

Ethical Dilemmas and Coping Strategies in the Application of AI in the Medical Field

Xinyi Liang

Chongqing Yucai Secondary School, Chongqing, 400050, China

Abstract. With the expanding application of Artificial Intelligence (AI) technology in medical diagnosis, treatment plan formulation, drug research and development, and other areas, AI has significantly improved the efficiency of diagnosis and treatment in the medical field. However, it has also brought about a series of practical challenges. Therefore, this study focuses on the ethical dilemmas arising from AI applications in healthcare and explores corresponding coping strategies. By extensively reviewing domestic and international academic papers, industry reports, policies, and regulations related to AI medical ethics, this research summarizes existing findings and perspectives, categorizing the ethical dilemmas into six key areas: issues of data privacy and security, algorithmic bias and fairness, responsibility attribution, clinical application challenges, lagging ethical review and legislation, and problems with patients' right to autonomy and informed consent. To address these practical issues, the study proposes targeted policy recommendations, including establishing a full-process data governance system, implementing full-lifecycle algorithm management, enhancing training for medical professionals, and improving the ethical and humanistic support system. The aim is to promote the healthy and sustainable development of AI technology in the medical industry while safeguarding patients' rights and interests.

Keywords: AI in Healthcare; Ethical Dilemmas; Coping Strategies; Privacy Protection; Algorithmic Bias.

1. Introduction

Amid the digital transformation driven by new-quality productive forces, "AI +" has been officially incorporated into the 2025 Government Work Report, becoming a national strategic tool to advance industrial intelligent upgrading and modernize social governance[1]. The Opinions of the State Council on Further Implementing the "AI +" Initiative also states that efforts should be made to "promote the orderly application of AI in scenarios such as auxiliary diagnosis and treatment, health management, and medical insurance services, significantly enhancing the capacity and efficiency of primary healthcare services"[2]. Fueled by both policy support and technological breakthroughs, AI is reshaping the core ecosystem of the medical industry with unprecedented depth and breadth.

From the sub-second accurate analysis of medical images by large language models to the sub-millimeter precision of surgical robots, AI has evolved from a mere auxiliary tool to a core force spanning the entire healthcare chain—encompassing disease prediction, diagnosis, treatment, and rehabilitation. For instance, Guangzhou is rapidly advancing the application of "AI + Healthcare" scenarios. Data shows that as of November 2024, 33 medical institutions in Guangzhou have implemented over 100 "AI + Healthcare" applications. Among these, the "AI Intelligent Guidance" and "AI Pre-Consultation" services at the Women and Children's Medical Center Affiliated to Guangzhou Medical University record a daily usage of 2,500 times, with a cumulative call volume exceeding 2.8 million—markedly improving service efficiency for patients. The "Yue Jing Jing" mini-program developed by Sun Yat-sen University Zhongshan Ophthalmic Center allows citizens to conduct preliminary screenings for infant eye diseases using their mobile phones at home. Additionally, the smart medical information system for lung cancer screening at Guangdong Provincial People's Hospital supports the development of precision medicine for early cancer detection and treatment[3]. This intelligent revolution is not only addressing existing challenges in healthcare but also redefining the future boundaries of precision medicine and patient experience[4].

Despite its contributions to improving diagnostic and treatment efficiency, AI applications in healthcare have raised several practical issues. For example, the "data silo" phenomenon is prominent:

heterogeneous systems across hospitals lead to inconsistent data formats and coding standards, while unstandardized medical records and missing key fields are common problems. Furthermore, the "black box" nature of AI algorithms makes their decision-making logic difficult to interpret, undermining trust between doctors and patients. These issues have prompted reflections on the ethical dilemmas of AI in healthcare and the need for corresponding coping strategies. In-depth analysis of AI medical ethics and the development of targeted strategies have become urgent requirements to promote compliant technological development and protect patients' rights.

The current situation—where "technology advances rapidly while ethics lags behind"—has left the prevention and control of ethical risks in AI healthcare in a passive state. In response, the academic community has begun to focus on this topic. Existing research primarily covers three areas: descriptions of the current state of AI applications in healthcare, analysis of risks and ethical dilemmas associated with these applications, and the design of systems to improve AI implementation in healthcare. Studies suggest that the deep integration of AI and healthcare is fraught with uncertainties, presenting both risks and opportunities. Neither stagnation nor blind enthusiasm is advisable; instead, ethical reflection and legal framework design should accompany technological development and application to ensure that AI truly serves the health needs of the public[5].

Based on this context, this study examines the current use of AI in healthcare and the associated ethical dilemmas. Drawing on system designs from pioneering domestic regions and advanced international experiences, it proposes corresponding coping strategies to provide policy references for the application of AI in China's medical field.

2. Current State of AI Applications in the Medical Field

2.1 Key Application Areas of AI in Healthcare

2.1.1 Auxiliary Diagnosis: Full-Chain Penetration from Image Recognition to Disease Prediction

In medical image diagnosis, AI has formed a complete technical system covering multi-modal data such as X-rays, CT scans, MRIs, and pathological sections. According to the 2024 Global AI Medical Imaging Report, the average accuracy of AI-assisted lung CT nodule detection reaches 97.3%, and the sensitivity of AI in breast cancer pathological section analysis exceeds 95%—with some indicators approaching or surpassing the level of senior radiologists. Domestic enterprises like InferVision and United Imaging Intelligence have integrated their AI imaging systems into over 3,000 hospitals nationwide. For example, InferVision's chest CT AI diagnostic product increased the detection rate of early lung cancer by 28% in primary medical institutions in 2023. Internationally, Google Health's AI system for diabetic retinopathy screening has a misdiagnosis rate of only 2.7% and has been incorporated into the clinical pathway of the UK's National Health Service (NHS).

In disease prediction and early screening, AI prediction models based on electronic health records (EHRs), genomics, and wearable device data are reshaping disease prevention models. For instance, an AI system developed by Stanford University analyzes electrocardiogram (ECG) data to predict arrhythmia with an accuracy of 98.5% and has been integrated into the health monitoring function of the Apple Watch. Domestically, LinkDoc's "AI Cancer Early Screening" platform combines fecal genetic testing with clinical data, increasing the sensitivity of early colorectal cancer screening to 92%—a 40% improvement in coverage compared to traditional colonoscopy.

2.1.2 Treatment Plan Formulation: Precision Breakthroughs from Standardization to Personalization

In precision oncology, AI applications have advanced from "auxiliary reference" to "personalized plan generation." IBM Watson Health's oncology treatment system has accumulated over 3 million cancer cases and can generate multi-line treatment plans within 20 minutes based on a patient's genetic profile, pathological type, and clinical indicators. In 2023, a team from Shanghai Jiao Tong

University developed an AI system that analyzes gene mutations and tumor immune microenvironment data of non-small cell lung cancer patients, controlling the prediction error of progression-free survival for targeted therapy within ± 1.2 months—providing precise references for clinical medication.

In chronic disease management, AI enables dynamic adjustments to treatment plans through real-time data collection for conditions such as diabetes and hypertension. Livongo's AI diabetes management system in the United States continuously monitors blood glucose, diet, and exercise data to automatically generate insulin dosage recommendations, offering references for patient treatment.

2.1.3 Healthcare Management: Data-Driven Efficiency Revolution and Resource Optimization

In intelligent electronic health record (EHR) management, AI applications have evolved from "structured processing" to "clinical decision support." Tencent Miying's AI EHR system automatically extracts diagnostic keywords, medication history, and other information from medical records, increasing documentation efficiency by 60%. Simultaneously, it identifies potential diagnostic inconsistencies through semantic analysis, reducing the error rate of EHR quality control by 75%. Nuance's Dragon Ambient Experience (DAX) system in the United States uses speech recognition and AI understanding to automatically generate outpatient records, saving doctors 40% of their documentation time.

In intelligent healthcare resource allocation, AI applications have significantly improved service efficiency. At the regional healthcare level, the AI-driven healthcare resource scheduling platform used in Shanghai during the pandemic analyzed real-time data on hospitals' treatment capacity, material reserves, and patient distribution, reducing the average emergency response time by 1.5 hours. Additionally, AI applications in scenarios such as medical imaging equipment booking and operating room scheduling have greatly improved the utilization rate of hospital equipment.

2.2 Summary of Current State and Future Trends

Overall, current AI applications in healthcare exhibit the characteristic of "stratified technical implementation." Areas involving structured data processing—such as image diagnosis and EHR management—have entered the mature commercialization stage. Complex scenarios like surgical robots and precision treatment are in the accelerated phase of clinical validation. In contrast, decision-making applications involving ethical sensitivities (e.g., end-of-life care plan formulation) remain in the exploratory stage.

AI is accelerating the reconstruction of the healthcare ecosystem, showing a trend of multi-dimensional in-depth development. It is driving healthcare transformation from disease treatment to health management and from experience-driven practice to data-driven practice.

3. Ethical Dilemmas in AI Applications in the Medical Field

3.1 Data Privacy and Security Issues

The in-depth application of medical AI relies on massive volumes of sensitive medical data. During technological iteration and scenario expansion, risks of privacy breaches and system security incidents continue to escalate—threatening both patient rights and the healthy development of the industry.

The most prominent risk is privacy leakage throughout the data lifecycle. For example, in the data collection phase, some institutions fail to fully implement informed consent: patients remain unaware of how their data will be used or shared. This issue is particularly prevalent in data-sharing scenarios, where ambiguous data flow chains often lead to compliance disputes.

Furthermore, security vulnerabilities in algorithms and systems can exacerbate data leakage risks. Medical data has high economic value, making it a prime target for criminals. Some cybercriminals aim to disrupt services by poisoning data or altering models—leading AI to output incorrect

diagnostic information. They may also continuously contaminate data pools through model iteration, posing severe threats to users' lives and health.

3.2 Algorithmic Bias and Fairness Issues

AI algorithms rely on learning from large-scale datasets. However, certain demographic groups have long been underrepresented in existing biomedical datasets, leading to discrimination against specific groups and widening inequalities in healthcare resource allocation.

Algorithm training depends on massive data; if the data is biased, the resulting AI system will also be biased. For instance, in an AI model for cardiovascular disease diagnosis, cases involving elderly or low-income groups account for less than 15% of the training sample—despite the fact that the elderly are a high-risk group for cardiovascular disease. This sample bias reduces the AI system's accuracy in identifying cardiovascular conditions by over 20%.

Similarly, training data for diabetes primarily comes from patients in tertiary hospitals in urban areas—individuals who tend to follow medical advice and monitor their blood glucose regularly. In contrast, rural patients are underrepresented in the data. Due to limited healthcare access and low health awareness, rural patients monitor their blood glucose less frequently. As a result, the AI system's ability to manage blood glucose for rural patients is far inferior to its performance for urban patients—further widening health disparities between groups. For rural diabetic patients who are inadequately represented in training datasets, AI algorithms may fail to accurately identify disease characteristics, leading to misdiagnosis or missed diagnosis.

Medical decisions based on biased algorithms may result in certain groups receiving fewer healthcare resources or treatment opportunities than others. Additionally, algorithmic bias may redirect more resources to groups deemed "more needy" by the algorithm—overlooking other groups that are equally or more in need. This further exacerbates inequalities in healthcare resource allocation.

3.3 Responsibility Attribution Issues

With the widespread adoption of AI in healthcare, the rate of AI usage among doctors has increased annually. However, the ambiguous boundaries of responsibility in AI medical applications mean that in the event of medical errors, disputes over liability often arise between AI developers, medical institutions, and healthcare providers—ultimately harming patients' legitimate rights.

Currently, doctors generally view AI as an auxiliary decision-making tool. Both doctors and AI participate in the decision-making process: doctors typically adjust AI-generated diagnostic results based on their clinical experience. Although doctors make the final medical decision, this decision is shaped by multiple factors—making it difficult to attribute responsibility to a single party.

If a medical institution fails to conduct strict validity verification on an AI system before deploying it—resulting in the use of a system with algorithmic flaws—the institution should bear corresponding responsibility. If an AI developer overlooks key clinical features during algorithm design—leading to algorithmic vulnerabilities—the developer should be held accountable. If a healthcare provider overrelies on AI-generated diagnostic results without conducting necessary manual reviews, the provider should also take responsibility.

Currently, AI functions as a complex algorithmic system and does not possess legal personality—preventing it from being an independent responsible entity. Even if AI were granted legal personality, it would lack independent assets and the ability to bear liability; ultimately, humans would still be held responsible. The traditional legal framework fails to clearly identify designers, developers, and users as responsible entities—creating a legal gap in responsibility attribution.

3.4 Clinical Application Challenges

Healthcare providers face dual challenges regarding their willingness and ability to use medical AI.

In terms of willingness to use: There is significant variation in healthcare providers' perceptions of AI. Some older doctors rely on traditional experience and refuse to use AI auxiliary tools—missing

opportunities to improve their medical skills with AI support. Conversely, some younger doctors over-rely on AI (e.g., following AI-generated medication plans without considering a patient's allergy history or special physical conditions)—increasing medication risks.

In terms of ability to use: Mismatches between medical AI and clinical scenarios, coupled with uneven AI proficiency among healthcare providers, pose hidden risks to medical safety. On one hand, AI technology is often disconnected from clinical needs. For example, when surgical robots operate on patients with complex anatomical structures, insufficient precision in pre-operative image data processing may lead to errors in surgical path planning. Similarly, AI-assisted emergency triage systems often have low accuracy in identifying symptoms in child patients due to limited sample sizes.

Emotional communication between doctors and patients is a deeply humanistic aspect of healthcare. Excessive AI involvement can make doctor-patient interactions mechanical and formulaic. Communication is also a process of building trust: doctors' descriptions of patients' conditions and positive feedback can boost patients' confidence in recovery. Over-reliance on AI undermines the empathy and trust that stem from face-to-face interactions, weakening the emotional bond between doctors and patients.

Furthermore, the high costs of introducing and maintaining medical AI mean that advanced AI healthcare technologies are often concentrated in large hospitals or high-income regions. This further widens gaps in healthcare accessibility between different regions and groups, exacerbating inequalities in healthcare resource allocation and challenging fairness and inclusivity in the medical field.

3.5 Lagging Ethical Review and Legislation

Medical ethical review is a critical step in the application of medical technologies. However, existing ethical review mechanisms are primarily designed for traditional medical scenarios and cannot effectively address emerging ethical challenges in AI healthcare.

Although the Interim Measures for Science and Technology Ethical Review explicitly includes science and technology activities involving data and algorithms in key review content, lax implementation remains common in some regions.

The legislative process in the field of medical AI lags behind technological development. The lack of clear legal boundaries not only triggers legal disputes but also hinders the standardized development of the industry. Currently, core issues—such as whether medical AI should be classified as a "medical tool" or an "auxiliary entity," and the legal validity of AI-generated diagnostic reports—remain unregulated. For example, when medical AI violates patient privacy through data leakage, disputes often arise over whether to pursue liability under the Cybersecurity Law or the Regulations on the Prevention and Handling of Medical Disputes—leaving patients struggling to defend their rights.

3.6 Issues with Patients' Right to Autonomy and Informed Consent

The inadequate protection of patients' right to autonomy and informed consent in AI healthcare violates core principles of medical ethics, triggering a crisis of trust in the technology.

In AI-assisted diagnosis and treatment, patients often lack understanding of the technology's working principles or decision-making basis. Informed consent is frequently a formality: when hospitals introduce AI diagnostic systems, they may only briefly mention "potential AI usage" in admission notices—without explaining key information such as the AI system's diagnostic logic, data usage scope, or error rate. In practice, patients remain unaware of the extent to which their treatment relies on AI.

Additionally, when personal sensitive medical data is used for AI training or sharing, institutions provide vague explanations of data usage. Patients may sign informed consent forms, but they are actually unaware of how their data flows. This not only violates patients' right to autonomy but also埋下 hidden risks for data ethics issues.

4. Coping Strategies for AI Application Issues in the Medical Field

4.1 Establish a Full-Process Data Governance System

Current issues—such as data silos, security vulnerabilities, and insufficient compliance—seriously undermine the reliability and credibility of medical AI. To address these challenges, a data security management system centered on "full-process coverage and full-element control" is needed to ensure that data is "usable but not visible, controllable, and traceable."

4.1.1 Standardize the Data Collection Phase: Fortify Compliance and Quality Foundations

Data collection is the starting point of governance; it requires balancing "data usability" and "privacy protection." Medical data is the "cornerstone" of AI model training and application—making a wide range of medical data sources crucial for model training and optimization. Therefore, data collection must prioritize patient privacy protection.

In the data collection phase, institutions must strictly comply with laws and regulations such as the Personal Information Protection Law of the People's Republic of China and the Measures for the Management of Medical Institution Data Security. A "informed consent + dynamic authorization" mechanism should be adopted to clarify data usage (e.g., "for training a specific disease AI model only"). "One-size-fits-all authorization" should be prohibited. For highly sensitive information (e.g., HIV status or mental illness history), "dual authorization" is required—needing approval from both the patient and the ethics committee.

To avoid model generalization limitations caused by single-source data, a "multi-center data alliance" should be established to integrate diagnostic data from hospitals of different levels (tertiary hospitals, community health centers), regions (eastern, western China), and demographic groups (varying by gender, age, ethnicity, and underlying conditions). Additionally, medical data cleaning standards should be developed to eliminate duplicate, erroneous, or incomplete data (with missing values exceeding 30%)—ensuring the accuracy of training data.

4.1.2 Strengthen the Data Storage Phase: Hierarchical Protection and Technical Encryption

Given the high sensitivity of medical data, a security system combining "hierarchical storage" and "multi-layer encryption" is essential. Hierarchical storage: Data is classified based on sensitivity: "Core sensitive data" (e.g., genetic information, medical records) is stored on local physical servers in hospitals. "Generally sensitive data" (e.g., test reports) uses a "local + cloud backup" model. "Non-sensitive data" (e.g., anonymized diagnostic statistics) can be used for public scientific research collaboration.

Multi-layer encryption: Encryption is applied across storage and transmission: National cryptographic algorithms (e.g., SM4) are used for static data encryption. Transport Layer Security (TLS 1.3) ensures secure data upload/download. Blockchain technology is introduced to enable data storage traceability: each data write generates an immutable timestamp, ensuring no unauthorized data tampering.

4.1.3 Strictly Control Data Flow and Application: Establish a "Closed-Loop Traceability" Mechanism

Data flow is a high-risk phase for security incidents; "permission control + dynamic monitoring" should be used to achieve full-link traceability.

In terms of permission control: The principle of "minimum necessary permissions" should be followed, and a "role-permission-data" mapping system established. For example: AI algorithm engineers only access anonymized training data. Clinicians only view data related to their own patients. All operations are recorded in a "data operation log," which is retained for at least 5 years.

In terms of dynamic monitoring: A medical data flow traceability platform should be built to clarify the full-chain information of data—including "source-handler-usage-destination." Regular data security audits should be conducted by third-party institutions to verify compliance with data

flow rules. Violations (e.g., unauthorized data export, over-range usage) should result in an "immediate disqualification" penalty, with relevant parties held accountable.

4.2 Implement Full-Lifecycle Algorithm Management

Algorithms are the "core engine" of medical AI. However, two issues—lack of interpretability due to the "black box" nature and unfairness caused by "data bias"—have become major sources of trust crises in AI-driven diagnostic decisions. To address these, a closed-loop management mechanism for algorithms (covering R&D to iteration) should be established, with the goals of "full-lifecycle coverage and dual improvement in interpretability and fairness."

4.2.1 Enhance Algorithmic Interpretability and Transparency: Break the "Black Box" Barrier

To solve the problem of AI diagnostic results being "understandable in outcome but not in process," algorithm "transparency" should be promoted through both technical and institutional measures.

At the technical level: Explainable AI (XAI) technologies should be introduced to translate AI decision-making processes into language understandable to clinicians. For critical care algorithms (e.g., AI-assisted surgical planning), a "decision tree visualization report" should be provided—clarifying the weight of key features (e.g., the impact of tumor size and location on surgical path selection).

At the institutional level: An algorithm "transparency disclosure" system should be established. Before launching a medical AI product, developers must submit an Algorithm Specification to regulatory authorities (e.g., the National Medical Products Administration). This document should include the algorithm's principles, training data sources and distribution, core parameters, interpretability methods, and limitations. During clinical application, patients should receive both the AI-assisted diagnostic result and a "summary explanation"—safeguarding their right to information.

4.2.2 Establish an Algorithmic Bias Governance Mechanism: Ensure Diagnostic Fairness

Algorithmic bias (e.g., higher misdiagnosis rates for women, the elderly, or ethnic minorities) is essentially an extension of "training data bias." It should be addressed through a three-step approach: "data balancing → bias detection → continuous optimization."

Data balancing: Training data for AI models must meet "demographic representativeness" requirements. For example, in an AI model for hypertension diagnosis: Data from elderly patients (aged 65+) should account for at least 30% of the sample. Data from female patients should account for at least 45%. This prevents diagnostic bias against other groups caused by over-reliance on data from young males. For rare disease models, multi-center rare disease case data should be included to ensure coverage of minority groups.

Bias detection and evaluation: A quantitative index system for algorithmic bias should be established—including metrics such as "misdiagnosis rate differences across groups" and "the weight of sensitive features (e.g., ethnicity, gender) in decision-making." Before clinical application, third-party institutions must conduct a "fairness test." If the misdiagnosis rate for any group exceeds the average by 15%, the model must be re-optimized.

Dynamic iteration and optimization: An "algorithm bias early warning → rapid response" mechanism should be established. During clinical application, real-time monitoring of demographic differences in AI decision results is required. If bias exceeds the threshold (e.g., abnormally high misdiagnosis rates for rural patients in a certain region), application in that scenario should be immediately suspended. The root cause (e.g., training data failing to cover common symptoms in rural areas) should be analyzed, the model retrained, and fairness testing re-conducted before resuming use.

4.3 Strengthen Training for Medical Professionals

The effectiveness of medical AI implementation ultimately depends on "human-machine collaboration capabilities." Currently, medical professionals face issues such as "insufficient AI awareness, weak application skills, and low trust in AI." To address these, a dual-track approach—combining "university education" and "on-the-job training"—should be adopted to build a healthcare workforce adapted to the AI era.

4.3.1 Optimize University Talent Training Models: Achieve Integrated "Medicine + AI" Education

To bridge the gap between university medical education and the needs of AI healthcare, curriculum systems and practical training models must be restructured.

In curriculum reform: Compulsory courses such as "Fundamentals of Medical AI," "Medical Data Science," and "Algorithmic Ethics and Security" should be added to majors like clinical medicine, medical imaging, and public health. Course content should cover AI diagnostic principles, data annotation methods, and algorithm limitations. Interdisciplinary "Medicine + AI" programs should be launched to cultivate compound talents who understand both medicine and AI tools.

In practical training enhancement: A tripartite collaborative practice base—integrating "universities, hospitals, and AI enterprises"—should be established. For example, partnerships with enterprises like United Imaging and InferVision can allow students to participate in real AI medical projects. "Case-based teaching" should be introduced: analyzing "AI misdiagnosis cases" to help students understand the limitations of AI and develop a mindset of "taking clinical needs as the core and AI as an auxiliary tool."

4.3.2 Intensify On-the-Job Training: Improve AI Application Skills and Risk Awareness

In-service doctors are direct users of medical AI. "Stratified, job-specific, and practice-oriented" training should be provided to address the problem of "being afraid to use AI or not knowing how to use it."

Stratified and job-specific training: Differentiated training programs should be developed based on job requirements: Radiologists focus on training in AI image analysis tools. General practitioners focus on applying AI chronic disease management systems to improve primary care efficiency. Hospital administrators focus on training in the compliance management of AI medical projects. Practice-oriented training: A hybrid training model combining "online micro-courses, offline hands-on practice, and case discussions" should be adopted. For example: AI application cases can be uploaded to the hospital's internal system, with doctors grouped to discuss "differences between AI recommendations and clinical judgments."

Regular AI application skill assessments should be conducted, with results linked to professional title evaluation and performance bonuses to motivate learning.

4.4 Improve the Ethical and Humanistic Support System

Upholding the bottom line of humanistic care, AI should be prohibited from fully replacing doctors in consultations. The time allocated to doctor-patient communication should account for no less than 50% of the total treatment time.

The essence of healthcare lies in "humanistic care + scientific treatment." As an auxiliary tool, AI must adhere to ethical bottom lines and maintain humanistic warmth—preventing "cold healthcare" caused by technological alienation.

4.4.1 Establish a Full-Process Ethical Review Mechanism

To prevent ethical risks from the source, a full-stage ethical review system—covering "R&D, clinical application, and post-application monitoring"—should be built.

R&D phase: University or enterprise ethics committees review the "data ethics" (e.g., compliance of data sources, protection of patient privacy) and "algorithmic ethics" (e.g., absence of discriminatory design) of AI projects. Projects failing the review cannot proceed with R&D.

Pre-clinical application phase: Applications must be submitted to the hospital ethics committee for approval, accompanied by an "algorithm performance report, risk assessment report, and patient informed consent plan." The review focuses on evaluating the impact of AI on patient rights (e.g., whether it increases misdiagnosis risks or undermines patients' right to information).

Clinical application phase: A "dynamic ethical monitoring mechanism" should be established. The ethics committee regularly reviews AI diagnostic cases and collects feedback from doctors and patients. If ethical issues are identified, AI application in the relevant scenario is immediately suspended, rectification is ordered, and improvements are made. Additionally, cross-institutional ethical collaboration should be promoted: a national expert database for AI medical ethics review should be established to provide guidance to local institutions and ensure unified review standards.

4.4.2 Uphold the Bottom Line of Humanistic Care

To prevent AI from replacing the "emotional value" of doctors, the "auxiliary role" of AI in healthcare should be clearly defined. AI must not fully replace doctors in core links—such as consultations, diagnosis, or treatment plan formulation. A clear rule should be established: "AI shall only serve as an auxiliary tool for doctors; the final diagnostic decision must be made by a qualified medical professional."

The requirement for "doctor-patient communication time" must be strictly enforced: in all treatment scenarios (outpatient consultations, hospitalization, surgery), the time allocated to face-to-face doctor-patient communication (including condition explanation, treatment plan discussion, and psychological counseling) must account for no less than 50% of the total treatment time. Doctors are prohibited from shortening communication time on the grounds that "AI has already provided a conclusion."

Furthermore, "humanistic care training" should be strengthened. Training programs should focus on improving doctors' communication skills and empathy—helping them balance AI usage with humanistic care. A "patient satisfaction evaluation mechanism for humanistic care" should be established: patient feedback on doctor-patient communication and humanistic care should be incorporated into doctors' performance evaluations. This forces healthcare services to return to the essence of "patient-centered care."

5. Research Conclusions and Discussion

Currently, the rapid development of AI technology has profoundly reshaped the healthcare ecosystem, playing a transformative role in the entire chain of diagnosis, treatment, and management. However, AI has also brought multiple risks—including data privacy leakage, increased misdiagnosis rates for specific groups due to algorithmic bias, ambiguous responsibility attribution, lagging ethical review and legislation, and the formalization of patient informed consent.

To address these issues, comprehensive measures are required: establishing a full-process data governance system, implementing full-lifecycle algorithm management, providing "Medicine + AI" training for medical professionals, and improving the ethical support system. The government should take the lead in improving relevant systems, while enterprises and medical institutions collaborate to advance implementation. By balancing technological efficiency with humanistic care, we can ensure that medical AI truly serves the goals of healthcare accessibility and public health improvement.

References

- [1] National Development and Reform Commission of the People's Republic of China(2025). Strengthening the Policy and Institutional Foundation for Deepening the "AI+" Initiative
- [2] From: https://www.ndrc.gov.cn/xxgk/jd/jd/202508/t20250828_1400103.html
- [3] The State Council(2025). Opinions of the State Council on Deeply Implementing the "AI+" Initiative, from https://www.gov.cn/zhengce/zhengceku/202508/content_7037862.htm?sharetype=link

- [4] Guangzhou Daily(2025). Focusing on Guangzhou's Health and Medical Data Management and Application, Deputies and Committee Members Offer Suggestions for "AI + Healthcare" from https://www.gz.gov.cn/zt/jj2025gzlh/lhgjc/msbz/content/post_10122397.html
- [5] Kang Xiaoyan, Li Jiafei & Sun Hongliang(2025). The Development System of Medical Artificial Intelligence in Europe and Its Implications. *Industry and Technology Forum*, 24(14), 46-48
- [6] Wu Bei & Lu Shucheng (2019). Reflection and Response: Ethical and Legal Risks of AI in Healthcare. *Medicine and Philosophy*, 40(21), 9-13
- [7] Liu Junlin & Feng Gong (2025). Ethical Paradoxes and Countermeasures in Medical Artificial Intelligence. *Journal of Naval Medical University*, 46(08), 982-988