Dai et al. Journal of Digital Management (2025) 1:8
https://doi.org/10.1007/s44362-025-00009-2

Journal of Digital Management

RESEARCH

Open Access

The paradox of AI assistance: enhancing quality while hindering efficiency in local hospitals

Siqi Dai[1], Zhiyi Xie[1], Zheshuai Yang[1*] and Wei Miao[2]

Abstract

Artificial intelligence (AI) is transforming the medical industry, with AI applications in healthcare expanding across clinical domains. By 2025, medical AI is expected to be adopted in 90% of hospitals to support doctors' work. Although AI has demonstrated proven capabilities in enhancing medical diagnosis and treatment efficacy, there remains a lack of in-depth research on its impact on doctors' work, particularly for doctors with average qualifications in small-scale hospitals. Through an analysis of chest CT diagnostic data from a local hospital in China, our analysis reveals that after the introduction of AI assistance, doctors' work quality improved, as evidenced by a 2.8% increase in the length of report conclusions and a 1.0% increase in the description length. However, work efficiency declined, with the average number of chest CT reports processed daily reduced by 4.3% for the overall department and 2.8% per doctor. Notably, over a six-month period following the adoption of AI, this trade-off became increasingly significant. Understanding the impact of AI assistance on doctors' work performance is crucial for optimizing healthcare resource allocation and management decisions, ultimately enhancing patient satisfaction and well-being. This study redirects attention from patient perceptions to clinician behaviors, offering actionable insights for AI implementation in small-scale hospitals.

Keywords Artificial intelligence (AI) assistance, Doctors' work performance, Work efficiency, Work quality, Efficiency-quality trade-off

1 Introduction

Advances in medical technology and improved healthcare access have led to exponential growth in medical data volume, while expectations for service efficiency and quality rise steadily. Propelled by both technological advancements and escalating medical demands, the healthcare industry has transitioned into the era of big data and artificial intelligence (AI). The application of medical AI are becoming increasingly diverse, spanning diverse healthcare domains from diagnosis and treatment to prevention. Currently, many major enterprises and healthcare organizations are investing heavily in the development and application of AI-assisted systems to optimize clinical workflows and enhance treatment outcomes. For instance, IBM has introduced the IBM Watson Oncology System in China (Zhou et al., 2019), assisting clinicians in developing more precise cancer diagnoses and treatment plans, thereby improving diagnostic accuracy and efficiency. Additionally, Google's DeepMind achieved breakthroughs using advanced algorithms to enable early diagnosis of diabetic retinopathy and other ocular diseases (Gulshan et al., 2016). Projections suggest 90% of hospitals will adopt physician-assistance AI by 2025, yielding over $150 billion in healthcare savings (Das, 2016; Sullivan, 2018).

*Correspondence:
Zheshuai Yang
yangzheshuai@zju.edu.cn
[1] Zhejiang University, Hangzhou, China
[2] University College London, London, UK

In this context, optimizing AI-assisted systems to enhance physician support and advance healthcare intelligence represents a critical research priority. Both clinical efficiency and care quality fundamentally shape patient evaluations of hospital care standards (Arocena & Garcia-Prado, 2007). Additionally, improving both healthcare efficiency and quality serve as primary motivators for physician adoption of AI tools (Li & Xiong, 2024). Therefore, exploring the impact of AI assistance on doctors' work efficiency and quality is essential for progressively optimizing and effectively utilizing these systems, and carries significant theoretical and practical implications.

Previous research on healthcare AI applications has focused primarily on clinical diagnostic accuracy, precision, scalability, and patient acceptance of AI-assisted systems (Chen, et al., 2021; Longoni et al., 2019). However, scant research examines AI's actual impact on physician workflows, with most studies confined to urban tertiary hospitals. Physicians constitute critical mediators between diagnostic processes and patients, directly influencing clinical outcomes. Therefore, this study redirects scholarly attention from patient to physician behaviors, systematically examining AI's effects on both workflow efficiency and clinical quality. The extant literature reveals no empirical consensus about AI's specific impacts on clinical workflows. Most scholars posit that AI enables faster, more accurate diagnoses, potentially improving efficiency and quality simultaneously (Esteva et al., 2017; Topol, 2019). However, most studies conceptualize efficiency and quality as isolated dimensions, rarely examining how AI assistance affects both dimensions simultaneously. This critical gap overlooks the fundamental efficiency-quality trade-off physicians face given clinical resource limitations and temporal constraints.

Our research posits that physicians encounter a novel efficiency-quality trade-off in AI collaboration, mediated by factors including AI trust and clinical experience. This trade-off constitutes a fundamental healthcare challenge, creating an inverse relationship where care quality improvements typically reduce work efficiency (operationalized as case throughput), and vice versa (Yang & Zeng, 2014). For instance, when physicians allocate more time per patient for diagnostic accuracy and detailed care, their overall patient volume decreases, impacting efficiency. We argue that while AI assistance improves work quality, it may concurrently decrease work efficiency. This trade-off is especially pronounced among less experienced physicians.

Accordingly, this study examines AI assistance's impact on physician performance, contributing to both scholarly research and clinical practice. First, understanding how AI assistance affects physician performance is crucial for optimizing healthcare resource allocation and management decisions. Our findings advance current knowledge about medical AI applications. While prior healthcare AI studies examined either efficiency gains (Ardila et al., 2019; Topol, 2019) or quality improvements (Esteva et al., 2017; Katwaroo et al., 2024) separately, we uniquely investigate their interaction and identify a new efficiency-quality trade-off. By analyzing these physician-AI collaboration trade-offs, developers can strategically optimize systems, and healthcare managers can make informed decisions about healthcare resource allocation or formulate more effective training strategies. This dual approach not only enhances the overall quality of healthcare services but also achieves optimal benefits under resource constraints, thereby promoting the sustainable development of the healthcare system. Second, contrasting with studies focused on well-resourced urban hospitals (McKinney et al., 2020; Raimondi et al., 2020), we examine AI assistance in small hospitals where doctors may have less clinical experience. This context is crucial as these hospitals serve large populations yet encounter distinct technology adoption challenges. Our findings challenge the direct applicability of large-hospital findings to smaller settings, revealing how hospital scale and physician experience influence AI outcomes. Finally, while existing literature has primarily examined immediate post-implementation effects (Li et al., 2022), our research tracks performance changes over an extended period (180 days pre- and post-implementation). This extended timeframe captures both initial impacts and adaptation patterns, offering a more comprehensive understanding of AI's performance effects.

2 Literature review

2.1 The application of artificial intelligence assistance in the medical field

Artificial intelligence (AI) is reshaping healthcare, generating extensive and profound impacts across the medical field. AI's advanced capacity to process medical data, text, images, and biological information has led to increasingly diverse and widespread healthcare applications. These encompass medical imaging analysis, personalized medicine, predictive analytics, virtual health assistants, surgical robots, and natural language processing, among others (Haleem et al., 2019). Although research has consistently demonstrated these robust AI tools' effectiveness, their clinical implementation remains limited (Wiens et al., 2019). Currently, AI demonstrates substantial impact in three primary domains—medical imaging analysis, personalized medicine, and predictive analytics (Ahmed & Adil, 2023)—all supporting physicians in diagnosis and treatment decisions.

The proliferation of AI assistance in medicine stems from its capacity to attain diagnostic accuracy comparable to medical experts (Longoni et al., 2019; Rajpurkar et al., 2022). For instance, AI demonstrates dermatological diagnostic accuracy through image analysis that matches or exceeds board-certified dermatologists (Leachman & Merlino, 2017). Additionally, deep learning enables AI to automatically identify and localize thoracic lesions in radiographs with performance rivaling seasoned radiologists (Rajpurkar et al., 2018). With expanding healthcare access, researchers prioritize developing clinically viable AI applications that optimize diagnostic workflows, improve accuracy and efficiency, and enhance patient outcomes and satisfaction (Khullar et al., 2022; Kulkarni & Singh, 2023; Nelson et al., 2020; Pierre et al., 2023).

Management and marketing scholars have extensively investigated patient attitudes and compliance regarding healthcare AI. Some scholars contend that uniqueness neglect makes consumers more reluctant to use AI providers versus human providers (Longoni et al., 2019). Neuroscience studies demonstrate this divergence, showing distinct brain activation patterns when patients receive identical personalized conversations from AI versus human providers (Yun et al., 2021). Practically, some research finds patients perceive no difference between AI-assisted and unaided physicians (Chen, et al., 2021. However, perceived physician effort may alter this perception (Chen et al., 2024). Thus, patients may report higher satisfaction when perceiving greater effort from AI-assisted physicians.Conversely, other studies indicate technological risks, ethical concerns, and communication barriers may decrease trust in AI-assisted versus unaided physicians (Esmaeilzadeh, 2020; Zhang et al., 2021).

While these findings illuminate patient perspectives, significant gaps persist in understanding AI's impact on healthcare providers. What specific impacts does AI assistance have on physician workflows? Management research has largely overlooked the crucial examination of AI's influence on physician work patterns and behaviors. Since physicians serve as both primary AI users and critical mediators between diagnosis and patients, comprehending AI's workflow impacts is essential for optimizing healthcare delivery and patient outcomes. This study redirects scholarly attention from patient to physician behaviors, systematically investigating AI's effects on clinical workflows.

2.2 The impact of artificial intelligence assistance on doctors' work performance

2.2.1 The efficiency-quality trade-off

The optimal medical practice model prioritizes high-quality, high-efficiency, patient-centered, and collaborative care (Moore, 2007). Arocena and Garcia-Prado (2007) contend that hospital performance improvements stem from simultaneous quality and efficiency enhancements. Moreover, physicians frequently adopt AI tools to augment both work output and quality (Li & Xiong, 2024). This highlights the dual importance of efficiency and quality in clinical practice.

However, resource-constrained healthcare systems inevitably face efficiency-quality trade-offs (Newhouse, 1970). Physicians confront a fundamental dilemma: enhancing diagnostic and treatment quality necessitates greater time investment, often reducing case throughput. Conversely, prioritizing higher output elevates medical error risks, compromising care quality and patient outcomes (Bao & Bardhan, 2022; Kane et al., 2007).

AI's emergence in medicine offers potential solutions to the efficiency-quality trade-off. This potential arises from AI's dual capacity for both substituting and complementing human capabilities (Krakowski et al., 2023). On one hand, AI assumes responsibility for repetitive clinical tasks including medical image interpretation, data analytics, and report generation (Imran et al., 2020; Jorg et al., 2024; McKinney et al., 2020). On the other hand, AI augments human limitations by detecting subtle anomalies and processing large datasets rapidly (Ardila et al., 2019; Attia et al., 2019; Doyle et al., 2020). This coexistence of substitution and complementarity seems to enable concurrent improvements in both clinical efficiency and quality.

However, AI's actual clinical impact depends on multiple factors: system performance, physician adaptation capacity, trust levels, and institutional management approaches. These factors may sustain efficiency-quality trade-offs in physician-AI collaboration. Prior research in tertiary hospitals suggests AI improves both physician efficiency and quality (Katwaroo et al., 2024; Li et al., 2022; Topol, 2019). However, these findings involve above-average qualified practitioners.AI's expansion to small-scale hospitals raises questions about replicability among less-qualified physicians. This study examines AI's effects on work behaviors in this physician cohort.

2.2.2 The impact of artificial intelligence assistance on doctors' work efficiency

A key rationale for AI integration in clinical practice involves efficiency enhancement. AI's productivity benefits have been empirically validated across domains, especially in complex healthcare settings (Kellogg et al., 2020). In a study involving over 300 doctors regarding their views on AI-assisted diagnosis, the results revealed that 81.16% of them believed AI-assisted diagnosis could "reduce their workload," while 78.55% felt it could "increase diagnostic efficiency" (Huang et al., 2021). These findings demonstrate widespread clinician recognition of AI's efficiency-enhancing potential.

Extensive research demonstrates that AI assistance significantly enhances physician efficiency, as evidenced by numerous empirical studies. For instance, in the realm of medical imaging analysis, where AI is most commonly applied, its ability to recognize complex patterns in images allows for a transformation of image interpretation from a purely qualitative and subjective task to a quantifiable and replicable process (Bi et al., 2019). AI-assisted screening accelerates positive case identification by approximately 33% compared to unaided evaluation (Nature Editorial Team, 2023). This accelerated throughput proves particularly valuable for time-sensitive cases, facilitating prompt diagnosis and intervention. Furthermore, AI enables faster, more accurate clinical decisions through automated processing of extensive patient datasets (Chen et al., 2021; Topol, 2019). For example, AI can scan hundreds of medical images and identify potential disease risks within minutes (Ardila et al., 2019), providing recommendations that are comparable to those of experts (McKinney et al., 2020), thereby directly improving the overall efficiency of the healthcare system. Moreover, AI's detection of subvisual abnormalities reduces clinician cognitive load, further optimizing workflow efficiency (Sathykumar & Munoz, 2020).

While extensive research supports AI's efficacy in improving physician efficiency, concerns persist that AI implementation may increase workload and impede efficiency gains (Yoo et al., 2023). In fact, AI's effectiveness is contingent upon multiple factors, particularly physician qualifications and trust levels (Tong et al., 2023; Wang et al., 2024) For instance, some studies have indicated that after collaborating with AI, the efficiency of producing diagnostic reports improved by 20.7% for junior doctors and 18.8% for senior doctors, with less experienced junior doctors benefiting more from AI assistance (Wei et al., 2022). However, it is noteworthy that AI-assisted junior physicians still exhibit longer average diagnostic times than their senior counterparts (Tong et al., 2023). Prior research has predominantly examined highly-qualified physicians in urban tertiary hospitals (Li et al., 2022), who demonstrate superior learning capacity, adaptability, and AI trust levels. As AI assistance becomes more widely adopted in healthcare settings, a crucial question arises: when AI assistance is introduced in small hospitals, can it still enhance the work efficiency of less qualified doctors?

Prior research has utilized diverse efficiency metrics across multiple levels to assess physician productivity. Patient-centered measures like length of hospitalization (Bao & Bardhan, 2022; Tasi et al., 2019) have been used to evaluate clinical efficiency. Alternatively, physician-centered metrics such as mean diagnosis time (Overhage & McCallie, 2020; Reuben et al., 2014) have quantified workflow efficiency. Our study measures physician efficiency through output quantification, specifically diagnostic report volume. With fixed working hours, higher daily report output per department or physician indicates reduced per-case time investment, thus demonstrating greater efficiency.

2.2.3 The impact of artificial intelligence assistance on doctors' work quality

Beyond efficiency gains, improving work quality represents another primary motivation for incorporating AI assistance in professional practice. In clinical settings, AI enhances physician work quality primarily through its expert-level diagnostic accuracy and precision (Esteva et al., 2017; Wu et al., 2023; Xu et al., 2024). On the one hand, AI improves abnormality detection sensitivity and accuracy, thereby reducing physician misdiagnosis and missed diagnosis rates (Katwaroo et al., 2024; Li et al., 2022). These improvements enhance care quality while facilitating earlier interventions, ultimately lowering healthcare expenditures and reducing government insurance costs (Khalifa & Albadawy, 2024). On the other hand, AI's advanced image recognition and data processing capabilities minimize time spent on routine tasks, enabling greater focus on personalized treatment planning and thereby enhancing overall work quality (Krishnan et al., 2023; Patankar, 2024).

The impact of AI assistance on physician work quality similarly depends on both system accuracy/performance and clinical utilization patterns. We posit that AI's quality-enhancing effects may be particularly significant for less experienced clinicians. Research demonstrates that AI systems significantly enhance junior radiologists' diagnostic performance, whereas mid-level and senior physicians show negligible differences between AI-assisted and independent assessments (Wang et al., 2022; Xu et al., 2024). Furthermore, AI assistance enables junior physicians to attain diagnostic accuracy comparable to senior colleagues (Wei et al., 2022). Consequently, this study focuses on less-qualified physicians to examine whether AI assistance disproportionately enhances their work quality.

However, work quality proves more conceptually complex and methodologically challenging to standardize than efficiency metrics (Yang & Zeng, 2014). Academic literature typically dichotomizes healthcare quality into technical quality and patient-perceived quality. Technical quality encompasses clinical diagnostic accuracy and treatment efficacy, whereas patient-perceived quality reflects service satisfaction (Navarro-Espigares & Torres, 2011). Prior studies have operationalized quality measurement through patient readmission rates (Bao & Bardhan, 2022; Janakiraman et al., 2023). However, this metric suffers from confounding by extraneous variables including baseline health status, treatment adherence, and socioeconomic factors, thereby providing only indirect performance assessment. This study employs diagnostic reports as a more direct physician quality metric.

Diagnostic reports constitute essential medical documentation that comprehensively records physicians' diagnostic conclusions (Zhang et al., 2021). As direct outputs of clinical diagnosis, these reports objectively reflect physician work quality. Growing evidence indicates patient demand for such documentation to facilitate self-management and shared decision-making (Ross et al., 2005). The technical composition of diagnostic reports - characterized by specialized terminology and standardized language (Rad-Insights, 2024)- enables report length to function as a proxy for diagnostic thoroughness and specificity. Consequently, this metric effectively captures diagnostic quality across the clinical evaluation process.

This study examines AI assistance's dual effects on work efficiency and quality among physicians in small-scale hospitals with limited experience. We aim to investigate potential efficiency-quality trade-offs in physician-AI collaboration within these settings. For average-qualified physicians, AI aids in detecting sub-visual abnormalities and provides expert-level diagnostic references as a "second opinion" (Li et al., 2023) enhancing diagnostic quality. However, AI implementation may increase cognitive burden, particularly during diagnostic disagreements with AI recommendations. Trust limitations may prompt additional verification time for AI results to ensure patient safety. Furthermore, adaptation periods may be needed to optimize physician-AI interaction protocol. Based on these considerations, we propose the following hypotheses:

H1: The introduction of AI assistance will lead to a decrease in doctors' work efficiency.
H2: The introduction of AI assistance will result in an improvement in doctors' work quality.

3 Context and data

3.1 Context

The International Agency for Research on Cancer (IARC/WHO) reports lung cancer as the most common malignancy worldwide, responsible for 1.8 million deaths (18.7% of cancer mortality) (World Health Organization, 2024). Low-dose computed tomography (CT) has emerged as the standard screening modality for pulmonary conditions. However, expanding CT utilization in both healthy populations and symptomatic individuals has substantially increased radiologists' workload. This has exacerbated quality-efficiency trade-offs, manifesting as reduced efficiency, elevated false-positive rates, and interpretation variability (Huang et al., 2021).

(Pesapane et al., 2018). AI's advanced image interpretation capabilities are anticipated to enhance radiologists' diagnostic accuracy and timeliness (Ahmed & Adil, 2023). Numerous hospitals have adopted AI diagnostic tools to improve radiologists' workflow efficiency and output quality. However, empirical evidence regarding AI's impact on radiologists' lung CT interpretation performance remains lacking.

In this analysis, we focus on CT scans from a focal hospital located in a county in China. As the county's primary medical center, the facility manages substantial daily patient throughput. Since May 2021, the hospital has implemented an AI system employing deep learning for image classification, lesion detection, and segmentation. This technology facilitates lesion identification and annotation, 3D (Three-Dimensional) reconstruction, and automatic target delineation, aiding in the screening, diagnosis, and treatment of diseases. Currently, the system providers connect their localized servers to the hospital's existing PACS (Picture Archiving and Communication System) system via the DICOM (Digital Imaging and Communications in Medicine) imaging protocol, displaying AI-generated disease detection results on doctors' workstations. Consequently, CT scans are now supported by this AI-assisted system.

In the process of lung CT diagnosis, doctors first need to import the CT images into the AI software, which then begins to identify potential lesions and lists the suspicious ones individually. Physicians can review each flagged lesion to make final diagnostic determinations. The system allows adjustment of detection sensitivity, typically configured to exclude nodules that are 3 mm or smaller.

. We would like to understand the causal effect of introducing AI assistance on the work of these radiologists: (1) extensive margin (i.e., the quantity of reports processed each day); and (2) intensive margin (i.e., the quality of the reports produced).

3.2 Data

Following the hospital's AI implementation on May 20, 2021, we collected chest CT diagnostic data spanning

November 21, 2020 to November 16, 2021 - encompassing 180-day pre- and post-implementation periods centered on the intervention date. The dataset comprises multiple variables: patient demographics (ID, sex, age), clinical information (requesting department, examination date/site), and physician documentation (description and conclusion).

To assess doctors' work efficiency, we aggregated the scan-level data into a department-day level, so that we can investigate whether, after AI was introduced, doctors could process more scan reports on a daily basis. We created two outcome variables for the extensive margin: (1) log(total_scan + 1), representing the logarithm of total number of CT scan reports per day; (2) log(avg_scan_per_doc + 1), representing the logarithm of average number of CT scan reports per doctor per day. The latter was calculated by dividing the total number of CT scan reports by the number of doctors working that day before taking the logarithm.

To assess doctors' work quality, we aggregated the scan-level data into a doctor-day level panel dataset. We focus on the lengths of CT scan reports as an indicator for the intensive margin, assuming longer reports may reflect more detailed analysis. Since the raw data included both the conclusion and the description of the report written by the doctor, we also created two outcome variables for the intensive margin: (1) log(report_lengh_concl), representing the logarithm of the length of the doctor's conclusion text; (2) log(report_length_desc), representing the logarithm of the length of the doctor's description text. These variables allow us to examine potential changes in the depth and detail of doctors' analyses following the introduction of AI assistance.

While we utilize CT report volume and length as efficiency and quality proxies respectively, we recognize their inherent limitations. For efficiency assessment, report volume may not fully reflect case complexity or diagnostic time investment, as simple counts cannot capture workload distribution nuances. Similarly, for quality evaluation, while report length indicates documentation completeness, it may not perfectly correlate with diagnostic accuracy or clinical utility. Although longer reports typically reflect more comprehensive documentation, they do not invariably signify superior diagnostic quality. Despite these limitations, these metrics offer quantifiable measures enabling examination of AI's impact on physician performance, particularly given the challenges in direct clinical measurement. Future research might benefit from incorporating additional metrics such as diagnostic accuracy rates, peer review scores, or patient outcomes to provide a more comprehensive evaluation of AI's effect on clinical work.

4 Main analysis

4.1 Empirical strategy

Given the absence of a true control group in this setting, we were unable to employ a difference-in-differences design. Instead, we adopted the regression discontinuity in time (RDiT) design (Hausman & Rapson, 2018). RDiT is an econometric method used for causal inference, particularly suited for analyzing time-series data. This method has been extensively applied in marketing and economics research examining policy changes (Busse et al., 2010; Chen et al., 2009; Liu & Cong, 2023). RDiT extends conventional regression discontinuity design (RDD) to temporal contexts where time serves as the running variable. This approach identifies causal effects by exploiting a discontinuity or threshold in a policy, intervention, or treatment that occurs at a specific point in time. In our setting, first, the introduction of AI in the hospital creates a discrete time threshold – a natural "before" and "after" comparison opportunity – which is crucial for RDiT analysis, allowing us to explore changes in doctor's work performance immediately around the focal event date. Second, RDiT is well-suited for our setting where finding a comparable control hospital would be challenging. Third, since all doctors in the radiology department were required to use the AI assistance after its introduction, RDiT helps address potential selection bias issues.

The RDiT regression we conducted is as follows:

$$Y_{it} = \beta_0 + \delta Post_t + Controls_{it} + \epsilon_{it}$$

where we take a short time window before and after the policy change. The coefficient of interest, δ, measures the before-after difference, that is, the effect of the introduction of AI assistance on doctors' work performance. The indicator variable $Post_t$ takes a value of 1 after the focal event date and a value 0 before the focal event date. $Controls_{it}$ are control variables such as fixed effects.

4.2 Treatment effects on extensive margin

Our department-day level dataset contains 10,206 observations (see Table 1). We conducted Regression Discontinuity in Time (RDiT) analyses using log(total_scan + 1) and log(avg_scan_per_doc + 1) as dependent variables, with results presented in Table 1. We found that the introduction of AI had an unexpected impact on doctors' work efficiency. After controlling for fixed effect of requesting departments, we discovered that after the introduction of AI, the average number of chest CT reports processed daily by the CT department significantly decreased by approximately 4.3% ($\delta = -0.043$, $p < 0.001$). Similarly, the average number of chest CT reports processed daily per

Table 1 Treatment Effects on Extensive Margin

	Lhs: log(total_scan + 1)	Lhs: log(avg_scan_per_doc + 1)
Post	-.043*** (.009)	-.028*** (.004)
Num. Obs	10,206	10,206
R^2	.686	.628
R^2 Adj	.683	.625
R^2 Within	.003	.004
R^2 Within Adj	.002	.004
AIC	10,578.8	−3234.2
BIC	11,229.5	−2583.5
RMSE	.40	.20
Std. Errors	Heteroskedasticity-robust	Heteroskedasticity-robust
Fixed Effect: Requesting department	×	×

Notes: This table reports estimates from two regression models based on the department-day level log-transformed average number of chest CT reports processed daily by the department and per doctor from 180 days before to 180 days after the focal event date. The coefficient of Post (× 100%) is the percentage change in a dependent variable after the focal event date. All specifications include fixed effect of requesting department

+ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

doctor in the department also significantly decreased by about 2.8% ($\delta = -0.028$, $p < 0.001$). These findings suggest that after the introduction of AI, there was a reduction in the quantity of reports (extensive margin of work). The two models explained 68.6% and 62.8% of the variance in the dependent variables, respectively, indicating a good model fit.

4.3 Treatment effects on intensive margin

Our doctor-day level panel dataset contains 46,690 observations (see Table 2). We conducted RDiT analyses using log(report_lengh_concl) and log(report_length_desc) as dependent variables, with results presented in Table 2. Consistent with our hypothesis, after controlling for fixed effects of doctors and requesting departments, we found that the introduction of AI assistance had a positive impact on the intensive margin (quality of report) of doctor's work. Specifically, after the introduction of AI, the length of the conclusion in CT reports significantly increased by about 2.8% ($\delta = 0.028$, $p < 0.001$). The length of the description section also showed a significant increase of about 1.0% ($\delta = 0.010$, $p < 0.01$). These findings suggest that after the introduction of AI, there was an improvement in the detail and potentially the quality of the reports, as indicated by their increased length. The two models explained 29.9% and 25.3% of the variance in the dependent variables, respectively. While these R-squared values are lower than those in our previous models, they still indicate that a substantial portion of the variation in report length is accounted for by our predictors.

4.4 Dynamic treatment effects

To better understand how the impact of AI assistance evolves over time, we conducted additional analyses examining the temporal effects of AI assistance adoption. We constructed month dummies for each post-treatment month and analyzed the dynamic treatment effects on both extensive and intensive margins.

For the extensive margin, our results reveal an increasingly negative impact on work efficiency over time (see Table 3). While the immediate effect of AI adoption (Post) showed no significant decrease in the total number of CT scan reports per day (or per doctor per day), the negative impacts gradually emerged and intensified over subsequent months. Specifically, during the first three months, the treatment effects were small and statistically insignificant. However, starting from the fourth month, we observed a significant decrease of 6.9% ($\delta = -0.069$, $p < 0.001$) in the average number of chest CT reports processed daily by the CT department. This negative effect progressively intensified, reaching 10.3% ($\delta = -0.103$, $p < 0.001$) in the sixth month. Similar patterns were observed in the average number of scans processed daily per doctor, with the negative effect becoming significant from the fourth month ($\delta = -0.028$ $p < 0.010$)

Table 2 Treatment Effects on Intensive Margin

	Lhs: log(report_lengh_concl)	Lhs: log(report_length_desc)
Post	.028*** (.006)	.010** (.003)
Num. Obs	46,690	46,690
R^2	.299	.253
R^2 Adj	.297	.251
R^2 Within	.000	.000
R^2 Within Adj	.000	.000
AIC	85,990.2	24,697.1
BIC	86,961.6	25,668.5
RMSE	.61	.31
Std. Errors	Heteroskedasticity-robust	Heteroskedasticity-robust
Fixed Effect: Doctor	×	×
Fixed Effect: Requesting department	×	×

Notes: This table reports estimates from two regression models based on the doctor-day level log-transformed average length of the doctor's conclusion text and description text from 180 days before to 180 days after the focal event date. The coefficient of Post (×100%) represents the percentage change in a dependent variable after the focal event date. All specifications include fixed effects for both doctor and requesting department

+ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

and increasing to 3.0% (δ = −0.030 $p < 0.010$) by the sixth month. The two models explained 68.7% and 62.9% of the variance in the dependent variables, respectively, demonstrating strong explanatory power.

Regarding the intensive margin, we observed a contrasting trend of gradually increasing positive effects on work quality (see Table 4). The length of the conclusion in CT reports showed a progressive increase over time, starting from an insignificant effect in the first month and reaching a significant increase of 7.3% (δ = 0.073, $p < 0.001$) by the sixth month. Similarly, the description length demonstrated a gradual improvement, with the effect becoming marginally significant in the fourth month and increasing to 2.6% (δ = 0.026, $p < 0.010$) by the sixth month. The two models explained 29.9% and 25.3% of the variance in the dependent variables, respectively.

These temporal patterns reveal a progressively intensifying efficiency-quality trade-off during the six-month post-implementation period. The unexpected efficiency decline - contrary to anticipated improvements with system familiarity - likely reflects physicians' heightened AI dependence and more rigorous verification practices. Meanwhile, sustained quality improvements demonstrate physicians' progressive mastery of leveraging AI for enhanced diagnostic documentation. These findings offer critical managerial implications, indicating AI integration may necessitate extended adaptation periods beyond initial expectations. While our six-month observation period reveals a persistent efficiency-quality trade-off that appears to intensify over time, longer-term studies are needed to determine whether this trade-off eventually stabilizes, intensifies further, or potentially diminishes as doctors and healthcare systems optimize their integration of AI assistance.

5 Robustness checks

To verify the robustness of our findings, we conducted several additional analyses.

First, we re-ran our main analyses without the log transformation of the dependent variables. The results remained consistent with our main findings. For the extensive margin (see Table 5), we found that the introduction of AI assistance led to a significant decrease in both the average number of chest CT reports processed daily by the CT department (δ = −0.299, $p < 0.001$) and the average number of reports processed daily per doctor (δ = −0.090, $p < 0.001$). Similarly, for the intensive margin (see Table 6), we observed significant increases in both the length of report conclusions (δ = 1.511, $p < 0.001$) and descriptions (δ = 1.318, $p < 0.010$). These findings

Table 3 Temporal Effects of AI Assistance Adoption on Extensive Margin

	Lhs: log(total_scan + 1)	Lhs: log(avg_scan_per_doc + 1)
Post	.000 (.015)	-.011 (.008)
Month 2	-.009 (.019)	-.011 (.010)
Month 3	-.028 (.019)	-.010 (.010)
Month 4	-.069*** (.020)	-.028** (.010)
Month 5	-.075*** (.019)	-.029** (.009)
Month 6	-.103*** (.019)	-.030** (.009)
Num. Obs	10,206	10,206
R^2	.687	.629
R^2 Adj	.684	.625
R^2 Within	.007	.006
R^2 Within Adj	.006	.005
AIC	10,547.6	−3239.1
BIC	11,234.5	−2552.2
RMSE	.40	.20
Std. Errors	Heteroskedasticity-robust	Heteroskedasticity-robust
Fixed Effect: Requesting department	×	×

Notes: This table reports estimates from two regression models based on the department-day level log-transformed average number of chest CT reports processed daily by the department and per doctor from 180 days before to 180 days after the focal event date. The coefficients of Post and Months 2–6 (× 100%) represent the percentage changes in a dependent variable immediately after the focal event date and in the following months, respectively. All specifications include fixed effect of requesting department

+ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

demonstrate that the directional effects and statistical significance of our main results were robust to alternative specifications of the dependent variables.

Second, considering that our count dependent variable, total number of CT scan reports per day, contained zeros, we conducted a Poisson Fixed Effects regression model as an alternative specification (Chen & Roth, 2024; see Table 7). The results from this model further corroborated our main findings. Specifically, the Poisson estimation showed that the introduction of AI was associated with a significant decrease of 6.4% ($\delta = -0.064$, $p < 0.001$) in the total number of CT scan reports processed daily, which aligned with the negative impact we observed in our main analysis.

These additional analyses demonstrate the robustness of our findings across different model specifications. The consistency of results across various estimation approaches – including OLS without log transformation and Poisson Fixed Effects regression – strengthens our confidence in the main conclusion that the introduction of AI led to a trade-off between improved work quality and decreased work efficiency.

6 Discussion and conclusion

Despite AI's advanced capabilities, its implementation may not fully resolve healthcare's persistent efficiency-quality trade-off. As a novel clinical resource, AI's effectiveness ultimately depends on human-AI collaboration dynamics. Therefore, AI's impact may differ by hospital size, physician qualifications, and medical specialty. We analyzed chest CT diagnostic data from a county-level central hospital in China. In this hospital, the overall qualifications of the doctors in the CT department are relatively average. Our analysis revealed that AI implementation improved work quality marginally but reduced efficiency. These findings demonstrate the enduring efficiency-quality trade-off despite AI integration. This fundamental healthcare challenge persists, where quality gains compromise efficiency. Contrary to initial expectations, AI may not simultaneously enhance both

Table 4 Temporal Effects of AI Assistance Adoption on Intensive Margin

	Lhs: log(report_lengh_concl)	**Lhs: log(report_length_desc)**
Post	.005 (.010)	.005 (.005)
Month 2	.008 (.014)	.000 (.007)
Month 3	.022 (.014)	-.005 (.007)
Month 4	.032* (.014)	.014+ (.007)
Month 5	.050*** (.015)	.015+ (.008)
Month 6	.073*** (.015)	.026** (.008)
Num. Obs	46,690	46,690
R^2	.299	.253
R^2 Adj	.298	.251
R^2 Within	.001	.001
R^2 Within Adj	.001	.001
AIC	85,971.7	24,687.8
BIC	86,986.9	25,702.9
RMSE	.61	.31
Std. Errors	Heteroskedasticity-robust	Heteroskedasticity-robust
Fixed Effect: Doctor	×	×
Fixed Effect: Requesting department	×	×

Notes: This table reports estimates from two regression models based on the doctor-day level log-transformed average length of the doctor's conclusion text and description text from 180 days before to 180 days after the focal event date. The coefficients of Post and Months 2–6 (× 100%) represent the percentage changes in a dependent variable immediately after the focal event date and in the following months, respectively. All specifications include fixed effects for both doctor and requesting department

+ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

dimensions, especially in small hospitals with less experienced physicians. In other words, AI appears to create a new efficiency-quality tension for physicians, intensifying over the six-month post-implementation period.

Although our study documents the efficiency-quality trade-off in AI-assisted diagnosis, methodological constraints prevent direct examination of its underlying mechanisms. Specifically, key psychological factors - including physicians' cognitive load during AI interaction, trust development in AI systems, and their temporal dynamics - remain unmeasured in our quantitative framework. This limitation reflects practical barriers to collecting granular psychobehavioral data in operational clinical environments. In light of these constraints, we propose several potential mechanisms that could explain the observed trade-off, which future research endeavors may aim to investigate and verify.

The quality-enhancing effects of AI assistance are well-documented in clinical research. Three primary mechanisms explain these quality improvements. First, AI's exceptional image processing precision enables detection of subvisual abnormalities and small lesions frequently missed in human interpretation. This reduces diagnostic omissions and improves detection thoroughness. For example, empirical evidence demonstrates particular efficacy in mammography and CT-based early lesion detection (Ardila et al., 2019; Gulshan et al., 2016). Second, AI's data-driven learning provides expert-level "second opinions," particularly benefiting junior clinicians (Li et al., 2023). Through multimodal data analysis, AI delivers clinical insights that enhance diagnostic accuracy in complex cases (Wang et al., 2022; Xu et al., 2024). Third, AI promotes protocol adherence, reducing practice irregularities and subjective errors (McKinney et al., 2020; Rajpurkar et al., 2018). Additionally, integrated tracking features enable retrospective decision analysis, facilitating continuous diagnostic improvement.

Table 5 OLS Results without Log Transformation for Extensive Margin

	Lhs: total_scan	Lhs: avg_scan_per_doc
Post	-.299*** (.064)	-.090*** (.016)
Num. Obs	10,206	10,206
R^2	.734	.677
R^2 Adj	.732	.674
R^2 Within	.002	.003
R^2 Within Adj	.002	.003
AIC	51,833.4	23,300.5
BIC	52,484.2	23,951.2
RMSE	3.04	.75
Std. Errors	Heteroskedasticity-robust	Heteroskedasticity-robust
Fixed Effect: Requesting department	×	×

Notes: This table reports estimates from two regression models based on the department-day level analysis of the average number of chest CT reports processed daily by the department and per doctor from 180 days before to 180 days after the focal event date. The coefficient of Post (× 100%) represents the percentage change in a dependent variable after the focal event date. All specifications include fixed effect of requesting department

+ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

Contrary to expectations, we observed AI implementation reduced radiologist efficiency in this hospital setting. Notably, this efficiency decline intensified temporally, contradicting anticipated improvements with system familiarity. Three key factors explain this efficiency reduction. First, AI increased cognitive burden for junior radiologists. Diagnostic conflicts with AI outputs induced self-doubt, prompting more cautious verification behaviors that prolonged diagnostic time (Hsieh, 2023). Second, limited system understanding reduced trust and accuracy assessment capacity (Wang et al., 2024). Consequently, this prompted excessive time investment in outcome verification for patient safety. Third, new collaboration requirements with AI disrupted existing clinical workflows. Optimal interaction patterns required extended practical exploration (Davenport & Glaser, 2022). Over time, while familiarity improved, it paradoxically increased AI dependence and verification time allocation. Junior radiologists' weaker foundational skills further slowed adaptation, compounding efficiency losses.

This study provides new insights into AI assistance's impact on physicians' work and offers clinical practice recommendations. While AI clearly enhances diagnostic quality, balancing efficiency and quality remains a critical implementation challenge. Healthcare institutions

Table 6 OLS Results without Log Transformation for Intensive Margin

	Lhs: report_lengh_concl	Lhs: report_length_desc
Post	1.511*** (.327)	1.318** (.415)
Num. Obs	46,690	46,690
R^2	.274	.247
R^2 Adj	.273	.245
R^2 Within	.000	.000
R^2 Within Adj	.000	.000
AIC	456,131.4	477,831.9
BIC	457,102.8	478,803.3
RMSE	31.92	40.28
Std. Errors	Heteroskedasticity-robust	Heteroskedasticity-robust
Fixed Effect: Doctor	×	×
Fixed Effect: Requesting department	×	×

Notes: This table reports estimates from two regression models based on the doctor-day level analysis of the average length of the doctor's conclusion text and description text from 180 days before to 180 days after the focal event date. The coefficient of Post (× 100%) represents the percentage change in a dependent variable after the focal event date. All specifications include fixed effects for both doctor and requesting department

+ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

Table 7 Results from Poisson Fixed Effects Regression

	Lhs: total_scan
Post	-.064*** (.014)
Num. Obs	10,206
R^2	.472
R^2 Adj	.470
R^2 Within	.001
R^2 Within Adj	.001
AIC	42,120.1
BIC	42,770.8
RMSE	3.04
Std. Errors	Heteroskedasticity-robust
Fixed Effect: Requesting department	×

Notes: This table reports estimates from a Poisson Fixed Effects regression model based on the department-day level count of chest CT scan reports processed daily by the department from 180 days before to 180 days after the focal event date. The coefficient of Post (× 100%) represents the percentage change in the dependent variable after the focal event date. The specification includes fixed effect of requesting department

$^{+}$ $p < 0.1$

* $p < .05$

** $p < .01$

*** $p < .001$

must anticipate this challenge and develop strategies to balance efficiency and quality during AI implementation. For instance, institutions should adopt case-specific AI usage strategies based on complexity and risk levels. Prioritize AI for complex/high-risk cases to ensure quality, while limiting AI for routine cases to preserve efficiency. Additionally, enhanced physician training is equally crucial. The training curriculum should not only cover how to operate AI systems efficiently, interpret AI recommendations quickly, and integrate those recommendations into their diagnostic processes, but it should also emphasize the importance of maintaining independent clinical judgment. Physicians should view AI as a decision-support tool, not a replacement, preserving clinical judgment in decision-making. Considering that our research was conducted in a small to mid-sized hospital, other similar healthcare institutions should adjust the findings to fit their specific contexts. Meanwhile, AI application providers should strive to develop more user-friendly interfaces, continuously updating and expanding AI databases, and optimizing algorithms to consistently enhance diagnostic accuracy. In summary, future efforts should optimize both AI systems and their clinical use to simultaneously improve efficiency and quality, ultimately benefiting healthcare delivery and patient outcomes.

Our research also has several limitations. First, our study is limited to data from a single local hospital and focuses exclusively on lung CT diagnosis, potentially limiting generalizability to other medical contexts. Therefore, future studies should validate these findings across diverse hospital settings and clinical contexts. Second, our relatively short observation period limits assessment of long-term effects. Extended observation periods would better assess AI's long-term impacts.Third, using report count and length as efficiency/quality indicators represents potentially inadequate proxies for physician performance. . Fourth, our study omits potential confounding factors including physician characteristics and work-related stress , which may affect the final application of AI assistance in practice. Exploring physician characteristics as potential moderators of AI's effects would be valuable. Lastly, the literature lacks research examining patient perceptions of physician-AI collaboration. Prior research identifies persistent patient concerns about privacy, liability, and trust in healthcare AI (Esmaeilzadeh et al., 2021). Given that AI-human collaboration is likely to become the norm in the foreseeable future, subsequent studies could also explore the potential impacts of AI-assisted healthcare from the patients' perspective.

Abbreviations

AI	Artificial Intelligence
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
IARC	International Agency for Research on Cancer
IBM	International Business Machines
PACS	Picture Archiving and Communication System
WHO	World Health Organization
3D	Three-Dimensional

Acknowledgements
The work was supported by the National Natural Science Foundation of China (NSFC72102207).

Clinical Trial number
Not applicable.

Authors' contributions
Siqi Dai and Zhiyi Xie performed the data analysis and wrote the main manuscript text. Zheshuai Yang and Wei Miao generated the idea, revised and polished the manuscript. All authors reviewed the manuscript.

Funding
The research is supported by a grant funded by the National Natural Science Foundation of China (NSFC72102207).

Data availability
The research data was provided by a local hospital in China. In accordance with Chinese laws, to protect patient privacy, the data must be retained by the hospital and cannot be shared. We apologize for any inconvenience and thank you for your understanding.

Declarations

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 18 October 2024 Revised: 13 January 2025 Accepted: 12 March 2025
Published online: 08 May 2025

References

Ahmed, A., & Adil, K. (2023). Transforming healthcare systems with artificial intelligence: Revolutionizing efficiency, quality, and patient care. *Research Square*. https://doi.org/10.21203/rs.3.rs-3175341/v1

Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., Naidich, D. P., & Shetty, S. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine, 25*(6), 954–961. https://doi.org/10.1038/s41591-019-0447-x

Arocena, P., & García-Prado, A. (2007). Accounting for quality in the measurement of hospital performance: Evidence from Costa Rica. *Health Economics, 16*(7), 667–685. https://doi.org/10.1002/hec.1204

Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., Asirvatham, S. J., Deshmukh, A. J., Gersh, B. J., Carter, R. E., Yao, X., Rabinstein, A. A., Erickson, B. J., Kapa, S., & Friedman, P. A. (2019). An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: A retrospective analysis of outcome prediction. *The Lancet, 394*(10201), 861–867. https://doi.org/10.1016/S0140-6736(19)31721-0

Bao, C., & Bardhan, I. R. (2022). Performance of accountable care organizations: Health information technology and quality–efficiency trade-offs. *Information Systems Research, 33*(2), 697–717. https://doi.org/10.1287/isre.2021.1080

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., ... & Aerts, H. J. (2019). Artificial intelligence in cancer imaging: clinical challenges and applications. *CA: A Cancer Journal for Clinicians*, *69*(2), 127–157. https://doi.org/10.3322/caac.21552

Busse, M. R., Simester, D. I., & Zettelmeyer, F. (2010). "The best price you'll ever get": The 2005 employee discount pricing promotions in the US automobile industry. *Marketing Science, 29*(2), 268–290. https://doi.org/10.1287/mksc.1090.0516

Chen, C., Liao, M., Walther, J. B., & Sundar, S. S. (2024). When an AI doctor gets personal: The effects of social and medical individuation in encounters with human and AI doctors. *Communication Research, 51*(7), 747–781. https://doi.org/10.1177/00936502241263482

Chen, J., Chen, C., B. Walther, J., & Sundar, S. S. (2021, May). Do you feel special when an AI doctor remembers you? Individuation effects of AI vs. human doctors on user experience. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–7). https://doi.org/10.1145/3411763.3451735

Chen, J., & Roth, J. (2024). Logs with zeros? Some problems and solutions. *The Quarterly Journal of Economics, 139*(2), 891–936. https://doi.org/10.1093/qje/qjad054

Chen, X., John, G., Hays, J. M., Hill, A. V., & Geurs, S. E. (2009). Learning from a service guarantee quasi experiment. *Journal of Marketing Research, 46*(5), 584–596. https://doi.org/10.1509/jmkr.46.5.584

Das, R. (2016, March 30). Five technologies that will disrupt healthcare by 2020. *Forbes*. Retrieved from https://www.forbes.com/sites/reenitadas/2016/03/30/top-5-technologies-disrupting-healthcare-by-2020/#4a56027e6826

Davenport, T. H., & Glaser, J. P. (2022). Factors governing the adoption of artificial intelligence in healthcare providers. *Discover Health Systems, 1*(1), 4. https://doi.org/10.1007/s44250-022-00004-8

Doyle, O. M., Leavitt, N., & Rigg, J. A. (2020). Finding undiagnosed patients with hepatitis C infection: An application of artificial intelligence to patient claims data. *Scientific Reports, 10*(1), 10521. https://doi.org/10.1038/s41598-020-67013-6

Esmaeilzadeh, P. (2020). Use of AI-based tools for healthcare purposes: A survey study from consumers' perspectives. *BMC Medical Informatics and Decision Making, 20*, 1–19. https://doi.org/10.1186/s12911-020-01191-1

Esmaeilzadeh, P., Mirzaei, T., & Dharanikota, S. (2021). Patients' perceptions toward human–artificial intelligence interaction in health care: Experimental study. *Journal of Medical Internet Research, 23*(11), e25856. https://doi.org/10.2196/25856

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118. https://doi.org/10.1038/nature21056

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA, 316*(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

Haleem, A., Javaid, M., & Khan, I. H. (2019). Current status and applications of Artificial Intelligence (AI) in medical field: An overview. *Current Medicine Research and Practice, 9*(6), 231–237. https://doi.org/10.1016/j.cmrp.2019.11.005

Hausman, C., & Rapson, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics, 10*(1), 533–552. https://doi.org/10.1146/annurev-resource-121517-033306

Hsieh, P.-J. (2023). Determinants of physicians' intention to use AI-assisted diagnosis: An integrated readiness perspective. *Computers in Human Behavior, 147*, 107868. https://doi.org/10.1016/j.chb.2023.107868

Huang, G., Wei, X., Tang, H., Bai, F., Lin, X., & Xue, D. (2021). A systematic review and meta-analysis of diagnostic performance and physicians' perceptions of artificial intelligence (AI)-assisted CT diagnostic technology for the classification of pulmonary nodules. *Journal of Thoracic Disease*, *13*(8), 4797–4811. https://doi.org/10.21037/jtd-21-810

Imran, A., Posokhova, I., Qureshi, H. N., Masood, U., Riaz, M. S., Ali, K., John, C. N., Hussain, M. I., & Nabeel, M. (2020). AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Informatics in Medicine Unlocked, 20*, 100378. https://doi.org/10.1016/j.imu.2020.100378

Janakiraman, R., Park, E., Demirezen, M., & E., & Kumar, S. (2023). The effects of health information exchange access on healthcare quality and efficiency: An empirical investigation. *Management Science, 69*(2), 791–811. https://doi.org/10.1287/mnsc.2022.4378

Jorg, T., Halfmann, M. C., Stoehr, F., Arnhold, G., Theobald, A., Mildenberger, P., & Müller, L. (2024). A novel reporting workflow for automated integration of artificial intelligence results into structured radiology reports. *Insights into Imaging, 15*(1), 80. https://doi.org/10.1186/s13244-024-01660-5

Kane, R. L., Shamliyan, T. A., Mueller, C., Duval, S., & Wilt, T. J. (2007). The association of registered nurse staffing levels and patient outcomes: Systematic review and meta-analysis. *Medical Care, 45*(12), 1195–1204. https://doi.org/10.1097/MLR.0b013e3181468ca3

Katwaroo, A. R., Adesh, V. S., Lowtan, A., & Umakanthan, S. (2024). The diagnostic, therapeutic, and ethical impact of artificial intelligence in modern medicine. *Postgraduate Medical Journal, 100*(1183), 289–296. https://doi.org/10.1093/postmj/qgad135

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals, 14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Khalifa, M., & Albadawy, M. (2024). AI in diagnostic imaging: Revolutionising accuracy and efficiency. *Computer Methods and Programs in Biomedicine Update, 5*, 100146. https://doi.org/10.1016/j.cmpbup.2024.100146

Khullar, D., Casalino, L. P., Qian, Y., Lu, Y., Krumholz, H. M., & Aneja, S. (2022). Perspectives of patients about artificial intelligence in health care. *JAMA Network Open, 5*(5), e2210309. https://doi.org/10.1001/jamanetworkopen.2022.10309

Krakowski, S., Luger, J., & Raisch, S. (2023). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal, 44*(6), 1425–1452. https://doi.org/10.1002/smj.3387

Krishnan, G., Singh, S., Pathania, M., Gosavi, S., Abhishek, S., Parchani, A., & Dhar, M. (2023). Artificial intelligence in clinical medicine: Catalyzing a sustainable global healthcare paradigm. *Frontiers in Artificial Intelligence, 6*, 1227091. https://doi.org/10.3389/frai.2023.1227091

Kulkarni, P. A., & Singh, H. (2023). Artificial intelligence in clinical diagnosis: Opportunities, challenges, and hype. *JAMA, 330*(4), 317–318. https://doi.org/10.1001/jama.2023.11440

Leachman, S. A., & Merlino, G. (2017). The final frontier in cancer diagnosis. *Nature, 542*(7639), 36–38. https://doi.org/10.1038/nature21492

Li, J., Zhou, L., Zhan, Y., Xu, H., Zhang, C., Shan, F., & Liu, L. (2022). How does the artificial intelligence-based image-assisted technique help physicians in diagnosis of pulmonary adenocarcinoma? A randomized controlled experiment of multicenter physicians in China. *Journal of the American Medical Informatics Association, 29*(12), 2041–2049. https://doi.org/10.1093/jamia/ocac179

Li, M., & Xiong, X. (2024). Attitudes and Perceptions of Chinese Oncologists Towards Artificial Intelligence in Healthcare: A Cross-Sectional Survey. *Frontiers in Digital Health, 6*, 1371302. https://doi.org/10.3389/fdgth.2024.1371302

Li, Z., Zhang, X., & Liu, Y. (2023). A preliminary exploration on the establishment of an AI-assisted remote imaging diagnostic system for major infectious diseases. *Chinese Journal of Infectious Diseases, 41*(7), 456. https://doi.org/10.3760/cma.j.cn118332-20230710-00456

Liu, J., & Cong, Z. (2023). The daily me versus the daily others: How do recommendation algorithms change user interests? Evidence from a knowledge-sharing platform. *Journal of Marketing Research, 60*(4), 767–791. https://doi.org/10.1177/0022243722113423

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., ... & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, *577*(7788), 89–94. https://doi.org/10.1038/s41586-019-1799-6

Moore, L. G. (2007). The ideal medical practice model: improving efficiency, quality and the doctor-patient relationship. *Family Practice Management*, *14*(8), 20–24. https://www.aafp.org/pubs/fpm/issues/2007/0900/p20.html

Nature Editorial Team. (2023, September 6). AI tools are revolutionizing the future of diagnostics. *Nature Portfolio*. https://www.nature.com/articles/d42473-023-00300-8

Navarro-Espigares, J. L., & Torres, E. H. (2011). Efficiency and quality in health services: A crucial link. *The Service Industries Journal, 31*(3), 385–403. https://doi.org/10.1080/02642060802712798

Nelson, C. A., Pérez-Chada, L. M., Creadore, A., Li, S. J., Lo, K., Manjaly, P., ... & Mostaghimi, A. (2020). Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. *JAMA Dermatology*, *156*(5), 501–512. https://doi.org/10.1001/jamadermatol.2019.5014

Newhouse, J. P. (1970). Toward a theory of nonprofit institutions: An economic model of a hospital. *The American Economic Review*, *60*(1), 64–74. https://www.jstor.org/stable/1807855

Overhage, J. M., & McCallie, D., Jr. (2020). Physician time spent using the electronic health record during outpatient encounters: A descriptive study. *Annals of Internal Medicine, 172*(3), 169–174. https://doi.org/10.7326/M18-3684

Patankar, A. (2024). AI in personalized treatment plans: Enhancing patient care. *Healthcare IT Today*. Retrieved from https://www.healthcareittoday.com

Pesapane, F., Codari, M., & Sardanelli, F. (2018). Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *European Radiology Experimental, 2*, 1–10. https://doi.org/10.1186/s41747-018-0061-6

Pierre, K., Haneberg, A. G., Kwak, S., Peters, K. R., Hochhegger, B., Sananmuang, T., Tunlayadechanont, P., Tighe, P. J., Mancuso, A., & Forghani, R. (2023). Applications of artificial intelligence in the radiology roundtrip: Process streamlining, workflow optimization, and beyond. *Seminars in Roentgenology, 58*(2), 158–169. https://doi.org/10.1053/j.ro.2023.02.003

Rad-Insights. (2024). *The radiology report: Everything you need to know*. https://rad-insights.com/the-radiology-report-everything-you-need-to-know/

Raimondi, F., Ochoa, M. S., Regan, D., & Komorowski, M. (2020). The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine, 26*(11), 1820–1827. https://doi.org/10.1038/s41591-020-1042-7

Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine, 28*(1), 31–38. https://doi.org/10.1038/s41591-021-01614-0

Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., ... & Lungren, M. P. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Medicine*, *15*(11), e1002686. https://doi.org/10.1371/journal.pmed.1002686

Reuben, D. B., Knudsen, J., Senelick, W., Glazier, E., & Koretz, B. K. (2014). The effect of a physician partner program on physician efficiency and patient satisfaction. *JAMA Internal Medicine, 174*(7), 1190–1193. https://doi.org/10.1001/jamainternmed.2014.1315

Ross, S. E., Todd, J., Moore, L. A., Beaty, B. L., Wittevrongel, L., & Lin, C. T. (2005). Expectations of patients and physicians regarding patient-accessible medical records. *Journal of Medical Internet Research, 7*(2), e399. https://doi.org/10.2196/jmir.7.2.e13

Sathykumar, K., Munoz, M., et al. (2020). Automated lung cancer detection using artificial intelligence (AI) deep convolutional neural networks: A narrative literature review. *Cureus, 12*(8), 1–9. https://doi.org/10.7759/cureus.10017

Sullivan, F. (2018). Artificial intelligence market—Key application areas for growth in healthcare IT, forecast to 2022. *Frost & Sullivan*. Retrieved from https://www.marketresearch.com/Frost-Sullivan-v383/Artificial-Intelligence-Key-Application-Areas-11839578.

Tasi, M. C., Keswani, A., & Bozic, K. J. (2019). Does physician leadership affect hospital quality, operational efficiency, and financial performance? *Health Care Management Review, 44*(3), 256–262. https://doi.org/10.1097/HMR.0000000000000173

Tong, W. J., Wu, S. H., Cheng, M. Q., Huang, H., Liang, J. Y., Li, C. Q., ... & Wang, W. (2023). Integration of artificial intelligence decision aids to reduce workload and enhance efficiency in thyroid nodule management. *JAMA Network Open*, *6*(5), e2313674. https://doi.org/10.1001/jamanetworkopen.2023.13674

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Wang, W., Gao, G., & Agarwal, R. (2024). Friend or foe? Teaming between artificial intelligence and workers with variation in experience. *Management Science, 70*(9), 5753–5775. https://doi.org/10.1287/mnsc.2021.00588

Wang, X., Zhou, B., Gong, P., Zhang, T., Mo, Y., Tang, J., ... & Wang, L. (2022). Artificial intelligence–assisted bone age assessment to improve the accuracy and consistency of physicians with different levels of experience. *Frontiers in Pediatrics*, *10*, 818061. https://doi.org/10.3389/fped.2022.818061

Wei, X., Jiang, J., Cao, W., Yu, H., Deng, H., Chen, J., Bai, S., & Zhou, Z. (2022). Artificial intelligence assistance improves the accuracy and efficiency of intracranial aneurysm detection with CT angiography. *European Journal of Radiology, 149*, 110169. https://doi.org/10.1016/j.ejrad.2022.110169

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., ... & Goldenberg, A. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature Medicine*, *25*(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-6

World Health Organization. (2024). Global cancer burden growing, amidst mounting need for services [EB/OL]. Retrieved from https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services

Wu, A., Xue, P., Abulizi, G., Tuerxun, D., Rezhake, R., & Qiao, Y. (2023). Artificial intelligence in colposcopic examination: A promising tool to assist junior colposcopists. *Frontiers in Medicine, 10*, 1060451. https://doi.org/10.3389/fmed.2023.1060451

Xu, D., Sui, L., Zhang, C., Xiong, J., Wang, V. Y., Zhou, Y., ... & Wang, L. (2024). The clinical value of artificial intelligence in assisting junior radiologists in thyroid ultrasound: a multicenter prospective study from real clinical practice. *BMC medicine*, *22*(1), 293. https://doi.org/10.1186/s12916-024-03510-z

Yang, J., & Zeng, W. (2014). The trade-offs between efficiency and quality in the hospital production: Some evidence from Shenzhen, China. *China Economic Review, 31*, 166–184. https://doi.org/10.1016/j.chieco.2014.09.005

Yoo, J., Hur, S., Hwang, W., & Cha, W. C. (2023). Healthcare professionals' expectations of medical artificial intelligence and strategies for its clinical implementation: A qualitative study. *Healthc Inform Res, 29*(1), 64–74. https://doi.org/10.4258/hir.2023.29.1.64

Yun, J. H., Lee, E. J., & Kim, D. H. (2021). Behavioral and neural evidence on consumer responses to human doctors and medical artificial intelligence. *Psychology & Marketing, 38*(4), 610–625. https://doi.org/10.1002/mar.21445

Zhang, Z., Genc, Y., Wang, D., Ahsen, M. E., & Fan, X. (2021). Effect of ai explanations on human perceptions of patient-facing ai-powered healthcare systems. *Journal of Medical Systems, 45*(6), 64. https://doi.org/10.1007/s10916-021-01743-6

Zhou, N., Zhang, C. T., Lv, H. Y., Hao, C. X., Li, T. J., Zhu, J. J., Zhu, H., Jiang, M., Liu, K. W., Hou, H. L., Liu, D., Li, A. Q., Zhang, G. Q., Tian, Z. B., & Zhang, X. C. (2019). Concordance study between IBM Watson for oncology and clinical practice for patients with cancer in China. *The Oncologist, 24*(6), 812–819. https://doi.org/10.1634/theoncologist.2018-0255

Publisher's Note