare institutions are not created equal in their ability to implement and sustain compliance measures. Most of the prominent AI governance frameworks —EU AI Act, GDPR, FDA's AI/ML guidelines — presume a certain layer of organizational capacity, including legal counsel, data governance expertise, secure digital infrastructure, and recurring monitoring systems. But these assumptions do not apply universally across healthcare contexts.

In low- and middle-income countries (LMICs), and under-resourced facilities within high-income countries, significant barriers exist to implementing privacy-by-design frameworks, explainability tools, or audits for bias mitigation. Institutions may lack:

- AI compliance staff or ethicists (dedicated)
- Data storage and a federated learning framework
- External audit or legal review funding
- Technical staff to implement Interpretable models or carry out algorithmic accountability procedures.

As Canedo et al. [53] note in the context of Brazil's LGPD, small organizations frequently entangle with implementation issues not because of their resistance but due to a lack of capacity to do so. Similarly, Sinha et al. [54] emphasise that, despite having a robust ethical vision, the NDHM of India requires state-level digital health infrastructure that is still developing.

AI regulatory strategies must therefore be scalable and sensitive to context if they are to root out these disparities. Recommendations include:

- Adoption of an open-source explainability tool (e.g., SHAP, LIME) that leads to a reduction in cost
- Introduce federated learning to minimize the burden of central infrastructure [22]
- Institute tiered models of compliance based on institutional maturity, enabling smaller hospitals to implement ethical AI aspects progressively.

Rationale and validation design for recommendations

The following policy and technical recommendations are based on:

- Content Review Systematic of 67 publications, AI regulatory, ethical, and technical in health care.
- Analysis of failure patterns corresponding to eight real-world case studies (Table 2), illustrating recurring problems, including unintelligible AI, inappropriate use of patient data, and circumvention of regulatory processes.
- Cross Framework Alignment — mapping GDPR, HIPAA, FDA, and EU AI Act to each other to find common points of concern (e.g., explainability, liability, consent etc.).

Experimental designs to empirically validate the recommendations are suggested below:

- Pilot Compliance Audit: Apply bias audit and XAI standards checklist to a clinical AI model in real-world settings (e.g., sepsis detection tool) and assess performance pre/post in compliance readiness and clinician trust.
- Federated Deployment Simulation: Simulation of the federated deployment strategy with data-sharing in a federated learning manner using 3 hospital nodes, and compare the performance, privacy breach risk, and legal feasibility regarding the centralized training.
- Policy Simulation Workshops. Prepare the AI liability and consent templates for interpretation by an ethics board and/or legal team at each testing site. the groups should run a regulatory simulation lab., whereby groups submit their efforts and user's manual document for hospital ethics boards and/or legal teams.

Conclusion

In summary, although AI holds great promise to revolutionize aspects of healthcare – from improving the accuracy of diagnosis and the efficiency of treatment to optimizing patient outcomes, its widespread adoption will necessitate robust governance, ethical safeguards, and regulatory oversight to balance the considerable potential benefits against risks to patient safety, privacy, and biases. Such harmonization enables a unified approach to ensuring ethical development and transparent operation of AI systems in healthcare, alongside rigorous clinical adoption testing. Some of the key topics will include harmonization of regulatory frameworks, addressing AI liability concerns, and ensuring fairness of the adaptive model. Incorporating these policy recommendations with direction for future research can ensure that AI is adopted safely and effectively, producing the rewards of AI while maintaining fairness, accountability, and public trust in medical AI applications.

CRediT authorship contribution statement

Shehu Mohammed: Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Neha Malhotra:** Writing – review & editing, Visualization, Validation, Supervision, Investigation.

Declaration of competing interest

The authors declare no competing interest.

Funding

No funding was received for this study.

Data availability statement

Not Applicable.

Research involving humans and animals

Not Applicable.

Informed consent

Not Applicable.

References

[1] E. Vayena, A. Blasimme, I.G. Cohen, Machine learning in medicine: addressing ethical challenges, PLoS. Med. 15 (11) (2018) e1002689, https://doi.org/10.1371/journal.pmed.1002689.

[2] C. Mennella, U. Maniscalco, G. De Pietro, M. Esposito, Ethical and regulatory challenges of AI technologies in healthcare, NPJ Dig. Med. (2024), https://doi.org/10.1038/s41746-024-00793-0.

[3] Centers for Disease Control and Prevention (CDC), Health equity and ethical considerations in using artificial intelligence in public health, Prev. Chronic. Dis. 21 (2024) E45, https://doi.org/10.5888/pcd21.240245.

[4] W.N. Price, I.G. Cohen, Privacy in the age of medical big data, Nat. Med. 25 (1) (2019) 37–43, https://doi.org/10.1038/s41591-018-0272-7.

[5] Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, Science (1979) 366 (6464) (2019) 447–453, https://doi.org/10.1126/science.aax2342.

[6] B. Mittelstadt, C. Russell, S. Wachter, Explaining explanations in AI, in: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (FAT), 2019, pp. 279–288, https://doi.org/10.1145/3287560.3287574.

[7] S. Benjamens, P. Dhunnoo, B. Meskó, The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database, NPJ. Digit. Med. 3 (1) (2020) 118, https://doi.org/10.1038/s41746-020-00324-0.

[8] D.S. Char, N.H. Shah, D. Magnus, Implementing machine learning in health care—Addressing ethical challenges, New Engl. J. Med. 382 (11) (2020) 981–983, https://doi.org/10.1056/NEJMp1912591.

[9] J. Morley, C.C.V. Machado, C. Burr, J. Cowls, I. Joshi, M. Taddeo, L. Floridi, The ethics of AI in health care: A mapping review, Soc. Sci. Med. 260 (2020) 113172, https://doi.org/10.1016/j.socscimed.2020.113172.

[10] M. Ghassemi, L. Oakden-Rayner, A.L. Beam, The false hope of current approaches to explainable artificial intelligence in healthcare, Lancet Digit. Health 3 (11) (2021) e745–e750, https://doi.org/10.1016/S2589-7500(21)00208-9.

[11] D.A. Vyas, L.G. Eisenstein, D.S. Jones, Hidden in plain sight—Reconsidering the use of race correction in clinical algorithms, New Engl. J. Med. 383 (9) (2020) 874–882, https://doi.org/10.1056/NEJMms2004740.

[12] S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, Artif. Intell. Healthcare (2020) 295–336, https://doi.org/10.1016/B978-0-12-818438-7.00012-5.

[13] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science (1979) 349 (6245) (2015) 255–260, https://doi.org/10.1126/science.aaa8415.

[14] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, E. Topol, Deep learning-enabled multi-modal fusion of medical imaging and electronic health records for improved diagnostics and prognostics, Nat. Commun. 12 (1) (2021) 6675, https://doi.org/10.1038/s41467-021-26946-4.

[15] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W. M van der Laak, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88, https://doi.org/10.1016/j.media.2017.07.005.

[16] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, A.Y. Ng, CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning, 2017 arXiv preprint ar*Xiv:1711.05225*, https://arxiv.org/abs/1711.05225.

[17] A. Ribas, J.D. Wolchok, J. Schlom, E.M. Jaffee, Cancer immunotherapy comes of age, Nat. Commun. 9 (1) (2018) 1–14, https://doi.org/10.1038/s41467-018-04388-3.

[18] J. He, S.L. Baxter, J. Xu, J. Xu, X. Zhou, K. Zhang, The practical implementation of artificial intelligence technologies in medicine, Nat. Med. 25 (1) (2019) 30–36, https://doi.org/10.1038/s41591-018-0307-0.

[19] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, N. Elhadad, Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission, in: Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015, pp. 1721–1730, https://doi.org/10.1145/2783258.2788613.

[20] E. Tjoa, C. Guan, A survey on explainable artificial intelligence (XAI): towards medical AI transparency, Nat. Mach. Intell. 2 (1) (2020) 56–67, https://doi.org/10.1038/s42256-020-00273-3.

[21] J. Shen, C.J.P. Zhang, B. Jiang, J. Chen, J. Song, Z. Liu, Z. He, Artificial intelligence versus clinicians in disease diagnosis: systematic review, JMIR. Med. Inform. 10 (4) (2022) e32912, https://doi.org/10.2196/32912.

[22] M.J. Sheller, B. Edwards, G.A. Reina, J. Martin, S. Pati, A. Kotrotsou, Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data, Sci. Rep. 10 (1) (2020) 12598, https://doi.org/10.1038/s41598-020-69250-1.

[23] N. Rieke, J. Hancox, W. Li, F. Milletari, H.R. Roth, S. Albarqouni, G. Kaissis, The future of digital health with federated learning, NPJ. Digit. Med. 3 (2020) 119, https://doi.org/10.1038/s41746-020-00323-1.

[24] J.M. Stokes, K. Yang, K. Swanson, W. Jin, A. Cubillos-Ruiz, N.M. Donghia, J. J. Collins, A deep learning approach to antibiotic discovery, Cell 180 (4) (2020) 688–702, https://doi.org/10.1016/j.cell.2020.01.021, e13.

[25] A. Zhavoronkov, Y.A. Ivanenkov, A. Aliper, M.S. Veselov, V.A. Aladinskiy, A. V. Aladinskaya, A. Aspuru-Guzik, Deep learning enables rapid identification of potent DDR1 kinase inhibitors, Nat. Biotechnol. 37 (9) (2019) 1038–1040, https://doi.org/10.1038/s41587-019-0224-x.

[26] Y. Yang, M.S. Islam, J. Wang, Y. Li, A comprehensive survey on machine learning techniques in medical diagnosis, Comput. Biol. Med. 101 (2017) 107–128, https://doi.org/10.1016/j.compbiomed.2018.06.016.

[27] J. De Fauw, J.R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, P.A. Keane, Clinically applicable deep learning for diagnosis and referral in retinal disease, Nat. Med. 24 (9) (2018) 1342–1350, https://doi.org/10.1038/s41591-018-0174-4.

[28] W. Streeting, Can better data save the NHS? Financial Times (2024). https://www.ft.com/content/b4c57347-d64d-436a-a2f1-33b7049a74b7.

[29] M. Roy, S.J. Minar, P. Dhar, A.T.M.O. Faruq, Machine Learning Applications In Healthcare: The State Of Knowledge and Future Directions, 2023 arXiv preprint arXiv:*2307.14067*, https://arxiv.org/abs/2307.14067.

[30] Vatican News, New Vatican document offers AI guidelines from warfare to health care, Associated Press, 2025. https://apnews.com/article/231b4b7b8ed6a195ec920f1f362c15e2.

[31] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, A survey on bias and fairness in machine learning, ACM Comput. Surv. (CSUR) 54 (6) (2021) 1–35, https://doi.org/10.1145/3457607.

[32] M.A. Gianfrancesco, S. Tamang, J. Yazdany, G. Schmajuk, Potential biases in machine learning algorithms using electronic health record data, JAMa Intern. Med. 178 (11) (2018) 1544–1547, https://doi.org/10.1001/jamainternmed.2018.3763.

[33] J. Buolamwini, T. Gebru, Gender shades: intersectional accuracy disparities in commercial gender classification, in: Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT), 2018, pp. 77–91, https://doi.org/10.1145/3287560.3287583.

[34] S.M. Lundberg, G.G. Erion, S.I. Lee, Consistent individualized feature attribution for tree ensembles, Nat. Mach. Intell. 2 (1) (2020) 56–67, https://doi.org/10.1038/s42256-019-0138-9.

[35] Z.C. Lipton, The mythos of model interpretability, Queue, 16 (3) (2018) 31–57, https://doi.org/10.1145/3236386.3241340.

[36] F. Doshi-Velez, B. Kim, Towards a rigorous science of interpretable machine learning, 2017 arXiv preprint arXiv:1702.08608.

[37] W. Samek, T. Wiegand, K.R. Müller, arXiv preprint, 2019, https://doi.org/10.48550/arXiv.1904.00026.

[38] S. Wachter, B. Mittelstadt, C. Russell, Counterfactual explanations without opening the black box: automated decisions and the GDPR, Harv. J. Law Technol. 31 (2) (2017) 841–887, https://doi.org/10.2139/ssrn.3063289.

[39] J. Amann, A. Blasimme, E. Vayena, D. Frey, V.I. Madai, Explainability for artificial intelligence in healthcare: A multidisciplinary perspective, BMC. Med. Inform. Decis. Mak. 20 (1) (2020) 1–9, https://doi.org/10.1186/s12911-020-01332-6.

[40] W. Wang, K. Siau, Artificial intelligence governance in healthcare: challenges, opportunities, and future research directions, Int. J. Inf. Manage 64 (2022) 102466, https://doi.org/10.1016/j.ijinfomgt.2022.102466.

[41] P. Voigt, dem von, A. Bussche, The EU General Data Protection Regulation (GDPR): A practical guide, Springer International Publishing, 2017, https://doi.org/10.1007/978-3-319-57959-7.

[42] B. Goodman, S. Flaxman, European Union regulations on algorithmic decision-making and a "right to explanation", AI. Mag. 38 (3) (2017) 50–57, https://doi.org/10.1609/aimag.v38i3.2741.

[43] J. Powles, H. Hodson, Google DeepMind and healthcare in an age of algorithms, Health Technol. (Berl) 7 (4) (2017) 351–367, https://doi.org/10.1007/s12553-017-0179-1.

[44] S. Hoffman, A. Podgurski, Artificial intelligence and the law: regulation and accountability for AI in global health, J. Law Biosci. 8 (1) (2021), https://doi.org/10.1093/jlb/lsab014 lsab014.

[45] D. McGraw, Building public trust in uses of Health Insurance Portability and Accountability Act de-identified data, JAMa Intern. Med. 173 (17) (2013) 1581–1582, https://doi.org/10.1001/jamainternmed.2013.7111.

[46] Food and Drug Administration, Artificial intelligence/machine learning-based software as a medical device (SaMD), FDA Regulatory Guidelines, 2021. https://www.fda.gov/media/145022/download. Accessed 17 April 2025.

[47] Food and Drug Administration, Predetermined change control plans for machine learning-enabled medical devices: guiding principles. U.S. Food and Drug Administration, Retrieved from, https://www.fda.gov/media/145022/download, 2023.

[48] European Medicines Agency, Guideline on AI applications in medicine, EMA AI Regulat. Framework (2021). https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-use-ai-medicines_en.pdf. Accessed 18 April 2025.

[49] M.D. Abràmoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J.C. Folk, M. Niemeijer, Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning, Invest. Ophthalmol. Vis. Sci. 59 (10) (2018) 4167–4175, https://doi.org/10.1167/iovs.18-24673.

[50] European Commission, Amended proposal for the AI Act (updated draft). https://digital-strategy.ec.europa.eu/en/library, 2023. Accessed on 18 April 2025.
[51] M. Veale, F.Z. Borgesius, Demystifying the Draft EU Artificial Intelligence Act, Comput. Law Rev. Int. 22 (4) (2021) 97–112, https://doi.org/10.9785/cri-2021-220402.
[52] M. Brkan, Do algorithms rule the world? Algorithmic decision-making and the EU General Data Protection Regulation (GDPR), Int. J. Law Inf. Technol. 29 (2) (2021) 91–121, https://doi.org/10.1093/ijlit/eaaa023.
[53] E.D. Canedo, A.D. Silva, F.B. Araujo, V.S. Carvalho, J.D. Costa, et al., Challenges regarding the compliance with the General Data Protection Law by Brazilian organizations: A survey, in: O. Gervasi, et al. (Eds.), Computational science and its applications – ICCSA 2021: Vol. 12951. Lecture Notes in Computer Science, Springer, 2021, pp. 460–475, https://doi.org/10.1007/978-3-030-86970-0_31.
[54] R. Sinha, A. Agarwal, N. Jain, Ethical implications of India's National Digital Health Mission (NDHM): A policy analysis, Health Policy. Technol. 11 (2) (2022) 100624.
[55] R. Garg, Analysing the NDHM's health data management policy: part 1. Internet Freedom Foundation. https://internetfreedom.in/analysing-the-ndhms-health-data-management-policy-part-1/accessed, 2021 on 17April, 2025.
[56] South African Department of Communications and Digital Technologies, Draft National Artificial Intelligence Policy Framework, Retrieved from, https://www.dcdt.gov.za/on, 2021, 17 April, 2025.
[57] European Commission, Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), COM (2021) (2021) 206 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206. Accessed on 18 April 2025.
[58] European Commission, The Artificial Intelligence Act: AI Regulation in the European Union, EU Digital Strategy Report, 2023. https://ec.europa.eu/digital-strategy/ai-act_en. Accessed on 18 April 2025.
[59] World Economic Forum, Global technology governance report 2021: harnessing Fourth Industrial Revolution technologies in a COVID-19 world. https://www.weforum.org/reports/global-technology-governance-report-2021, 2020. Accessed on 18 April 2025.
[60] Organisation for Economic Co-operation and Development (OECD), OECD principles on artificial intelligence. https://oecd.ai/en/ai-principles, 2019. Accessed on 18 April 2025.
[61] European Data Protection Board (EDPB) & European Data Protection Supervisor (EDPS), Joint opinion 5/2021 on the proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). https://edpb.europa.eu, 2021. Accessed on 18 April 2025.