International Business & Economics Studies
ISSN 2640-9852 (Print) ISSN 2640-9860 (Online)
Vol. 8, No. 3, 2026
www.scholink.org/ojs/index.php/ibes

Original Paper

Research on the Economic Value and Risk Governance of AI-Assisted Medical Diagnosis

Qiuchen Zhang[1]

[1] Shanghai University of Science and Technology, Shanghai, China

Received: April 16, 2025 Accepted: May 09, 2026 Online Published: June 1, 2026

doi:10.22158/ibes.v8n3p40 URL: http://dx.doi.org/10.22158/ibes.v8n3p40

Abstract

AI-assisted medical diagnosis is becoming an important part of digital healthcare. It can help doctors read images, identify possible disease signals, and support earlier clinical decisions. This paper studies the economic value and risk governance of AI-assisted medical diagnosis from the perspective of healthcare economics and hospital management. The study uses literature analysis and normative analysis. It reviews 12 published studies and reports on medical artificial intelligence, clinical validation, economic evaluation, and ethical governance. The findings show that AI-assisted diagnosis may improve diagnostic efficiency, reduce repeated work, support cost control, and help allocate medical resources more evenly. At the same time, its value depends on data quality, clinical validation, workflow integration, legal responsibility, and patient trust. The paper argues that AI should not replace doctors in diagnosis. It should work as a decision-support tool under human review. Hospitals and regulators should build clear rules for data use, performance testing, responsibility sharing, and long-term economic evaluation.

Keywords

ai-assisted medical diagnosis, healthcare economics, medical artificial intelligence, risk governance, healthcare efficiency

1. Introduction

AI-assisted medical diagnosis refers to the use of artificial intelligence systems to support clinical judgment. These systems often analyze medical images, laboratory data, electronic health records, or symptom information. In practice, the doctor remains the main decision maker. The AI system works as an assistant that highlights abnormal signals, gives possible diagnostic suggestions, or sorts clinical information. This topic has gained strong attention because many health systems face pressure from aging populations, growing chronic disease burden, and uneven distribution of skilled doctors. Topol

argues that the value of medical AI should be seen through the cooperation between human professionals and intelligent machines, not through a simple replacement of doctors (Topol, 2019). This view is important for this paper because economic value in healthcare is not only about speed. It is also about safety, trust, and the quality of care.

The use of AI in healthcare has also changed the way hospitals think about medical resources. Jiang et al. explain that AI can be applied to structured data, such as laboratory results, and to unstructured data, such as images and clinical notes (Jiang, Jiang, Zhi, Dong, Li, Ma, Wang, Dong, & Shen, 2017). This means that AI-assisted diagnosis is not limited to one medical department. It can be used in radiology, dermatology, ophthalmology, pathology, cardiology, and general clinical decision support. For hospitals, this creates a new management question. If AI can help doctors make faster and more stable judgments, hospitals may reduce waiting time, avoid some repeated examinations, and make better use of expert resources. Yet this value does not appear automatically. Hospitals need data, equipment, training, maintenance, and clear clinical processes. These requirements turn AI-assisted diagnosis into an economic and governance issue, not only a technical issue.

Current discussion about AI diagnosis often focuses on model accuracy. This focus is understandable because diagnosis must be safe. However, a high-performing model in a research setting may not bring the same result in a busy hospital. A hospital has different patient groups, different machines, different data habits, and different doctor workflows. A system that works well in one place may need further testing before it is used in another place. From an economic perspective, this gap matters. The hospital pays for software, hardware, training, data cleaning, and long-term monitoring. Patients also pay in time, money, and trust. If the system does not fit the real clinical environment, the cost may be higher than the benefit.

This paper studies AI-assisted medical diagnosis from the perspective of economic value and risk governance. It does not build an algorithm. It also does not claim new clinical trial results. The paper uses published literature to explain how AI-assisted diagnosis may create value, where the main risks come from, and what kind of governance path may make this value more stable. The research question is clear: how can AI-assisted medical diagnosis improve healthcare efficiency and resource allocation while keeping medical risk under control? This question is suitable for a small journal article because it connects technology, economy, hospital management, and public governance in one specific medical scenario.

The contribution of this paper is mainly practical. Many studies show that AI has technical potential in diagnosis, but hospitals still need a clear way to judge whether the system is worth adopting. This paper organizes the value of AI-assisted diagnosis into diagnostic efficiency, cost control, resource allocation, and service quality. It also organizes risk governance into data governance, evidence standards, legal responsibility, human-AI collaboration, and long-term evaluation. This structure can help hospitals, healthcare managers, and policy makers think about AI diagnosis in a more balanced way.

2. Literature Review and Research Questions

2.1 Clinical Application of AI-assisted Diagnosis

Existing literature shows that AI-assisted diagnosis has made clear progress in image-based medical tasks. Esteva et al. used deep neural networks for skin cancer image classification and compared model performance with dermatologists (Esteva, Kuprel, Novoa, Ko, Swetter, Blau, & Thrun, 2017). This study is often cited because it showed that AI could reach a high level of performance in a focused diagnostic task. The study also shows an important condition. AI diagnosis needs large and carefully labeled datasets. It does not come from simple software installation. For hospital managers, this point has economic meaning. Good data and expert labeling require time, labor, and money. These costs should be counted when hospitals discuss AI investment.

Another important example comes from breast cancer screening. McKinney et al. evaluated an AI system for breast cancer screening across international datasets (McKinney, Sieniek, Godbole, Godwin, Antropova, Ashrafian et al., 2020). This type of study matters because screening programs have high workload and strong public health value. If AI can help radiologists reduce missed signals or false alarms, it may bring both clinical and economic benefits. Yet the study also reminds readers that AI performance needs external evaluation. A model should be tested beyond the development environment. This is very relevant to hospitals that purchase AI diagnostic tools. They should not rely only on vendor claims. They need local validation and continuous review after deployment.

Medical imaging is one of the most mature fields for AI-assisted diagnosis. Litjens et al. reviewed deep learning in medical image analysis and showed that deep learning had been used for classification, detection, segmentation, registration, and other image tasks (Litjens, Kooi, Bejnordi, Setio, Ciompi, Ghafoorian et al., 2017). This broad application range explains why AI diagnosis is often linked with radiology and pathology. These departments produce large volumes of digital data, and doctors often spend much time on repeated reading tasks. AI can help mark suspicious areas and support work prioritization. However, image tasks are not the whole of diagnosis. Real diagnosis usually combines images, medical history, laboratory results, symptoms, and doctor-patient communication. So AI should be used as part of a wider clinical process.

2.2 Evidence Quality and Clinical Translation

The literature also shows that AI-assisted diagnosis faces evidence problems. Nagendran et al. reviewed studies that compared diagnostic deep learning algorithms with clinicians and found concerns about study design, reporting quality, and risk of bias (Nagendran, Chen, Lovejoy, Gordon, & Komorowski, 2020). This finding is important because economic value depends on reliable clinical evidence. If the evidence is weak, hospitals may overestimate the benefit of AI. A weak evidence base may also lead to unsafe adoption. In this sense, clinical evidence is not only a medical issue. It is also an investment risk issue.

Kelly et al. discuss the key challenges of turning AI research into clinical impact (Kelly,

Karthikesalingam, Suleyman, Corrado, & King, 2019). Their discussion helps explain why many AI systems look promising in research papers but face barriers in actual hospitals. These barriers include data shift, workflow mismatch, unclear responsibility, lack of prospective testing, and limited regulation. A hospital is not a laboratory. A diagnostic system needs to work with real doctors, real patients, and real time pressure. This is where risk governance becomes necessary. Governance should not be treated as a burden after the technology is developed. It should be part of the whole adoption process.

Bajwa et al. describe AI as a technology that may change healthcare delivery, but they also stress the need for safe and reliable systems (Bajwa, Munir, Nori, & Williams, 2021). Their discussion supports a balanced view. AI can improve healthcare, but it can also create new forms of risk. For example, doctors may trust a system too much, patients may not understand the role of AI, and hospitals may lack staff who can monitor model performance. These problems are not solved by better algorithms alone. They require management systems, training, and clear communication.

2.3 Reporting Standards, Economic Evaluation, and Governance

Clinical reporting standards are also important for AI diagnosis. Liu et al. proposed the CONSORT-AI extension for clinical trial reports involving AI interventions (Liu, Rivera, Moher, Calvert, & Denniston, 2020). Ibrahim et al. further discussed SPIRIT-AI and CONSORT-AI as reporting guidelines for clinical trials of AI systems (Ibrahim, Liu, Zariffa, & Morris, 2021). These guidelines are useful for this paper because they show that AI clinical studies need more detail than ordinary trials. Researchers need to explain how the AI system works in the clinical pathway, how humans interact with it, and how errors are handled. Without this information, hospitals cannot judge the real value and risk of AI adoption.

Economic evaluation is a newer but necessary part of this field. Wu et al. reviewed economic evaluations of AI-assisted healthcare technologies and pointed out that the economic value of AI remains uncertain in many settings because studies use different methods and often miss AI-specific costs (Wu, Chao, Lin, Huang, & Hsieh, 2025). This finding directly supports the theme of this paper. AI-assisted diagnosis may be cost-effective in some clinical scenarios, but the result cannot be copied from one hospital to another without careful review. Software license fees, data preparation, model maintenance, staff training, and workflow change may all affect the final cost. A hospital that ignores these costs may make a wrong investment decision.

The World Health Organization also provides an ethical and governance framework for AI in health (World Health Organization, 2021). The WHO report discusses issues such as human autonomy, transparency, responsibility, inclusiveness, and public interest. These principles are important because medical diagnosis touches life, safety, privacy, and trust. A hospital should not treat AI diagnosis as a normal office system. It should treat it as a clinical decision-support technology with public value and public risk. This paper uses these ideas to build a practical governance path for AI-assisted diagnosis.

2.4 Research Questions

Based on the literature above, this paper focuses on three research questions. The first question asks how AI-assisted medical diagnosis creates economic value in hospitals and health systems. The second question asks what risks may weaken or offset this value. The third question asks how hospitals and regulators can build a governance system that keeps doctors responsible, protects patients, and supports fair use of AI. These questions are connected. Economic value cannot be discussed without risk because a diagnostic mistake may create direct patient harm and large hidden costs. Risk also cannot be discussed without value because strict rules without practical value may slow down useful medical innovation.

3. Research Design

3.1 Research Method

This paper uses literature analysis and normative analysis. The literature analysis focuses on published studies about AI diagnosis, medical image analysis, clinical evidence standards, economic evaluation, and health AI governance. The normative analysis uses these studies to build a framework for value judgment and risk control. This design fits the purpose of the paper because the paper does not have access to hospital-level primary data or patient-level clinical data. It also avoids making unsupported empirical claims. The paper only discusses conclusions that can be supported by the cited literature and by reasonable analysis of hospital management practice.

The research object is AI-assisted medical diagnosis. The paper pays more attention to decision support than to fully automated diagnosis. This distinction is important. In most real clinical settings, AI tools are used to assist doctors. They may mark suspicious lesions, rank images by risk, generate alerts, or provide a suggested reading. Doctors still need to check the result and make the final clinical judgment. Because of this, the economic value of AI does not come from replacing doctors. It mainly comes from helping doctors handle repeated tasks, reducing information pressure, and improving the use of scarce expert time.

The paper uses the term economic value in a broad but clear way. Economic value includes improved diagnostic efficiency, lower avoidable cost, better resource allocation, reduced waiting time, and higher service quality. The paper does not use economic value to mean simple profit. Healthcare is not the same as ordinary commercial service. A useful medical technology should improve patient outcomes and system efficiency at the same time. If a system saves money but increases clinical risk, its value is not acceptable. If a system improves safety but creates costs that no hospital can bear, its adoption will also be difficult.

3.2 Analytical Framework

The analytical framework of this paper contains three connected parts. The first part is value creation. AI-assisted diagnosis may create value by improving efficiency, reducing repeated work, supporting

earlier detection, and allowing experts to focus on complex cases. The second part is risk formation. Risk may come from poor data quality, biased data, weak validation, unclear responsibility, privacy problems, and weak workflow design. The third part is governance response. Hospitals and regulators need data rules, clinical evidence standards, human review, legal responsibility design, and continuous economic evaluation. Table 1 summarizes this framework.

Table 1. Analytical Framework for AI-assisted Medical Diagnosis

Dimension	Main economic meaning	Key risk	Governance response
Diagnostic efficiency	Shorter reading time and better use of doctor labor	Overreliance on model output	Human review and clear clinical workflow
Cost control	Possible reduction of repeated work and avoidable errors	Hidden costs in data, training, maintenance, and updates	Full cost accounting and long-term evaluation
Resource allocation	Support for lower-level hospitals and remote care	Uneven data quality and digital divide	Shared standards, training, and regional support
Service quality	More stable screening and faster alerts	False positives, false negatives, and bias	External validation and continuous monitoring
Public trust	Clearer use of technology in diagnosis	Privacy concern and unclear responsibility	Transparent communication and accountability rules

This framework shows a simple idea. AI-assisted diagnosis should be evaluated as a socio-technical system. The system includes algorithm, data, doctors, patients, hospitals, payment rules, and regulation. A model may be technically strong, but the final result depends on how it is used. In a hospital with poor data management, weak training, and unclear responsibility, AI may add confusion. In a hospital with good workflow and governance, AI may help doctors deliver faster and more stable service. This difference explains why governance is part of value creation.

4. Economic Value of AI-assisted Medical Diagnosis

4.1 Improving Diagnostic Efficiency

The most direct economic value of AI-assisted diagnosis is efficiency improvement. Many diagnostic tasks require doctors to read large amounts of information. Radiologists read images. Pathologists check tissue slides. Clinicians review test results and patient histories. Some of this work is highly skilled, but part of it is repetitive. AI systems can help screen information, mark suspicious areas, or give risk scores. This support may reduce the time doctors spend on routine searching and allow them to focus on difficult judgment. From a hospital management view, this means the same number of

doctors may serve more patients or spend more time on complex cases.

Efficiency improvement should not be understood as faster work only. A faster but unsafe system has no real value. The more meaningful value is that AI may help doctors use attention more carefully. For example, in image reading, AI may flag a possible abnormal area that needs closer review. In a screening program, AI may help sort high-risk cases earlier. In such settings, doctors can spend more time on cases with greater need. This can reduce the pressure caused by large patient volume. It can also improve the stability of service when experienced doctors are limited.

The economic meaning of efficiency is especially clear in departments with heavy workload. Medical imaging departments often face growing demand because modern diagnosis uses more imaging tests. If doctor numbers do not grow at the same speed, waiting time may increase. AI-assisted diagnosis can be used as a work support system. It may help with preliminary sorting, lesion marking, and report preparation. Even if the AI does not replace any doctor, it may reduce some repeated steps. This type of value is small in each case but large at the system level when the number of examinations is high.

However, hospitals should measure this value carefully. A vendor may say that AI improves efficiency, but real efficiency depends on the workflow. If doctors need to open another system, check extra alerts, and correct many false alarms, the actual workload may rise. If the system is well connected with the existing hospital information system and picture archiving system, the tool may save time. This shows why workflow integration is an economic issue. A hospital should evaluate not only model accuracy but also time use, doctor acceptance, and the change in patient flow.

4.2 Supporting Cost Control

AI-assisted diagnosis may also help control medical costs. The possible cost value comes from several sources. The system may reduce repeated examinations by helping doctors make clearer judgments. It may reduce avoidable errors by reminding doctors of suspicious signals. It may improve screening efficiency by helping health systems focus resources on high-risk patients. These effects may reduce direct medical spending and indirect social cost. However, this paper needs to state the limit clearly. The degree of cost saving depends on the clinical scenario, local payment system, patient group, and hospital management practice. It cannot be assumed without evidence.

Cost control also has another side. AI systems create new costs. Hospitals need to buy software, upgrade hardware, prepare data, train staff, and maintain the system. Hospitals may also need technical staff who understand data security and model monitoring. If the system changes over time, hospitals need to check whether performance remains stable. These costs are often less visible than purchase price. Wu et al. note that existing economic evaluations of AI-assisted healthcare technologies do not always capture AI-specific costs in a consistent way (Wu, Chao, Lin, Huang, & Hsieh, 2025). This point is important for hospital decision making. A low purchase price does not mean a low total cost.

A sound cost analysis should include the whole life cycle of AI-assisted diagnosis. The cost begins before installation because hospitals need data cleaning and workflow design. The cost continues

during use because doctors need training and system performance needs monitoring. The cost may also appear when the system is updated, replaced, or audited. If the hospital ignores these long-term costs, it may overestimate return on investment. In practice, a hospital should compare the AI system with the current diagnostic process. It should ask whether the system reduces waiting time, repeated work, avoidable follow-up, and doctor workload enough to justify the total cost.

The cost question also connects to payment systems. If hospitals pay for AI tools but the payment system does not recognize the value of better screening or fewer errors, hospitals may lack incentive to adopt the technology. This is common in healthcare innovation. The party that pays for the technology may not be the same party that receives the main benefit. For example, a hospital may pay for AI, while patients and insurers may benefit from earlier detection and fewer repeated visits. This mismatch can slow adoption. Policy makers should consider this issue when designing digital health payment rules.

4.3 Optimizing medical Resource Allocation

AI-assisted diagnosis may help improve medical resource allocation. Many health systems have a mismatch between patient demand and expert supply. Large hospitals often attract many patients, while smaller hospitals may have limited specialist capacity. This pattern can create crowded outpatient services and long waiting time. AI diagnostic support may help lower-level medical institutions handle some preliminary diagnosis and screening tasks. If the tool is reliable and doctors are well trained, patients with lower risk may receive care closer to home. Patients with higher risk can be referred to higher-level hospitals earlier.

This value is especially important in image-based screening and chronic disease management. For example, some communities may lack enough specialists to read every image or test result in time. AI can provide a first layer of support by identifying cases that require urgent review. This does not mean that AI replaces specialists. It means that expert time may be used where it is most needed. From an economic perspective, this is a resource allocation gain. The health system can use scarce expert labor more effectively.

However, the same technology may also increase inequality if it is not governed well. Wealthier hospitals may buy better AI systems, have cleaner data, and hire stronger technical staff. Smaller hospitals may lack money, digital infrastructure, and training. If this happens, AI may widen the gap between institutions instead of narrowing it. The digital divide is therefore a governance problem. Public policy may need to support shared platforms, regional diagnostic centers, common data standards, and training programs. Without this support, the economic benefit of AI may remain concentrated in large hospitals.

Resource allocation also involves patient trust. Some patients may feel safer when a large hospital uses AI, but they may distrust AI support in a small hospital. This trust gap can weaken the role of AI in primary care. Hospitals and public health agencies should explain how AI tools are used, who makes the final decision, and how errors are handled. Clear communication can reduce misunderstanding. It

can also help patients see AI as a support tool rather than an unknown machine making decisions about their health.

4.4 Improving Service Quality and Patient Experience

AI-assisted diagnosis can also influence patient experience. Patients often care about waiting time, report speed, clear explanation, and confidence in diagnosis. If AI helps doctors read tests faster and reduce delayed reports, patient experience may improve. If AI helps doctors find possible risks earlier, patients may receive follow-up care sooner. These changes may not always appear as direct financial profit, but they have real economic meaning. Better patient experience can reduce repeated visits, improve trust, and support hospital reputation.

At the same time, patient experience may become worse if AI is used without clear explanation. Some patients may worry that doctors no longer take responsibility. Some patients may fear that their data will be used without consent. Some may think an AI result is always correct, while others may reject it completely. These reactions can affect doctor-patient communication. Hospitals should therefore treat communication as part of AI governance. Doctors should be able to explain that AI is a tool, the doctor reviews the result, and the final diagnosis remains a clinical decision.

Service quality also depends on fairness. If AI tools are trained on data that do not represent all patient groups, the system may perform better for some groups than others. This is a serious issue because diagnosis should be fair. A hospital should not assume that one model works equally well for all patients. It should check performance across age, sex, disease type, equipment source, and other relevant groups when possible. The goal is not to remove all uncertainty. The goal is to find visible risk early and reduce avoidable harm.

5. Practical Risks of AI-assisted Medical Diagnosis

5.1 Data Quality and Data Privacy Risks

Data quality is the foundation of AI-assisted diagnosis. If the training data are incomplete, biased, or poorly labeled, the system may learn wrong patterns. If the hospital data are different from the development data, model performance may decline. This risk is common in healthcare because hospitals differ in equipment, patient groups, clinical habits, and data standards. A model trained in one environment may not work equally well in another. This is why local validation is important before clinical deployment.

Data privacy is another major risk. Medical data are sensitive. They include disease history, images, test results, identity information, and sometimes genetic or family information. If hospitals use data for AI training or system operation, they need clear consent rules, access control, storage security, and audit records. A data leak can harm patients and damage hospital trust. It can also create legal and financial cost. Data privacy should not be treated as a technical detail. It is part of the economic risk of AI adoption.

Data governance should also consider data use after deployment. Some AI systems may improve through updates or continued learning. This can be useful, but it also raises questions. Who controls the data? Does the patient know how the data are used? Does the updated model need new validation? Can the hospital explain changes in model performance? These questions matter because a medical AI system is not always fixed. Governance needs to cover the whole data life cycle.

5.2 Diagnostic Accuracy and Model Generalization Risks

Diagnostic accuracy is the most visible risk. AI may produce false positives and false negatives. A false positive may lead to anxiety, extra tests, and unnecessary cost. A false negative may delay treatment and harm patient safety. Both types of error have economic consequences. Hospitals should not only ask whether an AI system is accurate on average. They should ask where it fails, for whom it fails, and what happens after it fails. This question is more useful for real governance.

Model generalization is also a serious issue. A model may perform well on one dataset but less well on data from another hospital. This can happen because of different imaging machines, disease prevalence, patient age structure, or clinical practice. External validation can reduce this risk, but it cannot remove it completely. Continuous monitoring is needed after deployment. Hospitals should track model output, doctor correction, missed cases, and user feedback. If performance changes, the hospital should have a clear process to pause, update, or restrict the system.

The reporting problem discussed by Nagendran et al. is closely related to this risk [6]. If studies do not report design and validation clearly, hospitals cannot judge whether a model is suitable for their setting. The reporting guidelines proposed by Liu et al. and Ibrahim et al. are useful because they ask researchers to explain the AI intervention and human-AI interaction more clearly [9][10]. For hospitals, this means that evidence standards are part of procurement. A hospital should ask vendors and researchers for transparent validation information before adoption.

5.3 Legal Responsibility and Accountability Risks

AI-assisted diagnosis creates difficult responsibility questions. If a doctor follows an AI suggestion and the result is wrong, who should take responsibility? If a doctor rejects an AI alert and harm occurs, how should the case be judged? If the system fails because of a software update or poor data quality, should the vendor, hospital, or doctor be responsible? These questions do not have one simple answer. They depend on law, contract, clinical practice, and the design of the AI system.

Unclear responsibility can affect adoption. Doctors may not trust AI if they fear that they will carry all responsibility for a system they cannot fully understand. Hospitals may hesitate to adopt AI if liability risk is unclear. Vendors may avoid transparency if they fear legal exposure. This creates a governance challenge. A good system should define the role of the doctor, hospital, and vendor before deployment. It should also keep audit logs so that decisions can be reviewed after an incident.

Accountability also relates to explainability. Doctors do not need every technical detail of a model, but they need enough information to use it safely. They need to know the intended use, suitable patient

group, known limits, and correct response to alerts. If the system only gives a result without context, doctors may either overtrust it or ignore it. Both reactions are risky. Human-centered design can reduce this problem by making AI output easier to understand and easier to check.

5.4 Organizational Dependence and Workflow Risks

AI-assisted diagnosis may create organizational dependence. When doctors use a system for a long time, they may become less careful in some routine tasks. This is sometimes called automation bias. It means that humans may accept machine output too easily. In diagnosis, this risk is serious because a small ignored detail can affect patient safety. Hospitals should design workflow rules that keep doctors actively involved. AI output should support thinking, not replace it.

Workflow risk may also come from poor system integration. If the AI system is separated from existing medical systems, doctors may need more clicks, more screens, and more manual checks. This can increase workload and reduce adoption. If alerts are too many or too unclear, doctors may suffer alert fatigue. They may stop paying attention. In such a case, a tool designed to reduce risk may create new risk. Hospitals should test usability before full deployment.

Training is also part of workflow governance. Doctors, nurses, technicians, and managers need different types of training. Doctors need to understand clinical use and limits. Technicians need to understand data quality and system operation. Managers need to understand cost, risk, and performance indicators. Without training, AI tools may be misunderstood. Good training can turn AI from an external product into a real part of clinical work.

5.5 Equity and Public Trust Risks

Equity risk appears when AI systems do not serve all patient groups fairly. This may happen because training data are not representative. It may also happen because some hospitals have better digital infrastructure than others. If AI improves service only in large urban hospitals, the technology may increase inequality. If AI performs poorly for underrepresented patient groups, it may create unfair clinical risk. This problem connects economic value with public ethics.

Public trust is also fragile. Patients may accept AI when they believe that doctors remain responsible and data are protected. Patients may reject AI when they feel that technology is used to reduce human care or save cost at their expense. Hospitals should avoid presenting AI as a magic solution. They should explain its role in simple language. Patients should know whether AI is used, what it supports, and who makes the final diagnosis. This type of transparency can support trust.

The WHO report stresses that AI in health should protect human autonomy, promote well-being, ensure transparency, and support fairness (World Health Organization, 2021). These principles are not abstract words. They can guide hospital policy. For example, a hospital can require human review for high-risk diagnosis. It can also disclose AI use in patient communication. It can audit model performance across patient groups. It can set up a complaint and review process when patients question AI-supported decisions.

Table 2. Main Risks and Governance Responses of AI-assisted Medical Diagnosis

Risk category	Main problem	Possible impact	Governance response
Data quality	Incomplete or biased training data	Lower diagnostic reliability	Data standards, labeling review, and local validation
Privacy	Sensitive medical data may be misused or leaked	Patient harm and legal cost	Consent rules, access control, encryption, and audit logs
Accuracy	False positives and false negatives	Extra tests, delayed treatment, and medical disputes	External validation and continuous performance monitoring
Responsibility	Unclear roles of doctor, hospital, and vendor	Lower adoption and higher liability risk	Contracts, clinical protocols, and decision records
Workflow	Poor integration with hospital systems	More workload and lower doctor acceptance	Usability testing and workflow redesign
Equity	Different performance across groups or hospitals	Unequal access and unfair diagnosis	Bias audits, regional support, and public reporting

6. Risk Governance and Optimization Path

6.1 Building a Human-in-the-loop Diagnosis Model

The most important governance rule is that AI-assisted diagnosis should keep humans in the loop. The doctor should remain the final clinical decision maker, especially in high-risk diagnosis. This does not reduce the value of AI. It makes the value safer. AI can process information quickly, but doctors understand patient context, symptoms, history, and communication. A safe diagnostic process should combine machine speed with human judgment.

A human-in-the-loop model needs clear workflow. The hospital should define when doctors must check AI results, when they can override AI suggestions, and how disagreements should be recorded. The system should not pressure doctors to accept the AI result. It should also not hide its uncertainty. If an AI system gives a risk score, doctors should know what the score means and what action is expected. These details can prevent blind trust and blind rejection.

Human review also supports patient trust. When patients know that doctors still make the final decision, they may feel more comfortable with AI use. Doctors can explain that AI is used to help check information and reduce missed signals. This explanation is easier to accept than telling patients that a machine gives a diagnosis. In medical service, trust itself has economic value because it affects compliance, follow-up, and the stability of doctor-patient relations.

6.2 Improving Data Governance and Privacy Protection

Hospitals should build clear data governance before adopting AI-assisted diagnosis. The governance system should cover data collection, labeling, storage, use, sharing, deletion, and audit. It should define who can access data, for what purpose, and under what conditions. It should also define how patient identity information is protected. A hospital that lacks basic data governance should be cautious about adopting complex AI tools.

Data quality management should be part of routine hospital work. Clinical data are often messy because doctors and departments record information in different ways. AI systems need stable input. If data standards are weak, model output may become unreliable. Hospitals should improve data coding, image quality control, and clinical documentation. These tasks may look basic, but they strongly affect AI performance. Good AI depends on good data habits.

Privacy protection should also be practical. Hospitals should use access control, encryption, de-identification when appropriate, and operation logs. They should also review contracts with vendors. The contract should state whether data can be used for model training, whether data can leave the hospital, how long data are stored, and what happens if a breach occurs. These rules protect patients and reduce future disputes.

6.3 Strengthening Clinical Evidence and Economic Evaluation

AI-assisted diagnosis should be evaluated through both clinical evidence and economic evidence. Clinical evidence asks whether the system is safe and useful for patients. Economic evidence asks whether the system brings value compared with its full cost. Both are needed. A system with good accuracy but very high total cost may not be suitable for all hospitals. A system that saves time but has weak safety evidence should not be widely used.

Hospitals should require evidence that matches their use scenario. If a system is used for breast cancer screening, evidence from that area is needed. If a system is used in a primary care setting, evidence from only a large specialist hospital may not be enough. Local validation is also useful because local data may differ from published datasets. Hospitals can begin with a pilot project and compare indicators before and after adoption. Useful indicators include report time, recall rate, doctor workload, patient waiting time, cost per case, and rate of clinical disagreement.

Economic evaluation should include hidden and long-term costs. These costs include software fees, hardware, data preparation, training, maintenance, updates, cybersecurity, legal review, and workflow redesign. Hospitals should also consider cost savings that may not appear immediately, such as fewer repeated tests or earlier risk detection. Because these effects may take time, evaluation should not stop at the purchase stage. It should continue after the system enters daily use.

Policy makers can help by building evaluation standards for medical AI. Standard indicators can make results more comparable. They can also prevent hospitals from adopting tools based only on marketing claims. The review by Wu et al. shows that current economic evaluations of AI-assisted healthcare

technologies need more standardized methods (Wu, Chao, Lin, Huang & Hsieh, 2025). This point is highly relevant to AI diagnosis. Without common evaluation methods, hospitals may find it hard to compare products and judge long-term value.

6.4 Designing Clear Responsibility Rules

Clear responsibility rules can reduce fear and confusion. Hospitals should define the role of the AI system in clinical protocols. The protocol should state that the system is a support tool and that doctors must make the final diagnosis according to clinical standards. It should also state when doctors need to document reasons for accepting or rejecting AI suggestions. This documentation can support learning and accountability.

Vendor responsibility should also be defined. Vendors should provide information about intended use, performance limits, update history, data requirements, and known failure conditions. They should help hospitals monitor performance and respond to technical problems. If a vendor updates a model, the hospital should know what changed and whether new validation is needed. These rules protect hospitals and patients from hidden system changes.

Responsibility should not be pushed only to doctors. Doctors use tools that hospitals purchase and vendors develop. A fair accountability system should consider all parties. The hospital is responsible for procurement, training, workflow design, and monitoring. The vendor is responsible for product quality, technical support, and truthful information. The doctor is responsible for clinical judgment within the defined workflow. This shared responsibility model is more realistic than blaming one party after a problem occurs.

6.5 Promoting Standardized and Staged Adoption

Hospitals should adopt AI-assisted diagnosis in stages. A staged approach is safer than full-scale rollout. The hospital can start with a clear task, such as image triage or lesion marking in one department. It can set baseline indicators before adoption and compare results after use. If the system works well, the hospital can expand the use scenario. If problems appear, the hospital can adjust or stop the project before risk spreads.

Standardization is also important. Hospitals should use standard data formats, standard reporting templates, and standard clinical protocols where possible. This can reduce errors and improve comparability. It can also help smaller hospitals join regional AI platforms. If every hospital uses different data rules, AI tools will be harder to validate and maintain. Standardization is therefore not only a technical issue. It is a condition for wider economic value.

Training should continue after deployment. Doctors may need time to learn how to use AI output. New staff may join the department. The system may change through updates. Hospitals should provide repeated training and collect user feedback. This can improve acceptance and reveal workflow problems early. A system that doctors do not trust or do not understand will not create stable value.

7. Conclusion

AI-assisted medical diagnosis is an important direction in digital healthcare. It may help health systems respond to doctor shortages, heavy diagnostic workload, long waiting time, and uneven medical resources. The economic value of AI-assisted diagnosis mainly appears in four areas: improved diagnostic efficiency, better cost control, more balanced resource allocation, and improved service quality. These values are meaningful, but they are not automatic. They depend on data quality, clinical validation, workflow integration, doctor training, and patient trust.

The main risks of AI-assisted diagnosis include data privacy risk, weak data quality, false positives and false negatives, unclear legal responsibility, workflow mismatch, automation bias, and unequal access. These risks can reduce or even offset economic value. A hospital that adopts AI without governance may pay for a system that creates more work, more disputes, and more uncertainty. For this reason, risk governance should be designed before adoption and continued during use.

This paper argues that the best development path is not full automation but human-AI collaboration. Doctors should remain responsible for final clinical judgment. AI should help doctors screen information, identify possible risks, and improve workflow efficiency. Hospitals should build data governance, require clinical and economic evidence, define responsibility rules, and evaluate performance over time. Regulators should support standards, post-market monitoring, and fair access. These measures can help AI-assisted diagnosis move from technical promise to practical healthcare value.

For future development, researchers should pay more attention to real-world evidence. Hospitals need evidence about cost, workflow, doctor acceptance, patient trust, and long-term safety. Policy makers also need better evaluation tools to compare AI products. If these problems are addressed, AI-assisted medical diagnosis may become a useful tool for improving healthcare efficiency and public health fairness. If these problems are ignored, the technology may remain impressive in research but limited in real clinical value.

References

Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthcare Journal*, *8*(2), e188-e194.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118.

Ibrahim, H., Liu, X., Zariffa, N., Morris, A. D., & Denniston, A. K. (2021). Reporting guidelines for clinical trials of artificial intelligence interventions: The SPIRIT-AI and CONSORT-AI guidelines. *Trials*, *22*(1), 11.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017).

Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230-243.

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*(1), 195.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M. et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60-88.

Liu, X., Rivera, S. C., Moher, D., Calvert, M. J., & Denniston, A. K. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nature Medicine*, *26*(9), 1364-1374.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H. et al. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, *577*(7788), 89-94.

Nagendran, M., Chen, Y., Lovejoy, C. A., Gordon, A. C., Komorowski, M., Harvey, H. et al. (2020). Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*, *368*, m689.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44-56.

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. Geneva: World Health Organization.

Wu, W. T., Chao, Y. W., Lin, T. K., Huang, C. K., & Hsieh, P. H. (2025). Economic evaluation of AI-assisted technologies in healthcare: A systematic review. *Journal of Food and Drug Analysis*, *33*(4), 487-500.