White Paper on AI Healthcare Governance

（2026）

医疗健康产业俱乐部
明博学社

上海社会科学院
SHANGHAI ACADEMY OF SOCIAL SCIENCES
1958

管理视野
MANAGEMENT INSIGHTS

January 2026

Acknowledgments

This report was made possible with the generous support of the United Nations Public Administration Network (UNPAN). The English version is available for download at: **https://unpan.un.org/**

Project Team Members

Author:

- **Zhao Fuchun,** Director of the Digital Economy Research Center, Institute of Information, Shanghai Academy of Social Sciences (SASS); Associate Professor, PhD

Contributors:

- **Xia Beili,** Assistant Professor, Institute of Information, SASS; PhD
- **Liu Shufeng,** Assistant Professor, Institute of Information, SASS; PhD
- **Huang Hao,** Chairman of Mingbo Medical Club; Senior Industry Marketing Expert
- **Gu Wenbing,** Board Member of Mingbo Medical Club; Senior Industry Consulting Expert
- **Yu Baoping,** Director of the Center for FBK, School of Management, Fudan University; PhD

Contents

Executive Summary

As Artificial Intelligence (AI) integrates deeply into the healthcare sector, a profound industrial transformation is accelerating. While AI healthcare significantly enhances clinical efficiency and diagnostic precision, it also introduces unprecedented complexities. Consequently, constructing a scientific, prudent, and forward-looking governance system has become an urgent task for the current era. This white paper aims to systematically review global AI healthcare governance trends, conduct an in-depth analysis of China's current governance landscape, and propose a framework and policy recommendations that blend international perspectives with local characteristics.

First, after clarifying the developmental background and the urgency of governance, this report conducts a comparative study of the governance models of developed economies—including the United States, the European Union, the United Kingdom, and Singapore. It extracts their core characteristics, categorized as **"Market-Driven," "Rule-Driven," "Value-Oriented," and "Agile Governance."** Building on this, the report establishes a comprehensive evaluation index system for governance levels across four dimensions: **Governance System, Governance Capability, Governance Effectiveness, and Continuous Improvement.**

Second, the report systematically outlines China's significant achievements in policies, regulations, and regulatory practices driven by top-level strategies. Utilizing the aforementioned evaluation framework, it objectively analyzes deep-seated challenges such as insufficient institutional granularity, an incomplete multi-stakeholder governance mechanism, and the lack of a closed-loop value feedback system. The report then performs a horizontal comparative analysis of the governance status of five major economies to clarify their respective strengths, weaknesses, and evolutionary paths.

To address these challenges, the report adopts a risk-based perspective to map out an

AI healthcare life-cycle risk profile covering four dimensions: technology, data, law, and ethics. Based on this, it proposes a "Troika" path to build a core regulatory system: **Multi-stakeholder Governance** to define subjects, **Differentiated Regulation** to clarify objects, and **Full-Process Coverage** to provide necessary tools.

Finally, the report puts forward specific policy recommendations and action initiatives across multiple levels—improving top-level design, strengthening technical support, optimizing regulatory sandboxes, and building feedback mechanisms. These aim to provide a decision-making reference for a new paradigm of safe, credible, efficient, and equitable AI healthcare governance, contributing Chinese wisdom and solutions to the global arena.

Key Insights of the Report

1. AI Healthcare: Governance Outweighs Technology. In the evolution of AI healthcare, governance is not an adjunct to technology but the cornerstone of its success. Healthcare involves high-stakes life-and-death decisions; algorithmic superiority cannot hedge against legal risks or trust deficits. While technology addresses efficiency, governance addresses market access and liability. Without scientific approval and traceability mechanisms, advanced algorithms cannot obtain institutional authorization. In the long run, governance is the lifeline for rebuilding physician-patient trust. The "black box" nature of AI requires governance to possess dynamic monitoring and corrective capabilities to prevent bias and safeguard fairness. Medicine is fundamentally a "humanistic" discipline; a people-centered governance architecture ensures that humans retain final decision-making power, protecting physician autonomy while upholding patient dignity.

2. Multi-stakeholder Governance is an Inevitable Trend. AI healthcare governance is not a single-dimension administrative regulation but a complex socio-technical system engineering project. A single government department cannot fully grasp the intricacies of algorithmic iterations; it must rely on the professional depth of enterprises and medical institutions to bridge regulatory gaps. As AI intervenes in decision-making, liability for harm extends from individual physicians to data providers and developers. These blurred liability boundaries require coordinated efforts from all stakeholders along the value chain to prevent legal vacuums. Governance is essentially a dynamic balance of multiple values—a negotiation between technological innovation, clinical safety, policy inclusivity, and privacy protection. Only through multi-stakeholder co-governance can we find the "greatest common denominator" between incentivizing innovation and controlling risk.

3. Dynamic Governance Mechanisms are the Key. AI healthcare is inherently "evolutionary," with algorithmic performance drifting in real-time based on the data environment. Technological iteration cycles are far faster than the update cycles of traditional regulations. Static regulation not only fails to capture hidden algorithmic risks but may also stifle innovation due to institutional lag. This contradiction necessitates a shift from **"Static Access Control" to "Dynamic Evolution,"** ensuring that the regulatory pace resonates with the pulse of technology. In practice, we should build a smart feedback system for the entire life cycle. By establishing "Regulatory Sandboxes," we provide elastic testing spaces for new technologies and promote the iterative cycling of policies and technical standards. Simultaneously, digital platforms should be used for real-time audits and continuous ethical assessments, allowing for immediate intervention when performance degrades or bias emerges. This flexible approach shifts the focus from "post-event accountability" to "whole-process monitoring," ensuring that AI healthcare operates safely within a dynamic equilibrium.

Chapter 1: AI Healthcare: Growth and Governance Challenges

1.1 The Rise of the AI Healthcare Boom

Over the past decade, as the application of Artificial Intelligence (AI) has become increasingly widespread in the medical field, China has inaugurated a new competitive track for AI healthcare[①]. In early 2025, the emergence of DeepSeek marked the arrival of the era of "democratized" large language models(LLMs) in China, making 2025 the "Year One" of AI healthcare in the country.

The immense potential of AI healthcare is translating into tangible market value, attracting global capital and tech giants. According to forecasts by EO Intelligence, the market size for medical large models in China will reach nearly 2 billion RMB by 2025 and is expected to surpass the 10 billion RMB mark by 2028[②], reflecting robust growth momentum under the wave of generative AI.

Globally, the outlook is equally expansive. A study by The Research Insights indicates that the global AI healthcare market will reach a staggering $187.7 billion by 2030, with a compound annual growth rate (CAGR) of 37%[③]. This growth is fueled by rising global healthcare demands, escalating costs, and active government promotion of digital transformation. From startups to multinational pharmaceutical companies, the entire industry chain is being reshaped by AI.

The rise of this boom is the result of a synergy between technological breakthroughs, market demand, and policy support:

- **Technological breakthroughs:** The current prosperity of AI stems from the revival of deep learning led by Geoffrey Hinton and others since 2006. The 2012 breakthrough of the AlexNet team in the ImageNet competition indirectly advanced medical image analysis. The release of ChatGPT in late 2022 further triggered a surge in large language models (LLMs), multimodal models, and multi-agent systems. AI algorithms have demonstrated diagnostic accuracy surpassing humans in imaging for diseases such as diabetic retinopathy, breast cancer, and lung cancer. Research shows that GPT-4's capabilities in medical

① This report defines "AI Healthcare" as the systematic and intelligent transformation of the healthcare industry using AI technology across R&D, clinical diagnosis, and health management.

② *EO Intelligence: 2025 China Medical Large Model Industry Analysis.* [2025-10-25] https://www.vzkoo.com/read/2025070375b28b3d8ed02cc9579f6c3b.html

③ *The Research Insights: Global AI Healthcare Market Forecast 2030.* https://www.cn-healthcare.com/articlewm/20250529/content-1651020.html examinations and clinical reasoning tasks have reached or exceeded human passing benchmarks.①

- **Market demand:** For pharmaceutical companies, AI accelerates R&D in molecular synthesis and target discovery, shortening cycles and reducing costs. For large hospitals, AI replaces routine tasks and assists in designing clinical trials using "real-world" data and electronic health records (EHR), predicting outcomes and optimizing dosages to improve patient compliance and service efficiency.
- **Patient needs:** Demand is driven by aging populations, the prevalence of chronic diseases (e.g., diabetes, obesity), and rising expectations for quality care. Meanwhile, there is an increasing shortage of skilled personnel in high-workload fields like radiology and pathology. AI medical devices are expected to provide expert-level diagnosis on demand, avoiding human error and inconsistency.
- **Policy support:** Authorities face challenges like budget deficits and uneven resource distribution. AI healthcare provides continuous care via remote monitoring and wearable devices, predicts disease risks through data analysis, and reduces waste. Consequently, governments are actively issuing regulations and development policies to encourage innovation.

1.2 The Evolution of AI Healthcare

The evolution of AI in medicine has progressed through three distinct stages:

Phase 1: Rule-based Expert Systems & Early Machine Learning (1960s – 2011)
Explorations centered on expert systems like INTERNIST-1 and Stanford's MYCIN. ②These systems simulated diagnostic reasoning via "if-then" rules. However, their inability to handle uncertainty or update knowledge bases limited their clinical use. Later, traditional machine learning (e.g., SVM, Random Forest) began learning patterns from structured laboratory data. This phase was characterized by Artificial Narrow Intelligence (ANI), which was highly dependent on manual knowledge engineering.

Phase 2: Deep Learning-driven Single-task Models (2012 – 2022.10)
Breakthroughs in Convolutional Neural Networks (CNNs) brought a "golden age" for imaging③. In 2017, a Stanford study published in Nature demonstrated a deep learning

① Nori H, King N, McKinney S M, et al. Capabilities of GPT-4 on medical challenge problems[J]. *arXiv preprint arXiv:2303.13375*, 2023.
② Hirani R, Noruzi K, Khuram H, et al. Artificial intelligence and healthcare: a journey through history, present innovations, and future possibilities[J/OL]. Life, 2024, 14(4): 557[2025-10-26]. https://pubmed.ncbi.nlm.nih.gov/38792579/
③ Lim J I, Regillo C D, Sadda S R, et al. Artificial intelligence detection of diabetic retinopathy: subgroup

model for skin cancer identification with accuracy comparable to dermatologists①. This era was marked by "specialized and precise" models that solved highly specific clinical problems but lacked synergy and required massive amounts of finely labeled data.

Phase 3: Cognitive Intelligence & The Emergence of Large Models (2022.11 – Present)

The era of Generative AI began with GPT-3, revolutionizing the processing of unstructured medical texts. Models like GPT-4 and Google's Med-PaLM 2 have gained unprecedented knowledge integration and reasoning capabilities through pre-training on massive medical datasets②. AI is evolving from a diagnostic tool into a "virtual physician assistant" and "medical research partner," capable of summarizing records, generating reports, and proposing scientific hypotheses.

1.3 Application Areas and Governance Risks of AI Healthcare

Artificial Intelligence (AI) is increasingly utilized across the healthcare system, primarily spanning four key domains: clinical practice, biomedical research, public health, and healthcare management.

1. Clinical Practice

AI holds immense potential in clinical practice, covering diagnostics, treatment decision-making, and clinical research. AI assists physicians in diagnosing and treating patients more efficiently by automating tasks such as image analysis (e.g., in radiology, ophthalmology, dermatology, and pathology) and signal processing (e.g., ECG, audiology, and EEG). For instance, deep learning models can automatically identify and label the boundaries of anatomical structures or lesions and prioritize findings that require immediate attention. AI can also integrate and analyze clinical data to streamline workflows. In fields such as digital pathology, emergency medicine, surgery, risk prediction, adaptive interventions, and home care, AI has demonstrated significant value in improving patient care, enhancing diagnostic accuracy, and optimizing resource allocation.

2. Biomedical Research

AI is playing an increasingly vital role in biomedical research, particularly in clinical knowledge retrieval, drug discovery, and clinical trials. AI optimizes retrieval results from biomedical literature, enhancing the efficiency and accuracy of information

---

comparison of the EyeArt system with ophthalmologists' dilated examinations[J]. *Ophthalmology Science*, 2023, 3(1): 100228. DOI:10.1016/j.xops.2022.100228.

① Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118.

② Singhal K, Tu T, Gottweis J, et al. Toward expert-level medical question answering with large language models[J]. *Nature Medicine*, 2025, 31(3): 943–950. DOI:10.1038/s41591-024-03423-7.

gathering. In drug design and discovery, AI analyzes vast compound databases to predict molecular activity and assess the efficacy and safety of drug candidates. Furthermore, AI helps conduct clinical trials more effectively by assisting in participant selection, endpoint assessment, and trial design optimization. Additionally, AI will drive the advancement of personalized medicine by analyzing patients' molecular and genetic profiles alongside clinical, demographic, social, and lifestyle data to provide more precise treatment plans.

3. Public Health

Applications in public health focus on disease surveillance, health promotion, and environmental health. AI helps identify high-risk populations or regions and predicts disease outbreaks. For example, digital epidemiological surveillance can integrate data from news, social media, sensors, and clinical reports to establish early warning systems. AI also assists in analyzing the impact of environmental factors on health and formulating preventive measures. Moreover, AI improves health system management by optimizing resource allocation, enhancing the performance of medical facilities, and reducing public safety incidents such as traffic accidents.

4. Healthcare Management

AI can optimize administrative processes in healthcare to improve efficiency and accuracy. For example, AI can automate medical coding, scheduling, fraud detection, and patient flow management. It also assists in medical audits to evaluate the quality of healthcare services and improvement measures. Furthermore, AI facilitates patient tracking and engagement—communicating via SMS and patient portals to provide personalized health management services. The application of AI can alleviate the burden on healthcare workers while improving the efficiency and accessibility of medical services.

Table 1-1: Typical Application Scenarios, Value, and Governance Risks of AI Healthcare

Scenario	Category	AI Technical Value	Governance Risks
Drug R&D	Research & Development	Used for target discovery and rapid compound matching; significantly shortens R&D cycles, reduces costs, and improves success rates.	Data privacy leaks, algorithmic bias, intellectual property (IP) ownership, and insufficient compliance for cross-border data transfer.
Clinical Trials	Research & Development	Enhances trial efficiency; predicts trial success and drug side effects; enables intelligent management of clinical data.	Patient privacy leaks, authenticity of trial data, lack of transparency in algorithmic decision-making, and ethical

			controversies.
Post-market Surveillance	Research & Development	Personalized efficacy assessment and patient care; early detection of safety events; automated adverse event reporting; strengthened pharmacovigilance.	Under-reporting due to incomplete pharmacovigilance data, false alarms in automated systems, misuse of patient data, and cross-border regulatory coordination.
Medical Imaging	Disease Screening	Improves screening efficiency; reduces rates of misdiagnosis and missed diagnosis.	Privacy leaks of imaging data, ambiguous liability for misdiagnosis, non-standardized data labeling, and compliance risks.
Imaging-Aided Diagnosis	Disease Screening	Analyzes imaging data to improve diagnostic accuracy and assist physicians in clinical decision-making.	Risks of misdiagnosis, data privacy leaks, and algorithmic bias.
Clinical Decision Support	Assisted Treatment	Decision support for treatment plans, preoperative planning, surgical navigation, and prognosis assessment.	Lack of personalized treatment plans, surgical navigation system failure, insufficient generalization of prognosis models, and unclear physician-patient liability.
Precision Medicine	Genetic Testing	Enhances testing efficiency and accuracy; provides better support for disease treatment.	Privacy leaks of genetic data, errors in interpreting test results, risks of genetic discrimination, and compliance challenges.
Health Management	Health Monitoring	Real-time monitoring of user health data via wearable devices; enables early warning of diseases.	Security vulnerabilities in device data, privacy infringement, false alarms due to equipment failure, and insufficient informed consent.
Health Intervention	Health Monitoring	Provides personalized intervention plans (e.g., diet, exercise advice) based on health data.	Lack of medical evidence, algorithmic opacity, and potential decline in users' autonomous health management capabilities.
Smart Diagnosis & Treatment	Virtual Assistants	Assists doctors in diagnosis; provides diagnostic suggestions and treatment plans.	Misdiagnosis risks, over-reliance on AI by physicians, and difficulties in liability attribution.

Smart Treatment	Smart Diagnosis	Assists in formulating personalized treatment plans (e.g., radiotherapy dose calculation, surgical path planning).	Medical accidents due to imprecise plans, "black box" algorithm issues, and insufficient interdisciplinary collaboration.
Surgery Scheduling	Medical Management	Optimizes resource allocation in operating rooms; improves scheduling efficiency and reduces patient wait times.	Unfair resource allocation due to algorithmic opacity, surgical delays due to system failure, and unclear responsible entities.
Medical Robotics	Surgical Robots	Enhances surgical precision; reduces trauma; assists doctors in performing complex surgeries.	Equipment failure, increased operator dependency, and difficulties in liability attribution.
Rehabilitation & Nursing Robots	Medical Robotics	Alleviates the scarcity of rehabilitation resources; improves the quality of patient recovery.	Insufficient equipment safety, unscientific evaluation of rehabilitation effects, and unclear responsible entities.
Medical Service Robots	Medical Robotics	Used for intelligent guidance, disinfection, and sterilization; reduces costs and increases efficiency.	Potential to mislead patients, lack of data transparency, equipment failure, and difficulties in liability attribution.
Public Health	Infectious Disease Management	Analyzes disease prediction and contact tracing by integrating population and mobility data.	Privacy leaks, bias in algorithmic predictions, data source reliability, and barriers to cross-departmental data sharing.

Sources: World Health Organization (2024), complied by authors

1.4 Definition and Scope of AI Healthcare Governance

The concept of AI healthcare governance is derived from "AI Governance," which refers to the design, development, deployment, and oversight of AI systems by diverse stakeholders (public, private, and social) to ensure risks are controlled and benefits align with societal values[①].

The World Health Organization (WHO) defines AI Governance for Health as the application of ethical principles, regulatory rules, clinical standards, and technical controls specifically to AI systems in health (e.g., diagnostic tools, clinical decision

① Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. AI and Ethics, 5, 3265–3279. https://doi.org/10.1007/s43681-024-00653-w; World Economic Forum & Info-communications Media Development Authority. (2020). *Companion to the Model AI Governance Framework: Implementation and Self-Assessment Guide for Organizations.* Geneva, Switzerland: World Economic Forum.

support). Its goal is to safeguard patient safety, ensure clinical effectiveness, protect privacy, promote equity, and maintain professional responsibility throughout the AI lifecycle.[①]

1.5 Core Issues in AI Healthcare Governance

AI healthcare governance faces the following five sets of tensions and challenges, which constitute its core issues:

Contradiction 1: High-Speed Technological Development vs. Institutional Lag

AI technology, especially large model technology, is iterating on a "monthly" basis, with new capabilities and application scenarios emerging constantly. In contrast, the formulation and revision of laws and regulations typically undergo lengthy research, demonstration, and legislative procedures, measured in "years." This stark contrast between "fast" and "slow" results in significant regulatory gaps. For instance, when an AI chatbot capable of providing diagnostic suggestions emerges, are our existing medical device approval systems, medical liability frameworks, and data privacy regulations sufficient to respond? This institutional lag not only exposes patients to unknown risks but also causes innovative enterprises to hesitate due to unclear compliance pathways.

Contradiction 2: Systemic Amplification of Medical Risks in Intelligent Diagnosis

AI, particularly large model technology, is increasingly intervening in critical clinical stages such as diagnostic support and medication recommendations. Due to the inherent uncertainty and "hallucination" characteristics of its outputs, errors are no longer merely technical deviations; they can directly influence clinical decisions and lead to irreversible consequences. Without institutionalized constraints and validation mechanisms tailored to these risks, patient safety faces new systemic challenges.

Contradiction 3: Blurred Liability Boundaries Impacting the Existing Medical Professional Order

In AI-assisted diagnosis and treatment scenarios, the rights and responsibilities among physicians, medical institutions, and technology providers have yet to be clarified. Once a diagnostic deviation occurs, the existing liability system struggles to address novel questions such as "should the AI's suggestion have been adopted?" or "how to define the division of responsibility between human and machine." This may lead to an improper downward shift of liability onto frontline medical staff, affecting their

[①] World Health Organization (2021), Ethics and Governance of Artificial Intelligence for Health: WHO Guidance, Geneva: World Health Organization.

professional judgment and occupational safety.

Contradiction 4: Unclear Compliance Pathways Restricting Standardized Technological Innovation

The current regulatory framework has not yet clearly defined the product attributes, market access methods, or compliance boundaries for the medical application of large models. This leaves innovators facing high levels of uncertainty during R&D and application. Such institutional ambiguity may amplify potential risks and hinder the guidance of technology toward a safe, controllable, and sustainable direction.

Contradiction 5: Global Innovation Synergy vs. Regional Regulatory Divergence

The R&D and application of AI in healthcare are essentially global, driven by top-tier research teams, massive data resources, and multinational clinical trials. However, at the regulatory level, different countries and regions exhibit significant discrepancies. This "Tower of Babel" phenomenon in regulation means that an AI medical product entering the global market must navigate vastly different compliance requirements, drastically increasing R&D and compliance costs and potentially widening the "Global AI Healthcare Divide."

The purpose of this white paper is to confront these challenges directly by systematically constructing a governance framework that both stimulates innovative vitality and reinforces safety bottom lines. To this end, the following sections will first review the pioneering experiences of major global economies to draw wisdom; then, we will propose a life-cycle governance methodology for AI medical models "from data to application" to ensure risks remain controllable. Finally, we will focus on building a core rule system that is both forward-looking and actionable, aiming to provide a clear guide for all stakeholders.

Chapter 2: AI Governance Dynamics in Developed Economies

The challenges and opportunities of AI in healthcare are global in nature. Governments and international organizations are actively exploring governance paths tailored to their specific national conditions. Due to significant differences in healthcare systems, technological development levels, legal traditions, cultural values, and market environments, global AI healthcare governance presents a diversified and multi-layered dynamic landscape. This chapter selects four developed economies—the United States, the United Kingdom, the European Union, and Singapore—to conduct an in-depth analysis of their core philosophies, key measures, unique strengths, and challenges in AI healthcare governance. It extracts common global trends such as "prudent inclusiveness," "risk classification," and "life-cycle management", while proposing a comprehensive framework of evaluation indicators.

2.1 The United States: Market-Driven and Innovation-First

As the global leader in technological innovation and the source of AI development, the United States holds a dominant position in AI healthcare. It has adopted a flexible "soft law" approach primarily based on administrative guidance, with the Food and Drug Administration (FDA) at its regulatory core. Its governance model profoundly reflects the values of "market-driven growth and innovation priority." Meanwhile, following several medical technology safety incidents, its regulatory system has matured, emphasizing a delicate balance between encouraging innovation and controlling risks.

2.1.1 Core Philosophy and Governance Framework

The core philosophy of U.S. AI healthcare (or digital health) regulation can be summarized as: building upon existing regulatory frameworks while using adaptive adjustments and innovative tools to achieve precise and agile oversight, while vigorously promoting industry self-regulation and multi-stakeholder governance.

The U.S. government recognizes that regulatory agencies alone cannot cope with the complexity and rapid iteration of AI. Consequently, it strongly advocates for a governance model involving industry self-discipline and multi-party participation. U.S. AI healthcare governance is not governed by a single, unified "AI Act," but rather exhibits "decentralized" and "multi-layered" characteristics.

Regarding the division of labor among agencies within the regulatory system:

- The White House National AI Initiative Office (NAIIO): Serves as the general coordinating body.

- Department of Health and Human Services (HHS): Acts as the lead department, responsible for the life-cycle regulation of medical AI, development strategies, reimbursement (payment), and compliance guidelines.
- Office of Science and Technology Policy (OSTP): Responsible for launching the National AI R&D Strategic Plan and coordinating federal medical AI policies and timelines.
- Federal Trade Commission (FTC): Focuses on anti-deception, anti-discrimination, and preventing privacy data leaks.
- National Institute of Standards and Technology (NIST): Leads the formulation of trustworthy AI technical standards and evaluation metrics.
- National Institutes of Health (NIH): Invests heavily in foundational research, dataset construction, and talent cultivation through specialized programs like "Bridge2AI," driving high-quality, responsible AI innovation from the source.

The U.S. Department of Health and Human Services (HHS) introduced five core principles for healthcare AI governance in 2023, known as the FAVES Principles. These require that AI-generated results must be:[①]

1. **Fair:** Model outputs must not exhibit bias or favoritism based on inherent or acquired characteristics such as race, gender, age, or socioeconomic status.
2. **Appropriate:** The model and its outputs must be highly compatible with the target use cases and populations, ensuring applicability within specific clinical or public health contexts.
3. **Valid:** The model must be proven to accurately estimate intended values across both internal and external datasets, backed by reproducible and verifiable scientific evidence.
4. **Effective:** In real-world deployment, AI tools must produce measurable positive benefits for diagnosis, care, or administrative processes, such as improving diagnostic accuracy, reducing costs, or enhancing patient experience.
5. **Safe:** Model results must not present known and unacceptable risks, and their expected benefits must outweigh potential harms; meanwhile, risk monitoring and mitigation mechanisms must be established to ensure the safety of patients and users.

The FAVES principles have been codified into federal rules (such as ONC's HTI-1 Final Rule), becoming core requirements for AI modules within certified health IT (including

① https://www.healthit.gov/sites/default/files/2023-12/Health_Sector_AI_Commitments_FINAL_120923.pdf

Electronic Health Records). They have also been adopted by numerous hospitals, insurers, and AI companies as "guardrails" for internal AI governance.

In January 2025, HHS further released the Strategic Plan for the Use of Artificial Intelligence in Health, Human Services, and Public Health. [①]As an action blueprint, it encompasses four major strategic goals:

- Catalyze Health AI Innovation: Rapidly incubate new applications that improve patient outcomes through modernized data infrastructure, public-private collaboration, and clear regulatory pathways.
- Promote Trustworthy AI: Establish "Federal Standards for Trustworthy AI Use" to continuously evaluate and mitigate risks such as algorithmic bias, privacy leaks, and biosafety threats.
- Democratize AI Technology and Resources: Develop open-source, customizable, and user-friendly AI tools, providing funding and technical support to community organizations and local health departments to bridge the "AI Divide."
- Nurture an AI-Enabled Workforce: Establish a federal-academic-industrial joint training system and create the role of Chief AI Officer (CAIO) to alleviate healthcare personnel shortages and burnout while fostering a culture of continuous learning.

Regarding risk governance and compliance, HHS explicitly stated it would synchronize updates to HIPAA, requiring privacy impact assessments and patient informed consent before AI processes sensitive health data. It also introduced external red-teaming audits, algorithmic transparency reports, and continuous monitoring to ensure systems are "explainable, traceable, and reversible." Social equity is now integrated into life-cycle reviews, specifically monitoring differential impacts on ethnic minorities, rural populations, and vulnerable groups.

The Center for Devices and Radiological Health (CDRH) under the FDA remains the core body for AI medical device regulation, housing a specialized Digital Health Center of Excellence to review AI/ML algorithms. The FDA not only handles approvals but also actively publishes guidance, organizes workshops, and collaborates with academia and industry to advance regulatory science.

The Office of the National Coordinator for Health Information Technology (ONC) focuses on the promotion and standardization of health IT. Its 21st Century Cures Act final rule mandates that medical institutions open all electronic health information to patients and prohibits "Information Blocking." This serves as critical infrastructure for

[①] U.S. Department of Health and Human Services. (2025). Strategic Plan for the Use of Artificial Intelligence in Health, Human Services, and Public Health (January 2025 version).

AI systems to access high-quality, standardized clinical data.

As the largest payer, the Centers for Medicare & Medicaid Services (CMS) wields decisive influence through reimbursement policies. Currently, CMS remains cautious, requiring substantial evidence of improved outcomes or cost reduction before reimbursing AI services. How CMS designs innovative payment models to incentivize high-value AI applications will be a key area to watch in U.S. governance.

Additionally, industry alliances and non-profits, such as the Coalition for Health AI (CHAI) and the Mayo Clinic Platform, play vital roles in developing evaluation standards, sharing data resources, and establishing best practices for industry self-regulation.

2.1.2 Relevant AI Healthcare Laws

At the federal level, several regulations have been issued, including the Health Insurance Portability and Accountability Act (HIPAA) (revised in 2021) and others (see Table 2-1 for details).

Table 2-1: U.S. Federal AI Healthcare Regulations (Until September 2025)

#	Title / Reference	Date	Key Points
1	Section 520(o) of the 21st Century Cures Act (P.L. 114-255)	2017-12-13	Excludes specific Clinical Decision Support (CDS) software from the definition of a medical device; defines regulatory boundaries for AI-CDS.
2	HIPAA Security & Privacy Rules (45 CFR Part 160/162/164)	2021-01-01*	Mandates "minimum necessary" access, encryption, auditing, and breach notification for AI handling Protected Health Information (PHI).
3	HITECH Act 2021 Amendment (HIPAA Safe Harbor, P.L. 116-321)	2021-01-01	Offers potential leniency in enforcement for entities following recognized security practices; impacts AI cloud storage and security design.
4	Section 1557 Nondiscrimination Final Rule (86 FR 27984)	2021-08-04	Prohibits federal financial assistance recipients from using AI decision-making tools that result in discrimination.
5	FTC Health Breach Notification Rule (16 CFR Part 318)	2021-09-15**	Requires non-HIPAA entities (health apps, AI wearables) to notify users and the FTC within 60 days of a data breach.
6	FDA AI/ML-SaMD Action Plan (2025 Update)	2021-01-12	Introduces the PCCP (Predetermined Change Control Plan); allows AI devices to self-update within pre-approved bounds without a new 510(k).
7	HTI-1 & HTI-2 Final Rules (88	2024-	Mandates certified EHRs and Health IT to

	FR 27956 / 89 FR 101773)	01-22	disclose AI-CDS algorithm sources, training data characteristics, and performance metrics (Transparency).
8	CMS 2024 IPPS Final Rule (88 FR 58962)	2024-10-01	Formally includes AI stroke software under NTAP payments; requires hospitals to maintain public clinical validation documentation for AI.
9	CMS 2025 MPFS Final Rule (89 FR 79224)	2025-01-01	Establishes independent CPT codes for billing AI-DSA (Digital Subtraction Angiography); requires submission of Real-World Evidence (RWE).
10	OCR-ONC Joint Final Rule: Telehealth & AI Discrimination (89 FR 34567)	2025-04-26	Clarifies that AI-driven care decisions resulting in discrimination violate ACA Section 1557; penalties include fines and suspension of federal funding.

Notes:

* Refers to the effective date of the HIPAA Safe Harbor law.

** Refers to the effective date of the 2020/2021 updated version.

Source: Compiled and updated from the *Bipartisan Policy Center (2023)*.[①]

In 2017, the FDA launched the Software Pre-Certification (Pre-Cert) Pilot Program, attempting to shift its regulatory focus from "reviewing products" to "reviewing companies." Under this model, developers who demonstrate "excellence" in their quality management systems, culture, and processes can obtain "Pre-Cert" status. This allows their lower-risk AI products to enjoy faster review pathways or streamlined data requirements upon market entry. The underlying logic is that a trustworthy developer is more likely to produce safe and effective products.

To accommodate the "adaptive learning" and "continuous evolution" characteristics of AI, the FDA implemented a "Proposed Regulatory Framework for Modifications to AI/ML-based SaMD" in April 2019. Under this framework, developers are responsible for the real-world performance of their AI systems and must provide updates to the FDA regarding changes in performance and inputs. This signifies a shift in the regulatory paradigm from "one-time market access" to "continuous oversight." In January 2021, the FDA released the "AI/ML-Based Software as a Medical Device (SaMD) Action Plan," which includes fostering consensus on Good Machine Learning Practice (GMLP), promoting Real-World Performance (RWP) monitoring, and updating regulatory frameworks to adapt to AI's autonomous modification capabilities.

[①] Bipartisan Policy Center. (2023). *Current regulatory landscape of AI in public health & health care: A brief overview*. https://bipartisanpolicy.org/report/current-regulatory-landscape-of-ai-in-public-health-health-care/

The Predetermined Change Control Plan (PCCP) is one of the most innovative regulatory tools introduced following the conclusion of the Pre-Cert pilot in 2022. Drawing from pilot experiences, the PCCP requires developers to submit a detailed plan before market authorization, explaining how the AI algorithm will autonomously learn and modify (e.g., adapting to new data sources or optimizing parameters) and how these modifications will be validated and controlled. Once approved, the AI system can perform autonomous updates within the predefined scope without the need for a new 510(k) submission for every minor change. This significantly accelerates iterative optimization while keeping risks within predictable bounds.

The FDA's ISTAND (Innovative Science and Technology Approaches for New Drugs) program is a hallmark of its support for AI in drug development. This program provides qualification for novel drug development tools, including AI algorithms. If an AI algorithm receives ISTAND qualification, its scientific validity and reliability are recognized by the FDA, allowing it to be used in specific stages of drug development (such as serving as a surrogate endpoint in clinical trials), thereby greatly accelerating the R&D process. This practice demonstrates that the FDA is valuing AI not just as a "product," but as a certified "tool."

2.1.3 FDA Classification and Management for AI-Enabled Medical Devices

Under the traditional medical device classification system, AI-driven devices generally follow established pathways (Class I/II/III via 510(k), De Novo, or PMA). However, the FDA has made specific adjustments to address AI/ML characteristics.

1. Evolution of Guidance Documents and Systems

The FDA maintains a dedicated "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices" page on its official website to outline its regulatory philosophy and resources.[①]

Following its 2019 discussion paper on software modifications[②], the FDA released a draft guidance in 2025 titled "Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions." This document provides recommendations on how AI-driven software functions should submit data, manage modifications, and monitor performance throughout their lifecycle.[③] Additionally, the FDA co-developed the Good Machine Learning Practice (GMLP) guiding principles to steer AI/ML medical

① U.S. Food and Drug Administration.[2025-10-12] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device

② https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device?utm_source=chatgpt.com

③ https://www.fda.gov/regulatory-information/search-fda-guidance-documents/artificial-intelligence-enabled-device-software-functions-lifecycle-management-and-marketing?utm_source=chatgpt.com software in model development, validation, and maintenance.[①]

2. Classification Criteria and Considerations for AI-Enabled Devices

In practical approval and classification determinations, the FDA typically considers several factors to decide the risk category, pre-market pathway, and data requirements for AI-enabled software (Table 2-2).

Based on these factors, the FDA conducts risk assessments for AI-driven software functions during the specific device approval process to determine their classification and required review pathway.

The FDA categorizes most medical AI applications as "Medical Devices" and applies varying degrees of regulatory intensity based on their risk level (Class I, II, or III). For "locked" AI products (where the algorithm remains static after approval), the FDA typically follows traditional medical device pathways, such as the 510(k) pre-market notification or the PMA (Pre-market Approval) program.

For high-risk AI-aided diagnostic software, the FDA generally requires pre-market approval or 510(k) clearance, supported by robust evidence of clinical effectiveness and safety. This "extended regulatory" model, built upon existing frameworks, ensures regulatory continuity and stability, avoiding the uncertainties that entirely new legislation might introduce.

Table 2-2: FDA Regulatory Consideration Factors for AI-Enabled Devices

#	Consideration Factor	Description / Impact
1	Intended Use	The claimed purpose of the AI software (diagnosis, prediction, monitoring, decision support, alerts, etc.) directly determines its risk level. AI used for final diagnosis or treatment recommendations is subject to stricter regulation than that used for prompts or auxiliary support.
2	Consequences of Erroneous Output	If an algorithm produces incorrect outputs that could lead to misdiagnosis, delays, injury, or incorrect treatment, the risk level is likely to be higher.
3	Model Stability / Adaptability	A key distinction is whether the AI/ML model is Static (unchanged after training) or Dynamic (continues to learn/update after deployment). Adaptive models introduce risks of "model drift" and uncertainty, requiring specific design controls, validation, and change management.
4	Training/Validation Data &	The representativeness of the training data, coverage of sufficient sample sizes, and generalization across diverse populations (gender,

① https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles?utm_source=chatgpt.com

	Generalization	age, ethnicity, device variations, etc.) are vital dimensions for assessing safety and effectiveness.
5	Bias / Fairness Risks	AI systems may exhibit bias against certain subgroups, leading to poor performance for specific populations (e.g., ethnic minorities). FDA guidance emphasizes designing strategies to detect and mitigate such biases.
6	Transparency / Explainability / Traceability	The FDA encourages describing model inputs, outputs, decision logic, limitations, and boundary conditions during submission to help users understand its behavior. Transparency aids regulatory assessment, clinical adoption, and safety monitoring.
7	Lifecycle Management	AI performance may degrade over time due to model drift or changes in input data distribution. A clear post-market performance monitoring plan, change control mechanism, and re-validation strategy are required. The FDA emphasizes risk management throughout the entire product lifecycle.
8	Safety / Verification & Validation	AI software must meet standard V&V requirements for Software as a Medical Device (SaMD), while also accounting for cybersecurity, software vulnerabilities, and potential adversarial attack risks.
9	Change / Upgrade Strategy	For future algorithm upgrades or modifications, the FDA requires a change control mechanism to determine if a new submission is necessary. The Predetermined Change Control Plan (PCCP) allows manufacturers to pre-define the scope and mechanisms for such changes during the initial submission.

In terms of international cooperation, the medical device regulatory authorities of the United States, the United Kingdom, and Canada maintain close collaboration. In 2021, the three nations jointly released the "Good Machine Learning Practice (GMLP) for Medical Device Development: Guiding Principles." This was the first multi-national initiative to propose practical principles for AI/ML medical device development, providing the industry with an ethical and safety baseline, a quality management framework spanning design to deployment, and a foundation for international regulatory harmonization and standardization.[①]

In October 2023, the three countries released the "Predetermined Change Control Plans for Machine Learning-Enabled Medical Devices: Guiding Principles," proposing five core guiding principles: Focused and Bounded, Risk-based, Evidence-Based, Transparent, and a Total Product Lifecycle (TPLC) Perspective.[②]

① U.S. Food and Drug Administration, Health Canada, MHRA. Good machine learning practice for medical device development: guiding principles[EB/OL]. 2021. https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development.

② U.S. Food and Drug Administration, Health Canada, MHRA. Predetermined change control plans for machine learning-enabled medical devices: guiding principles[EB/OL]. 2023. https://www.fda.gov/medical-

In 2024, the regulators further introduced "Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles"[①]. This represents the first systematic guidance on transparency for MLMDs, emphasizing that transparency is not merely information disclosure but a core mechanism for ensuring patient safety, promoting health equity, and building public trust. Its principles are shown in the table below:

Table 2-3: Six Guiding Principles for Transparency of MLMD (US, UK, Canada)

Principle	Description	Key Guidance Points
Who: Relevant Audience	Transparency is aimed at all parties involved in patient care, including: those who use or receive device services (e.g., physicians, patients, caregivers); and those involved in device decision-making (e.g., administrators, payers).	• Identify the specific information needs of different audiences. • Consider their technical backgrounds and usage contexts.
Why: Rationale for Transparency	Transparency helps to: ensure safe and effective device use; achieve patient-centered care; identify and assess risks vs. benefits; support informed decision-making and risk management; facilitate maintenance and performance monitoring; promote health equity; and enhance user trust and proficiency.	• Treat transparency as a core tool for risk management and patient safety. • Use transparency to promote technical acceptance and equitable use.
What: Key Information	Essential disclosures include: medical purpose and functionality; target diseases or conditions; intended users and environments; the device's role in the clinical workflow (inputs, outputs, impact on decisions); performance, benefits, and risks; training/testing data characteristics; clinical validation summaries; model logic (where applicable); known biases, failure modes, and confidence intervals; data gaps and limitations; and lifecycle management/update strategies.	• Information must be accurate, complete, and easy to understand. • Emphasize known limitations and uncertainties.
Where: Information Placement	Information should be conveyed via user interfaces (UI), labeling, training materials, software dashboards, and documentation. Software interfaces should be optimized to be responsive, personalized, and adaptive.	• Utilize diverse media (text, graphics, audio, video). • Ensure information is easily accessible and

devices/software-medical-device-samd/predetermined-change-control-plans-ml-enabled-medical-devices.

① U.S. Food and Drug Administration, Health Canada, MHRA. Transparency for machine learning-enabled medical devices: guiding principles[EB/OL]. 2024. https://www.fda.gov/medical-devices/software-medical-device-samd/transparency-machine-learning-enabled-medical-devices

		embedded within the user workflow.
When: Timing of Communication	Disclosure should consider key points across the entire product lifecycle, such as: prior to procurement or implementation; during first-time or every-time use; when the device is updated or issues arise; and providing real-time prompts during high-risk steps or specific triggers.	• Provide timely notifications and dynamic updates. • Offer necessary reminders to users at critical decision-making points.
How: Methodology	Methods to support transparency include: adopting human-centered design (HCD) principles; providing level of detail appropriate for the audience; organizing content by importance or cognitive sequence; using plain language or technical terminology as appropriate; and continuously validating and optimizing communication effectiveness.	• Design should be iterative, participatory, and centered on user experience (UX). • Consider user diversity (culture, language, professional background).

2.1.4 Strengths and Challenges of the U.S. Governance Model

The U.S. governance model for AI healthcare presents several key advantages:

1. **Robust Innovation Vitality:** A market-driven model, coupled with abundant capital, top-tier talent, and flexible regulatory exploration, has fostered the world's most active AI healthcare innovation ecosystem.
2. **Rich Regulatory Experience:** The FDA possesses the world's most mature medical device regulatory system, and its adaptive adjustments for AI are highly professional and forward-looking.
3. **Deep Multi-stakeholder Participation:** Practices such as the NAM (National Academy of Medicine) guidelines demonstrate that U.S. society has formed a healthy atmosphere where government, industry, academia, and patients deeply participate in governance.

In terms of challenges, beyond the typical risks of data privacy, bias, and health inequality, there is the issue of regulatory fragmentation. The division of power between federal and state levels, as well as among different agencies (e.g., FDA, FTC, OCR), is sometimes unclear, leading to overlaps or gaps. Furthermore, on December 11, 2025, President Trump signed an Executive Order titled "Establishing a National Policy Framework for Artificial Intelligence." This order asserts that U.S. global leadership in AI is threatened by fragmented state-level regulations and mandates a unified, "minimum-burden" national standard to replace conflicting state laws. This signals a shift toward a more innovation-centric, technical, and "de-ethicized" regulatory environment.

2.2 The European Union: Human-Centric and Rule-Driven

Through its General Data Protection Regulation (GDPR) and the AI Act, the European Union has established the world's most influential and rigorous "rule-driven" model. The core of its governance philosophy is "Human-Centric," placing the protection of fundamental rights, safety, and democratic values at absolute priority over technological development, favoring a comprehensive "hard law" framework.

2.2.1 Core Philosophy and Governance Framework

The EU's core philosophy is to build a trustworthy AI environment by establishing unified, legally binding rules that ensure AI applications—especially in high-risk medical fields—are safe, transparent, fair, and accountable.

- **European Commission (EC):** The executive arm that proposes major legislation like the AI Act and the European Health Data Space (EHDS).
- **European AI Board:** A newly established body responsible for ensuring the consistent implementation of the AI Act.
- **European Medicines Agency (EMA):** While medical devices are primarily managed by member state notified bodies (via CE marking), the EMA plays an increasingly vital role in AI-driven drug development.
- **National Competent Authorities (NCAs):** Such as Germany's BfArM and France's ANSM, these bodies serve as direct regulators in their respective markets, performing specific approvals and market surveillance.

2.2.2 Regulatory Logic and Key Measures

The GDPR (2016) set an extremely high bar for data compliance, establishing principles like "data minimization" and the "right to object to automated decision-making."

To facilitate health data sharing while ensuring security, the EU launched the EHDS, which became effective in 2025.[①] It utilizes a "dual-track" concept:

1. **Primary Use:** Empowering individuals to control their data for cross-border healthcare services.
2. **Secondary Use:** Providing a trusted system for re-using health data for research, innovation (including training AI algorithms for medical devices and clinical decision support), and policy-making.

In terms of innovation, the EU AI Act (Article 57) mandates that all member states

[①] European Health Data Space Regulation (EHDS). [2025-10-25] https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en establish at least one national AI Regulatory Sandbox by August 2, 2026. Spain led the way, launching its operations in May 2025 with an initial batch of 12 projects. The EU provides infrastructure like TEFs (Testing and Experimentation Facilities), which received 220 million euros in investment to provide real-world data and supercomputing resources for developers.

2.2.3 Major Regulations for AI-Enabled Medical Devices in the EU

In the EU, AI-driven medical devices are subject to intersecting regulations, as shown in Table 2-4. The Medical Device Regulation (MDR (EU) 2017/745) provides a broad definition of medical devices that includes software designed for diagnosis, prevention, monitoring, and prediction of diseases.

Table 2-4: Key EU Regulations for AI-Enabled Medical Devices

Regulatory Framework	Role / Target Audience	Relationship with AI Medical Devices
Medical Device Regulation (MDR) (Regulation (EU) 2017/745)	Applies to medical devices (hardware and Software as a Medical Device).	If AI-driven software has a medical purpose, it is considered a medical device and must meet MDR requirements for safety, performance, clinical evaluation, surveillance, and lifecycle management.
In Vitro Diagnostic Regulation (IVDR) (Regulation (EU) 2017/746)	Applies to in vitro diagnostic (IVD) medical devices.	AI-driven software used for IVD purposes (e.g., pathology slides, genetic testing) falls under IVDR control.
Clinical Trials Regulation (CTR) (Regulation (EU) No 536/2014)*	All "clinical trials"—covering pharmaceuticals, biologics, Advanced Therapy Medicinal Products (ATMPs), and drug-device combinations.	Applies not only to traditional drugs but also to advanced therapies and combination products involving AI. Any AI medical device involved in clinical trials must register in CTIS and comply with CTR transparency and safety rules.
AI Act (Regulation (EU) 2024/1689)	Includes medical devices, diagnostic tools, and regulated clinical support tools, ranging from standalone software (SaMD) to complex integrated systems.	Effective since August 2024, it introduces risk classification and additional compliance obligations specifically for AI systems. AI systems for medical use are typically categorized as "High-Risk AI Systems."
Medical Device Coordination	While not an enforcement body, its guidelines are	For AI medical devices, MDCG guidelines clarify the cross-regulatory

Group (MDCG) Guidance	universally adopted by member state authorities and Notified Bodies as the de facto "Unified Enforcement Manual."	boundaries between MDR/IVDR and the AI Act. They provide transitional technical solutions before harmonized standards are finalized, serving as essential "soft law" for manufacturers.

Notes: * Applicable since 2022-01-31; fully covers all trials as of 2025-01-31.

If software has a medical purpose (such as for diagnosis, monitoring, prediction, or auxiliary treatment decision-making), it may be considered a medical device. The MDCG's "Guidance on Qualification and Classification of Software" (MDCG 2019-11 and its revised/expanded versions) specifies the conditions under which software qualifies as a device.① Recent revised guidance more explicitly incorporates AI-driven software, stating that AI modules must be classified according to their specific use cases.②

Under the MDR framework, software (whether embedded, cloud-based, or mobile) categorized as a medical device is subject to the classification principles and rules outlined in Annex IX.

- **Risk Classification:** If an AI module is responsible for diagnosis/treatment decisions or critical monitoring and poses a high risk to patients, the software may be classified as Class IIb or III. If it serves auxiliary, low-risk purposes, it may be categorized as Class I or IIa. AI software that provides recommendations for decision-making carries higher risk than that which merely displays information.
- **Infrastructure Risks:** AI software dependent on Electronic Health Records (EHR), networked services, cloud platforms, or external databases must undergo risk assessments regarding interoperability, interface risks, and data reliability.

The EU AI Act (2024) adopts a "risk-based" tiered regulatory approach, dividing AI applications into four levels: "Unacceptable Risk," "High Risk," "Limited Risk," and "Minimal Risk." AI applications in the medical field (e.g., diagnostic systems, surgical robots, risk assessment tools) are almost exclusively classified as "High-Risk AI." This imposes strict obligations throughout the AI lifecycle:

1. **Mandatory Conformity Assessment:** Prior to market entry, a rigorous assessment by a third-party (Notified Body) is required. Developers must

① https://health.ec.europa.eu/medical-devices-sector/new-regulations/guidance-mdcg-endorsed-documents-and-other-guidance_en?utm_source=chatgpt.com
② https://www.biosliceblog.com/2025/07/revised-guidance-on-classification-of-medical-device-software-in-the-eu/?utm_source=chatgpt.com

implement a continuous risk management system.

2. **High-Quality Datasets:** Training, validation, and testing datasets must be relevant, representative, and unbiased. Under Article 17, manufacturers must establish an AI Quality Management System (AI QMS) integrating documentation (Art. 18), compliance logs (Art. 20), and post-market surveillance (Art. 61).
3. **Technical Documentation:** Requirements are exhaustive, covering design specifications, system architecture, computational resources, and performance metrics (Annex IV).
4. **Transparency and Explainability:** Operations must be transparent enough for users to understand and monitor decisions, and explain the AI's purpose and limitations to affected individuals.
5. **Human Oversight:** Systems must be designed with measures allowing humans to intervene or override AI decisions at any time.
6. **Robustness and Cybersecurity:** Systems must remain resilient against varied conditions and potential cyberattacks.

2.2.4 Strengths and Challenges of the EU Governance Model

The EU emphasizes "Ex-ante approval" (pre-market approval). An AI imaging tool must obtain a CE mark by proving compliance with both the MDR and the AI Act. While rigorous, some critics argue this process is slow and costly.

Strengths:

- **Strong Protection of Citizens' Rights:** The "human-centric" philosophy provides robust legal safeguards for data privacy and fundamental rights.
- **High Legal Certainty:** Once fully implemented, the AI Act provides a unified, legally binding rulebook, reducing ambiguity for enterprises.
- **Global Impact:** Due to the "Brussels Effect," EU regulations often become de facto global standards.
- **Incentivizing "Trustworthy AI":** Encourages a competitive advantage based on high safety, ethical, and transparency standards.

Challenges:

- **High Compliance Costs:** The financial and human resource burden may stifle SMEs and startups.
- **Potential Innovation Lag:** Cumbersome regulatory processes may delay market entry compared to the US or UK.
- **Implementation Complexity:** Consistent enforcement across member states amidst rapid technological shifts poses a significant challenge to regulatory

capacity.

2.3 The United Kingdom: Value-Oriented, Pragmatic Exploration

The United Kingdom has forged a distinct path in AI healthcare governance. Its approach is characterized by strong government leadership, clear national strategies, and integrated, pragmatic exploration leveraging its unique National Health Service (NHS).

2.3.1 Core Philosophy and Governance Framework

The core philosophy of UK AI healthcare governance is to treat Artificial Intelligence as a "systemic capability" of the public health system. Under the unified guidance of national strategy, AI is deployed and regulated through the NHS—a single, authoritative public platform. Within a framework of strict data governance and ethical constraints, the primary goal is to enhance system efficiency, accessibility, and health outcomes, rather than focusing solely on market expansion or technology for its own sake.

In the *National AI Strategy* released in 2021, "Health and Life Sciences" was designated as one of the four key areas for AI application. This strategy emphasizes building an entire ecosystem—data infrastructure, talent, ethical standards, and a regulatory environment—to support innovation.

The UK governance structure involves several key departments:

1. **Medicines and Healthcare products Regulatory Agency (MHRA):** The primary regulator ensuring the safety, quality, and effectiveness of medicines and medical devices (including AI-MD). It sets legally binding regulations but often relies on third-party "Approved Bodies" for individual device assessments.
2. **National Institute for Health and Care Excellence (NICE):** Evaluates the clinical effectiveness and cost-benefit of medical technologies to provide procurement and reimbursement recommendations to the NHS.
3. **Care Quality Commission (CQC):** Oversees the service quality and safety of healthcare providers, including the clinical compliance of AI device usage.
4. **Health Research Authority (HRA):** Approves and regulates the ethics of clinical research, ensuring AI device trials protect participant rights.
5. **Regulatory Horizon Council (RHC):** A high-level think tank providing strategic advice to the government on regulating emerging technologies (like AI) to inform future legislation.

2.3.2 Regulatory Logic and Key Measures

The UK utilizes a multi-layered, collaborative system where multiple authorities

(MHRA, NICE, RHC) address clinical value, macro-policy, and core regulatory reform simultaneously.

1. Medicines and Healthcare products Regulatory Agency (MHRA)

- **ILAP (Innovative Licensing and Access Pathway):** Originally launched in 2021 to accelerate patient access to innovative medicines, this pathway was updated in **January 2025**. It now explicitly includes **"Drug-Device Combination Products"** and **"Emerging Digital Health Technologies."** It emphasizes early assessment of "NHS operational readiness" to bridge the gap between regulatory approval and clinical adoption [1].
- **Software and AI as a Medical Device Change Programme:** Launched in April 2024, this is a systematic reform project designed to clarify the qualification, classification, pre-market requirements, and post-market surveillance of AIaMD (AI as a Medical Device).

Table 2-5: Key Work Packages of the UK MHRA Future-Focused AI/ML Medical Device Regulatory Reform[①]

No.	Work Package	Focus Issues	Key Deliverables / Objectives
WP1	Qualification	Ambiguity in defining whether software qualifies as a Medical Device (SaMD).	Clarify SaMD scope; guidance on writing "Intended Use"; defining manufacturer identities.
WP2	Classification	Existing risk classification rules are unsuitable for modern software.	Update classification regulations; explore the "Airlock" mechanism; publish classification guidance.
WP3	Pre-market Requirements	Pre-market requirements are either inapplicable or unclear for software.	Revise essential requirements; publish development best practices; clarify clinical evidence requirements.
WP4	Post-market Surveillance	Weak signal detection post-market; lagging change management.	Optimize adverse event reporting; launch Predetermined Change Control Plans (PCCP); strengthen change management.
WP5	Cybersecurity for Medical Devices	Lack of cybersecurity requirements; high risk for legacy devices.	Cybersecurity legislation; vulnerability reporting mechanisms; guidance for legacy device management.
WP9	AI Rigour	Lack of a systemic regulatory framework for AI medical devices.	Publish GMLP (Good Machine Learning Practice) guidance; AI development best practices; bias

① Medicines and Healthcare products Regulatory Agency. Software and AI as a medical device change programme roadmap[EB/OL]. 2023. https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap.

			detection and mitigation tools.
WP10	Project Glass Box (AI Interpretability)	The "Black Box" nature of AI models impacts safety and usability.	Publish Human-Centered AI guidance; trustworthy AI frameworks; standards for interpretability.
WP11	Project Ship of Theseus (AI Adaptability)	Adaptive AI (e.g., continuous learning) is difficult to manage via traditional change control.	Categorized management of AI model updates; research methods for concept drift detection; develop AI-specific PCCP processes.

Note: Work Packages 6-8 (covering mobile health, Apps, and cross-cutting issues) have been integrated into other WPs and do not have independent deliverables.

This program advances through two parallel workstreams—**Software as a Medical Device (SaMD)** and **AI as a Medical Device (AIaMD)**—addressing AI-specific challenges such as model interpretability and adaptive algorithms.

In 2024, the MHRA launched AI Airlock, the world's first regulatory sandbox specifically for artificial intelligence medical devices. It brings together regulators, hospitals, experts, and enterprises to collaborate in simulating 4–6 real-world medical scenarios. Within this controlled environment, the safety, effectiveness, and compliance of AI products are tested to inform future AI medical device regulatory policies. As of August 2025, the "AI Airlock" sandbox has completed its first pilot round and launched a second round of recruitment. The experiences gained from this sandbox will be directly translated into the MHRA's upcoming specialized AIaMD guidance documents, establishing it as the UK's core regulatory tool for accelerating the safe introduction of AI medical devices.

Table 2-6: Comparison Between the Airlock Mechanism and Traditional Approval

Project	Traditional Approval	Airlock Mechanism
Evidence Source	Clinical Trials (Prospective)	Real-World Evidence (RWE)
Time to Market	After evidence is sufficient	Pilot possible even when evidence is insufficient
Risk Tolerance	Low	Medium (Controlled)
Regulatory Style	One-time approval	Continuous monitoring + Dynamic assessment
Applicable Products	Mature technologies	Innovative, difficult-to-validate products

Source: Medicines and Healthcare products Regulatory Agency(2023) [①]

① Medicines and Healthcare products Regulatory Agency. Software and AI as a medical device change programme roadmap[EB/OL]. 2023: WP2-02. https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap.

2. National Institute for Health and Care Excellence (NICE)

NICE plays a dual role in the UK healthcare system as both a "gatekeeper" and a "compass." Its primary responsibility is to evaluate the **clinical effectiveness** and **cost-effectiveness** of medical technologies—whether already on the market or upcoming—to provide procurement and reimbursement recommendations to the NHS. Put simply, NICE determines where NHS funds should be allocated.

This indicates that UK regulation of medical AI is closely tied to clinical practice and health economics from the outset. Even if an AI product is technologically advanced, if NICE deems it lacks significant patient benefit or carries a price tag that is too high (low value for money), it will not be procured or used on a large scale by the NHS. This "value-based" regulatory framework ensures that AI is not only safe but also truly valuable to patients and the healthcare system.

NICE's **"Evidence Standards Framework (ESF) for Digital Health Technologies"** establishes a unified evidence standard for various digital health technologies (e.g., apps, software, online platforms). Developers must collect and submit evidence according to these standards to prove their product's effectiveness and value. This framework addresses the early challenges of the digital health sector—such as inconsistent quality and mixed results—providing a clear yardstick for developers, the NHS, and patients to ensure that AI products entering the system are validated and genuinely beneficial.

3. Regulatory Horizons Council (RHC)

As a high-level think tank, the RHC provides strategic advice to the UK government on the regulation of frontier technologies like AI, regularly publishing influential reports. In November 2022, the RHC released *"The Regulation of Artificial Intelligence as a Medical Device,"* proposing several policy recommendations that have since been implemented into subsequent regulations (see Table 2-7). This work has shaped the UK's policy direction, providing the "intellectual ammunition" for legislative and policy adjustments.

Table 2-7: Implementation Status of RHC 2022 Policy Recommendations (as of 2025)

Recommendation	Implementation Status as of 2025
Rec 1: Strengthen MHRA's long-term funding and capacity	Government "accepted in principle," providing specific funding through the AI Airlock project to support MHRA's AI device oversight; AI regulatory team capacity continues to expand.
Rec 2: Enhance training for regulators and evaluators	MHRA published guidance through the "Software & AI as a Medical Device Change Programme" to upskill both the regulator and manufacturers.

Rec 3: Build a "legislatively light, guidance heavy" framework	This philosophy continues into 2025, prioritizing guidelines, standards, and sandbox mechanisms over rigid mandatory legislation to adapt to rapid AI evolution.
Rec 4: Require assessments of AI bias and generalization	Technical evidence for bias mitigation and generalization validation is now explicitly required within the AI Airlock and Change Programme.
Rec 5: Require interpretability for AI devices	The "Project Glass Box" workstream within the Change Programme is driving the development of AI interpretability guidance, though it is not yet a mandatory regulatory requirement.
Rec 6: Support innovative access mechanisms	The "AI Airlock" regulatory sandbox has been implemented, allowing AI devices to generate real-world evidence and accelerate market entry.
Rec 7: Validation before local deployment	While not a mandatory law, manufacturers are encouraged through the AI Airlock and PMS guidance to collaborate with healthcare institutions on local validation and "silent trials."
Rec 8: Establish proactive, systemic post-market surveillance (PMS)	The PMS Regulations were fully implemented in June 2025, requiring manufacturers to establish PMS plans, trend reporting, and rapid incident notification.
Rec 9: Support model update mechanisms (PCCP)	The Predetermined Change Control Plan (PCCP) has been advanced within the AI Airlock and Change Programme, allowing AI models to update within preset boundaries.
Rec 10: Develop device exit and contingency plans (Plan B)	Manufacturers are required to develop risk mitigation and corrective actions under PMS regulations, though a mandatory "Plan B" exit mechanism is not yet explicitly enforced.
Rec 11: Harmonize regulatory language and pathways (MAAS)	MHRA continues to collaborate with NICE and CQC to promote unified guidance and terminology, though the "single point of entry" mechanism is not yet fully realized.
Rec 12: Establish a public registry for AI devices	MHRA has established the Public Access Registration Database (PARD), allowing the public to search for device risk classes, manufacturers, and other key info.
Rec 13: Promote Patient and Public Involvement (PPI)	A specific workstream in the Change Programme encourages incorporating patient and public perspectives into the design, evaluation, and regulation of AI devices.
Rec 14: Drive international coordination and standardization	The 2025 policy explicitly introduced the "International Reliance Route," recognizing approval results from regulators such as the US FDA, Health Canada, and Australia's TGA.
Rec 15: Push for mutual recognition with the EU/US	Currently, only a **"Unilateral Recognition Route"** is in place, where the UK recognizes FDA and other results, but reciprocal recognition has not yet been established.

In addition, the UK established the **NHS AI Lab** in 2019, supported by £250 million in government funding. Its mission is to serve as the "engine" and "hub" for AI innovation within the NHS, responsible for identifying and evaluating the most promising AI technologies, funding their piloting and scaling across the NHS, and resolving barriers to clinical integration. Its **"AI in Health and Care Award"** program has funded dozens of AI projects for real-world testing in NHS primary care, hospitals, and communities, covering diverse fields ranging from mental health screening to early cancer diagnosis.

2.3.3 Strengths and Challenges of the UK Governance Model

The UK possesses several significant advantages in AI healthcare governance:

1. **Clear Strategy and Strong Execution:** Leadership from national strategies combined with the unified NHS system ensures high consistency and powerful execution in the UK's AI healthcare governance.
2. **Unique Data and Application Scenarios:** The NHS provides standardized health data covering the entire population and offers a massive application landscape as the world's largest single-payer healthcare system.
3. **Explicit Value Orientation:** NICE ensures that only AI technologies proven to be "value for money" are promoted. This prevents resource waste and guides innovation toward solving practical clinical problems.
4. **Comprehensive Ethical and Safety Frameworks:** From national strategy to internal NHS guidelines, the UK has built a robust ethical and safety governance framework, providing a safeguard for the responsible application of AI.

However, challenges remain. As a massive, long-standing public system, the NHS may face issues such as **bureaucracy and rigid processes** during technological integration. Furthermore, the NHS is under chronic pressure from **funding and staffing shortages**, which may hinder its speed and capacity to adopt and deploy new technologies. Additionally, while the state-led model is effective for scaling applications, it can sometimes **stifle market vitality and disruptive innovation**. Attracting and retaining top AI startups—and helping them successfully commercialize their innovations—remains a critical question for the UK.

2.4 Singapore: National Vision, Agile Governance

Singapore is a global leader in the construction of a "Smart Nation." Its AI healthcare governance is unfolded under the ambitious **"Smart Nation 2025"** strategic blueprint, characterized by strong government leadership, a focus on pragmatism, an emphasis on agility, and the pursuit of international cooperation. Singapore's model aims to create a "regulatory sandbox" environment friendly to the rapid development of AI while ensuring alignment with national interests and public safety.

2.4.1 Core Philosophy and Governance Framework

The core philosophy is guided by the "Smart Nation" strategy: the government plays a leading role through forward-looking policy planning, active investment, world-class data infrastructure, and agile tools like regulatory sandboxes to rapidly and safely apply AI to key areas like healthcare, enhancing national competitiveness and public well-being.

1. **Strategic Planning:** The **National AI Strategy** (updated in 2019) outlines a blueprint for AI in healthcare to address aging populations and chronic disease management.
2. **Key Agencies:**
    - Smart Nation and Digital Government Group (SNDGG): Coordinates cross-sector digital policies.
    - **Ministry of Health (MOH) & Health Sciences Authority (HSA):** Manage healthcare policy and market access regulation. HSA actively aligns with international standards (FDA, CE) while exploring flexible pathways.
    - **AI Singapore (AISG):** A national program funded by the National Research Foundation to anchor AI talent, research, and industrial transformation.
    - Integrated Health Information Systems (IHiS): Operates the National Electronic Health Record (NEHR) system, managing the secure sharing of national health data.

2.4.2 Regulatory Logic and Key Measures

Singapore has invested heavily in the **NEHR**, integrating patient data from all public (and increasingly private) institutions to provide a high-quality data source for AI research under strict privacy controls.

Legal Basis and Guidance:

The HSA regulates AI medical devices under “the Health Products Act 2007” and “Health Products (Medical Devices) Regulations”(2010).

1. **Change Management Programme (CMP):** Implemented in December 2024 (GN-37-R1), this allows manufacturers to plan **"pre-specified changes."** This means certain updates within an approved scope can be made during the lifecycle without submitting a full new change application every time.
2. **New Guidelines (July 2025):** The revised "Guidance on Risk Classification of SaMD and Qualification of CDSS" clarified the criteria for Clinical Decision Support Systems (CDSS) and mandated lifecycle management for AI/ML-

driven software, covering design, verification, validation, and performance monitoring.

Agile Innovation:

1. **Regulatory Sandboxes:** HSA allows innovative AI products to be tested in real-world clinical settings with temporary regulatory exemptions and close supervision. This "learning by doing" approach helps regulators understand new technologies while letting companies validate innovations with lower risk.
2. **AI Verify:** Launched in May 2022 by IMDA and PDPC, this is the world's first government-led AI governance testing framework and software toolkit. It covers 11 international principles (e.g., transparency, fairness, robustness) and aligns with the EU AI Act and OECD principles. It allows companies to perform self-verification and generate standardized reports to prove their AI is "responsible and trustworthy," facilitating a "test once, use anywhere" approach.

2.4.3 Strengths and Challenges of the Singapore Model

Strengths:

1. **Strong Leadership & High Efficiency:** The centralized administrative system allows for rapid decision-making and unified execution of national strategies.
2. **World-Class Data Infrastructure:** The unified NEHR and a culture of open (yet secure) data sharing provide an unparalleled data advantage for AI development.
3. **Agile Governance:** Tools like regulatory sandboxes attract international tech giants and startups by providing a flexible testing environment.
4. **Global Vision:** Active international cooperation helps Singapore absorb global best practices and influence global AI governance.

Challenges:

1. **Privacy vs. Security:** Balancing widespread data sharing for AI development with the protection of citizen privacy and national security remains a persistent challenge.
2. **Private Sector Vitality:** A strong government-led model can sometimes lead to an over-reliance on the state. Singapore must continue to stimulate independent innovation within the private sector.
3. **Talent Competition:** As a small nation, Singapore faces immense pressure from global powers like the US and China in attracting and retaining top-tier AI talent.

2.5 Evaluation Index System for AI Healthcare Governance

Based on an in-depth analysis of AI healthcare governance models across major global economies, this section seeks to construct a scientific, objective, and practically relevant evaluation framework to systematically measure the effectiveness of governance. Recent pioneering efforts in China, such as the local standard released by Shenzhen, **"Comprehensive Evaluation Index System for Medical Artificial Intelligence Governance" (DB4403/T 634—2025)**, have provided valuable conceptual references. Building upon these practical insights, we have developed an evaluation index system that is logically rigorous, clearly layered, and balances international vision with local characteristics.

Our proposed evaluation framework consists of **four primary indices** and **11 secondary indices**, characterized by two distinct features:

1. **Logical Closed-Loop:** The four dimensions of **"System - Capability - Efficacy - Improvement"** are interlinked, viewing governance as a dynamic, evolving, and continuously optimizing lifecycle rather than a static collection of regulations.
2. **Value-Oriented:** Particularly within the "Continuous Improvement" dimension, we have introduced core evaluation methodologies from **health economics**. By utilizing "Efficiency" and "Health Outcomes" as critical feedback signals to drive governance upgrades, we ensure that the development of AI technology ultimately serves the fundamental goals of enhancing healthcare system efficiency and improving public health and well-being.

Table 2-8: AI Healthcare Governance Level Evaluation Index System

Dimensions	Index	Tertiary Index (Core Evaluation Points)
1. Governance System	1.1 Institutional Norms	• Laws, Regulations & Ethical Guidelines: Assess the completeness of laws regarding AI medical data privacy, algorithmic liability, and product market access. • Safety & Risk Definition: Examine how institutional frameworks define and cover medical safety, privacy security, and ethical risks.
	1.2 Organizational Management	• Organizational Structure & Synergy: Evaluate the effectiveness of cross-departmental coordination mechanisms and multi-stakeholder governance models.
	1.3 Technical Standards	• Data & Algorithm Standards: Assess technical requirements for data quality, model testing, algorithmic transparency, and explainability. • Risk Prevention Standards: Examine standardized requirements for identifying and mitigating algorithmic risks (e.g., bias, discrimination) and training data risks.
2. Governance	2.1 Technical Support	• Infrastructure & Tools: Evaluate the application level of computing power infrastructure, public data platforms, and

Capability		trustworthy AI technologies. • Risk Handling Capability: Examine the ability to identify, warn, and emergency-respond to risks in generated content (specifically for Generative AI).
	2.2 Talent Assurance	• Interdisciplinary Talent Pool: Assess the scale and cultivation of talent possessing cross-domain knowledge in AI, medicine, and law. • Public Awareness & Communication: Examine public trust in AI healthcare and the ability of governing bodies to communicate and guide user experience and social safety risks.
	2.3 Execution Capability	• Regulatory Execution Efficiency: Evaluate the clarity and efficiency of product approval processes and the effectiveness of post-market surveillance.
3. Governance Efficacy	3.1 Application & Promotion	• Clinical Penetration & Integration: Assess the breadth and depth of AI penetration in core clinical scenarios (e.g., imaging, pathology). • Real-World Performance: Measure the effectiveness (e.g., accuracy, recall) and stability of models in real clinical environments. • User Satisfaction: Evaluate the acceptance and satisfaction levels of medical staff and patients through user experience surveys.
	3.2 Industry Promotion	• Economic Value Creation: Assess the economic benefits brought by the AI healthcare industry, such as market size and industrial chain ripple effects.
	3.3 Safety Assurance	• Social Risk Control: Monitor the actual occurrence of social and ethical risks and evaluate the effectiveness of control measures.
4. Continuous Improvement	4.1 Efficiency Evaluation	• Cost-Benefit Analysis: Introduce efficiency metrics such as cost efficiency, scale efficiency, and allocative efficiency to assess the economic rationality of AI applications and guide resource optimization.
	4.2 Value Evaluation	• Health Outcome Measurement: Introduce benefit metrics such as health benefits and social benefits to measure the ultimate contribution of AI to national health levels and public health, driving further technical iteration and policy adjustment.

Chapter 3: Development of AI Healthcare Governance in China

In the global wave of AI healthcare, China has become an indispensable key player, characterized by its massive market scale, rich medical data resources, strong government propulsion, and a rapidly developing AI industry. China's AI healthcare governance presents features significantly different from those of Western nations: driven by powerful top-level national strategies and dual goals of "promoting development" and "ensuring safety," it has rapidly constructed a governance system with Chinese characteristics through a path of "policy first, pilot exploration, and gradual standardization." This chapter provides an in-depth analysis of the strategic background, policy and regulatory systems, regulatory practices, achievements, and deep-seated challenges facing China's AI healthcare governance.

3.1 Top-Level Drive: AI Healthcare as a National Strategy

Unlike the spontaneous market evolution of the United States or the rights-driven model of the European Union, China's AI healthcare development has carried a strong "national strategy" flavor from the outset. The government acts not only as a regulator but also as an active promoter, planner, and investor. This top-down design provides the fundamental momentum and direction for the rapid development of AI healthcare and the construction of its governance system.

3.1.1 Macro-Guidance of National-Level Strategies

In recent years, the Chinese government has intensively issued a series of national strategic plans, elevating AI—particularly its application in health—to an unprecedented strategic height.

- **"Healthy China 2030" Planning Outline (2016):** This provided broad scenarios and clear demand orientation. It proposed "promoting health science and technology innovation" and "building a health informatics service system," providing a policy basis for AI in optimizing resource allocation and enhancing service efficiency.
- **New Generation Artificial Intelligence Development Plan (2017):** A milestone document and the first national-level plan for AI. It explicitly listed "Intelligent Medicine" as one of the eight key areas for AI application. It set a "three-step" goal: by 2025, the AI industry will reach the high end of the global value chain; by 2030, China aims to become a major world AI innovation center.

Specific tasks included smart hospitals, surgical robots, intelligent clinical decision support, and AI-driven drug R&D.

- **14th Five-Year Plan and 2035 Vision Outline (2021):** Focused on deep informatics application and digital empowerment, aiming to build a unified, authoritative, and interconnected national health information platform to realize "Internet + Healthcare" across hospital levels.
- **Proposals for the 15th Five-Year Plan (2025):** The Communist Party of China's central committee proposed to "fully implement the 'AI+' action" and lead a paradigm shift in scientific research with AI. It positions AI as a core supporting force for the Healthy China strategy, emphasizing the construction of a safe, controllable, and standard-compliant digital health service system while strengthening ethics, safety, and law-based regulation.

3.1.2 Relevant Authorities

Under national strategic guidance, various ministries have issued specific operational policies, forming a multi-layered support system.

Table 3-1 Regulatory Departments and theirs Responsibilities

Department	Primary Responsibilities in AI Healthcare
National Health Commission (NHC)	The core medical regulator. Oversees clinical application management, medical institutions, and data standards. Issued regulations for health big data security and the "Internet + Healthcare" service actions to clear policy barriers for AI-assisted diagnosis and remote medicine.
National Medical Products Administration (NMPA)	The core regulator for AI medical devices. Responsible for registration, approval, and quality supervision of AI medical products, playing a critical role in market access and post-market surveillance.
National Healthcare Security Administration (NHSA)	Manages medical insurance funds and service pricing. Negotiates prices and includes AI-assisted diagnostic technologies into the medical service price structure to ensure efficient use of insurance funds.
National Administration of Traditional Chinese Medicine (NATCM)	Formulates policies and oversees safety for AI in Traditional Chinese Medicine (TCM), ensuring the integration of technology serves the inheritance and innovation of TCM.
Ministry of Industry and Information Technology (MIIT)	The industry lead. Focuses on technological breakthroughs (e.g., surgical robots, imaging systems), industrial cultivation, and standard-setting. Initiated projects to tackle "bottleneck" core technologies in AI healthcare.
Cyberspace Administration of China	Monitors data flows and protects privacy and national security based on the Personal Information Protection Law and Measures for

(CAC)	Security Assessment of Cross-border Data Transfer.

3.2 Governance Policy and Regulatory System

Driven by top-level strategies, China's policy and regulatory system for AI healthcare governance is rapidly taking shape. Its core characteristic is the **"equal emphasis on development and regulation"**—actively encouraging technological innovation and application promotion while gradually establishing a normative framework covering the entire chain, including data, algorithms, products, applications, and security.

3.2.1 Overview

In recent years, various national ministries have issued a substantial number of AI healthcare policies. Based on the characteristics of AI healthcare, this paper categorizes them into four groups:

1. Strategic Planning and Resource Foundation

- **Macro Strategy:** the "AI+" Action, the "14th Five-Year Plan for National Health Informatics," and the "Digital and Intelligent Transformation Scheme for the Pharmaceutical Industry."
- **Data Resources:** Data Classification and Grading Standards, the *Data Security Law*, and the *Personal Information Protection Law*.
- **Technical Capability:** contents regarding the construction of computing power and platforms within the "National Health Informatics Plan."

2. Scenario Application and Value Realization

- **Scenario Guidance:** "Reference Guidelines for AI Application Scenarios."
- **Disease Prevention and Control:** e.g., Prevention and Control Schemes for cancer/diabetes/respiratory diseases, and "AI-Enabled Health Services for the Elderly."
- **Payment and Pricing (The Value Loop):** "Guidelines for the Establishment of Medical Service Pricing Projects for Radiological Examinations" and sections involving costs and payments in the "Economic Management Year."

3. Product Access and Industry Promotion

- **Product Classification:** "Guiding Principles for the Classification and Definition of AI Medical Software Products" and "Standardizing the Classification and Definition of Medical Devices."
- **Registration and Review Guidance:** "Guiding Principles for the Registration Review of AI Medical Devices" and "Guiding Principles for the Clinical Evaluation of AI-Assisted Detection Software."
- **Industrial Development:** sections regarding product R&D and industrialization in the "Digital and Intelligent Transformation Scheme for the

Pharmaceutical Industry."

4. Supervision, Governance, and Ethical Norms

- **Industry Supervision:** "Detailed Rules for the Supervision of Internet Diagnosis and Treatment" and "Strengthening Cross-departmental Law Enforcement Linkage for Medical Supervision."
- **Ethical Norms:** "Ethical Norms for the New Generation of Artificial Intelligence," as well as requirements regarding fairness, explainability, and informed consent within various guiding principles.
- **Liability and Enforcement:** "Cross-departmental Law Enforcement Linkage for Medical Supervision."

Table 3-2 Summary of AI-Healthcare Related Documents Issued at the National Level (2021–2025.09)

No.	Document Title	Issuing Department	Date	Group	Key Content
1	*Opinions on Deeply Implementing the "AI+" Action*	State Council	2025-08-21	1	First national-level deployment of the "AI+ Healthcare" action; aims to build a full-process AI healthcare application system by 2027.
2	Three Standards including *Health Industry Data Classification and Grading Standards*	National Health Commission (NHC)	2025-08-28	1	Initiated three mandatory standards for AI medical data classification, application management, and scenario classification to provide unified technical specifications for data governance.
3	*Implementation Plan for Digital & Intelligent Transformation of the Pharmaceutical Industry (2025–2030)*	MIIT, NMPA, et al. (7 depts)	2025-04-24	1	Deploys "AI Empowerment for the Entire Pharmaceutical Value Chain" pilots; supports AI-aided innovative drug R&D, smart production lines, and industrialization of urgent digital medical products.
4	*"14th Five-Year Plan" for National Health Informatization*	NHC and other ministries	2022-11	1	Proposes building a unified national health information platform by 2025 and launching AI medical pilots.
5	*Personal Information Protection Law (PIPL)*	NPC Standing Committee	2021-11	1	Stipulates that processing medical personal info requires "separate consent" and restricts the mining of personal info by AI algorithms.
6	*Data Security Law*	NPC Standing Committee	2021-09	1	Establishes a classification and grading protection system for sensitive data, setting "red lines" for AI medical data usage.
7	*Reply to Suggestion No. 3178 of the 3rd Session of the 14th NPC (AI-empowered Elder Care)*	NHC (with Civil Affairs & NHSA)	2025-08-21	2	Incorporates mature services like AI chronic disease management and post-diagnosis rehab into the elder care system; extends medical insurance coverage while strengthening privacy and ethical review.
8	*Reference Guidelines for AI Application Scenarios in the Health Industry*	General Office of the NHC	2024-11	2	Systematically proposes a catalog of 84 AI medical scenarios as top-level guidance for industry-wide planning and data sharing.
9	*Guidelines for Price Categories of Radiology Imaging Medical Services*	NHSA	2024-11	2	Adds an "AI-Aided Diagnosis" extension under main radiology items and initiates medical insurance payment pilots.
10	*Guidelines on Establishing a Smart, Multi-point Trigger Monitoring & Early Warning System for Infectious Diseases*	National Disease Control Bureau, et al.	2024-08	2	Mandates building an AI/Big Data-based monitoring platform by 2030, defining data sharing and emergency response workflows.
11	*Action Plan for Prevention and Treatment of Chronic Respiratory Diseases (2024–2030)*	NHC, et al.	2024-07	2	Promotes AI and Big Data to establish standardized diagnostic aid systems for respiratory diseases at the primary healthcare level.
12	*Action Plan for Diabetes Prevention and Treatment (2024–2030)*	NHC, et al.	2024-07	2	Utilizes AI, wearables, and the internet for diabetic retinopathy screening and CGM, enhancing efficiency while ensuring informed consent.

13	*"Economic Management Year" Activities for Public Medical Institutions (2024–2025)*	NHC, NATCM	2024-06	2	Requires public hospitals to build AI-driven business-finance integration platforms to achieve internal system interoperability and data sharing.
14	*Promoting the Construction of Tight County-level Medical Communities*	NHC, et al.	2023-12	2	Promotes unified information systems in county medical communities, AI-aided diagnosis, and telemedicine, while reinforcing cybersecurity.
15	*Action Plan for Cancer Prevention and Treatment (2023–2030)*	NHC, et al.	2023-01	2	Uses AI and the internet for remote MDT (Multi-Disciplinary Team) and standardized cancer diagnosis aid systems to improve primary care capacity.
16	*Provisions for the Management of Traditional-Chinese-Medicine Standards*	NMPA	2024-07	3	Encourages the use of Big Data and AI in setting TCM standards to improve quality controllability.
17	*Regulating the Classification and Definition of Medical Device Products*	Dept. of Device Registration, NMPA	2024-05	3	Mandates that medical devices containing AI algorithms disclose algorithm types and whether they provide diagnostic conclusions to unify classification.
18	*Guiding Principles for Clinical Evaluation and Registration of AI-Aided Detection Medical Devices (Software)*	CMDE, NMPA	2023-11	3	Refines technical regulatory requirements for AI device clinical trial design, evaluation indicators, and MRMC statistical methods.
19	*Guiding Principles for Registration Review of AI Medical Devices*	CMDE, NMPA	2022-03	3	Defines full-lifecycle quality control for AI medical devices: risk management, data quality, validation/verification, update control, and post-market monitoring.
20	*Guiding Principles for Classification and Definition of AI Medical Software Products*	NMPA	2021-07	3	Clarifies that AI medical software is managed as Class III medical devices; provides classification boundaries and review points.
21	*Strengthening Cross-departmental Law Enforcement for Medical Supervision*	NHC, et al.	2024-02	4	Establishes provincial medical supervision platforms using AI and to break information silos and achieve cross-departmental data sharing.
22	*Notice on Administrative Measures for Internet Diagnosis and Treatment (Trial)*	General Offices of NHC & NATCM	2022-02	4	systematic regulatory requirements for internet diagnosis, access, quality control, AI-aided diagnosis liability, and informed consent.
23	*Ethical Norms for the New Generation Artificial Intelligence*	National Governance Committee for the New Generation AI	2021-09	4	Embeds ethical requirements such as "Fairness, Safety, and Explainability" into the entire AI lifecycle.

Source: Compiled by the author, as of Sept 1, 2025.

3.2.2 Data Governance: From "Data Silos" to "Ordered Flow"

Data is the "fuel" of AI, and data governance is the cornerstone of AI healthcare governance. In terms of medical data governance, China has undergone a transformation from emphasizing information security to encouraging data sharing and value release.

1. Constraints of Laws and Regulations:

Control over data is primarily based on the following three laws:

- **Cybersecurity Law (2017):** Established the security protection obligations of network operators and proposed principled requirements for the security of personal information and important data, including medical data.
- **Data Security Law (2021):** A foundational law in China's data field, it clarifies security obligations for data processing activities, establishes a data classification and grading protection system, and stipulates security management for cross-border transfer of important data. As vital data related to national welfare and people's livelihood, healthcare data processing is strictly regulated under this law.
- **Personal Information Protection Law (2021):** Similar to the EU's GDPR, it provides comprehensive and strict protection for personal information processing. Medical health information belongs to the category of "sensitive personal information," and its processing requires "separate consent" from the individual, along with disclosure of necessary information such as the purpose and method of processing. This sets a clear legal red line for the processing of patient data during the training and use of AI models.

2. Guidance from Departmental Regulations and Standards:

The Administrative Measures on Standards, Security, and Services of National Health and Medical Big Data (Trial) (2018), led by the National Health Commission, defines the standards, security management, and service norms for medical big data. It encourages open sharing under safe and controllable premises to support scientific research and service innovation. Furthermore, the establishment of Data Exchanges in cities like Beijing, Shanghai, Shenzhen, and Guiyang explores the market-based allocation of data elements, providing new channels for AI medical enterprises to obtain compliant data.

3. Technology and Product Regulation: From "No Basis" to "Standardized Procedures"

For AI medical products, especially software used as medical devices, the NMPA's regulation is the core link. China has followed a path from following international trends to independent innovation.

- **Classification:** The NMPA clarifies that if the intended use of AI software meets the definition of a medical device (diagnosis, prevention, monitoring, treatment, etc.), it must be managed as such. Currently, most of approved AI-driven medical software and devices are regulated under the strictest **Class III** category.
- **The "Green Channel":** To encourage innovation, the NMPA established the **"Special Review Procedure for Innovative Medical Devices."** AI products

with core invention patents, domestic originality, international leadership, and significant clinical value can apply for this priority channel, significantly shortening time-to-market.

- **Technical Guidance:** The NMPA has released various specific review guiding principles, such as the Guiding Principles for the Registration Review of Artificial Intelligence Medical Software and Key Points for the Review of Deep Learning Assisted Decision-Making Medical Devices. These provide detailed technical requirements for algorithm performance, datasets, software updates, and explainability.

4. Application Ethics and Safety Governance: From "Technology-Oriented" to "Value-Oriented"

As applications deepen, the Chinese government and society have shifted focus toward the ethical, fair, and safe dimensions of AI.

- **National Ethical Framework:** The *Governance Principles for a New Generation of Artificial Intelligence* (2019) proposed the goal of "developing responsible AI" based on eight principles, including harmony, fairness, and respect for privacy.
- **Ethical Norms:** The *Ethical Norms for a New Generation of Artificial Intelligence* (2021) further refined these requirements, emphasizing the integration of ethics into the entire AI lifecycle.
- **Professional Consensus:** Authoritative medical organizations like the Chinese Medical Association have released "Expert Consensuses" to provide guidance on clinical scope, usage processes, and risk communication, bridging the gap between technical standards and clinical practice.
- **Hospital Ethics Committees:** In large Grade-A tertiary hospitals, these committees act as "gatekeepers," strictly reviewing scientific value, risk-benefit ratios, and privacy measures before AI projects enter clinical phases.

3.3 Comparative Analysis of Global AI Healthcare Governance Levels

Applying the evaluation index system developed in the previous chapter to the governance models of China, the U.S., the EU, the UK, and Singapore, we can perform a qualitative comparative analysis. The following table provides a multidimensional contrast of their governance efficacy.

Overall, China's AI healthcare governance has forged a path characterized by strategic leadership, policy propulsion, and an equal emphasis on development and regulation. It has achieved remarkable results, contributing "Chinese wisdom" and "Chinese solutions" to the global governance of AI in healthcare. A governance system—with national strategy at the top level, laws and regulations as the framework, departmental rules as support, and standards as guidance—has begun to take shape.

However, China must also confront significant challenges on the road ahead. The lag in legislation, data silos, regulatory fragmentation, ethical ambiguities, and weaknesses in core technologies remain formidable mountains to climb. Future AI healthcare governance needs to build upon existing achievements and address these deep-seated challenges systematically, with even greater determination and wisdom.

Table 3-3 Assessment of AI Healthcare Governance Levels Across Five Major Economies

Economy	1. Governance System (Institutions, Org, Standards)	2. Governance Capability (Tech, Talent, Execution)	3. Governance Efficacy (Application, Industry, Safety)	4. Continuous Improvement (Efficiency, Value Feedback)	Comprehensive Summary
EU	**A** • **Regs:** Highly complete (e.g., *AI Act*); global rule-setter. • **Org:** Complex transnational coordination but unified framework. • **Standards:** Mature medical device regs (CE-MDR) with strict AI requirements.	**B+** • **Tech:** Strong basic research; commercial compute platforms lag behind US/China. • **Talent:** Rich reserves but facing brain drain pressure. • **Execution:** Varying rigor across member states; generally cautious.	**B** • **Apps:** Conservative clinical adoption; focus on safety validation. • **Industry:** Strong startup ecosystem but fragmented market. • **Safety:** World's highest legal requirements for safety and ethics.	**A-** • **Efficiency:** Less focus on commercial speed, more on social equity and safety. • **Value:** Strong ethical feedback; emphasizes social and health benefits.	**Rule Leader, Safety First.** The most rigorous framework setting "Gold Standards" globally, potentially at the cost of innovation speed and market efficiency.
USA	**B+** • **Regs:** Relies on existing regs (e.g., FDA Pre-Cert); flexible but lacks a unified AI Law. • **Org:** FDA-centric with multi-agency synergy; market-driven. • **Standards:** FDA guidelines are influential but often non-binding.	**A** • **Tech:** World-leading compute, data, and AI tech giants. • **Talent:** Global hub for top-tier talent. • **Execution:** Experienced FDA approval process; highly efficient but challenged by new tech.	**A** • **Apps:** Most extensive and deep clinical application; high scenario penetration. • **Industry:** Most active industrial ecosystem and capital market. • **Safety:** Relies on post-hoc liability and self-regulation; less focus on systemic risk.	**B+** • **Efficiency:** Extreme emphasis on cost-efficiency and economic returns. • **Value:** Driven by health and economic benefits; weaker social feedback loops.	**Innovation Driver, Market First.** Leads in application and industry; governance is "patchwork-style," with systemic risk mechanisms still in exploration.
UK	**A-** • **Regs:** Clear system (e.g., UKCA); actively pursues flexibility and innovation.	**B+** • **Tech:** Strong research; lacks US commercial scale. • **Talent:** High-quality	**B+** • **Apps:** NHS provides a vast testing ground but integration is slow.	**A-** • **Efficiency:** Focuses on allocative efficiency within the NHS framework.	**Value-Oriented, Pragmatic Explorer.** Navigates a "Third Way" between EU rigor and US flexibility; pragmatic but

	• **Org:** Specialized AI bodies (e.g., CDEI); efficient cross-dept synergy. • **Standards:** Active in international standard-setting; practical guidelines.	talent, especially in AI ethics. • **Execution:** Clear regulatory pathways; seeks balance between safety and speed.	• **Industry:** Active govt support and vibrant startup ecosystem. • **Safety:** Leading focus on ethical risks.	• **Value:** Strong sandbox feedback; emphasizes health and social value.	limited by national scale.
Singapore	**A-** • **Regs:** Strategy-driven; rapid policy rollout; lean and efficient system. • **Org:** Government-led; exceptionally smooth cross-dept synergy. • **Standards:** Adopts international standards with localized guidelines.	**A-** • **Tech:** Large investment in national infrastructure; leading data sharing. • **Talent:** Proactive global recruitment and robust training. • **Execution:** Agile regulation; efficient approvals; mature sandbox mechanisms.	**B+** • **Apps:** Rapid rollout in public health; fast scenario penetration. • **Industry:** A regional AI hub attracting multinational corporations. • **Safety:** Complete legal framework; strict enforcement; high social trust.	**A** • **Efficiency:** Extreme focus on governance and application efficiency. • **Value:** Agile feedback; integrates health/economy into national vision.	**Agile Governor, Vision-Driven.** Exemplary "Whole-of-Nation" governance; highly efficient and lean, though limited by domestic market size.
China	**B+** • **Regs:** Clear top-level strategy; rapid policy rollout; laws still being perfected. • **Org:** State-led; multi-dept synergy; power/responsibility boundaries need clarity. • **Standards:** Active national standard-setting; implementation/interoperability are challenges.	**B+** • **Tech:** Massive data/compute scale; rich application scenarios. • **Talent:** Large talent base; shortage of top-tier interdisciplinary experts. • **Execution:** Exploring "Sandboxes"; approval processes are accelerating.	**B+** • **Apps:** Wide scenarios; high penetration in areas like imaging; uneven development. • **Industry:** Explosive growth and active capital. • **Safety:** High focus on data security; ethical/algorithmic risk mechanisms in progress.	**B** • **Efficiency:** Starting to value Health Tech Assessment (HTA); lacks systematic feedback. • **Value:** Emphasizes social/health benefits; value-driven loops are forming.	**Fast Developer, Scale-Driven.** Rapidly catching up via market size and policy support; governance construction speed is impressive, but needs more precision and feedback maturity.

Source: Compiled by authors

Chapter 4: AI Healthcare Full-Lifecycle Risk Atlas

While Artificial Intelligence brings revolutionary opportunities to the medical field, it is accompanied by complex and intertwined risks. These risks permeate the entire lifecycle of AI medical products—from design and R&D to deployment and application—posing profound challenges across technical, legal, ethical, and social dimensions. This chapter aims to construct a systematized **AI Healthcare Full-Lifecycle Risk Atlas**, providing a deep deconstruction and concrete analysis of potential risks from four core dimensions: technology, law, ethics, and data.

4.1 AI Healthcare Full-Lifecycle Risk Framework

No AI medical technology exists as an isolated, static product; rather, it is a dynamic evolution from concept to retirement. Risks do not merely emerge during technical deployment; they are latent in every link of the lifecycle, interrelated and dynamically evolving.[①] Therefore, the primary task in constructing a risk atlas is to identify and locate key risk points along its life trajectory.

A complete AI medical lifecycle can be divided into four core stages:

4.1.1 Design and Training Phase

This is the incubation period for AI models, where quality directly determines the product's ceiling and floor. The core task is to perform data collection, labeling, algorithm selection, and model training based on specific clinical needs (e.g., lung CT nodule screening).

- **Technical Risk:** The focus is on the rationality of the algorithm architecture and the effectiveness of the training process, avoiding the aggravation of "black box" characteristics or insufficient generalization caused by design flaws.
- **Data Risk:** This is the root of all data risks. Lack of representativeness in training data introduces bias, poor labeling causes the model to "learn incorrectly," and non-compliant data collection creates legal liabilities.
- **Ethical Risk:** This is particularly prominent here. If data contains historical diagnostic biases against specific groups (e.g., ethnic minorities, women, or the impoverished), the algorithm will not only replicate but amplify this injustice, leading to systemic "garbage in, garbage out" discrimination.

4.1.2 Deployment and Integration Phase

When the trained model leaves the lab and enters the real hospital environment, it enters the deployment and integration phase. This is the first intense collision between theory and reality.

- **Technical Risk:** The primary challenge is system compatibility and environmental adaptability. The AI engine must interface seamlessly with

---

① Reddy S, Allan S, Coghlan S, et al. A governance model for the application of AI in health care[J]. *Journal of the American Medical Informatics Association*, 2020, 27(3): 491–497. https://doi.org/10.1093/jamia/ocz192 existing hospital information systems (HIS, PACS, EMR). Any mismatch in interface protocols or data formats may lead to deployment failure.

- **Legal Risk:** The core risks are qualification and liability. Products must obtain registration certification (e.g., China's NMPA, US FDA) before clinical application. Contracts with hospitals must clearly define the allocation of responsibility for system failures.
- **Ethical Risk:** This phase triggers concerns about equitable accessibility. If AI systems are only deployed in resource-rich tertiary hospitals, it may exacerbate the "Matthew Effect" and widen the healthcare gap.

4.1.3 Clinical Operation and Monitoring Phase

This is the long phase where the AI product delivers clinical value while risks are directly exposed to patients. The focus shifts from "building well" to "using well" and "managing well."

- **Technical Risk:** Models face the inevitable challenge of performance degradation because disease patterns and medical technologies are constantly changing, leading to "data drift." A continuous monitoring system must be established to track accuracy and recall in real-time.
- **Legal Risk:** The core issue is the definition of medical liability. If an AI suggestion leads to misdiagnosis, is the liable party the physician following the AI, the hospital, or the manufacturer? This remains a legal gray area.
- **Ethical Risk:** The human-AI relationship becomes the focus. Physicians may develop over-dependence, leading to skill atrophy or blind compliance with algorithms while ignoring individual patient contexts, thereby eroding professional autonomy and the physician-patient trust bond.[①]

4.1.4 Retirement and Iteration Phase

The retirement and iteration phase marks the end of an old model and the start of a new cycle. Retirement may occur because performance no longer meets clinical needs or is replaced by a superior version.

- **Technical Risk:** Challenges include smooth transition, data migration and archiving, and ensuring service continuity.
- **Legal Risk:** Manufacturers must clarify retirement obligations, such as notifying users in advance, disposing of sensitive patient data compliantly, and determining if retrospective liability exists for historical results.
- **Ethical Risk:** Iteration is a new, streamlined restart. The risk lies in "minor" iterations having "major" cumulative effects, which might bypass regulatory re-approval for "significant changes" and introduce unknown risks.

As demonstrated by the full-lifecycle framework, AI healthcare risks are interlinked and continuous. A bias introduced at the data stage may be amplified during model training, cause injustice during clinical deployment, and ultimately trigger litigation at the legal level. Consequently, we will use the four core dimensions—Technology, Law,

① Reis, M., Reis, F., & Kunde, W. (2024). Influence of believed AI involvement on the perception of digital medical advice. *Nature Medicine*. https://doi.org/10.1038/s41591-024-03180-7

Ethics, and Data—as our primary threads to examine the specific risks at each stage, constructing a comprehensive, multi-dimensional, and dynamic AI healthcare risk atlas.

Table 4-1: AI Healthcare Full-Lifecycle Risk Point Analysis

Dimension	1. Design & Training	2. Deployment & Integration	3. Clinical Operation & Monitoring	4. Retirement & Iteration
Technical / Data Risks	• Design flaws leading to "Black Box" opacity • Insufficient generalization • Low labeling quality	• Poor system compatibility • Weak environmental adaptability	• Model performance drift • Lack of real-time monitoring and feedback mechanisms	• Unstable transition during phase-out • Massive cumulative effects from "minor" iterations
Legal Risks	• Non-compliant data collection	• Use without regulatory certification/registration • Unclear rights and responsibilities in hospital contracts	• Ambiguous liability subjects • Data security and privacy breaches	• Unclear retirement obligations • Evasion of regulatory approval for "significant changes"
Ethical Risks	• Algorithmic bias and discrimination	• Disruption to existing medical workflows	• Erosion of professional autonomy • Damage to physician-patient trust	• Lack of fairness in iterations • Absence of sustainability considerations

4.2 The Technical Dimension

Technical risks are rooted in the algorithmic models themselves and serve as the source of all other risks. These risks are often hidden, dynamic, and transmissible, posing unprecedented challenges to clinical applications.

1. **Algorithmic Black Boxes and Lack of Explainability:** Modern AI, represented by deep neural networks and Large Language Models (LLMs), features extremely complex decision-making processes that are difficult to explain in human-understandable terms. If a physician receives a "high suspicion of lung cancer" diagnosis from an AI but cannot discern the basis (whether it is a minute feature in the image or a keyword in a text report), they cannot fully trust or adopt the suggestion. This "knowing *what* but not *why*" hinder effective human-AI collaboration and creates obstacles for retrospective error analysis.
2. **Model Drift and Performance Decay:** An AI model's performance is not static. Once deployed in the real world, if it encounters data distributions different from its training data (e.g., new imaging equipment, different demographic characteristics, or epidemiological shifts), its performance may gradually decline—a phenomenon known as "Model Drift." This "silent decay" is dangerous because it is hard to detect and can lead to incorrect diagnostic or treatment decisions without warning.
3. **Robustness and Adversarial Attacks:** Robustness refers to a model's ability to

remain stable in the face of noise or interference. In healthcare, minute variations in imaging (such as different scanning parameters or slight motion artifacts) can cause an AI to output entirely different results. More severely, attackers could use "Adversarial Attacks"—adding perturbations invisible to the human eye to "trick" the AI into a wrong judgment (e.g., modifying a benign tumor image so the AI identifies it as malignant).

4.3 The Data Dimension

Data is the "fuel" of AI, but its associated risks are a Sword of Damocles hanging over the processes of collection, storage, and sharing.

1. **Privacy Leaks and Security Risks:** Healthcare data contains the most sensitive personal information (genetics, history, habits). The massive concentration of data required for AI training increases the risk of leaks from hackers, insider malpractice, or third-party vulnerabilities. Even after de-identification, the risk of "re-identification" via multi-dimensional data correlation remains a threat.
2. **Data Silos and Inconsistent Quality:** High-quality models require diverse, standardized data. In reality, data is fragmented across institutions into "data silos" with inconsistent formats. Furthermore, "dirty data"—errors or missing values in unstructured electronic health records—inevitably compromises the reliability of the resulting AI models.
3. **Data Ownership and Benefit Distribution:** Who owns medical data? Is it the patient, the hospital, or the AI company that labeled it? This lack of clear legal definition creates friction. Patients fear exploitation, hospitals fear losing core assets, and AI companies struggle with "starving" for data. Establishing fair distribution of value is a primary obstacle to a healthy AI ecosystem.

Summary: AI healthcare risks are multi-dimensional and cross-pollinating. Technical risks are foundational; data risks act as the ultimate amplifier. Only by clearly recognizing this atlas can we implement the precise strategies needed to ensure the "great ship" of AI healthcare navigates safely.

4.4 The Legal Dimension

As AI becomes deeply integrated into medical practice, traditional legal relationships and liability frameworks are facing severe disruption, giving rise to a series of urgent legal challenges. In China, current regulations explicitly state that AI software must not impersonate or replace physicians, and prescriptions must be issued by the attending physician themselves—prohibiting the use of AI to automatically generate prescriptions. While this has prevented major litigation so far, the eventual encounter with "Black Box" issues and diagnostic disputes is inevitable. Legal preparation must therefore be proactive.

1. **The "Liability Black Box":** This is the core legal dilemma. When an AI-assisted or AI-driven diagnosis leads to patient harm, who is liable?
   - The tech company that developed the algorithm?
   - The hospital that provided the (potentially biased) data?
   - The medical institution that procured and deployed the system?

- The physician who ultimately accepted the AI's suggestion?

This ambiguity creates a "responsibility vacuum," where victims find it difficult to seek compensation, and stakeholders become hesitant to adopt AI for fear of unlimited liability.

2. **Complexity of Intellectual Property (IP):**
    - **Data Rights:** How do we define the ownership and usage rights of the massive datasets (de-identified EMRs, medical images) used for training?
    - **Patentability:** Should AI models be patentable? If so, does the patent belong to the developer, or the collaborative entity that provided data and compute?
    - **AI-Generated Content (AIGC):** Who owns the copyright for new diagnostic plans or drug molecular structures generated by AI? The lack of clarity here significantly inhibits collaborative innovation.
3. **Lag in Regulatory Compliance and Approval Standards:** Traditional medical device regulations are designed for "locked" algorithms. They struggle to adapt to AI products capable of continuous learning. How should regulators oversee every minor update? Determining whether a fine-tuned Large Language Model (LLM) is a "medical device" or a "software service" remains a gray area, creating compliance uncertainty for enterprises.

4.5 The Ethical Dimension

The ethical risks of AI healthcare touch the core values of medicine: fairness, justice, and respect for persons. If mishandled, technology could become a tool that exacerbates social inequality.

1. **Algorithmic Bias and Medical Injustice:** AI decision-making is highly dependent on training data. If a skin cancer detection AI is trained primarily on images of fair-skinned populations, its accuracy for dark-skinned patients may drop significantly. This "digital divide" marginalized vulnerable groups, violating the principle of medical equity.
2. **Impact on Physician-Patient Relationships and Dehumanization:** Medicine is not just science; it is a human-centric practice. Over-reliance on AI risks the **"de-skilling"** of physicians and erodes their communication with patients. There is a danger of "technological supremacy," where patients feel treated as a string of data rather than a whole person, damaging the sacred bond of trust.
3. **Boundaries of Autonomous Decision-Making and Human Dignity:** This raises a profound question: To what extent should we cede life-and-death decisions to an emotionless, unaccountable machine? Even if AI decisions are statistically superior, bypassing a physician's professional autonomy and a patient's informed consent may infringe upon fundamental human dignity.

Chapter 5: The Regulatory System for AI Healthcare Governance

To address the full-lifecycle risks of AI healthcare, it is essential to construct a scientific, systematic, and actionable core regulatory system. This chapter proposes a forward-looking and implementable governance framework centered around three fundamental questions: **"Who governs?"**, **"What is governed?"**, and **"How to govern?"**
This multi-dimensional system aims to define clear "runways" and "guardrails" for the healthy development of AI healthcare, balancing the incentivization of innovation with risk prevention.

5.1 Multi-Stakeholder Collaborative Governance

Recent research suggests that the current development of medical AI focuses too heavily on technology while neglecting critical issues within the **socio-technical system**, particularly stakeholder participation and trust-building. While not every stakeholder needs to be involved in every stage, it is crucial to define "who should participate, when, and how"[①]. To this end, a collaborative governance pattern led by the government, with industry self-regulation, corporate responsibility, and hospital participation, must be established.

5.1.1 Government

As the leader of industry development, government departments are responsible for national strategies, industrial policies, regulatory frameworks, market access, and safety monitoring.

- **Legislation & Standard Setting:** Enact specialized laws to clarify the legal status, access conditions, and liability division for AI medical technology.
- **Market Access & Risk Supervision:** Establish scientific approval processes through agencies like the NMPA to strictly manage high-risk products while maintaining real-time risk warning mechanisms.
- **Infrastructure Construction:** Lead the creation of national medical big data platforms and algorithm testing centers to lower the threshold for innovation.
- **International Cooperation:** Participate in global rule-making to promote mutual regulatory recognition.

5.1.2 AI Healthcare Enterprises

As the engines of innovation and the source of algorithmic bias and safety risks, enterprises are the first line of defense in governance.

- **Internal Ethics & Compliance Committees:** Establish multi-disciplinary committees (technical, legal, medical, ethical) for pre-emptive and continuous review.
- **"Responsible AI" Principles:** Embed fairness, transparency, and

[①] Kaye J, Shah N, Kogetsu A, et al. Moving beyond technical issues to stakeholder involvement: Key areas for consideration in the development of human-centred and trusted AI in healthcare[J]. *Asian Bioethics Review*, 2024, 16: 501–511. https://doi.org/10.1007/s41649-024-00255-7 explainability into the design process (**Ethics by Design**). Explicitly recognize that humans, not AI, are the subjects of legal liability.

- **Data Security & Quality:** Implement rigorous data governance, including qualified labeling personnel and the exploration of **Privacy-Preserving Computation**.
- **Transparency:** Disclose algorithmic principles and performance limitations to regulators and the public to foster trust.

5.1.3 Healthcare Institutions

As the end-users and those directly exposed to clinical risks, hospitals must manage the application side.

- **Access Evaluation:** Strictly assess clinical effectiveness and workflow compatibility before procurement.
- **Human-AI Synergy Standards:** Formulate clear clinical guidelines defining AI as an "assistant" and specifying how doctors should review AI suggestions.
- **Training & Education:** Prevent "automation bias" or misuse by training staff on AI's capability boundaries.
- **Adverse Event Reporting:** Establish a feedback loop to report any discovered issues to developers and regulators.

5.1.4 Health & Commercial Insurance

- **Public Health Insurance (Social Security):** Responsible for basic protection. It defines the payment scope, pushes for the inclusion of AI in reimbursement catalogs, and monitors the cost-effectiveness of AI to prevent resource waste caused by "model drift."
- **Commercial Insurance:** Focuses on differentiated protection and innovation support. It can develop specific products like **"AI Diagnostic Error Insurance"** or **"Data Breach Insurance"** to mitigate legal risks. By utilizing "value-based payment" (paying by clinical outcome), it incentivizes companies to improve technical quality.
- **Synergy:** Both sectors should collaborate to fight insurance fraud and use AI data (e.g., from wearables) to optimize health management and control costs.

5.1.5 Industry Organizations

Groups such as medical device associations and AI societies connect the government with the market.

- **Industry Conventions:** Formulate detailed self-regulatory codes and ethical best practices.
- **Certification:** Conduct independent third-party evaluations of AI performance and ethical compliance.
- **Mediation:** Reflect industry concerns to the government and help resolve disputes to maintain fair competition.

5.2 Differentiated Regulation

A "one-size-fits-all" approach is unsuitable for AI healthcare. Regulation must be precision-targeted based on the magnitude of potential risk. Drawing from global experience, AI medical applications can be classified into three levels with corresponding regulatory requirements:

5.2.1 Low-Risk AI

Products such as health education tools or administrative workflow optimization systems do not directly influence clinical decisions.

1. **Focus:** Accuracy of information and user privacy.
2. **Regulatory Measures:** Primarily based on a filing system and industry self-regulation. Third-party content quality certification is encouraged.

5.2.2 Medium-Risk AI

Products providing assisted diagnosis or risk prediction directly impact clinical outcomes.

1. **Focus:** Clinical validity, algorithmic robustness, bias control, and transparency.
2. **Regulatory Checklist:**
    - **Performance Verification:** Mandatory reports on independent, representative test sets (Sensitivity, Specificity, etc.).
    - **Algorithmic Auditing:** Independent audits for bias and stability.
    - **Data Compliance:** Proof of legal source, quality, and de-biasing measures.
    - **Human Oversight:** Product design must mandate physician review of AI outputs and log the review process.
3. **Regulatory Measures:** Strict **registration and approval** system, with "Green Channels" for innovative products. Periodic performance update reports are required post-approval.

5.2.3 High-Risk AI

These products possess autonomous diagnostic or decision-making functions, forming a closed-loop treatment system where humans are largely excluded from the immediate decision.

1. **Focus:** Absolute safety, liability traceability, and controllability in extreme scenarios.
2. **Regulatory Checklist** (In addition to Medium-Risk requirements):
    - **Mandatory Clinical Trials:** Large-scale, multi-center, randomized controlled trials (RCTs) to prove non-inferiority or superiority.
    - **High Explainability:** Mandatory highly interpretable decision paths.
    - **"Kill Switch" & Redundancy:** Reliable manual override/emergency stop mechanisms and redundant systems to prevent single points of failure.
    - **Mandatory Liability Insurance:** High-value insurance coverage for potential damages.

3. **Regulatory Measures:** The strictest approval system utilizing "Conditional Approval." Products may launch within a restricted scope (specific hospitals/diseases) and receive full approval only after validating performance with real-world data.

5.3 Full-Process Coverage

Effective governance must span the entire lifecycle—from concept to retirement—to ensure continuous compliance.

5.3.1 Design & Training Phase

- **Stakeholder Involvement:** Data strategy meetings must include patient organizations and Data Protection Officers (DPOs).
- **Key Tools:** Utilize "Dynamic Consent" platforms to allow data subjects to authorize or withdraw data at any time, aligning with the *Personal Information Protection Law*.
- **Action:** Clinical experts and patient representatives should co-create a "Label Dictionary" and use Fairness Red-line Workshops to set acceptable performance variance thresholds for sensitive attributes (gender, age, race).

5.3.2 Deployment & Integration Phase

- **Methodology:** Conduct "Tabletop Exercises + Live Simulations" involving frontline clinicians, IT, and management.
- **Human-AI Division:** The principle is "Non-delegable Human Decision-making." Interfaces should retain "Double Confirmation" hard buttons and visualize AI confidence using a "Traffic Light" system.
- **Patient Rights:** Waiting areas should display videos explaining the AI's purpose and how patients can exercise their Right to Refuse.

5.3.3 Clinical Operation & Monitoring Phase

- **Monitoring:** Technical teams should build Role-Based Dashboards. Clinicians track sensitivity/specificity, while Quality Control monitors correlations with medical disputes.
- **Feedback Loop:** Any breach of "Fairness Red-lines" should trigger an automatic alert to the Ethics Committee. An "One-click Feedback" button in apps allows patients to upload queries, creating a "Public Issue Pool" for closed-loop resolution.

5.3.4 Retirement & Iteration Phase

- **Action:** Establish a smooth transition plan to prevent service interruption.
- **Knowledge Management:** Post-retirement, hospitals or developers should publish a summary of benefits, harms, and lessons learned. This forms a Public Knowledge Base to prevent future errors and build social capital for the next AI implementation.

By constructing this core system of "Multi-stakeholder Collaboration, Differentiated Regulation, and Full-process Coverage," we can weave a governance net that is comprehensive yet flexible—providing "fuel" for innovation while effectively containing the risks it may ignite.

Chapter 6: Policy Recommendations and Action Initiatives

This chapter translates theoretical reflection and practical insights into concrete, actionable policy recommendations and initiatives. It proposes differentiated responsibility lists and action guides for stakeholders including governments, enterprises, and healthcare institutions. Finally, with the vision of building a "Global AI Healthcare Governance Community," it presents a series of initiatives to promote cross-border data collaboration, mutual recognition of standards, and experience sharing.

6.1 Policy Recommendations for the Government

The government plays multiple roles as a provider of systems and a promoter of technology. The scientific rigor, foresight, and execution of its policies directly determine the health of the entire industrial ecosystem.

6.1.1 Accelerate Specialized Legislation

Policy and regulation are the bedrocks of governance. As China's framework is still in its infancy, it is recommended that the legislative departments incorporate the "Administrative Regulations on the Application of Artificial Intelligence Medical Technology" into the formal legislative plan as a priority. This foundational law should address the following:

1. **Clear Legal Definitions & Classification:** Define the legal attributes of AI medical products and establish a dynamic classification catalog based on risk levels.
2. **Build a Multi-Layered Shared Responsibility Framework:** Clarify that developers, producers, medical institutions, and users assume responsibility proportional to their degree of fault. Explore the introduction of mandatory liability insurance and compensation fund systems.
3. **Regulate Data Rights & Interests:** Establish fundamental principles for medical data ownership and regulate collection, storage, and utilization, with enhanced protection for sensitive personal information.
4. **Establish a National AI Medical Technical Appraisal & Arbitration Center:** Create an independent third-party institution responsible for technical appraisal, ethical review, and dispute arbitration to provide authoritative support for judicial and administrative decisions.

6.1.2 Refine the Standards System

Led by the NMPA, NHC, and SAMR, and in collaboration with industry associations and leading enterprises, a full-chain standard system should be constructed:

1. **Fundamental General Standards:** Define terminology, data formats, interface protocols, performance evaluation, and security requirements to provide a "unified language."
2. **Industry Application Standards:** Tailor technical requirements and clinical validation methods for specific scenarios like medical imaging, drug R&D, and robotic surgery.
3. **International Standardization:** Actively participate in organizations like ISO,

IEC, and ITU to align Chinese standards with international ones, enhancing China's global "voice" in governance.

6.1.3 Innovate Regulatory Toolkits

The government should modernize its regulatory models to improve efficiency:

1. **Deepen "Regulatory Sandboxes":** Expand the scope of sandbox pilots and optimize entry/exit mechanisms to make them vital "incubators" for innovation.
2. **Construct a National Intelligent Supervision Platform:** Plan a "Regulatory Brain for AI Healthcare" using AI and Big Data to achieve dynamic, intelligent monitoring across the product lifecycle.
3. **Drive Health Insurance Payment Reform:** Establish a value evaluation system for AI medical services to clear the "last mile" for innovative applications.

6.1.4 Improve Data Governance

To break data silos and improve quality, priority should be given to:

1. **National Medical Data Sharing Platforms:** Facilitate interconnectivity between hospitals, research institutes, and enterprises.
2. **Quality Standardization:** Set standards for data cleaning and labeling to ensure accuracy and consistency for model training.
3. **Security Management:** Employ encryption, de-identification, and access control to prevent leakage or tampering.

6.1.5 Strengthen Ethical Review

To address privacy, bias, and fairness, the following mechanisms are prioritized:

1. **Establish Ethical Review Committees:** Require AI medical application entities to set up internal or joint committees for ethical oversight.
2. **Formulate Ethical Review Guidelines:** Issue the **"Guidelines for the Ethical Review of AI Medical Technology"** to provide operational instructions on principles, procedures, and responsibilities.
3. **Enhance Ethics Training:** Improve the ethical awareness and review capabilities of R&D personnel, clinicians, and committee members.

6.2 Action Initiatives for AI Enterprises

As the primary subjects of innovation, AI enterprises must also be the primary bearers of governance responsibility. Integrating ethics and compliance into every facet of corporate strategy and operations is the only path to sustainable development.

1. **Establish Internal Ethics Review Committees:** We advocate for all AI healthcare R&D enterprises—especially those above a certain scale—to set up a standing committee comprising technical, medical, legal, and ethical experts. This committee should hold **"veto power"** over R&D projects to ensure products meet ethical requirements from the concept stage.
2. **Promote Algorithm Transparency and Explainability:** Enterprises should move away from the "Black Box" mindset. While protecting trade secrets, they should disclose the basic principles, training data overviews, and performance limitations

of their algorithms to regulators and users. Adopting **Explainable AI (XAI)** is essential to enhancing decision-making transparency.

3. **Construct a Full-Lifecycle Quality Management System:** Quality control must span design, development, testing, deployment, updates, and retirement. Enterprises should establish robust adverse event monitoring and reporting mechanisms to continuously track and optimize post-market performance.
4. **Strengthen Data Governance and Privacy Protection:** Data security must be treated as the corporate lifeline. By utilizing **Privacy-Preserving Computation** and blockchain, firms should ensure compliance at every stage of data use and provide users with clear, accessible privacy management options.

6.3 Action Guidelines for Healthcare Institutions

Healthcare institutions represent the "last mile" of contact between AI and patients; their application standards directly impact patient safety.

1. **Define Internal AI Norms and Responsibility Lists:** Institutions should establish detailed regulations for the procurement, evaluation, and management of AI. Crucially, they must clarify the **boundary of responsibility**: AI is an auxiliary tool; final diagnostic and treatment decisions must be made by qualified physicians, who bear independent responsibility for reviewing AI suggestions.
2. **Enhance AI Literacy Training for Staff:** AI knowledge should be integrated into continuing medical education. Training should focus not just on software operation, but on AI principles, capability boundaries, and the **critical interpretation of outputs** to prevent "automation bias."
3. **Establish a Feedback Loop for Adverse Events:** Create internal channels for staff to report issues without fear of reprisal. Institutions should summarize and analyze this information, providing timely feedback to developers and regulators to form a "Clinical-Enterprise-Regulatory" interaction.

6.4 Promoting Broad Public Participation

Public understanding and supervision are vital for the legitimacy and safety of AI healthcare.

1. **Establish Public Participation Mechanisms:** Implement public consultations, hearings, and opinion solicitation processes to listen to societal views on AI development.
2. **Strengthen Science Communication:** Use media, lectures, and exhibitions to popularize knowledge regarding AI's prospects and risks, enhancing the public's scientific literacy and risk awareness.
3. **Encourage Social Supervision:** Empower social organizations and the media to monitor R&D and application, ensuring improper behaviors are corrected and public rights are protected.

6.5 Action Initiatives Under a Global Vision

The challenges of AI healthcare are global, and no single nation can solve them in isolation[①]. We propose the following to build a "Global AI Healthcare Governance Community":

1. **Global Data Security & Trustworthy Sharing Initiative:** Leveraging platforms like the **WHO**, China can advocate for international frameworks for cross-border medical data collaboration. This includes promoting the standardization of **Federated Learning** and Multi-party Computation to solve global health crises through compliant data flow.
2. **Mutual Recognition of Regulatory Standards and Test Results:** Call upon the **WHO** and **IMDRF** to coordinate global alignment on AI definitions, risk classifications, and core testing methods. Establishing bilateral or multilateral recognition of review reports will significantly reduce global compliance costs for innovators.
3. **Global Experience-Sharing Platform:** Establish an open online platform for regulators and scholars to share best practices, policy tools, and case studies.
4. **Addressing Ethical and Social Challenges Collectively:** Jointly develop universal AI healthcare ethical codes to address algorithmic bias and the Digital Divide, ensuring technology serves the health of all humanity rather than exacerbating inequality.

The wave of AI healthcare has arrived, carrying with it the hopes for a healthier future. Effective governance is not meant to bind this wave, but to build the levees that guide it toward the correct destination. Through the collective wisdom of governments, enterprises, and the international community, we can begin a new era of an AI-empowered healthy world!

① Pesapane, F.; Bracchi, D.A.; Mulligan, J.F.; et al. Legal and Regulatory Framework for AI Solutions in Healthcare in EU, US, China, and Russia: New Scenarios after a Pandemic. *Radiation* **2021**, *1*, 261–276. https://doi.org/10.3390/ radiation1040022

References

[1] Batool, A., D. Zowghi and M. Bano (2025), "AI governance: A systematic literature review", *AI and Ethics*, Vol. 5, pp. 3265–3279, https://doi.org/10.1007/s43681-024-00461-x.

[2] Bipartisan Policy Center. (2023). *Current regulatory landscape of AI in public health & health care: A brief overview*. https://bipartisanpolicy.org/report/current-regulatory-landscape-of-ai-in-public-health-health-care/

[3] *EO Intelligence: 2025 China Medical Large Model Industry Analysis.* https://www.vzkoo.com/read/2025070375b28b3d8ed02cc9579f6c3b.html[2025-10-25][Chinese]

[4] Esteva, A., B. Kuprel, R. A. Novoa, et al. (2017), "Dermatologist-level classification of skin cancer with deep neural networks", *Nature*, Vol. 542/7639, pp. 115-118, https://doi.org/10.1038/nature21056.

[5] Hirani, R., K. Noruzi, H. Khuram, et al. (2024), "Artificial intelligence and healthcare: a journey through history, present innovations, and future possibilities", *Life*, Vol. 14/4, pp. 557, https://doi.org/10.3390/life14040557.

[6] Kaye, J., N. Shah, A. Kogetsu, et al. (2024), "Moving beyond technical issues to stakeholder involvement: Key areas for consideration in the development of human-centred and trusted AI in healthcare", *Asian Bioethics Review*, Vol. 16, pp. 501–511, https://doi.org/10.1007/s41649-024-00255-7.

[7] Lim, J. I., C. D. Regillo, S. R. Sadda, et al. (2023), "Artificial intelligence detection of diabetic retinopathy: subgroup comparison of the EyeArt system with ophthalmologists' dilated examinations", *Ophthalmology Science*, Vol. 3/1, pp. 100228, https://doi.org/10.1016/j.ogla.2022.100228.

[8] National Academy of Medicine (2025), *An Artificial Intelligence Code of Conduct for Health and Medicine: Essential Guidance for Aligned Action*, Washington, DC: The National Academies Press, https://doi.org/10.17226/27536.

[9] Nori, H., N. King, S. M. McKinney, et al. (2023), "Capabilities of GPT-4 on medical challenge problems", *arXiv:2303.13375*, https://doi.org/10.48550/arXiv.2303.13375.

[10] Pesapane, F.; Bracchi, D.A.; Mulligan, J.F.; et al. Legal and Regulatory Framework for AI Solutions in Healthcare in EU, US, China, and Russia: New Scenarios after a Pandemic. *Radiation* **2021**, *1*, 261–276. https://doi.org/10.3390/ radiation1040022

[11] Reddy S, Allan S, Coghlan S, et al. A governance model for the application of AI in health care[J]. *Journal of the American Medical Informatics Association*, 2020, 27(3): 491–497. https://doi.org/10.1093/jamia/ocz192

[12] Reis, M., Reis, F., & Kunde, W. (2024). Influence of believed AI involvement on the perception of digital medical advice. *Nature Medicine*. https://doi.org/10.1038/s41591-024-03180-7

[13] Singhal, K., T. Tu, J. Gottweis, et al. (2025), "Toward expert-level medical question answering with large language models", *Nature Medicine*, Vol. 31/3, pp. 943–950, https://doi.org/10.1038/s41591-023-02448-8.

[14] *The Research Insights: Global AI Healthcare Market Forecast 2030.* https://www.cn-healthcare.com/articlewm/20250529/content-1651020.html

[15] World Economic Forum and Info-communications Media Development Authority (2020), *Companion to the Model AI Governance Framework: Implementation and Self-Assessment Guide for Organizations*, Geneva: World Economic Forum.

[16] World Health Organization (2021), Ethics and Governance of Artificial Intelligence for Health: WHO Guidance, Geneva: World Health Organization.