Proceedings of the 4th International Conference on International Law and Legal Policy
DOI: 10.54254/2753-7048/2025.LD31098

Legal Protection of Patients' Right to Know in Artificial Intelligence Diagnosis and Treatment

Ruixuan Liang

College of Liberal Arts and Law, Beijing University of Chemical Technology, Beijing,China
1405888071@qq.com

Abstract. With the rapid development of artificial intelligence technology in the field of medical diagnosis and treatment, the legal protection of patients' right to know is facing new opportunities and challenges. However, under the traditional medical model, patients' understanding of the right to know is limited to the clinician's obligation to inform, and it is difficult to adapt to the new diagnosis and treatment model with highly opaque algorithm decision-making and diversified responsible subjects. In artificial intelligence diagnosis and treatment, patients' right to know should be extended to knowledge of data sources, algorithm logic, differences in man-machine judgment and other dimensions. However, in practice, they face dilemmas such as algorithm black boxes, ownership of rights and responsibilities, and professional barriers. In this regard, it is necessary to establish a hierarchical disclosure system for diagnosis and treatment algorithm information, clarify the rules for dividing the notification obligations of developers, institutions and doctors, and improve the right relief mechanism of patients' right to know.

Keywords: artificial intelligence diagnosis and treatment, Patient's right to know, Algorithm transparency, Obligation to inform

1. Introduction

Patients' right to informed consent lies in ensuring that patients can independently make decisions to accept or reject a certain medical measure after obtaining sufficient and understandable medical information. Under the traditional diagnosis and treatment mode, this right is mainly realized by clinicians, but under the new mode of artificial intelligence intervention, the logic of information generation and decision-making has changed. Diagnosis and treatment decisions are no longer only derived from doctors' professional knowledge and clinical experience, but more deeply dependent on complex algorithm models. This change raises a series of urgent questions to be answered. What do patients need to know? Who will inform? How to tell? The neglect of these problems may lead to the overhead of patients' independent decision-making rights, and even fall into the dilemma of safeguarding rights when diagnosis and treatment are damaged. Therefore, an in-depth discussion of the core connotation, dilemma and protection path of patients' right to know in artificial intelligence diagnosis and treatment is an important guarantee to ensure that this cutting-edge technology can truly be people-oriented and benefit society.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7048/2025.LD31098

2. The core demands of patients' right to know in artificial intelligence diagnosis and treatment

In the context of artificial intelligence technology intervening in diagnosis and treatment, the connotation of patients' right to know needs to be expanded from traditional risks, rights and responsibilities, alternative solutions, etc. to the cognition of decision-making tools themselves. This expansion is not disruptive, but ensures that patients' consent is based on understanding all the substantive factors that affect their health decisions.

2.1. Informed demands for the source and scope of use of artificial intelligence diagnosis and treatment data

The performance of artificial intelligence models is closely related to the quality and representativeness of their training data, which is the foundation of model fairness and universality. Patients have the right to know the training data on which the artificial intelligence model supporting diagnosis and treatment decisions is based, including its size, type and geographical population composition [1]. Secondly, patients have the right to know the specific use of their personal health data in this diagnosis and treatment process and the subsequent processing rules. Will these data only be used for this diagnosis and treatment decision analysis, or will they be anonymized and used for iterative optimization of the model? What level of security is in place during storage, transmission and processing of data to prevent disclosure or misuse? In addition to the above-mentioned right to know at the data level, patients should also enjoy the right to have a comprehensive understanding of their own diagnosis and treatment decision-making process during the process of artificial intelligence-assisted diagnosis and treatment. When there is a disagreement between man and machine judgment, the patient has the right to know the point of disagreement, and the right to know the professional considerations of the attending physician in finally adopting or revising the artificial intelligence recommendation.

2.2. Informed demands for decision logic and error threshold of artificial intelligence algorithm

Patients should also enjoy the right to know the algorithm decision logic and its potential error threshold during the process of artificial intelligence-assisted diagnosis and treatment. It is unrealistic to require the average patient to fully understand the complex principles of deep neural networks, but this does not mean that the decision-making process of the algorithm can become a completely closed "black box" without any explanation [2]. The core of informed consent is to make patients understand the diagnosis or treatment recommendations, rather than exhaustion of technical details. It is also critical that patients should be informed of the AI model's prediction confidence, possible error margin, or its recommended risk class under different scenarios. Only by understanding these error thresholds or uncertainty information can patients realize that artificial intelligence-assisted decision-making is not 100% accurate, so as to better weigh the potential risks and success rates of diagnosis and treatment plans.

2.3. Informed demands for the difference between artificial intelligence diagnosis and treatment suggestions and clinicians' judgments

At present, the mainstream usage of artificial intelligence in the medical field is "human-machine collaboration" [3], that is, artificial intelligence provides auxiliary suggestions, and qualified

DOI: 10.54254/2753-7048/2025.LD31098

clinicians finally confirm and are responsible for the diagnosis and treatment plan. In the practical application of artificial intelligence-assisted diagnosis and treatment, when the suggestions given by the intelligent system are inconsistent with the clinical judgment of the attending physician, the patient's right to know about such differences is particularly critical. This includes not only clearly pointing out the specific content of the inconsistent recommendations, but more importantly, patients have the right to know the underlying causes of this disagreement. For example, models may draw conclusions based on large-scale statistical data, while physicians may make comprehensive judgments based on the patient's individual situation, clinical experience and patient's personal wishes. In this process, the attending physician is obliged to clearly explain to the patient the professional considerations and clinical basis for adopting, revising or rejecting artificial intelligence recommendations.

3. Specific dilemmas in exercising patients' right to know in artificial intelligence diagnosis and treatment

Although in the era of artificial intelligence, the core demands of patients' right to know can be basically clarified, in practice, the satisfaction of these demands faces practical obstacles from the technical, institutional and cognitive levels, which leads to the effective exercise of rights. Faces multiple dilemmas.

3.1. Obstacles to information acquisition caused by algorithm black box

The internal operation mechanism of modern intelligent technology often presents a high degree of complexity and opacity. This characteristic is not subjective, but determined by its technical nature. A complex neural network may contain millions or even billions of parameters, and its decision logic is not based on explicit rules. Instead, the artificial intelligence model uses pattern learning and feature extraction on massive data to form a decision mechanism that is difficult to fully clarify in intuitive language. This inherent explanatory challenge to technology constitutes a fundamental obstacle to clearly explaining the algorithmic decision-making process to patients. In a word, the information barrier built by the "black box" of the algorithm prevents patients from obtaining the key information needed to evaluate the algorithm, which makes their right to "informed consent" essentially overhead and places them at a disadvantage of passive acceptance.

3.2. Confusion in the performance of notification obligations caused by ambiguity of responsible subjects

The artificial intelligence diagnosis and treatment model involves multiple participants, usually including upstream algorithm developers or software manufacturers, midstream medical equipment manufacturers, and downstream application-side medical institutions and front-line clinicians. Under this mode of coexistence of multiple subjects, the existing laws and regulations have not yet clearly defined the distribution of the obligation to inform information related to intelligent systems. Algorithm developers have the most comprehensive and in-depth technical information, but they are usually far away from clinical practice and lack a direct legal relationship with patients, so it is difficult to directly fulfill the obligation to inform. As the introducer and manager of intelligent system, medical institutions have the responsibility of managing and supervising internal medical staff, and should bear corresponding management responsibilities. However, medical institutions themselves may lack sufficient technical understanding to fully digest and interpret the complex

DOI: 10.54254/2753-7048/2025.LD31098

details of algorithms. Clinicians who are at the end of fulfilling their obligations to inform are legal subjects who communicate directly with patients, but they are also users of artificial intelligence themselves, and their understanding of the internal mechanism and potential risks of algorithms may be limited.

3.3. Understanding dilemma caused by information overload and professional barriers

Information about algorithm performance, data distribution characteristics, model uncertainty ranges, etc. is usually riddled with highly specialized terminology and complex statistical concepts. If it is directly presented to patients who lack medical and information technology background, it will not only not achieve effective notification, but may lead to information overload. Transforming this highly complex technical information into common language and appropriate examples that patients can understand, absorb, and translate it into a huge challenge to how to fulfill the obligation to inform. This gap in information understanding actually constitutes another form of knowledge isolation, which is contrary to the fundamental purpose of informed consent to promote patients' independent decision-making. The traditional notification method, which mainly relies on formatted written materials, faces severe challenges in its effectiveness in intelligent assisted diagnosis and treatment scenarios. Just asking the patient to sign a notice listing the technical parameters of the intelligent system cannot be regarded as a truly valid consent in the legal sense.

4. Legal protection path of patients' right to know in artificial intelligence diagnosis and treatment

Faced with the above-mentioned dilemma brought about by artificial intelligence diagnosis and treatment, the construction of legal protection system must adopt systematic and multi-level coping strategies based on the traditional framework to ensure that patients' right to know is substantially maintained in the new technological environment. and realization.

4.1. Establish a hierarchical disclosure system for diagnosis and treatment algorithm information

In order to effectively solve the contradiction between the technical barrier of "algorithm black box" and the understanding dilemma of "information overload"[], we should learn from the mature experience in the field of product responsibility and information disclosure, and construct a set of hierarchical disclosure system of diagnosis and treatment algorithm information for different subjects.

The first level is directly facing the patient. The law should force medical institutions to provide a summary notification document before implementing key artificial intelligence diagnosis and treatment. This document clearly explains the basic functions, intended uses and auxiliary roles of the artificial intelligence system in layman's language.

The second level is for medical staff as professional users and medical institutions that purchase the system. Algorithm developers must provide them with detailed technical documentation and systematic professional training. This information should thoroughly explain the design principles and core technical details of the algorithm, clarify its verified clinical indication scope, and provide complete performance data including accuracy, sensitivity, specificity, etc.

The third level is for administrative regulatory agencies and judicial organs such as drug supervision and administration departments. In order to achieve effective market access supervision

DOI: 10.54254/2753-7048/2025.LD31098

and post-event accountability, the law should strictly stipulate the filing obligations of algorithm developers. In the event of algorithm-related medical damage litigation or major security incident investigation, the judicial organ or its authorized regulatory agency has the right to obtain in-depth information on the record, and entrust a third party to conduct code audit or algorithm reproduction when necessary, so as to Conduct independent and professional review and responsibility determination.

4.2. Clarify the rules for the division of notification obligations of multiple subjects

In view of the characteristics of multiple subjects in the artificial intelligence diagnosis and treatment model, the law must build a closed loop of responsibilities with a clear division of labor and smooth connection to clearly divide the notification obligations of each participant, thereby avoiding the problem of responsibility attribution and assumption.

First, the developer or manufacturer of the algorithm should bear the obligation of generating and transmitting the source information. As designers and providers of artificial intelligence products, developers must produce and provide accurate and comprehensive information materials according to the aforementioned hierarchical disclosure system. The law should clearly stipulate that if there are defects, errors or critical concealment in the original information provided by the developer, which leads to improper notification by the downstream entity and damage to the patient, the developer shall bear corresponding legal responsibility for this information loophole.

Second, medical institutions, as purchasers and service providers of artificial intelligence systems, should undertake the obligations of access management and transformation implementation. First, medical institutions must establish a strict technical access review mechanism, conduct a professional assessment of the safety and effectiveness of the system and the completeness of information disclosure materials before purchasing, and prudently fulfill their duty of care as service providers. Second, the institution has the responsibility to translate the technical information provided by the developer into the specific clinical practice specifications of the hospital. This includes combining one's own medical conditions with the characteristics of patient groups, formulating standardized clinical application guidelines and patient notification processes, and organizing targeted physician training.

Third, clinicians bear the final obligation of informing and communicating directly to patients. Physicians are the last subjects to fulfill the obligation of informing, and their roles cannot be replaced by other subjects. The doctor's responsibility to inform is not simply scripted, but refers to the doctor's organic integration of the core content of artificial intelligence diagnosis and treatment abstract into the traditional doctor-patient communication in combination with the patient's individual condition, education level and psychological state in the specific diagnosis and treatment interaction.

4.3. Improve the right relief mechanism of patients' right to know

Effective rights protection requires not only careful regulation beforehand, but also strong relief guarantee afterwards. When patients suffer personal or property damage due to notification loopholes in artificial intelligence diagnosis and treatment, the law must provide them with smooth and effective relief channels.

First, the rules for determining the burden of proof and causality should be optimized in judicial practice. Referring to the spirit of Article 1,222 of the Civil Code of the People's Republic of China [], the burden of proof can be allocated obliquely, and the medical institution as the information

DOI: 10.54254/2753-7048/2025.LD31098

advantage party should bear the burden of proof to prove that it has followed the statutory standards and fulfilled the notification in a way that patients can understand. In the determination of causality, the "substantive impact criterion" [] can be introduced. Patients only need to prove that if they know the concealed information, rational patients may make different choices in the same situation.

The second is to build a clear liability sharing and recovery mechanism for damages. When the court confirms that the patient's damage is caused by the notification loophole, it should establish the principle that the medical institution directly facing the patient should bear the liability for compensation in advance. If a medical institution can prove that the root of its notification loophole lies in the wrong, incomplete or misleading information provided by the upstream algorithm developer, the law should clearly give the medical institution the right to recover from the faulty developer after compensating the patient.

Third, we should actively explore and establish a professional multi-dispute resolution mechanism. The establishment of a third-party mediation and appraisal committee composed of senior medical experts, legal experts and information technology experts should be encouraged. Such professional institutions can specialize in handling related disputes, and issue authoritative appraisal opinions on core technical disputes by providing neutral and professional mediation services, or accepting the entrustment of the court in litigation, thereby significantly improving the professionalism, efficiency and judicial credibility of dispute resolution.

5. Conclusion

The impact of artificial intelligence technology on the medical field is systematic, and its legal response must also be systematic. The core argument of this article is that in order to cope with the challenge of the dilemma of patients' right to know caused by the application of artificial intelligence in the field of medical diagnosis and treatment, we need to coordinate the advancement from multiple levels such as legislation, justice and industry practice. The fundamental purpose of this series of system design is to firmly safeguard the fundamental value of medical law, which is patients' independent decision-making right and personal dignity, while prudently embracing technological innovation. This is not only the responsibility for the rights and interests of individual patients, but also the active maintenance of social public interests.

References

[1] Sun Yunliang, Wang Can. Research on the result attribution of medical artificialintelligence intervention in negligent diagnosis and treatment behavior [J]. Journal ofJiangxi Normal University (Philosophy and Social Sciences Edition), 2025, 58 (05): 137-148.

[2] Li Runsheng. On the coping strategies and regulatory approaches to the "blackbox" problem of medical artificial intelligence [J]. Journal of Southeast University(Philosophy and Social Sciences Edition), 2021, 23 (06): 83-92.

[3] Zheng Yue, Jing Xiaobei, Li Guanglin. Application of human-computer intelligentcollaboration in the field of medical rehabilitation robots [J]. Journal ofInstrumentation, 2017, 38 (10): 2373-2380.

[4] Gao Songyu, Shang Jing, Wang Huaping, et al. Research status of healthinformation overload between doctors and patients under the backgroundofdigitalization [J]. Journal of Nursing Management, 2025, 25 (06): 480-485.

[5] Yang Lixin. Proof of causality and burden of proof of medical damage liability [J]. Legal Science, 2009, (01): 35-44.

[6] Li Zhixing, Lee Jihaeng, et al. Research on the Current Situation, Problemsand Countermeasures of the Legal Mechanism for Resolving Medical Disputesin China: A Comparative Interpretation Based on the "Regulations on MedicalAccidents" and the "Regulations on the Prevention and Handling of Medical Disputes" [J]. International Exchange and Law, 2022, (36): DOI: 10.31839/IBT.2022.01.36.227.