






Clinical Artificial Intelligence Implementation in Routine Care: Real-World Operational Outcomes from a Provincial Health System in China

Jin Tian ^{1*}, Yongzhao Song ^{2*}, Longmei Tang ³, Lejiao An ⁴, Jin Zhou ⁵

¹Hospital Management Innovation Research Center, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China; ²Public Health Department, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China; ³Department of Social Medicine and Health Services Management, Faculty of Public Health, Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China; ⁴Office of the President, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China; ⁵Health Management Center, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jin Tian, (Primary), Hospital Management Innovation Research Center, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China, Tel +86 19930739708, Email jtia716@aucklanduni.ac.nz; Jin Zhou, Health Management Center, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei, People's Republic of China, Tel +86 8633889116, Email 57101298@hebm.edu.cn

Background: Clinical artificial intelligence (AI) technologies are increasingly being introduced into hospital practice, yet evidence describing their operational integration and performance after deployment in routine clinical settings remains limited. This study examined the real-world implementation and operational integration of clinical AI within a provincial tertiary health system in China over an 18-month observation period.

Methods: This retrospective longitudinal observational study used aggregated institutional data generated during routine platform deployment, including electronic medical record-linked system logs, deployment records, quality-monitoring summaries, and operational reports. The analysis focused on implementation patterns, workflow integration, selected operational indicators, and user acceptance during routine clinical use. The study evaluated implementation and operational integration rather than algorithmic accuracy, diagnostic performance, or patient-level clinical effectiveness.

Results: Three AI-supported clinical pathways were included: an intelligent pre-consultation system, a multidisciplinary tumor decision-support system, and a duloxetine therapeutic drug-monitoring pathway. During the observation period, 127 clinicians across 53 specialties participated in AI-assisted clinical activities involving more than 27,000 patient encounters, 850 multidisciplinary tumor decision-support cases, and 320 therapeutic drug-monitoring episodes. Patient waiting time decreased from 18 to 13 minutes, patient satisfaction increased from 95.40% to 98.92%, consultation efficiency improved by approximately 40%, and documentation completion efficiency improved by approximately 80%. Implementation patterns differed substantially across pathways, reflecting differences in workflow position and clinical accountability rather than deployment effort alone.

Conclusion: In this provincial tertiary health system, clinical AI implementation was associated with sustained operational use, cross-specialty workflow integration, and measurable changes in selected workflow indicators. These findings suggest that technical functionality alone is insufficient for sustained clinical AI use and that workflow compatibility and organizational readiness are central to routine implementation. Because this was a single-site observational study using aggregated operational data, the findings should be interpreted as implementation evidence rather than proof of clinical effectiveness.

Plain Language Summary: Many hospital AI systems are tested in pilot settings but are not subsequently integrated into routine clinical operations. This study examined how a clinical AI platform was deployed in routine care at a large teaching hospital in Hebei Province, China, over an 18-month period.

The study used routine hospital data, including electronic medical record-linked system logs, deployment records, quality-monitoring summaries, operational reports, and patient satisfaction information. No identifiable patient information was accessed. The analysis examined how widely the AI tools were used, which clinical services adopted them, and whether selected workflow indicators changed following routine deployment.

Three AI-supported tools were included: an intelligent pre-consultation system, a multidisciplinary tumor decision-support system, and a duloxetine therapeutic drug-monitoring system. In total, 127 clinicians across 53 specialties used the platform, involving more than 27,000 patient encounters, 850 tumor decision-support cases, and 320 drug-monitoring episodes. Patient waiting time decreased from 18 to 13 minutes, patient satisfaction increased from 95.40% to 98.92%, consultation efficiency improved by approximately 40%, and documentation efficiency improved by approximately 80%.

The findings suggest that clinical AI can move from pilot use into routine hospital workflows when deployment is grounded in existing clinical workflows and backed by sustained organizational support. Because this was an observational study from a single hospital, the findings should be interpreted cautiously and confirmed in other settings.

Keywords: clinical artificial intelligence, real-world implementation, workflow integration, operational outcomes, implementation science

Introduction

Clinical artificial intelligence (AI) is increasingly being introduced into hospital practice, with applications spanning diagnostic support, clinical decision-making, patient triage, medication safety, documentation assistance, and operational coordination.^{1–5} Advances in large language models, multimodal data processing, and clinical decision-support systems have further expanded the range of AI-enabled functions that may be embedded within hospital workflows.^{6–8} Despite these developments, the transition from technical feasibility to routine clinical use remains difficult. Many AI systems continue to be evaluated in controlled, experimental, or pilot settings, while less is known about their operational performance after deployment in everyday clinical care.^{9–11}

Routine implementation differs fundamentally from algorithm validation. Validation studies typically emphasize accuracy, discrimination, calibration, or diagnostic performance under predefined technical conditions.^{12–14} Real-world implementation, however, requires sustained clinician participation, compatibility with electronic medical records, fit with departmental workflows, operational continuity, and monitoring of service-level indicators during routine use.^{15–17} An AI system that performs well technically may still fail to influence clinical work if it increases workload, disrupts routine processes, or lacks adequate organizational support.¹⁸ Implementation evidence therefore needs to describe not only technical performance, but also adoption, sustainability, and workflow integration in routine care.

This gap restricts the ability of hospital managers and policymakers to judge whether AI can be maintained as a routine operational capability rather than a short-term innovation project. Existing studies and reviews have highlighted barriers related to workflow fit, professional trust, infrastructure readiness, accountability, and organizational change, but empirical reports describing multi-pathway deployment across large hospital systems remain scarce.^{9–11,18} Many implementation reports focus on single tools, isolated departments, or early adoption phases, limiting understanding of how AI-supported services operate after broader institutional rollout. Similar post-deployment challenges have also been reported in health systems outside China, including North America, Europe, and other Asian settings, where promising AI tools may remain confined to pilot or early implementation phases rather than becoming stable components of routine care.^{9,10,18}

China provides an important and underrepresented setting for examining these questions. National efforts to develop smart hospitals, digital-health infrastructure, and AI-supported healthcare services have accelerated the introduction of clinical AI in tertiary hospitals.¹⁸ International governance guidance emphasizes transparency, accountability, human oversight, and lifecycle monitoring as foundational requirements for responsible AI deployment in healthcare settings. Provincial tertiary hospitals represent a particularly informative context: they combine high clinical service volume, mature information systems, and regulatory alignment with national and international AI standards, creating conditions in which multi-pathway deployment across heterogeneous clinical services is both feasible and observable. Empirical evidence describing operational outcomes after routine deployment in such settings, however, remains limited.

This study examined the real-world implementation and operational integration of a clinical AI platform at a large provincial tertiary hospital in Hebei Province, China. A related governance framework derived from the same institutional implementation context has been reported elsewhere.¹⁹ The present study differs from that viewpoint by focusing on empirical operational outcomes, pathway-level adoption, workflow integration, and selected service indicators during routine deployment.

Methods

Study Design and Objectives

This retrospective longitudinal observational study examined the real-world implementation and operational integration of a clinical artificial intelligence (AI) platform within routine hospital practice. Conducted from an implementation and operational evaluation perspective, it used aggregated operational data generated during routine deployment and did not aim to evaluate algorithmic accuracy, clinical effectiveness, or the previously reported governance framework.¹⁹ The objectives were to describe deployment characteristics, adoption and utilization patterns across clinical pathways and specialties, integration of AI-supported functions into routine workflows, and selected operational indicators and user acceptance during routine clinical use.

Setting and Observation Period

The study was conducted at a large provincial tertiary teaching hospital in Hebei Province, China. The hospital has approximately 3,300 beds and serves as a regional referral center. The observation period extended from August 2024 to January 2026, covering 18 months from initial clinical deployment to completion of operational data extraction.

The six-month pre-implementation baseline period was February 2024 to July 2024. This baseline period was used where institutional quality-monitoring indicators were available for comparison.

Clinical AI Pathways

Three AI-supported clinical pathways were included. The intelligent pre-consultation system was deployed in August 2024 and supported structured patient history collection, preliminary symptom assessment, and outpatient triage preparation before physician consultation. Initial deployment was conducted in gynecology, endocrinology, cardiology, otolaryngology, and neurology because these departments had high outpatient volumes, relatively standardized consultation processes, stable electronic medical record use, and sufficient readiness for workflow integration.

The multidisciplinary tumor decision-support system was deployed in November 2024. It supported oncology-related multidisciplinary team discussions by organizing patient information and providing decision-support outputs for clinician review. Oncology decision-making is guideline-intensive, multidisciplinary, and dependent on integration of patient-specific clinical information. Final clinical decisions remained the responsibility of treating clinicians and multidisciplinary teams.

The duloxetine therapeutic drug-monitoring pathway was deployed in December 2024. It supported individualized medication management by integrating patient characteristics and therapeutic drug-monitoring information into routine neuropsychiatric medication-management workflows. This pathway represented a specialized medication-support use case with narrower clinical scope and more structured data inputs.

Data Sources and Institutional Access

The study used aggregated institutional data generated during routine platform deployment. Data sources included electronic medical record-linked system logs, deployment records, quality-monitoring summaries, institutional operational reports, patient satisfaction summaries, and platform utilization records. Patient satisfaction data were derived from routine institutional service-evaluation summaries.

The study team did not directly participate in algorithm development, software engineering, or routine technical operation of the platform. Data were accessed from the perspective of hospital management and implementation evaluation. As part of the hospital's administrative and operational-management structure, aggregated summaries related to digital-health implementation, workflow operation, quality monitoring, and service performance were routinely compiled for institutional review. No identifiable patient-level or clinician-level data were accessed. All data extraction and aggregation procedures were conducted within institutional information-management systems under existing hospital data-governance policies.

Outcome Domains and Measures

Implementation outcomes were assessed across four domains. Deployment characteristics included the number and type of AI-supported pathways, deployment timing, participating departments, specialty coverage, and duration of continuous operation.

Clinical adoption was assessed using system utilization records, including active clinician users, number of specialties with sustained use, patient encounters, consultation sessions, tumor decision-support cases, therapeutic drug-monitoring episodes, and mean daily utilization.

Workflow integration was assessed according to whether AI-supported functions were incorporated into existing clinical processes, including outpatient pre-consultation workflows, multidisciplinary tumor board processes, medication-management workflows, documentation support, and quality-monitoring activities.

Operational indicators included patient waiting time, consultation efficiency, documentation completion efficiency, documentation standardization, AI-assisted quality-control coverage, and patient satisfaction. Where available, post-deployment values were compared with institutional baseline summaries from the six months before implementation.

Data Validation and Missing Data Handling

Aggregate indicators were cross-checked against source system logs and institutional reporting records to ensure consistency of deployment dates, pathway-level utilization counts, specialty coverage, and aggregate performance indicators; discrepancies were reviewed against institutional reporting records before final aggregation. Because the study used aggregated operational data, indicators without reliable baseline values were reported descriptively and interpreted as institutional monitoring estimates rather than causal effectiveness measures.

Statistical Analysis

Descriptive statistics were used to summarize deployment characteristics, adoption patterns, utilization volume, and operational indicators. Continuous indicators are presented as absolute values or percentage changes where appropriate. Categorical indicators are presented as counts and percentages.

Because this was a single-site observational study using aggregated institutional data without a control group, no causal inference was attempted. No inferential testing was performed for most operational indicators because patient-level or encounter-level data and complete pre-implementation denominators were not available. Observed changes were therefore interpreted as descriptive implementation observations rather than causal estimates of clinical or operational effectiveness.

Results

Platform Deployment and Infrastructure

During the 18-month observation period, three AI-supported clinical pathways were deployed within routine hospital practice: an intelligent pre-consultation system, a multidisciplinary tumor decision-support system, and a duloxetine therapeutic drug-monitoring pathway. The intelligent pre-consultation system was deployed in August 2024, followed by the multidisciplinary tumor decision-support system in November 2024 and the duloxetine therapeutic drug-monitoring pathway in December 2024.

By January 2026, the platform was supported by institutional graphics processing unit computing infrastructure integrating Huawei Ascend and Nvidia clusters. The hospital held Electronic Medical Record System Functionality Level 5 certification and Health Information Interoperability Level 5B certification. Regulatory approval for the oncology decision-support pathway was completed before routine clinical deployment. [Figure 1](#) illustrates the deployment timeline and operational rollout sequence of the three clinical AI pathways.

Clinical Adoption and Utilization

A total of 127 clinicians across 53 specialties participated in AI-assisted clinical activities during the study period. Platform utilization involved more than 27,000 patient encounters, 850 multidisciplinary tumor decision-support cases, and 320 therapeutic drug-monitoring episodes.

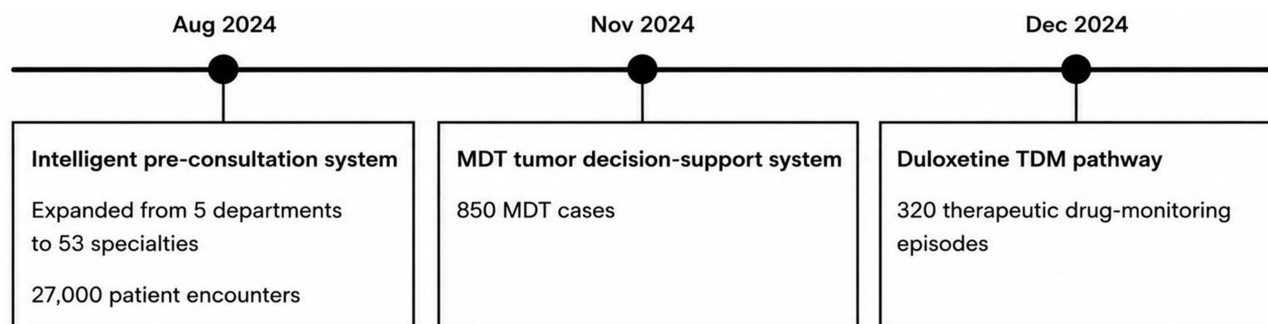


Figure 1 Timeline of clinical AI pathway deployment during routine hospital implementation. The intelligent pre-consultation system was deployed in August 2024, followed by the multidisciplinary tumor decision-support system in November 2024 and the duloxetine therapeutic drug-monitoring pathway in December 2024.

The intelligent pre-consultation system accounted for the largest utilization volume and expanded from five initial departments to broader outpatient service coverage. Mean daily utilization exceeded 100 encounters per day after phased operational rollout. The multidisciplinary tumor decision-support pathway was used in oncology-related multidisciplinary discussions, whereas the duloxetine therapeutic drug-monitoring pathway remained limited to a more specialized neuropsychiatric medication-management setting. All three pathways remained in active operational use throughout the observation period. [Table 1](#) summarizes pathway-level adoption and utilization.

Workflow Integration and Operational Indicators

AI-supported functions were incorporated into outpatient pre-consultation workflows, multidisciplinary tumor board discussions, medication-management workflows, clinical documentation support, and institutional quality-monitoring activities.

Selected operational indicators changed during the observation period. Mean patient waiting time decreased from 18 minutes during the baseline period to 13 minutes after deployment. Consultation efficiency improved by approximately 40%, and documentation completion efficiency improved by approximately 80% according to institutional operational monitoring records. Documentation standardization monitoring showed a 25-percentage-point increase relative to institutional baseline summaries. AI-assisted quality-control functions were incorporated into institutional monitoring activities during the later implementation period. [Table 2](#) summarizes operational indicators before and after deployment.

User Acceptance

Institutional patient satisfaction increased from 95.40% during the pre-implementation period to 98.92% after deployment. The post-implementation satisfaction sample included 3,847 survey responses. The pre-implementation satisfaction

Table 1 Adoption and Utilization Metrics Across Three Clinical AI Pathways (August 2024–January 2026)

Clinical AI Pathway	Initial Deployment	Operational Duration	Active Clinicians	Total Clinical Events	Organizational Coverage	Mean Daily Use
Intelligent pre-consultation system	Aug-24	18 months	127	>27,000 patient encounters	5 initial departments; 53 specialties	>100 encounters/day
Multidisciplinary tumor decision-support system	Nov-24	15 months	45	850 cases	8 oncology-related departments	1.9 cases/day
Duloxetine therapeutic drug-monitoring pathway	Dec-24	14 months	23	320 episodes	3 neuropsychiatric departments	0.8 episodes/day

Notes: Specialty coverage was defined as specialties with at least 10 system uses per month during the observation period. The intelligent pre-consultation system recorded more than 27,000 patient encounters across 24,154 clinician consultation sessions. Mean daily use for the intelligent pre-consultation system refers to the post-rollout operational phase after phased departmental onboarding, rather than a uniform average across the entire 18-month observation period. Mean daily use was derived from institutional platform utilization summaries and should be interpreted as an operational utilization indicator rather than a patient-level outcome measure.

Table 2 Selected Operational Indicators Before and After AI Deployment

Performance Indicator	Baseline Period (Feb 2024–Jul 2024)	Post-Deployment Period (Aug 2024–Jan 2026)
Consultation efficiency	Institutional baseline reference	Approximately 40% improvement
Patient waiting time (registration to consultation)	18 min (estimated)	13 min
Documentation completion efficiency	Institutional baseline reference	Approximately 80% improvement
Documentation standardization compliance	Institutional baseline reference	25-percentage-point increase
Quality-control coverage	Limited manual review	Automated system-supported review
Patient satisfaction	95.40% (baseline denominator unavailable)	98.92% (n = 3,847)

Notes: Consultation efficiency, documentation completion efficiency, and documentation standardization compliance were derived from institutional operational monitoring summaries rather than patient-level measurements. Complete pre-implementation denominators were unavailable for several operational indicators because the study used aggregated institutional operational data generated during routine deployment. Reported percentage improvements should therefore be interpreted as descriptive institutional monitoring estimates rather than causal estimates of intervention effectiveness. Patient waiting time refers to the interval from registration to physician consultation in departments with active AI deployment.

value was derived from institutional quality-monitoring summaries, and individual survey response counts for that period were not available at the aggregate reporting level.

Routine clinician use continued across participating departments without mandatory participation requirements or direct financial incentives. No formal institutional complaints related to AI system operation were recorded during the observation period.

Discussion

Pathway-Stratified Adoption as the Central Finding

The central finding of this study is not simply that clinical AI was adopted, but that adoption followed pathway-specific implementation patterns shaped by workflow position, clinical risk, and accountability requirements. Clinical AI did not enter routine care as a uniform category of technology. Instead, each pathway became embedded through a distinct operational route, depending on where it was positioned in the clinical workflow and what type of clinical responsibility it introduced.

The WenYi intelligent pre-consultation system represented a high-frequency, low-risk, upstream workflow-support tool. Positioned before the physician encounter, it supported patient information collection and preliminary structuring across outpatient specialties. Because it reduced repetitive information-gathering work without redistributing final clinical authority, the threshold for use was relatively low. This likely contributed to its expansion from five initial departments to 53 specialties and its use in more than 27,000 patient encounters.

The multidisciplinary tumor decision-support system followed a different implementation pattern. Its use volume was lower, but each interaction occurred within complex multidisciplinary oncology decision-making, where imaging findings, pathology results, guideline recommendations, and patient-specific information had to be interpreted under specialist authority. Implementation in this context was expressed not as broad diffusion, but as stable incorporation into a high-accountability decision process.

The duloxetine therapeutic drug-monitoring pathway represented a third pattern: narrow, technically specialized, and medication-sensitive integration into individualized pharmacotherapy management. Its purpose was not to increase throughput or support broad outpatient flow, but to inform bounded specialist decisions in a clinically sensitive area. Its relatively narrow deployment scope should therefore not be interpreted as incomplete implementation.

Clinical AI systems cannot be meaningfully evaluated through a single volume-based adoption metric.^{6,10–13} High-volume workflow-support systems and low-volume specialist decision-support systems generate different forms of operational relevance. A pathway supporting thousands of routine outpatient encounters and a pathway supporting a smaller number of complex specialist decisions may each represent successful implementation when decision sensitivity, responsibility structure, and operational complexity are considered. In this sense, maximal diffusion may be an inappropriate benchmark for some forms of clinical AI.

Workflow Position as a Structural Determinant of Diffusion

Implementation behavior in this study was structurally shaped by workflow position. AI systems entering upstream workflow stages, particularly those involving repetitive information collection, documentation support, or preliminary structuring, appeared to diffuse more easily because they reduced operational burden without substantially changing clinical authority. By contrast, AI systems operating closer to high-stakes interpretation, medication adjustment, or multidisciplinary decision-making diffused more selectively and remained more tightly bounded within specialist workflows.

Existing implementation literature often treats adoption as a general endpoint.^{6,10–13,16,17} The present findings suggest that adoption is pathway-dependent. Different AI systems occupy different operational positions and therefore encounter different diffusion ceilings, interpretive burdens, and responsibility boundaries. For specialist decision-support systems, limited diffusion may represent appropriate implementation rather than failure. Systems embedded within oncology decision-making or individualized pharmacotherapy require stronger interpretive oversight, narrower user groups, and more bounded oversight structures than low-risk workflow-support applications.^{20–22} Under these conditions, careful containment reflects risk-sensitive implementation.

The temporary operational suspension of one pathway pending clarification of model update authority deserves particular attention. Operational interruptions are often treated as implementation failures, but in clinical AI they may indicate functioning risk control.^{23,24} Suspension in this case allowed accountability boundaries and update authority to be reviewed before routine use continued under defined oversight arrangements. In this setting, temporary suspension functioned as part of operational oversight during routine deployment, illustrating that mature implementation requires not only mechanisms for activation and use, but also operational procedures for pausing, reviewing, and restarting AI-supported functions when responsibility boundaries remain unclear.

Operational Capacity as the Enabling Condition for Sustained Use

A recurring problem in clinical AI implementation is that technically functional systems fail to persist beyond early deployment.^{3,5–7} Transition into routine use in the present study depended less on technical deployment alone than on the formation of operational capacity around the deployed pathways. None of the three pathways functioned as a standalone AI product. Each was connected to established electronic medical record workflows, departmental documentation processes, specialist decision structures, or quality-monitoring routines.^{25,26}

Diffusion of the WenYi system to 53 specialties occurred without mandatory participation or direct financial incentives. Routine use became more sustainable once workflow-adjustment burden was reduced through prior standardization and data integration.²⁷ This pattern suggests that voluntary diffusion at scale may be achievable when organizational readiness lowers the adoption cost sufficiently, a condition rarely examined in implementation research.

Workflow compatibility, data readiness, operational coordination, and continuity of support created the environment in which each pathway stabilized, rather than technical deployment alone.¹⁹

Situating the Findings in the Implementation Literature

Most clinical AI literature continues to emphasize algorithm development, validation, or diagnostic performance.^{1–4} Comparatively fewer studies examine how AI systems behave after entering routine clinical operations.^{9–11} Implementation-oriented work has highlighted workflow integration, clinician trust, organizational readiness, infrastructure maturity, and accountability as central determinants of sustained use.^{6,10–13,16,17,25–27}

The present study extends this literature by showing that implementation itself is heterogeneous across pathway types. The same institutional environment produced broad cross-specialty expansion in one pathway and deliberately bounded specialist use in others. Implementation research designed around system-wide adoption metrics may obscure important differences between AI pathway types.

The workflow improvements observed during the implementation period were consistent with the operational role of the deployed systems. The pre-consultation pathway was designed to reduce information-collection burden before physician encounters, which provides a plausible mechanism for shorter waiting time and improved consultation efficiency. These improvements occurred alongside broader smart-hospital development, workflow standardization, and institutional digitalization, and should not be interpreted as isolated AI effects.

Documentation-related improvements may reflect a data-quality spillover from AI integration.^{24,25} The 80% improvement in documentation completion efficiency and the 25-percentage-point increase in standardization compliance are unlikely to reflect AI output quality alone. Each pathway required structured data inputs compatible with institutional information systems, which may have reinforced formatting, completeness, and standardization requirements at the point of care. AI deployment may therefore influence surrounding documentation practices not only through direct functional outputs, but also through the data conditions its integration imposes on clinical workflows. This mechanism has received limited attention in implementation studies, where AI-associated improvements are often attributed mainly to system outputs rather than to the structural demands of workflow integration. Describing this mechanism may help explain implementation-associated improvements that cannot be attributed to AI functional output alone.

Confounding Influences and Alternative Explanations

The observed operational improvements occurred within a hospital simultaneously advancing smart-hospital certification, workflow redesign, data governance, and digital quality-control systems. These parallel institutional changes may independently have contributed to improvements in consultation efficiency, documentation quality, and patient satisfaction. The results therefore show favorable movement during AI implementation, not isolated effects attributable to AI alone.

The initial departments were selected partly because they had high outpatient volume, relatively standardized workflows, established electronic documentation practices, and sufficient data readiness. These departments may have been more capable of integrating AI-supported workflows than less standardized clinical units. Observed efficiency gains may therefore partly reflect pre-existing departmental readiness rather than AI deployment alone.

Patient satisfaction increased from 95.40% to 98.92%, but this change emerged from an already high baseline and likely reflects multiple service-related influences, including waiting time, communication quality, institutional expectations, and broader service improvements.²⁸ The available aggregate data do not permit these effects to be separated from AI-supported workflow changes.

Strengths and Practical Implications

The 18-month observation period captured transition from initial deployment into sustained routine operation rather than short-term pilot use. Inclusion of three clinically distinct AI pathways allowed comparison across different workflow positions, accountability structures, and implementation patterns within the same institutional environment. Reliance on routine operational data rather than research-specific deployment conditions reflects implementation under ordinary clinical circumstances.

These findings suggest that implementation evaluation should consider where an AI system enters the workflow, what level of clinical accountability it carries, and what level of oversight and organizational support that accountability requires. Broad diffusion may be appropriate for low-risk workflow-support systems. Bounded specialist deployment may be more appropriate for high-complexity decision-support applications. Treating all clinical AI systems as though they should scale identically risks misunderstanding both their operational role and the oversight they require.

Large-scale clinical AI deployment requires continuing investment in computing resources, cybersecurity, data governance, maintenance, and specialized personnel. These requirements may exceed the independent capacity of smaller institutions.^{29,30} Shared regional infrastructure, cloud-based deployment under institutional data-governance agreements, and staged implementation focused on high-value workflows may provide more feasible pathways for broader adoption.

Limitations

This study was conducted at a single well-resourced tertiary teaching hospital with substantial digital infrastructure and organizational support capacity. Generalizability to smaller or resource-constrained institutions remains uncertain. The observational design did not include randomization, control groups, or patient-level outcome data, precluding causal inference.

Some operational indicators were available only as institutional quality-monitoring summaries rather than raw event-level datasets. Baseline values for consultation efficiency, documentation completion efficiency, and documentation standardization compliance were reported descriptively rather than analyzed through formal comparative statistical approaches.

The analysis focused on operational adoption, workflow indicators, documentation processes, quality-control coverage, and patient satisfaction. Diagnostic accuracy, clinical outcomes, safety events, and cost-effectiveness were not evaluated because the available dataset consisted of de-identified aggregate operational indicators rather than patient-level clinical outcome data.

User acceptance was inferred from sustained utilization patterns, patient satisfaction, and absence of formal complaints rather than validated attitudinal instruments. Future studies incorporating structured clinician and patient evaluation tools would provide more detailed understanding of trust, perceived usefulness, workload impact, and workflow adaptation.

Directions for Future Research

Future implementation research should move beyond binary notions of adoption toward pathway-specific models of operational embedding. Different categories of clinical AI diffuse under different workflow constraints, accountability structures, and organizational conditions. Comparative multi-site studies could determine whether the pathway-stratified adoption patterns observed here are reproduced across hospitals with different levels of digital maturity and organizational capacity.³¹

Further research should evaluate patient-level clinical outcomes, safety indicators, economic impact, and cost-effectiveness.^{29,30,32} Mixed-methods approaches combining operational data with clinician and patient interviews would clarify how different forms of AI integration reshape clinical work over time and how organizational change processes influence sustained adoption.^{33–35} The central question for the field is not only whether clinical AI is adopted, but what kind of AI is adopted, where it enters the workflow, and what operational conditions allow routine use to continue over time.

Conclusion

Routine clinical AI implementation should be understood not as a uniform act of technical deployment, but as a pathway-specific process shaped by workflow position, clinical risk, and accountability requirements. Over 18 months in a provincial tertiary hospital in China, three AI pathways achieved sustained operational use, with broad cross-specialty diffusion for an upstream pre-consultation tool and more bounded specialist integration for tumor decision support and therapeutic drug monitoring. Implementation success therefore should not be judged by scale alone; clinical AI systems require evaluation according to their workflow role, responsibility structure, and operational purpose. The single-site observational design and reliance on aggregate operational data limit causal interpretation, and further multi-site studies are needed to examine clinical outcomes, safety, and cost-effectiveness.

Clinical Trial Registration

This study does not meet the definition of a clinical trial because no participants were prospectively assigned to interventions; registration was therefore not required.

Data Sharing Statement

The data supporting the findings of this study are presented within the manuscript and tables. The study reports de-identified aggregate implementation data. Raw operational data are not publicly available due to institutional data governance policies and restrictions related to clinical information system access.

Ethics Approval and Consent to Participate

This study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Research Ethics Committee of The First Hospital of Hebei Medical University (Approval No. 2026032; February 2026). The requirement for informed consent was waived by the ethics committee because no identifiable patient information was accessed and the study involved secondary analysis of de-identified aggregate operational data.

Acknowledgments

The authors thank the clinical departments, information management teams, quality-management personnel, and platform operation teams of The First Hospital of Hebei Medical University for their support in clinical AI implementation and aggregate operational data management.

Author Contributions

Jin Tian and Jin Zhou share corresponding authorship. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Disclosure

The authors report no conflicts of interest in this work.

References

- Obermeyer Z, Emanuel EJ. Predicting the future — big data, machine learning, and clinical medicine. *N Engl J Med.* 2016;375(13):1216–1219. doi:10.1056/NEJMp1606181
- He J, Baxter SL, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med.* 2019;25(1):30–36. doi:10.1038/s41591-018-0307-0
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7
- Wang F, Casalino LP, Khullar D. Deep learning in medicine — promise, progress, and challenges. *JAMA Intern Med.* 2019;179(3):293–294. doi:10.1001/jamainternmed.2018.7117
- Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging. *Lancet Digit Health.* 2019;1(6):e271–e297. doi:10.1016/S2589-7500(19)30123-2
- Bates DW, Auerbach A, Schulam P, Wright A, Saria S. Reporting and implementing interventions involving machine learning and artificial intelligence. *Ann Intern Med.* 2020;172(11 Suppl):S137–S144. doi:10.7326/M19-0872
- Artsi Y, Sorin V, Glicksberg BS, Korfiatis P, Nadkarni GN, Klang E. Large language models in real-world clinical workflows: a systematic review of applications and implementation. *Front Digit Health.* 2025;7:1659134. doi:10.3389/fgdth.2025.1659134
- Gaber F, Shaik M, Allegra F, et al. Evaluating large language model workflows in clinical decision support for triage and referral and diagnosis. *NPJ Digit Med.* 2025;8(1):263. doi:10.1038/s41746-025-01684-1
- Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. *JMIR Hum Factors.* 2024;11:e48633. doi:10.2196/48633
- Wenderott K, Krups J, Zaruchas F, Weigl M. Effects of artificial intelligence implementation on efficiency in medical imaging: a systematic literature review and meta-analysis. *NPJ Digit Med.* 2024;7(1):265. doi:10.1038/s41746-024-01248-9
- Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet.* 2019;393(10181):1577–1579. doi:10.1016/S0140-6736(19)30037-6
- Norgeot B, Quer G, Beaulieu-Jones BK, et al. Minimum information about clinical artificial intelligence modeling. *Nat Med.* 2020;26(9):1320–1324. doi:10.1038/s41591-020-1041-y
- Park Y, Jackson GP, Foreman MA, Gruen D, Hu J, Das AK. Evaluating artificial intelligence in medicine: phases of clinical research. *JAMIA Open.* 2020;3(3):326–331. doi:10.1093/jamiaopen/ooaa033
- Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc.* 2020;27(3):491–497. doi:10.1093/jamia/ocz192
- Nilsen P, Svedberg P, Neher M, et al. A framework to guide implementation of AI in health care: protocol for a cocreation research project. *JMIR Res Protoc.* 2023;12:e50216. doi:10.2196/50216
- Liao F, Adelaide S, Afshar M, Patterson BW. Governance of clinical AI applications to facilitate safe and equitable deployment in a large health system: key elements and early successes. *Front Digit Health.* 2022;4:931439. doi:10.3389/fgdth.2022.931439
- Abdelwanis M, Simsekler MCE, Gabor AF, Sleptchenko A, Omar M. Artificial intelligence adoption challenges from healthcare providers' perspectives: a comprehensive review and future directions. *Saf Sci.* 2026;193:107028. doi:10.1016/j.ssci.2025.107028
- Wang Z. Artificial intelligence in Chinese healthcare: a review of applications and future prospects. *Biomed Eng Lett.* 2025;15(6):1065–1072. doi:10.1007/s13534-025-00515-2
- Tian J, Zhao Z, Tang L, Song Y, Li Y, Jiang N. From pilot trap to institutional capacity: a governance framework for sustainable clinical AI implementation in health systems. *J Med Internet Res.* 2026;28:e92680. doi:10.2196/92680
- Trinkley KE, An R, Maw AM, Glasgow RE, Brownson RC. Leveraging AI to advance implementation science: potential opportunities and challenges. *Implement Sci.* 2024;19(1):17. doi:10.1186/s13012-024-01346-y
- Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med.* 2018;378(11):981–983. doi:10.1056/NEJMp1714229
- Wang F, Beecy A. Implementing AI models in clinical workflows: a roadmap. *BMJ Evid Based Med.* 2025;30(5):285–287. doi:10.1136/bmjebm-2023-112727
- Liu Y, Yu W, Dillon T. Regulatory responses and approval status of artificial intelligence medical devices with a focus on China. *NPJ Digit Med.* 2024;7(1):255. doi:10.1038/s41746-024-01254-x

24. Feng J, Phillips RV, Malenica I, et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. *NPJ Digit Med.* 2022;5(1):66. doi:10.1038/s41746-022-00611-y
25. Cresswell K, Rigby M, Magrabi F, et al. Strengthening the evaluation of digital clinical decision support interventions through greater consideration of safety and unintended consequences. *Health Policy.* 2023;136:104889. doi:10.1016/j.healthpol.2023.104889
26. Sendak MP, Ratliff W, Sarro D, et al. Real-world integration of a sepsis deep learning technology into routine clinical care: implementation study. *JMIR Med Inform.* 2020;8(7):e15182. doi:10.2196/15182
27. Sandhu S, Lin AL, Brajer N, et al. Integrating a machine learning system into clinical workflows: qualitative study. *JMIR Med Inform.* 2020;8(11):e22421. doi:10.2196/22421
28. Batbaatar E, Dorjdagva J, Luvsannyam A, Savino MM, Amenta P. Determinants of patient satisfaction: a systematic review. *Perspect Public Health.* 2017;137(2):89–101. doi:10.1177/17579139166634136
29. Alami H, Lehoux P, Auclair Y, et al. Artificial intelligence and health technology assessment: anticipating a new level of complexity. *J Med Internet Res.* 2020;22(7):e17707. doi:10.2196/17707
30. Bates DW, Levine DM, Syrowatka A, et al. The potential of artificial intelligence to improve patient safety: a scoping review. *NPJ Digit Med.* 2021;4(1):54. doi:10.1038/s41746-021-00423-6
31. Matheny ME, Whicher D, Thadanev Israni S. Artificial intelligence in health care: a report from the National Academy of Medicine. *JAMA.* 2020;323(6):509–510. doi:10.1001/jama.2019.21579
32. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Digit Health.* 2020;2(10):e537–e548. doi:10.1016/S2589-7500(20)30218-1
33. Nilsen P, Seing I, Ericsson C, Birken SA, Schildmeijer K. Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. *BMC Health Serv Res.* 2020;20(1):147. doi:10.1186/s12913-020-4999-8
34. Fang J, Liu C, Fang P. What is the most important factor affecting patient satisfaction – a study based on gamma coefficient. *Patient Prefer Adherence.* 2019;13:515–525. doi:10.2147/PPA.S197015
35. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res.* 2017;19(11):e367. doi:10.2196/jmir.8775

Risk Management and Healthcare Policy

Publish your work in this journal

Risk Management and Healthcare Policy is an international, peer-reviewed, open access journal focusing on all aspects of public health, policy, and preventative measures to promote good health and improve morbidity and mortality in the population. The journal welcomes submitted papers covering original research, basic science, clinical & epidemiological studies, reviews and evaluations, guidelines, expert opinion and commentary, case reports and extended reports. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/risk-management-and-healthcare-policy-journal>

Dovepress
Taylor & Francis Group