WEP Warwick Evans Publishing

International Journal of Public Health and Medical Research

ISSN: 3006-7707 (Print), ISSN: 3006-3752 (Online)
Volume 6, Number 8, Year 2026
DOI: https://doi.org/10.62051/ijphmr.v6n8.04



Research on Ethical Issues and Countermeasures in Artificial Intelligence-Assisted Healthcare

Xinyu Zhou

Shanghai University of Medical & Health Sciences, Shanghai, China
xiaoyuuu820Zxy@163.com

ABSTRACT

Since emerging in the 1950s, medical artificial intelligence (AI) has been deeply integrated into all segments of healthcare systems. It delivers prominent advantages in medical image diagnosis, pathological analysis, intelligent surgery, drug research and development and other fields, effectively lifting the efficiency, accuracy and overall quality of medical services, and becoming a core driving force for the digital and intelligent transformation of the healthcare industry. Nevertheless, the extensive application and technological iteration of AI have also brought new governance challenges. Focusing on ethical risks arising from AI-assisted medical services, this paper analyzes the relevant ethical dilemmas, root causes and corresponding solutions. When diagnostic and therapeutic errors occur in AI-aided medical activities, liabilities involve three stakeholders: technology developers, medical institutions and patients. The complicated liability structure renders the traditional single-subject liability model inapplicable, creating obstacles to liability identification. In addition, medical AI features complex technical frameworks and the "black box" algorithm effect: its internal processing logic cannot be fully interpreted by humans, leading to untraceable decision-making processes. Meanwhile, safety and ethical problems including medical data leakage, algorithmic bias and clinicians' over-reliance on technology have become increasingly prominent in China. To address the above ethical risks, this paper compares and draws on mature regulatory practices of the United States, the European Union and Japan for medical AI. It puts forward targeted suggestions to improve the liability system and regulatory mechanism of medical AI, clarify the rights and obligations of all stakeholders, standardize the clinical application of artificial intelligence, and strike a balance between technological innovation and medical safety. This research provides references for the standardized development of AI-assisted healthcare in China.

KEYWORDS

Medical artificial intelligence; Ethical issues; Countermeasure research

1. RESEARCH BACKGROUND AND SIGNIFICANCE

In the medical field, the integration of artificial intelligence and healthcare combines machine learning, deep learning and reinforcement learning algorithms, accelerating the in-depth integration of AI and medical services, optimizing patients' medical experience, and simultaneously improving the quality and efficiency of medical care. AI-assisted diagnosis, widely applied in medical imaging, pathological detection, robotic surgery and drug R&D in recent years, has become a mainstream clinical practice. Medical artificial intelligence has gradually evolved into a core engine for the transformation and upgrading of the healthcare industry, and ethical controversies triggered by this technology have drawn widespread attention from academia and industrial practitioners. From the perspective of legal and normative analysis, medical AI presents four major types of risks: difficulties in liability identification for medical accidents, medical data leakage, algorithmic bias, and clinicians'

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

over-reliance on technology. The application of medical AI involves multiple responsible parties, yet current legal norms fail to form an effective connection mechanism. It is critical to learn from foreign regulatory systems and construct a new legal supervision model suitable for China's AI-assisted medical industry, as well as establish a sound legal and regulatory framework. Such a system can not only safeguard the legitimate rights and interests of patients, but also standardize the behaviors of medical staff, technology R&D enterprises and other relevant entities. Guiding all stakeholders to use medical AI in compliance can promote steady and orderly industrial development, and ultimately achieve a sound balance between technological value and medical safety.

2. APPLICATION OF ARTIFICIAL INTELLIGENCE IN THE MEDICAL INDUSTRY

2.1. Medical Imaging

AI is applied in medical imaging mainly to enhance image quality and deliver auxiliary diagnosis [1]. It improves the clarity of medical images, automatically identifies lesions and pathological features, and enables precise recognition of cancer cell morphological characteristics during tumor screening, greatly boosting the efficiency of pathological image interpretation [2]. It assists physicians in completing disease diagnosis faster and more accurately, and provides objective evidence for tumor staging, typing and efficacy evaluation. In 2021, the U.S. Food and Drug Administration (FDA) approved Paige Prostate, a clinical AI tool for auxiliary detection, grading and quantitative assessment of prostate cancer for pathologists.

2.2. Intelligent Surgical Robots

AI technology has transformed the whole process of precise treatment and surgical operations, covering auxiliary surgical navigation, specialized precise surgery and perioperative anesthetic regulation [3]. Equipped with machine learning and sensor technologies, intelligent surgical robots learn human anatomical structures through deep learning, accurately locate surgical sites, and automatically adjust the position and posture of surgical instruments according to surgical objectives. Remote operation is supported to facilitate complex surgeries in secondary hospitals. The Da Vinci Surgical System, developed by Intuitive Surgical, optimizes the motion path of mechanical arms via machine learning algorithms with powerful tremor filtering capability, enabling submillimeter-level precision in suturing.

2.3. Drug Research and Development

Traditional drug R&D features a long cycle of 12 to 14 years from lead compound screening to clinical launch, accompanied by high costs and unpredictable efficacy and side effects, which requires four phases of clinical trials for full verification [4]. In contrast, AI technology can be adopted to predict drug targets, design protein structures, screen drug molecules, optimize synthetic routes, simulate pharmacokinetic properties and forecast adverse reactions, filtering out potent compounds with mild side effects. It drastically shortens pharmaceutical development cycles and reduces clinical risks. Launched in 2021, AlphaFold, an AI-based protein prediction tool, learns the correlation between amino acid sequences and known protein structures to improve the accuracy and efficiency of protein structure prediction, generating high-quality structural models and advancing protein research.

3. ANALYSIS OF LIABILITY-RELATED ISSUES

3.1. Decentralized Responsible Parties

When misdiagnosis caused by artificial intelligence leads to patient injury, disputes arise over the subject of legal liability. Medical accidents stemming from AI-assisted diagnosis form a three-party liability chain of "developer-medical institution-patient" [5]. Developers bear product liability for technical defects in algorithm design and system construction; medical institutions and their clinical staff are obligated to operate, supervise and control technical risks; patients' personal conditions and cooperation levels also affect liability determination. Such multi-stakeholder liability breaks the traditional single-subject liability model and creates barriers to liability definition.

Under China's medical device regulatory rules, all AI medical applications are classified as Software as a Medical Device (SaMD), with high-risk products categorized as Class III medical devices. Within the civil legal system, such software only serves as an object of conduct and cannot independently bear legal liability. In accordance with the Civil Code of the People's Republic of China, tort liabilities arising from medical AI injuries fall into two categories. The first is medical product liability under the strict liability rule, where developers and manufacturers shall be held accountable for product defects and insufficient warnings as stipulated in Article 1223. The second is medical treatment damage liability under the fault liability rule, where medical institutions and medical personnel shall be liable for diagnostic negligence in accordance with Article 1218. The overlap of the two liability forms results in confusion during liability identification. After medical disputes occur, all parties tend to evade obligations: R&D enterprises claim they only provide technical tools, clinicians attribute misdiagnosis to AI errors, and medical institutions argue their internal management procedures are compliant. The separation of liabilities among multiple parties makes it difficult for injured patients to protect their legal rights through legal channels.

3.2. The "Black Box" Dilemma

The "black box" feature means that the internal decision-making process of AI systems, especially deep learning large models is highly opaque. Users and even model designers can only observe input data and output results, without fully understanding the logical path leading to outcomes.

The adoption of AI essentially disconnects users from the logical reasoning behind diagnostic results [6]. Its sophisticated operation rules and data processing methods cannot be fully interpreted by humans. Such lack of interpretability hinders the identification of causal links between injury outcomes and accident triggers in judicial proceedings. Data anomalies, model updates and improper operation may all lead to medical damage, yet the exact cause of accidents is hard to pinpoint. The black box effect blocks traceability, and further intensifies conflicts regarding informed consent. Severe information asymmetry undermines patients' right to independently choose diagnostic and therapeutic plans.

3.3. Unbalanced Risk Allocation

Medical artificial intelligence undergoes rapid iterative updates, with a much faster technological upgrade rhythm than the 5–8 year conventional revision cycle of domestic laws and regulations. Legal provisions lag far behind industrial realities, creating a supply gap in institutional rules [7]. Meanwhile, risks generated by technological applications are disproportionately borne by patients. According to Article 5 of the Judicial Interpretation on the Application of Laws in Hearing Medical Damage Disputes issued by the Supreme People's Court of China, the scope of medical institutions' review obligations is clearly defined, yet no specialized and clear review standards are formulated for medical AI systems.

Regulatory gaps cause medical institutions to prioritize costs and benefits when introducing intelligent medical technologies, while risk prevention measures are often neglected. Current laws governing product quality and medical damage liability fail to adapt to the technical attributes of medical AI, making it impossible to build a scientific risk-sharing mechanism and worsening the imbalance of risk allocation.

4. PRIVACY ISSUES AND SECURITY CHALLENGES

4.1. Severe Risks of Data Leakage

Medical data falls under the category of highly sensitive personal privacy information, which is prone to leakage during online application and circulation. Data leakage incidents enable malicious actors to resell information for targeted marketing, fraud or extortion, and may trigger long-term and severe social consequences such as employment and insurance discrimination.

4.2. Deficient Informed Consent

Patients often lack knowledge of which data is collected and who can access their personal information. User profiling algorithms driven by big data and the black box feature further create obstacles to the implementation of informed consent, violating patients' right to informed consent [8].

4.3. Structural Bias

Training datasets of AI models often suffer from structural imbalance in terms of race, gender, age and other dimensions [9], which may generate discriminatory diagnostic and therapeutic recommendations for specific minority groups. This leads to cognitive bias and misdiagnosis for rare or cross-regional cases, further aggravating health inequity in practice.

4.4. Hidden Medical Safety Hazards

Excessive reliance on AI assistance may cause clinical physicians to develop intellectual inertia and degrade professional capabilities. Once AI systems malfunction, output errors or crash, the loss of effective manual intervention will directly threaten patients' lives.

5. REFERENCE OF INTERNATIONAL REGULATORY EXPERIENCE

5.1. Algorithm Transparency Supervision by the U.S. FDA

The FDA attaches great importance to data quality and privacy security. It requires datasets to be complete, traceable, relevant, repeatable and reproducible, and mandates consistent and accurate data management. All activities regarding the collection, utilization, storage and transmission of personal health information must strictly comply with relevant regulations to reduce legal risks and sustain trust among all stakeholders [10].

In addition, the FDA highlights the issue of algorithm transparency for artificial intelligence models. It advocates that R&D enterprises shall disclose the operational logic and data processing methods of algorithms, so that the judgments and decisions generated by medical AI can be traced. This measure helps break the "black box dilemma" and enables human supervisors to better understand the reasoning logic of AI.

5.2. Tiered Risk Supervision Model of the European Union

The European Union (EU) has implemented a tiered risk supervision model for medical artificial intelligence. This framework first defines the risk scope of all types of medical AI products. Regulators categorize products into three tiers (high, medium and low) based on the potential harm they may cause. Authorities conduct comprehensive assessments of all medical AI applications by integrating multiple indicators including clinical risks, technical operation risks and data processing risks, thus establishing a complete risk control system. If medical AI infringes patients' rights and triggers medical disputes during diagnosis and treatment, medical institutions and technical developers shall bear joint legal liabilities. To a certain extent, this mechanism strengthens mutual supervision and restriction between technology developers and clinical medical practice.

5.3. AI-Assisted Clinical Decision-Making Regulation in Japan

The presumption-of-fault rule for medical AI in Japan's judicial practice emphasizes the defense principle based on technical maturity. The rule stipulates that when the output of a medical AI system differs significantly from physicians' professional judgment, courts shall determine liability allocation according to the technical maturity and known risks of the AI system. Specifically, licensed physicians hold the ultimate power over clinical decisions, and AI is merely positioned as an auxiliary tool. This model establishes the dominant role of physicians and prevents their professional diagnostic capabilities from being undermined by AI.

Table 1. Comparison of Medical AI Supervision Systems

Comparison Dimension	U.S. FDA	European Union	Japan
Supervision Model	Guarantee the security of information transmission and processing; advocate algorithm transparency	Implement tiered risk supervision	Define AI as auxiliary tool; physicians hold final clinical decision-making power
Implementation Mechanism	Require disclosure of algorithm operational logic	Conduct risk assessment and impose corresponding supervision obligations based on risk tiers	Adjust liability allocation according to the technical maturity of AI systems
Penalty Intensity	Heavy civil compensation for product defects, product removal from market; huge fines for data leakage	Medical institutions and technical developers bear joint legal liabilities	Pursue liability against medical staff; monitor the reliability of algorithms
Implementation Effect	High algorithm transparency and sound protection of data privacy, yet risk of leakage of confidential information of AI R&D enterprises	Accurate risk classification and clear liability division for all parties	Clarify physicians' core status and reduce over-reliance on AI

6. RESEARCH ON COUNTERMEASURES

6.1. Improve Legal and Regulatory Systems

Regulatory authorities can build a dedicated full-lifecycle supervision system for medical AI products. On the one hand, unified constraints shall be imposed on the optimization, upgrading and retraining

of algorithm models [11]. Special inspection procedures for algorithm models should be formulated to ensure the reliability and representativeness of datasets, consolidate data foundations, eliminate the "black box effect" in medical practice, and strengthen the interpretability of AI algorithms [12]. On the other hand, regular inspections shall be arranged to identify AI operational failures and hidden risks, so that timely responses can be taken.

6.2. Popularize Ethical Accountability

When training AI models, comprehensive and sufficient datasets covering various races, age groups and rare diseases shall be provided to avoid biased data sources that may lead to misdiagnosis or missed diagnosis. In addition, patients shall be fully informed and give informed consent before data collection. The "data minimization" principle shall be promoted in data utilization to maximize the protection of personal privacy. Furthermore, courses on AI ethics and humanities should be integrated into medical education at universities [13], encouraging physicians to explore and summarize optimal human-machine collaborative models in clinical practice.

6.3. Clarify the Scope of Liability

All liable parties including designers, manufacturers, administrative departments and users shall fully recognize their respective obligations, and apply AI in compliance with basic ethical principles during development and clinical use. Meanwhile, a "joint liability system" should be introduced to construct a full-process legal liability traceability system.

6.4. Rational Application of Artificial Intelligence

The core function of AI remains auxiliary basic work such as information collection and preliminary data judgment, which frees physicians from repetitive tasks and improves clinical efficiency. Physicians shall formulate diagnosis and treatment plans, and professional human experts shall maintain control over key decision-making and risk management links [14]. AI is prohibited from independently issuing prescriptions or making final clinical decisions, so as to guard against medical safety risks caused by improper overuse of AI.

Table 2. Correspondence Table of Core Risks and Targeted Countermeasures

Core Risks	Corresponding Countermeasures
Decentralized liable subjects	All three parties shall clarify their obligations; introduce a joint liability system
"Black box" dilemma	Establish inspection procedures to guarantee the reliability and interpretability of algorithms
Leakage of private medical data	Implement the data minimization principle
Structural bias in datasets	Adopt comprehensive and sufficient training datasets
Hidden medical safety hazards	Professional physicians make all final clinical decisions

7. CONCLUSION

Medical artificial intelligence drives the digital transformation of the healthcare industry yet breaks the traditional medical liability system, giving rise to multiple challenges including algorithmic black boxes, fragmented liable parties, data leakage and hidden medical safety risks. This paper compares and draws lessons from international AI regulatory models. To boost the sound development of domestic AI-assisted medical services, China should enhance the reliability of AI algorithms, implement full-lifecycle supervision, establish a tripartite joint liability system, protect data privacy and security, and stick to the auxiliary positioning of AI. Only in this way can AI serve medical

development in a healthy, efficient and rational manner, better safeguard public health and benefit humanity.

REFERENCES

[1] Xu, L. L., Huang, D. Q., Liao, Y. L., & Dai, H. M. (2026). Application status of artificial intelligence in the medical field. Clinical Medical Engineering, 33(4), 503–508.

[2] Jan, Z., El Assadi, F., Abd-Alrazaq, A., & Jithesh, P. V. (2023). Artificial intelligence for the prediction and early diagnosis of pancreatic cancer: Scoping review. Journal of Medical Internet Research, 25, e44248.

[3] Tang, H. R., & Huang, H. (2026). Application of artificial intelligence technology in the medical field. Clinical Research and Practice, 11(4), 195–198.

[4] Chen, X. P., Cao, L., & Chen, Y. C. (2026). Innovative applications and challenges of artificial intelligence medical devices. China Plant Engineering, (8), 31–34.

[5] Chen, Z. L. (2026). Research on responsibility allocation of medical artificial intelligence [Master's thesis, Jiangxi University of Finance and Economics].

[6] Zeng, W. Y., Zheng, S. Y., Zhao, X. Q., Wang, Z., Zhang, Y. J., & Ma, Z. Y. (2026). Advances in generative biology and bio-inspired artificial intelligence. Chinese Bulletin of Life Sciences, 38(2), 248–267.

[7] Al Kuwaiti, A., Nazer, K., Al-Reedy, A., Al-Shehri, S., Al-Muhanna, A., Subbarayalu, A. V., Al Muhanna, D., & Al-Muhanna, F. A. (2023). A review of the role of artificial intelligence in healthcare. Journal of Personalized Medicine, 13(6), 951.

[8] Weidener, L., & Fischer, M. (2023). Teaching AI ethics in medical education: A scoping review of current literature and practices. Perspectives on Medical Education, 12(1), 399–410.

[9] Chatterjee, S., Fruhling, A., Kotiadis, K., & Gartner, D. (2024). Towards new frontiers of healthcare systems research using artificial intelligence and generative AI. Health Systems (Basingstoke, England), 13(4), 263–273.

[10] Cao, Y., Guo, Y. D., & Song, H. L. (2026). Supervision of artificial intelligence-aided drug R&D by the U.S. FDA and its enlightenment. Chinese Journal of New Drugs and Clinical Remedies, 45(4), 257–263.

[11] Wang, G., & Deng, C. D. (2026). Value alignment of generative artificial intelligence: Generation logic, practical obstacles and systematic governance. Journal of Hainan University (Humanities & Social Sciences), 1–10.

[12] Duan, W. X. (2020). Research on liability risks of medical artificial intelligence [Master's thesis, Tianjin Medical University].

[13] Sung, J. (2023). Artificial intelligence in medicine: Ethical, social and legal perspectives. Annals of the Academy of Medicine, Singapore, 52(12), 695–699.

[14] Wu, X. L., Wang, L. T., & Zeng, E. Q. (2026). Analysis of ethical dilemmas and governance paths of generative artificial intelligence in medical decision-making. Policy & Scientific Consult, (10), 17–21.