

Reconstructing the Duty of Care in Medical Artificial Intelligence Applications

Zidan Ye*

School of Law, Ningbo University, Ningbo 315211, Zhejiang, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Medical artificial intelligence is becoming deeply integrated into the entire clinical workflow, driving the rapid development of intelligent healthcare while giving rise to new risks of medical tort liability. In the current scenarios of medical AI, physicians' duty of reasonable care faces challenges from the weakening of independent clinical judgment; the duty of disclosure is challenged by notification and blurred boundaries; and patients' personal information is exposed to risks throughout the entire data lifecycle. To address these issues, officials should establish prudent diagnostic responsibility rules under human-machine collaboration, a tiered disclosure and shared decision-making mechanism, a full-lifecycle data protection framework, and a multi-stakeholder collaborative oversight and technical support system. This will enable a reasonable reconstruction and effective regulation of duty of care obligations among responsible parties for the application of medical artificial intelligence.

Keywords: Medical artificial intelligence; Duty of reasonable care; Duty of disclosure; Duty of confidentiality

Online publication: October 21, 2026

1. Analysis of current difficulties in the duty of care for smart healthcare

1.1. Reasonable care in diagnosis and treatment: weakening of independent physician judgment and the emerging duty of re-evaluation

Article 1221 of the Civil Code stipulates that medical personnel shall offer a duty of reasonable care commensurate with the then-prevailing level of medical practice in diagnostic and treatment activities. However, the widespread adoption of medical AI, driven by two key technical attributes, has disrupted the traditional rules of clinical care. Firstly, the intrinsic technical logic of algorithmic processing of complex data renders the internal mechanisms of AI decision-making a "black box," lacking transparency and explainability^[1]. Secondly, AI possesses far superior capabilities in autonomous learning and integrating massive datasets, enabling it to rapidly translate the latest medical research into clinical recommendations. Consequently, when AI-assisted diagnosis achieves a high degree of accuracy, two problems arise. On the one hand, medical practitioners are prone to develop technological over-reliance, relaxing or even abandoning

their own independent professional judgment—a phenomenon known as “alertness degradation.” On the other hand, when facing complex conditions, practitioners may only receive treatment options that exceed their prior knowledge base, yet without adequate supporting evidence, they cannot quickly make independent decisions or effectively validate the AI’s output. These issues severely undermine and destabilize the dominant role of physicians in clinical practice, exerting a certain negative effect on professional autonomy, narrowing discretionary authority, and reducing their participation and decision-making power in medical practice ^[2]. Given that medical procedures are inherently high-risk and invasive, affecting citizens’ rights to life, health, and bodily integrity, there is an urgent need, in AI-intensive medical activities, to establish a duty for physicians to re-evaluate AI-generated conclusions. Current domestic legislation is silent on this new duty of care, creating a legislative gap. Moreover, under the human-machine collaborative medical model, whether the standard of reasonable care for physicians should take into account the AI’s own medical capability level is also a crucial issue that demands practical resolution.

1.2. Duty of disclosure: Expansion of disclosure content under the transformation of care models

Article 1219 of the Civil Code establishes a tiered duty of disclosure for medical personnel. A general duty of disclosure applies to routine diagnostic and treatment activities, requiring medical personnel to explain the patient’s condition and standard medical measures. A special duty of disclosure applies to high-risk scenarios such as surgery, special examinations, and special treatments, requiring detailed disclosure of medical risks, alternative treatment options, and other relevant information, as well as obtaining the patient’s explicit consent. The core purpose of the duty of disclosure is to ensure that patients can autonomously choose appropriate medical actions based on full understanding, which is a crucial prerequisite for the legitimacy of medical conduct and a manifestation of respect for patients’ autonomy and bodily rights ^[3]. In contrast to purely instrumental medical devices, medical AI, driven by algorithms, possesses a high degree of autonomy and has deeply penetrated core medical processes such as diagnosis, treatment decision-making, and treatment planning. Because of this, the following difficulties in implementing physicians’ disclosure obligations occur. Firstly, when medical AI is introduced into diagnostic procedures, its inherent technical risks may arouse reasonable concerns among patients. It remains unclear in current law whether physicians are required to disclose the use of AI in care and to allow patients to decide on its involvement. Secondly, if a patient explicitly chooses to accept AI participation, should a uniform special-duty disclosure apply, or should disclosure be differentiated based on specific scenarios, such as the risk level of AI involvement and the stage of intervention? Thirdly, when the AI-generated treatment plan diverges or conflicts with the physician’s professional judgment, how should the content and scope of disclosure be defined? Fourthly, the performance of the disclosure duty must take into account the patient’s level of understanding and professional knowledge, so as to substantially safeguard informed consent. However, given the inherent black-box defect of medical AI, how can medical personnel provide adequate explanations that enable patients to fully comprehend diagnostic bases and medical risks?

1.3. Duty of confidentiality: Extension of the confidentiality boundary in the context of data informatization

Article 1226 of the Civil Code explicitly imposes on medical institutions and their personnel a statutory duty of confidentiality regarding patients’ privacy and personal information. Patients’ truthful disclosure of

personal health data is a prerequisite for medical practitioners to achieve accurate diagnosis and treatment. However, in the course of the deep advancement of nationwide health informatization, intelligent medical algorithmic automated decision-making technologies have been widely applied in core scenarios such as medical imaging, disease diagnosis, and health management. Such automated decision-making heavily relies on vast quantities of health data. The issues of algorithmic black boxes and insufficient data transparency continue to challenge traditional confidentiality rules, causing multiple dilemmas for the protection of personal health information ^[4]. Various medical databases aggregate large amounts of personal identity and sensitive disease information; combined with the high value, vulnerability, and covert nature of privacy rights, this further amplifies the risk of privacy leakage in intelligent medical settings. In particular, violations of confidentiality obligations and patients' privacy rights by medical personnel and institutions are additionally highlighted among the three key stages of data collection, use, and sharing. Firstly, in practice, the cases of unauthorized collection of sensitive medical information without patient-specific authorization and batch collection of redundant data beyond the necessary scope of treatment are universal. Some institutions deliberately conceal the actual purposes of data usage—such as algorithm training or commercial cooperation—rendering patient informed consent a mere formality and undermining substantive protection of patients' autonomous choice ^[5]. Secondly, at the data use stage, de-identification is often superficial, failing to adequately anonymize information, making it easy to reverse-engineer original personal data through intelligent algorithms. Meanwhile, generative medical AI has memory retention characteristics, which may inadvertently retain or leak patient information during iterative model training. Some entities also misuse lawfully collected medical information for non-treatment-related purposes, gravely violating medical confidentiality principles and clinical ethics ^[6]. Thirdly, at the data sharing and circulation stage, regulatory mechanisms are evidently deficient. Some medical institutions transmit raw core medical data to AI technology service providers without conducting compliance risk audits or qualification verifications. When entrusting third parties with data processing, they fail to perform supervisory duties or establish long-term binding mechanisms, leading to uncontrolled and wide dissemination of sensitive patient health information, resulting in irreversible privacy harms ^[7]. Therefore, it is imperative, against the backdrop of medical informatization and AI development, to redefine the scope of confidentiality obligations and reasonably extend the applicable boundaries for both medical institutions and their personnel.

2. A New structural framework for the duty of care in intelligent healthcare

2.1. Establishing prudent judgment and responsibility boundaries for human-machine collaboration

To ensure the dominant role of medical staff in clinical activity, efforts can be made in both technology and regulation. On the technical side, Research and Development staff of technology enterprises can optimize the human-computer interaction design of diagnostic AI by requiring, at critical decision points, that physicians manually input their reasoning or diagnostic justifications before the AI proceeds to the next step, thereby blocking over-reliance on technology. On the regulatory side, healthcare institutions may refer to the Application Scenario Reference Guide to set minimum verification requirements for medical staff in different scenarios and stages. Whether a physician has fulfilled the duty of reasonable care should not be judged only by treatment outcomes, but also by the average developmental level of medical AI should be incorporated into the “then-prevailing medical level” consideration, and assessments should be differentiated by specific

scenarios to reconcile the promotion of technological innovation with the burden on medical personnel. On the one hand, if the AI-generated treatment plan falls within the physician's ordinary knowledge scope and contains no obvious errors discernible by general medical knowledge, the reasonable adoption of that plan should be deemed as satisfying the duty of care; conversely, if a physician adopts a plan with manifest errors that causes patient harm, the duty is breached. On the other hand, if the AI provides an innovative treatment plan beyond the physician's ordinary knowledge scope, the physician should not only consult the literature for preliminary verification but also organize multidisciplinary clinical consultations, conduct an in-depth "risk-benefit" assessment, and engage in inter-institutional exchanges if it's necessary ^[8]. If the physician adopts the plan after thorough multidisciplinary discussion and prudent evaluation, the duty of care is satisfied. However, if the physician blindly rejects or adopts the plan out of excessive conservatism or recklessness, this may be deemed a failure to meet the duty of care.

2.2. Establishing a tiered disclosure and physician-patient shared decision-making communication mechanism

Firstly, given the unique technical attributes of medical AI, the human-machine collaborative treatment model should be incorporated into the statutory disclosure obligation, introducing a "shared decision-making" mechanism whereby specific treatment plans are determined jointly by physicians and patients, taking into account not only therapeutic efficacy but also patients' personal preferences, economic costs, and other practical factors, thereby endowing patients with the autonomous right to choose AI involvement in their care. Secondly, to address the debate on the application of AI-related disclosure duties, a risk-level-based tiered disclosure rule should be established: for low-risk scenarios such as AI-assisted basic screening and routine condition sorting, general disclosure via outpatient notices or written information suffices. For medium-risk scenarios like AI-assisted imaging diagnosis and routine medication recommendations, standardized written disclosure is required, explaining AI technical limitations and treatment risks. For high-risk core diagnostic scenarios such as AI involvement in surgical planning and critical care decision-making, the special duty of disclosure applies, requiring full disclosure of risks and alternatives and obtaining explicit patient consent. Thirdly, if the patient selects AI-involved care, the physician should inform the patient of the AI's diagnostic conclusions. If the physician's and AI's diagnoses and treatment plans are consistent, the physician should explain them to facilitate patient cooperation. If they are inconsistent, the physician should disclose the AI's alternative plan, including its rationale and the pros and cons ^[9]. Fourthly, to tackle the problem of black-box opacity leading to unclear disclosure methods, a collaborative explanation mechanism between healthcare institutions and technology enterprises should be established. AI system developers should produce a user-friendly explanatory guide—through text or visualizations—on the system's decision-making and operational logic. Healthcare institutions should then set up specialized departments to study these materials and explain the AI's decision logic and risk points in plain language when patients have questions to enhance patient trust.

2.3. Building a full-lifecycle patient information protection system

In the intelligent healthcare context, the confidentiality obligations of medical personnel and institutions should shift from passive prevention of leakage to active prevention. This means not only refraining from unauthorized disclosure but also proactively adopting technical and organizational measures to prevent third-party unauthorized access and use of data, while balancing patient privacy protection and the developmental

needs of medical AI within the existing legal framework^[10]. Firstly, at the data collection stage, medical health information constitutes sensitive personal data and requires separate explicit consent from patients, with blanket authorization prohibited. The principle of minimal necessity must be strictly observed. For non-treatment purposes—such as algorithm training or commercial use—separate specific informed consent must be obtained, with clear disclosure of data flow, purposes, and potential risks. Secondly, at the data use stage, to counter the risks of superficial de-identification being reversible and generative AI memory retaining leaked information, a dual de-identification system combining legal standards and quantitative models should be implemented. Based on existing data classification and grading rules, high-sensitivity information should be restricted from non-treatment use, with reference to the Technical Specifications for Anonymization of Health and Medical Data (Trial) for deep de-identification. Additionally, a usage approval system should be instituted, prohibiting medical personnel from arbitrarily inputting raw patient data into public generative AI platforms. Thirdly, at the data sharing stage, a closed-loop mechanism of pre-sharing risk assessment, full-process supervision and monitoring, and long-term traceability should be established. Before transmitting data to third parties, qualification verification should be legally conducted, confidentiality agreements signed with clear usage boundaries and prohibitions on secondary sharing, and institutions should maintain chain-of-custody logs enabling traceability of data destinations and accountability for violations.

3. Conclusion

Medical artificial intelligence has been deeply integrated into the entire clinical workflow. While driving the development of intelligent healthcare, it also triggers new risks of medical tort liability by impairing physicians' independent clinical judgment, blurring the boundaries of informed notification, and exposing patients' personal information to leakage risks throughout the entire data life cycle. The current rules of duty of care, which are centered on human physicians and traditional clinical scenarios, are no longer adequate for human-AI collaborative healthcare. For the reasonable reconstruction of the duty of care, it is necessary to establish prudent clinical practice liability rules under human-AI collaboration, a hierarchical notification and doctor-patient shared decision-making mechanism, and a whole-life-cycle protection system for patients' personal information. Only through such reconstruction can physicians retain professional judgment while using artificial intelligence in a prudent and responsible manner, enable patients to truly understand and consent to the participation of artificial intelligence in diagnosis and treatment, and allow medical institutions to proactively prevent data abuse and leakage. This reconstruction not only facilitates the responsible innovation of intelligent healthcare, but also protects patients' lives, health, personal dignity and personal information, and guides the development of medical artificial intelligence toward a credible, controllable and people-centered direction.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Feng YJ, Shen HY, 2025, Research on the Obligation of Interpretability in Judicial AI under the Background

- of XAI—A Theoretical Perspective Based on Judicial Sincerity. *Journal of Beihang University (Social Science Edition)*, 38(4): 29–41. <http://doi.org/10.13766/j.bhsk.1008-2204.2024.2400>
- [2] Kang YQ, He XS, 2025, Challenges and Responses to Doctors' Professional Autonomy from Medical Artificial Intelligence. *Medicine and Philosophy*, 46(9): 18–23. <http://doi.org/10.12014/j.issn.1002-0772.2025.09.04>
- [3] Zhu SY, 2022, Research on the Duty of Care in Applying Artificial Intelligence for Diagnosis and Treatment, thesis, Southeast University. <http://doi.org/10.27014/d.cnki.gdnau.2022.000657>
- [4] Xu ZY, Yue YL, 2025, Research on Personal Health Information Protection under Automated Decision-Making by Intelligent Medical Algorithms. *Medicine and Philosophy*, 46(1): 30–34. <http://doi.org/10.12014/j.issn.1002-0772.2025.01.06>
- [5] Chen QY, 2026, Medical AI: Data Compliance and Privacy Protection Are the Primary Legal Challenges. *People's Posts and Telecommunications News*, 1.
- [6] Mo LF, Li Z, Gan HL, et al., 2025, Patient Privacy and Data Security in Medical AI from a Global Perspective: Focus and Strategies. *Journal of Naval Medical University*, 46(8): 989–999. <http://doi.org/10.16781/j.CN31-2187/R.20250363>
- [7] Wei Y, Yao ZR, Wang YH, 2025, Challenges and Countermeasures for Personal Health and Medical Data Sharing from the Perspective of Collaborative Governance. *Medicine and Society*, 38(8): 50–56. <http://doi.org/10.13723/j.yxysh.2025.08.007>
- [8] Zheng ZF, 2023, Medical Liability for Medical Damage in Diagnosis and Treatment of AI. *China Law Science*, 2023(1): 203–221. <http://doi.org/10.14111/j.cnki.zgfx.2023.01.011>
- [9] Zhu SY, 2022, Research on the Duty of Care in Applying Artificial Intelligence for Diagnosis and Treatment, thesis, Southeast University. <http://doi.org/10.27014/d.cnki.gdnau.2022.000657>
- [10] Ma XR, Lü XM, Deng R, 2026, Challenges and Countermeasures of Privacy Protection in Medical AI Applications. *Medicine and Philosophy*, 47(12): 36–40. <http://doi.org/10.12014/j.issn.1002-0772.2026.12.07>

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.