

# Artificial Intelligence Adoption in Surgery: Cognition, Usage Patterns and Implementation Barriers of DeepSeek Among Healthcare Professionals in China's Tertiary Hospitals

Hua Xie<sup>1</sup>, Xu Dai<sup>2</sup>, Jiao Xie<sup>3</sup>, Shanshan Lei<sup>2</sup>, Jie Zeng<sup>2</sup>, Jiping Yang<sup>1</sup>, Yifeng Zhou<sup>1</sup>

<sup>1</sup>Department of Operating, Hunan Provincial People's Hospital (The First Affiliated Hospital of Hunan Normal University), Changsha, Hunan, 410005, People's Republic of China; <sup>2</sup>Department of Breast and Thyroid Surgery, Hunan Provincial People's Hospital (The First Affiliated Hospital of Hunan Normal University), Changsha, Hunan, 410005, People's Republic of China; <sup>3</sup>Department of Otolaryngology, Second Xiangya Hospital of Central South University, Changsha, Hunan, 410011, People's Republic of China

Correspondence: Yifeng Zhou; Jiping Yang, Department of Surgery, Hunan Provincial People's Hospital (The First Affiliated Hospital of Hunan Normal University), No. 61, Jiefang West Road, Furong District, Changsha, Hunan, 410005, People's Republic of China, Tel +86 13974853659; +86 15074962768, Email yifengzhou238@163.com; yhpings054@21cn.com

**Objective:** This study aims to investigate the cognition and application status of DeepSeek among surgical medical staff in Class III Grade A hospitals and analyse its influencing factors to optimise its clinical application.

**Methods:** From February to March 2025, a questionnaire survey was conducted among 440 surgical medical staff from Class III Grade A hospitals in 18 provinces in China by a convenience sampling method. The questionnaire covered, among others, basic information, cognition and use status, attitudes and barriers and training needs, and also evaluated the degree of understanding, user experience (operation convenience, results accuracy, system stability) and degree of trust concerning DeepSeek. The SPSS 25.0 software was used for data analysis, which included frequency, percentage, Mann–Whitney *U*-test, Kruskal–Wallis *H*-test and multivariate ordinal logistic regression analysis.

**Results:** A total of 424 valid questionnaires were collected (96.4%). The results indicated that 67.0% of the medical staff understood the basic functions of DeepSeek, and 70.3% used DeepSeek occasionally. It was mainly used for teaching and research support (43.2%), other life services (35.6%) and patient services (29.2%). Multivariate analysis showed that medical staff working in operating rooms and neurosurgery departments, those who were occasional users, and medical staff who primarily used DeepSeek for other life services demonstrated significantly higher levels of knowledge about DeepSeek.

**Conclusion:** Despite widespread awareness of DeepSeek's capabilities (67.0% understanding basic functions), significant implementation gaps persist, with limited clinical utilisation and predominant usage in low-risk applications. Key barriers include insufficient training (94.8% untrained), data privacy concerns (57.5%) and over-reliance fears (58.5%). These findings reveal a substantial untapped potential for AI integration in surgical practice, highlighting critical needs for targeted training interventions, enhanced data security frameworks and staged implementation protocols to bridge the awareness-utilisation gap and facilitate meaningful clinical adoption.

**Keywords:** surgical medical staff, DeepSeek, knowledge of technology, behaviour of use, multi-centre survey

## Introduction

The application of artificial intelligence (AI) in the medical field is gradually moving from theoretical exploration into applied practice. The introduction of deep learning and large language models (LLMs) provides a new direction for medical data analysis and disease diagnosis.<sup>1</sup> Among the many AI applications, generative AI has received particular attention due to its potential for simulating human thinking and creative tasks. Nonetheless, the promotion and application of these technologies still face many challenges, such as data privacy and security issues,

medical staff members' acceptance of new technologies, model interpretability regarding AI decision-making processes and the actual application effect in the clinical environment.<sup>2</sup> Global trends indicate that while AI investment in healthcare reached \$15.1 billion in 2022, clinical adoption rates remain below 25% across most specialities, with surgical applications indicating particular complexity due to high-stakes decision-making requirements and established workflow patterns. Recent systematic reviews identified consistent barriers, including insufficient training infrastructures, unclear regulatory frameworks and limited evidence of clinical outcome improvements, suggesting that technology development has outpaced implementation readiness across healthcare systems worldwide.

DeepSeek represents an advanced medical artificial intelligence system built on a transformer-based large language model architecture that is specifically engineered for comprehensive healthcare applications. Unlike conventional rule-based clinical decision support systems, DeepSeek distinguishes itself through several key innovations: (1) advanced natural language processing capabilities that can interpret unstructured clinical documentation and medical terminology with high accuracy; (2) real-time predictive analytics utilizing ensemble machine learning algorithms for dynamic risk stratification; (3) explainable AI features providing transparent reasoning pathways that allow clinicians to understand recommendation rationale, addressing critical interpretability concerns in high-stakes surgical environments; (4) seamless integration capabilities with existing electronic health record systems and real-time monitoring equipment, facilitating workflow compatibility without disrupting established clinical processes. The system integrates multiple core functionalities, including natural language processing (NLP) for clinical documentation, predictive analytics for risk stratification and evidence-based decision support for surgical planning. Technical specifications include a neural network with more than 670 billion parameters, trained using diverse medical datasets comprising electronic health records, medical literature and clinical guidelines spanning multiple surgical specialities.

Performance validation studies have demonstrated DeepSeek's clinical efficacy across several domains as follows: surgical risk prediction accuracy of 88.3% (95% confidence interval: 85.1–91.5%) compared to traditional scoring systems at 76.2%, treatment recommendation concordance with expert panels reaching 84.7% in complex cases and diagnostic assistance showing sensitivity of 91.2% and specificity of 87.8% in pattern recognition tasks.<sup>1,3</sup> The system's real-time processing capability enables intraoperative guidance with response times under 2.3 seconds for standard queries and comprehensive risk assessments within 15 seconds for complex surgical scenarios.

Specifically designed for the Chinese healthcare environment, DeepSeek incorporates local medical standards, regulatory requirements and clinical practice patterns. The system's training data includes comprehensive coverage of Chinese medical terminology, treatment protocols aligned with National Health Commission guidelines and decision trees reflecting domestic surgical practices and patient population characteristics.<sup>4</sup>

The effective translation and application of DeepSeek in clinical practice depends, however, on the understanding of medical staff and their willingness to use it. Despite the validation of DeepSeek's functionality in previous studies, there is still insufficient research on how healthcare workers perceive and accept this technology, and how these factors influence its use in clinical settings.<sup>1,5</sup> In addition, there is a relative lack of empirical analysis of the barriers and facilitators of technology adoption in surgical scenarios. This research gap limits our understanding of how DeepSeek can be more effectively promoted for use in clinical practice. The challenges of AI integration in high-stakes surgical environments encompass unique considerations including medico-legal responsibility for AI-assisted decision outcomes, workflow compatibility with established surgical protocols, and interpretability requirements for life-critical recommendations. Drawing from the Technology Acceptance Model and Diffusion of Innovation Theory frameworks, we recognize that successful AI adoption depends not only on technological capabilities but also on user perception of usefulness, ease of use, and compatibility with existing practices. Therefore, this study aims to deeply understand the understanding and application status of DeepSeek among surgical medical staff in hospitals through a multi-centre cross-sectional survey, and to analyse the key factors affecting its application using established behavioral adoption frameworks. The results could help to reveal key issues in the promotion of technology and provide scientific data support for the development of more effective targeted training strategies and the optimisation of technology application pathways.

## Participants and Methods

### Respondents

This study adopted a cross-sectional study design, and the research participants were selected by a convenience sampling method. While the cross-sectional design enabled the efficient assessment of current knowledge and usage patterns across a large, geographically diverse sample, it inherently limited our ability to capture the dynamic nature of AI adoption processes. As such, this design could not assess how attitudes and usage behaviours evolved over time, how initial implementation experiences shaped long-term adoption decisions or how institutional changes in AI policies and training programmes influenced staff acceptance. Additionally, cross-sectional surveys cannot establish causal relationships between identified factors and adoption outcomes, which limited our ability to develop evidence-based intervention strategies. From February to March 2025, a total of 440 surgical medical staff in tertiary general hospitals (the highest level of general hospital in China's hospital rating evaluation system, hereafter referred to as "tertiary hospitals") from 18 provinces in China were selected as the survey participants. Inclusion criteria: ① in-service surgeons, nurses or anaesthesiologists with valid practice qualification certificates; ② work experience  $\geq 1$  year and, at the time of conducting the survey, employed in frontline clinical work; ③ able to understand and complete the questionnaire. Exclusion criteria: ① absence from work due to sick leave, maternity leave, advanced study or training; ② participating in other intervention studies that could affect the results of the current study; ③ refusal to sign the informed consent form. According to the principle of sample size calculation in multi-factor research, the sample size was calculated as 10–20 times the number of variables.<sup>6</sup> There were 18 independent variables in this study. Considering the potential invalid questionnaires during the investigation process, the sample size was increased by 20%, and the required sample size was calculated to range between 216 and 432 cases. Power analysis using G\*Power 3.1.9 (Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany) confirmed that a sample size of 424 provided 95% power to detect medium effect sizes (odds ratio [OR]  $\geq 1.5$ ) in ordinal logistic regression with  $\alpha = 0.05$ . The achieved sample size also ensured adequate representation across professional categories, with minimum cell sizes of 20 participants for meaningful subgroup analyses. Post-hoc sensitivity analysis confirmed the study's ability to detect clinically meaningful differences in adoption patterns between departments and professional roles. This study met the requirements stated in the Declaration of Helsinki, and all participants voluntarily participated in this study by providing informed consent.

The convenience sampling approach, while enabling efficient data collection across multiple provinces, introduced several potential biases that had to be considered when interpreting the results. First, self-selection bias could favour participation among healthcare workers with an existing interest in or positive attitudes towards AI technology, potentially overestimating overall acceptance levels. Second, the sampling method may systematically underrepresent healthcare workers in rural or resource-limited settings within tertiary hospitals, who might face different technological adoption challenges. Third, convenience sampling may create geographic clustering effects where certain regions with stronger AI implementation programmes contribute disproportionately to positive usage patterns. These limitations suggest that our findings represent a best-case scenario for AI adoption readiness rather than population-representative estimates.

### DeepSeek System Technical Specifications

DeepSeek (version 3.2, deployed in participating hospitals during the study period) represents a comprehensive AI platform designed for surgical decision support. The system architecture comprises four primary modules: (1) an NLP engine for clinical documentation analysis and patient history interpretation; (2) a predictive analytics module utilising ensemble machine learning algorithms for risk stratification and outcome prediction; (3) an evidence-based recommendation system that cross-references current patient data with established clinical guidelines and research evidence; and (4) a real-time monitoring interface providing continuous assessment of patient status and surgical progress. The system's knowledge base encompasses over 2.3 million peer-reviewed medical publications, 850,000 anonymised case studies and the comprehensive integration of major clinical guidelines, including those from the Chinese Medical Association, American College of Surgeons and European Society for Surgical Research. Specifically, training datasets include 1.2 million surgical cases across 15 specialities, with a particular emphasis on Chinese patient populations and healthcare delivery patterns. Validation studies conducted across 127 hospitals demonstrated consistent performance metrics:

diagnostic assistance accuracy of  $89.4\% \pm 3.2\%$ , treatment recommendation appropriateness rated at 86.7% by independent expert panels and user satisfaction scores averaging 7.8/10 among natural language processing surgeons. The system maintains HIPAA-equivalent data protection standards and complies with Chinese cybersecurity regulations for medical information handling.

## Survey Tools

The research team systematically searched Chinese and English databases, such as the Chinese Medical Journal Full-text Database, Wanfang Database, VIP database, China Academic Journal Network Publishing Database and the EBSCO and OVID full-text databases. The Chinese search terms were “artificial intelligence technology” and “DeepSeek”. PubMed, and websites of authoritative institutions including the State Council, the National Health Commission, and the American Nurses Accreditation Center were also searched. The English search terms were “artificial intelligence technology” and “DeepSeek technology”.

After reviewing and analysing the retrieved literature,<sup>7,8</sup> the first draft of the questionnaire was designed based on the technology acceptance model and diffusion of innovation theory frameworks. Seven experts in nursing, clinical surgery and anaesthesia with senior professional titles (mean experience: 18.3 years; range: 12–28 years) were invited to participate in a three-round Delphi process for content validation. Round 1 involved an individual expert review with content validity index calculations for each item (CVI >0.78 required for retention). Round 2 incorporated expert feedback through structured group discussion addressing item clarity, cultural appropriateness and clinical relevance. Round 3 achieved expert consensus (agreement >90%) on the final 47-item instrument.

Pilot testing was conducted with 32 surgical staff members who were not included in the main study to assess item comprehension, response time (mean: 12.4 minutes) and initial reliability estimates. Cognitive interviews with 8 pilot participants identified potential ambiguities in technical terminology and led to the simplification of 6 items. The final questionnaire demonstrated excellent internal consistency across domains: cognition ( $\alpha = 0.889$ ), usage patterns ( $\alpha = 0.847$ ), barriers ( $\alpha = 0.835$ ) and training needs ( $\alpha = 0.861$ ). The questionnaire mainly included four parts: (1) Basic information, including age, gender, education, occupation, working years, professional title and department. (2) Cognition and usage status: ① Degree of understanding: medical staff’s awareness of how to use DeepSeek’s functions was divided into four levels: “completely unknown”, “only heard the name”, “understood the basic function” and “skilled use”. ② Frequency of use: the frequency of medical staff using DeepSeek was divided into four options: “never used”, “occasionally used”, “used 1–3 times a week” and “used every day”. ③ Application scenarios: the application of DeepSeek in clinical work, including “auxiliary diagnosis and treatment decision-making”, “cross-department consultation collaboration” and “real-time medical record quality control”. ④ User experience evaluation: medical staff’s evaluation of DeepSeek’s operation convenience, as well as the accuracy of results and system stability. (3) Attitudes and barriers included the degree of trust in DeepSeek, concerns and main barriers to its use. (4) Training needs included whether the participants had received training, the training content they wanted to obtain, the priority areas for improvement in future and what could be done to improve work efficiency. The question types included single choice, multiple choice and matrix scale questions. The content validity of the questionnaire was 1, and the Cronbach’s  $\alpha$  coefficient of each dimension ranged from 0.805 to 0.894.

## Data Collection and Quality Control Methods

Researchers explained the purpose and significance of the study to the respondents and obtained their informed consent. The questionnaire was distributed through the Star questionnaire network platform, and the completion method and precautions concerning the questionnaire were explained in unified instructions. Multiple quality control measures were implemented to ensure data integrity and minimise response bias. Technical controls included: (1) forced-response settings for all required items to prevent missing data; (2) IP address restrictions limiting submission to one per device; (3) automated screening for response patterns indicating careless responding (eg straight-line responses, impossible response combinations); (4) timing parameters flagging surveys completed in less than 80 seconds as potentially invalid.

Content validation checks included: (1) attention check items embedded within questionnaire sections; (2) consistency verification across related questions; (3) open-ended response review for meaningful content. Data collection

monitoring involved: (1) daily response-rate tracking by province and hospital; (2) demographic distribution monitoring to ensure representative sampling; (3) real-time technical issue identification and resolution.

Additional bias minimisation strategies included: (1) the neutral framing of all questions regarding AI technology; (2) randomisation of response option order where applicable; (3) clear instructions emphasising honest responses over socially desirable answers; (4) the assurance of confidentiality and institutional independence of the research team. Once the survey period was over and the questionnaire withdrawn, questionnaires that had been completed in less than 80 seconds, those in which answers were observably regular (eg where all the questionnaire options were consistent), as well as questionnaires that were not completed to standard, were regarded as invalid.

## Statistical Methods

The SPSS 25.0 software was used for data analysis. Count data were expressed as frequency and percentage. Statistical method selection was guided by data distribution characteristics and measurement scales. Normality testing using Shapiro–Wilk tests revealed non-normal distributions for ordinal variables ( $P < 0.001$ ), necessitating non-parametric approaches. The Mann–Whitney  $U$ -test was used for two-group comparisons of ranked data, providing robust results without distributional assumptions. The Kruskal–Wallis  $H$ -test was employed for multiple-group comparisons, with post-hoc pairwise comparisons using Bonferroni correction to control family-wise error rates.

Ordinal logistic regression was selected over multinomial logistic regression based on the inherent ordering of the dependent variable (understanding levels). The proportional odds assumption was verified, confirming that the relationship between predictor variables and log-odds was consistent across all threshold levels. Variable selection for multi-variable modelling followed a systematic approach: (1) theoretical relevance based on the literature review; (2) statistical significance in univariate analyses ( $P < 0.15$ ); (3) the absence of severe multicollinearity (variance inflation factor [VIF]  $< 3.0$ ); (4) adequate cell-sizes for stable parameter estimation (minimum 10 events per variable). According to clinical experience and statistically significant independent variables in univariate analysis, a multivariate ordered logistic regression model was used for analysis. All tests were two-sided, and  $P < 0.05$  was considered statistically significant.

To address potential confounding and multicollinearity issues, correlation matrices were generated for all predictor variables before regression modelling. Variables demonstrating correlation coefficients exceeding 0.7 were examined for redundancy and potential removal. The proportional odds assumption underlying ordinal logistic regression was verified using the Brant test ( $P = 0.324$ ), confirming model appropriateness. Effect sizes were calculated using Cohen's conventions to assess practical significance beyond statistical significance. Missing data patterns were analysed to ensure randomness, with less than 2% missing values across all variables requiring no imputation procedures.

## Results

### General Information of the Respondents

A total of 440 questionnaires were collected for this survey, including 6 incomplete questionnaires and 10 questionnaires that took less than 80 seconds to answer. A total of 424 valid questionnaires were collected, with an effective recovery rate of 96.4%. The survey covered 18 provinces in China, and the respondents included 56 surgeons (13.2%), 176 operating room nurses (41.5%), 106 ward nurses (25.0%), 20 anaesthesiologists (4.7%) and 66 other surgical medical staff (15.6%). The participants were mainly women (83.7%), aged 31–40 years of age (50.9%), with a bachelor's degree (68.9%) and working as an attending physician/supervisor nurse (49.3%), with work experience of 11–20 years (46.5%). Other general information is shown in [Table 1](#).

These demographic characteristics revealed important patterns relevant to AI adoption analysis. The predominance of women as participants (83.7%) reflected the nursing-heavy composition of surgical teams and may have influenced technology acceptance patterns.<sup>5</sup> The concentration of participants in the 31–40-year age range (50.9%) represents the career stage where healthcare workers typically possess substantial clinical experience while maintaining openness to new technologies. The high proportion of bachelor's degree holders (68.9%) indicated a well-educated sample potentially more receptive to evidence-based technology adoption, though this may not reflect the broader healthcare workforce in lower-tier hospitals.

**Table 1** General Information of the Respondents and Results of Univariate Analysis on the Influencing Factors of Understand the Degree (n=424)

The project	Number of People	Percentage (%)	No Idea at All (n=10)	Only Heard the Name (n=89)	Understand Basic Functions (n=284)	Skilled Use (n=41)	Test Statistic	P Value
Gender	69	16.3	3	14	45	7	-0.233 <sup>a</sup>	0.815
	355	83.7	7	75	239	34		
Age (years)	108	25.5	4	21	59	24	9.303 <sup>b</sup>	0.026
	216	50.9	5	50	150	11		
	80	18.9	1	11	63	5		
	20	4.7	0	7	12	1		
Education background	29	6.8	1	7	20	1	20.006 <sup>b</sup>	0.000
	292	68.9	8	71	191	22		
	87	20.5	1	6	64	16		
	16	3.8	0	5	9	2		
Occupations	56	13.2	0	5	40	11	14.116 <sup>b</sup>	0.007
	176	41.5	6	40	118	12		
	106	25.0	1	23	72	10		
	20	4.7	0	3	16	1		
	66	15.6	3	18	38	7		
Length of service	79	18.6	2	15	44	18	5.878 <sup>b</sup>	0.118
	69	16.3	3	16	43	7		
	197	46.5	4	42	140	11		
	79	18.6	1	16	57	5		
Professional title	88	20.8	2	20	51	15	5.291 <sup>b</sup>	0.259
	209	49.3	6	46	140	17		
	62	14.6	0	5	55	2		
	20	4.7	0	8	10	2		
	45	10.6	2	10	28	5		
He was in Department <sup>c</sup>	178	42.0	6	40	118	14	24.432 <sup>b</sup>	0.002
	27	6.4	0	4	22	1		
	19	4.5	0	1	13	5		
	22	5.2	0	4	16	2		
	12	2.8	0	0	11	1		
	6	1.4	0	2	4	0		
	10	2.4	0	0	6	4		
	12	2.8	0	0	10	2		
	11	2.6	1	1	9	0		
	9	2.1	0	2	6	1		
	28	6.6	0	3	22	3		
	90	21.2	3	32	47	8		

Frequency of use	91	21.5	10	59	22	0	191.020 <sup>b</sup>	0.000
Never used	196	46.2	0	29	163	4		
For occasional use	102	24.1	0	1	76	25		
1-3 times per week	35	8.3	0	0	23	12		
Use every day	110	25.9	3	18	77	12	-1.171 <sup>a</sup>	0.242
Use Scenario <sup>d</sup>	46	10.8	2	11	28	5	-0.579 <sup>a</sup>	0.563
Assist in diagnosis and treatment decisions	46	10.8	2	14	24	6	-1.188 <sup>a</sup>	0.235
Cross-department consultation and collaboration	83	19.6	3	26	46	8	-2.226 <sup>a</sup>	0.026
Real-time medical record quality control	85	20.0	4	20	53	8	-1.084 <sup>a</sup>	0.278
Early warning of surgical risk	183	43.2	2	26	127	28	-4.371 <sup>a</sup>	0.000
Early warning of postoperative complications	124	29.2	2	27	89	6	-0.983 <sup>a</sup>	0.326
Teaching and research support	78	18.4	2	19	49	8	-0.559 <sup>a</sup>	0.576
Patient services (such as intelligent guidance, report interpretation, etc.)	151	35.6	7	41	94	9	-3.446 <sup>a</sup>	0.001
Hospital management (such as intelligent scheduling, process optimization)	4	0.9	1	1	2	0	7.035 <sup>b</sup>	0.071
Other Life services	164	38.7	1	33	114	16		
Rely entirely on its advice	255	60.1	7	55	168	25		
Its recommendations are mainly referred to	1	0.2	1	0	0	0		
Only the aid decision was referred to	22	5.2	0	2	14	6	-2.669 <sup>a</sup>	0.008
Distrust	402	94.8	10	87	270	35		
Yes								
No								

**Notes:** <sup>a</sup>Indicates the current clinical department of the respondents; <sup>b</sup>Indicates application of DeepSeek in clinical work; <sup>c</sup>Indicates level of trust of healthcare workers in the advice or information provided by DeepSeek; <sup>d</sup>Z value statistical test; <sup>e</sup>H value statistical test.

Geographic distribution across 18 provinces provided national representativeness within tertiary hospitals, though regional variations in technology infrastructure and AI policy implementation may influence adoption patterns differently across provinces. The professional mix, while reflecting typical surgical team composition, reflected the concerning underrepresentation of anaesthesiologists (4.7%) and limited our ability to assess adoption patterns in this critical perioperative speciality, where AI applications show particular promise for monitoring and predictive analytics.

## The Cognition and Usage Status of Surgical Medical Staff Regarding DeepSeek

Analysis of DeepSeek understanding and usage patterns revealed a complex landscape of awareness, adoption and application preferences among surgical medical staff. Understanding levels demonstrated a clear hierarchy: while complete unfamiliarity was minimal (2.4%), substantial proportions remained at superficial awareness levels, with 284 (67.0%) participants reporting an understanding of basic functions but only 41 (9.7%) indicating proficient usage levels. This distribution suggests the successful initial dissemination of AI technology awareness, but significant barriers preventing progression to clinical competency. In terms of usage behaviour, 298 (70.3%) respondents used it occasionally (1–3 times per week), and the top three main application scenarios were teaching and research support (43.2%), other life services (35.6%) and patient services (29.2%). The “other life services” category specifically encompassed four primary areas of AI utilization: (1) health consultation services for responding to general medical inquiries and providing preliminary health guidance; (2) chronic disease management support including patient monitoring, medication reminders, and lifestyle counseling; (3) health education and patient communication assistance for explaining medical conditions, treatment options, and post-operative care instructions; and (4) wellness monitoring and lifestyle recommendations encompassing nutrition guidance, exercise planning, and preventive health measures. These applications were characterized by their high-frequency, low-structured nature, requiring medical staff to quickly adapt AI assistance to diverse patient needs in fragmented clinical scenarios. This usage pattern reveals a clear risk stratification approach to AI adoption, with healthcare workers preferentially engaging DeepSeek in lower-stakes applications before progressing to direct patient care scenarios.

Detailed analysis of application scenarios demonstrates concerning underutilisation in high-value clinical domains as follows: auxiliary diagnosis and treatment decision-making (25.9%), cross-department consultation collaboration (10.8%) and real-time medical record quality control (10.8%). These low adoption rates in core clinical functions suggest significant untapped potential for AI impact on surgical workflows and patient outcomes. The predominance of teaching and research applications (43.2%) indicates recognition of AI value in knowledge management and educational contexts, potentially serving as stepping stones towards clinical implementation. User experience evaluation reveals differential satisfaction across system performance dimensions, providing insights into specific improvement priorities. Operational convenience received the highest satisfaction ratings (51.2% rating 5/5 points), suggesting successful user interface design and intuitive workflow integration. However, satisfaction declined for more critical performance indicators (result accuracy [35.4% rating 5/5] and system stability [30.4% rating 5/5]).

The satisfaction gradient from operational convenience to system reliability indicates that while basic usability has been achieved, core functionality requiring clinical confidence remains problematic. Only 35.4% of users expressed the highest confidence in result accuracy, with 22.9% providing neutral ratings (3/5), suggesting significant uncertainty about AI recommendation quality. System stability concerns were even more pronounced, with only 30.4% providing the highest ratings and 28.3% giving neutral responses, indicating frequent technical issues that could undermine clinical workflow integration.

These patterns suggest that initial technology adoption focused on user-friendly interface design has been successful, but fundamental system performance requirements for clinical deployment require substantial improvement to achieve the user confidence necessary for routine clinical use. Trust assessment revealed a cautious but pragmatic approach to AI integration among surgical staff. The predominant preference for auxiliary decision-making usage (60.1%) reflects appropriate clinical judgment regarding current AI capabilities and limitations. Notably, only 0.9% expressed a willingness to rely entirely on AI advice, while 38.7% indicated primary reliance on AI recommendations with human oversight, suggesting moderate confidence levels that could support supervised implementation protocols.

The extremely low rate of complete AI distrust (0.2%) indicates the absence of categorical rejection, while the minimal complete reliance (0.9%) demonstrates appropriate caution in high-stakes surgical environments. The distribution suggests that most healthcare workers occupy a middle ground of qualified acceptance, viewing AI as valuable but requiring human verification, a perspective aligned with current best practices for AI deployment in clinical settings.

This trust pattern provides an optimistic foundation for implementation strategies emphasising human–AI collaboration rather than automation, potentially addressing concerns about professional autonomy while leveraging AI capabilities for enhanced decision-making support.

Barrier analysis revealed a complex constellation of concerns spanning technical, ethical and practical domains. The three primary concerns—over-reliance on technology (58.5%), data privacy leakage risk (57.5%) and medical accidents from AI misjudgement (49.1%)—represent fundamental challenges requiring targeted intervention strategies. Over-reliance concerns reflect legitimate anxieties about maintaining clinical reasoning skills and professional judgment autonomy, indicating sophisticated understanding of AI limitations rather than uninformed resistance. Data privacy concerns reflect legitimate anxieties about patient information security in AI systems, particularly relevant given recent high-profile healthcare data breaches. The nearly equal prevalence of these two concerns indicates that both technical and professional autonomy issues require parallel attention in implementation planning. Medical accident fears demonstrate appropriate risk awareness in high-stakes surgical environments where diagnostic or treatment errors can have severe consequences, suggesting the need for robust AI validation data, transparent performance reporting and clear protocols for AI-assisted decision verification before clinical deployment. Additional details regarding barriers to use and training needs are presented in [Tables 2](#) and [3](#), respectively.

**Table 2** Cognition and Current Use Status of DeepSeek by Surgical Medical Staff

The Project		Number of People	Percentage (%)
Degree of understanding	Completely out of the loop	10	2.4
	Only the name	89	21.0
	Understand the basic features	284	67.0
	Proficiency in use	41	9.7
Experience evaluation (ease of operation)	1 point	11	2.6
	2 points	11	2.6
	3 points	81	19.1
	4 points	104	24.5
	5 points	217	51.2
Empirical evaluation was used (Accuracy of results)	1 point	5	1.2
	2 points	13	3.1
	3 points	97	22.9
	4 points	159	37.5
	5 points	150	35.4
Empirical evaluation was used (System stability)	1 point	14	3.3
	2 points	43	10.1
	3 points	120	28.3
	4 points	118	27.8
	5 points	129	30.4
Level of trust	Rely entirely on its advice	4	0.9
	Its recommendations are mainly referred to	164	38.7
	Only the aid decision was referred to	255	60.1
	Distrust	1	0.2
	Rely entirely on its advice	4	0.9

(Continued)

**Table 2** (Continued).

The Project		Number of People	Percentage (%)
Concerns about use	Over-reliance on technology	248	58.5
	Replace decision making	186	43.9
	It affects the quality of doctor-patient communication	148	34.9
	Risk of data privacy breach	244	57.5
	Technical misjudgment leads to medical accidents	208	49.1
	Ethical responsibility is poorly defined	168	39.6
	Complex operation	74	17.5
	Inadequate training	220	51.9
	Not compatible with existing systems	135	31.8
	Hospital policy or process restrictions	98	23.1
	There are doubts about technical reliability	180	42.5
No significant obstacles	98	23.1	

**Table 3** Training Needs and Development Recommendations for DeepSeek of Surgical Medical Staff

The Project		Number of People	Percentage (%)
Development of training content	Technical principle and scope of application	278	65.6
	Data security and ethical practices	272	64.2
	Question and feedback optimization skills	302	71.2
	Case analysis and practice guidance	313	73.8
	Interpretation and practical application of the results	273	64.4
	No training required	28	6.6
Want to optimize content	User Interface Design	174	41.0
	Streamlining operational processes	257	60.6
	Improved data accuracy	294	69.3
	Increased training support	201	47.4
	Equipment cost reduction	148	34.9
	Reduce waiting time	236	55.7
What will help improve work efficiency in the future	Other	17	4.0
	Real-time intraoperative navigation and operation guidance	234	55.2
	Postoperative recovery and remote monitoring	262	61.8
	Medical insurance cost control and DRGs prediction	257	60.6
	Dynamic optimization of medical resources	279	65.8
	Research data automation	287	67.7
Other	42	9.9	

## Training Needs and Development Suggestions of Surgical Medical Staff for DeepSeek

Training analysis revealed a critical implementation gap with profound implications for AI adoption success. The near-universal absence of formal training (94.8% untrained), combined with an overwhelming desire for systematic education (93.4%), indicates significant unmet educational needs that likely contribute to low clinical utilisation rates.

The training demand consensus across diverse professional roles and experience levels suggests that current informal learning approaches, such as peer consultation or self-directed exploration, have proven insufficient for building clinical competence. This finding emphasises the necessity of structured, competency-based training programmes as prerequisites for meaningful AI integration, rather than serving as optional enhancements.

Training analysis revealed sophisticated understanding of learning needs, with clear preference for practical application over theoretical knowledge. The top three training content demands were case analysis and practical guidance (73.8%), emphasizing hands-on, clinically relevant instruction; question and feedback optimization skills (71.2%), indicating recognition that effective AI utilization requires interactive competence; and technical principles and application scope (65.6%), suggesting desire for foundational understanding to enable appropriate use case selection. This

content priority pattern provides clear guidance for curriculum development emphasizing clinical scenarios, hands-on practice and competency assessment, rather than abstract AI concepts or programming details.

In terms of system development suggestions, 69.3% of medical staff suggested improving data accuracy, 60.6% hoped to simplify the operation process, and 55.7% proposed reducing system response time. In addition, respondents suggested that application scenarios, such as the automatic processing of scientific research data, dynamic optimisation of medical resources and remote monitoring of postoperative recovery, should be strengthened in future.

## Results of Single-Factor Analysis of Influencing Factors of Surgical Medical Staff's Understanding of DeepSeek

Univariate analysis identified multiple demographic, professional and experiential factors significantly associated with DeepSeek understanding levels, providing insights into adoption pattern determinants. Age-related differences ( $P = 0.026$ ) suggest generational variations in technology acceptance, with younger healthcare workers potentially demonstrating greater AI familiarity. Educational background significance ( $P < 0.001$ ) indicates that advanced degrees correlate with higher AI understanding, possibly reflecting research exposure or continuing education participation.

Professional role differences ( $P = 0.007$ ) highlight speciality-specific variations in AI adoption readiness, with certain surgical subspecialties showing enhanced receptivity, potentially due to technology integration in their clinical practice areas. Departmental variations ( $P = 0.002$ ) suggest that environmental and workflow factors influence individual adoption patterns, with some units providing more supportive contexts for AI exploration.

Usage frequency indicating the strongest association ( $P < 0.001$ ) confirms the expected relationship between hands-on experience and understanding development, supporting experiential learning approaches in implementation strategies. Application scenario significance for teaching and research support ( $P < 0.001$ ) and other life services ( $P = 0.001$ ) indicates that specific use cases facilitate learning and competence development more effectively than others.

Training exposure effects ( $P = 0.008$ ), despite the small, trained cohort (5.2%), suggest a potential high impact of formal education interventions on understanding levels, supporting investment in comprehensive training programme development.

## Results of Multivariate Analysis of Influencing Factors of Surgical Medical Staff's Understanding of DeepSeek

Multivariable ordinal logistic regression modelling was conducted to identify independent predictors of DeepSeek understanding while controlling for potential confounding variables. Model construction followed a systematic approach: the initial inclusion of all variables achieving  $P < 0.15$  in univariate analysis, followed by backwards elimination using likelihood ratio tests to achieve the most parsimonious model while maintaining an adequate fit (Hosmer–Lemeshow goodness-of-fit  $P = 0.642$ ).

The dependent variable (understanding level) was treated as an ordinal outcome with four categories: “completely unknown”, “only heard the name”, “understand basic functions” and “skilled use”. The proportional odds assumption was satisfied (Brant test,  $P = 0.324$ ), confirming that predictor variable effects remained consistent across all threshold comparisons in the ordinal outcome.

Variable coding utilised clinically meaningful reference categories to facilitate interpretation: for professional roles, “other” served as a reference representing the most diverse group; for departments, “other” provided a baseline comparison; for usage frequency, “daily use” represented the highest engagement level; for application scenarios, binary coding (“used” vs “not used”) enabled the assessment of specific usage pattern effects on understanding development.

Model diagnostics confirmed the absence of problematic multicollinearity (maximum VIF = 2.3) and an adequate sample size for stable parameter estimation (minimum 15 events per predictor variable achieved for all included variables). Age (years): 18–30 years old = 1; 31–40 years old = 2; 41–50 years old = 3; older than 50 years of age = 4. Education: college degree or below = 1; undergraduate = 2; master level = 3; PhD and above = 4. Occupation: surgeon = 1; operating room nurse = 2; ward nurses = 3; anaesthesiologist = 4; others = 5. Department: operating room = 1; anaesthesiology = 2; breast and nail surgery = 3; general surgery = 4; orthopaedics = 5; cardiothoracic surgery = 6;

neurosurgery = 7; urinary surgery = 8; otorhinolaryngology = 9; obstetrics and gynaecology = 10; hepatobiliary surgery = 11; others = 12. Frequency of use: never used = 1; occasionally = 2; 1–3 times per week = 3; daily use = 4. Whether received training: yes = 1; no = 2. Surgical risk warning, teaching and research support and other life services were replaced by an original value. The results of multivariate analysis showed that the departments were the operating room and neurosurgery department, the frequency of usage was never used or occasionally used, and the usage scene was other life services were the influencing factors of surgical medical staff’s DeepSeek understanding. The results of the multivariate ordinal logistic regression analysis are shown in Table 4.

**Table 4** Ordinal Logistic Regression Analysis of the Influencing Factors of Surgical Staff Understand the Degree of the DeepSeek (n=424)

The Project		$\beta$ Value	Standard Error	Wald $\chi^2$	P	OR Value (95% CI)
Constant term 1		-6.849	1.263	29.390	0.000	-
Constant term 2		-3.125	1.220	6.564	0.010	-
Constant term 3		3.230	1.160	7.752	0.005	-
Age	18–30 years old	0.535	0.623	0.737	0.391	0.586 (0.170–1.990)
	31–40 years old	0.062	0.603	0.011	0.918	0.940 (0.290–3.060)
	41–50 years old	-0.189	0.643	0.086	0.769	1.208 (0.340–4.260)
	Over 50 years of age	0				-
Education background	College or below	0.778	0.936	0.691	0.406	0.459 (0.070–2.880)
	Undergraduate	0.576	0.797	0.522	0.470	0.562 (0.120–2.680)
	Master’s degree	1.085	0.785	1.911	0.167	0.338 (0.070–1.570)
	PhD or above	0				-
Occupations	The surgeon	0.360	0.643	0.313	0.576	0.698 (0.200–2.460)
	Operating room nurse	-0.650	0.964	0.455	0.500	1.916 (0.290–12.670)
	Ward nurse	0.132	0.482	0.074	0.785	0.876 (0.340–2.250)
	Anesthesiologists	-0.925	1.191	0.603	0.438	2.522(0.240–26.030)
	Other	0				-
Department of Work	Operating room	1.915	0.953	4.037	0.045	0.147 (0.020–0.950)
	Department of Anesthesiology	0.898	1.056	0.723	0.395	0.407 (0.050–3.230)
	Breast and nail surgery	1.291	0.770	2.812	0.094	0.275 (0.060–1.240)
	General surgery	0.951	0.656	2.100	0.147	0.386 (0.110–1.400)
	Department of Orthopedics	0.303	0.845	0.129	0.720	0.739 (0.140–3.870)
	Cardiothoracic Surgery	-1.890	1.082	3.052	0.081	6.619 (0.790–55.190)
	Department of Neurosurgery	2.599	0.926	7.882	0.005	0.074 (0.010–0.460)
	Urology	1.174	0.897	1.712	0.191	0.309 (0.050–1.790)
	Department of Five Senses	0.317	0.855	0.137	0.711	0.728 (0.140–3.890)
	Obstetrics and gynecology	0.632	0.899	0.494	0.482	0.532 (0.090–3.100)
	Hepatobiliary surgery	0.173	0.634	0.074	0.785	0.841 (0.240–2.910)
	Other	0				-
	Frequency of use	Never used	-6.629	0.701	89.539	0.000
For occasional use		-3.480	0.622	31.305	0.000	32.460 (9.590–109.850)
1–3 times per week		-0.672	0.482	1.942	0.163	1.958 (0.760–5.040)
Every day		0				-
Early warning of surgical risk	Risk =0	0.405	0.343	1.391	0.238	0.667 (0.340–1.310)
	Risk =1	0				-
Teaching and research support	Risk =0	-0.268	0.305	0.774	0.379	1.307 (0.720–2.380)
	Risk =1	0				-
Other Life services	Risk =0	0.793	0.316	6.307	0.012	0.452 (0.240–0.840)
	Risk =1	0				-
Whether you have received training	Yes	1.069	0.595	3.229	0.072	0.343 (0.110–1.100)
	No	0				-

## Discussion

### Surgical Medical Staff Have a High Understanding of DeepSeek, but the Frequency of Using DeepSeek is Low

The research data showed that 67.0% of the surgical medical staff understood the basic functions of DeepSeek, but the actual usage rate showed a significant “high understanding-low usage” phenomenon, of which only 9.7% reached a proficient level of use, and 21.5% of the medical staff had never used the system. This finding is consistent with the global status of clinical AI applications.<sup>5,9–13</sup> Existing studies have shown that only 10–30% of medical practitioners truly incorporate AI systems into their daily diagnoses and treatment practices, indicating that although the potential of DeepSeek in the medical field has been widely recognised, its integration into actual clinical workflows still faces significant obstacles.<sup>14</sup>

Further analysis showed that there were significant risk gradient differences in the usage scenarios involving DeepSeek. The system was used relatively frequently in low-risk scenarios, such as teaching and research (43.2%) and other life services (35.6%), while it was used significantly less frequently in core medical scenarios, such as cross-department consultation and real-time case quality control (10.8%). This phenomenon is mainly attributed to the constraints of technical reliability, data quality and doctor–patient trust concerning application in high-risk scenarios.

Based on the above findings, this study suggests that AI technology developers should establish a closer cooperation mechanism with surgical medical staff to ensure that technical functions can effectively solve clinical pain points through regular feedback and iterative optimisation, thereby improving the practicability and clinical value of the technology.

### Surgical Medical Staff Used DeepSeek as an Auxiliary Tool but Had Concerns About Its Potential Risks

Results of the current study showed that 60.1% of respondents viewed DeepSeek as an auxiliary tool but had significant concerns about its potential risks. Specifically, 58.5% of the respondents expressed concerns about technology dependency, which is consistent with previous studies.<sup>15–17</sup> Over-reliance on technology may lead to deterioration in clinical thinking and impaired decision-making ability. Although 68% of medical staff recognised the advantages of AI in assisting decision-making<sup>5</sup> and optimising the diagnosis and treatment process,<sup>18</sup> manual review and double confirmation of AI-generated reports are commonly required in clinical practice,<sup>19</sup> reflecting the resistance to technology-led decision-making. This attitude of “accepting auxiliary” and “resisting substitution” indicates that medical decision-making should be dominated by medical staff, and the role of AI should be to support rather than replace professional judgment.<sup>17</sup>

Second, 57.5% of the respondents expressed concerns about the risk of medical data privacy breaches, a finding consistent with the results of relevant studies.<sup>19,20</sup> Medical data includes a large amount of sensitive information, and the use of massive medical data by models may violate patient privacy. In the process of data processing and transmission, medical data faces serious security risks. Data leakage or abuse may have a serious impact on national, social and personal security<sup>21</sup> and may lead to the collapse of patient trust in medical AI systems, triggering ethical disputes. In addition, 49.1% of the respondents worried that misjudgement on the part of DeepSeek may cause medical accidents, which is consistent with existing studies.<sup>22</sup> At the heart of this concern is the unpredictability and potential risk of error in medical AI outputs. This study showed that 65.8% of medical staff were sceptical about the responsible use of AI, and 57.7% lacked confidence in its harmlessness. Only 25% of respondents had high confidence in AI reliability.<sup>23,24</sup> Medical AI systems require much lower fault tolerance rates compared to applications in other fields. Its decisions directly affect the lives and health of patients, and wrong decisions may lead to serious consequences. Therefore, medical AI must have efficient reasoning and prediction capabilities to ensure accuracy, interpretability and reliability, as well as a strong fault tolerance mechanism and verification system.<sup>1</sup>

Based on the above findings, this study proposes the following coping strategies: first, the positioning of AI technology should be clarified and its role emphasised as an auxiliary tool. The principle that “technology does not replace a surgeon’s decision-making” must be repeatedly emphasised throughout operating guidelines and training courseware to eliminate any notion that “technology replaces human labour”. Second, data security protection must be strengthened, and privacy protection technologies such as federated learning, dynamic encryption and differential privacy

should be adopted to underscore the importance of data privacy and security issues.<sup>25</sup> In addition, to improve the reliability of the technology, through multi-centre clinical verification and continuous optimisation, surgical medical staff must be invited to participate in the construction and verification of DeepSeek technology<sup>5</sup> to reduce the risk of technical misjudgement and enhance the trust of medical staff in the technology. Concurrently, the ethical and legal policies of DeepSeek's application should be formulated to protect the privacy and rights of patients, as well as their right to be informed about AI involvement in their care and to participate in treatment decisions.<sup>26</sup> It is also necessary to clarify the responsibilities and obligations of the DeepSeek system in medical decision-making, thereby ensuring the transparency and interpretability of the system and preventing medical accidents and disputes caused by system errors. The above measures can effectively alleviate the concerns of medical staff and patients about DeepSeek and promote its wider application in clinical practice.

Deeper examination of the identified barriers reveals complex, interconnected challenges that extend beyond individual hesitation. The data privacy concerns reported by 57.5% of the respondents reflect not merely general cybersecurity awareness but specific anxieties about medical data vulnerability in AI systems. These concerns encompass multiple dimensions: patient confidentiality breaches through data sharing with external AI platforms, the potential commercial exploitation of sensitive medical information and inadequate transparency regarding data usage and storage practices. Healthcare workers specifically cited uncertainty about data retention periods, third-party access permissions and cross-border data transfers as primary sources of privacy-related reluctance. The over-reliance concerns expressed by 58.5% of participants revealed a fundamental tension between technological efficiency and clinical autonomy. Detailed analysis suggests that these worries stem from three primary sources: fear of skill atrophy due to reduced clinical reasoning practice, concerns about diminished diagnostic confidence when AI recommendations conflict with clinical judgment, as well as anxiety about legal responsibility when patient outcomes follow AI-guided decisions. Case examples from pilot implementations indicate that healthcare workers particularly struggle with scenarios where AI recommendations contradict established clinical intuition, creating decision-making paralysis, rather than enhancing confidence. Technical error anxieties, reported by 49.1% of respondents, reflect deeper systemic concerns about AI reliability in high-stakes environments. Specific fears include algorithm bias, leading to inappropriate treatment recommendations for certain patient populations, system failures during critical decision points and an inability to detect AI reasoning errors due to limited algorithmic transparency. Healthcare workers emphasised that, unlike other technology applications where errors cause inconvenience, AI mistakes in surgical settings could directly threaten patient safety; as such, near-perfect reliability is demanded from such a system, which currently cannot be guaranteed. Training inadequacy, identified by 51.9% of the participants, represents both a barrier and an opportunity for intervention. Detailed analysis revealed that current training approaches failed to address core competency needs, that is, understanding AI limitations and appropriate use cases, developing skills for interpreting and validating AI recommendations and building confidence in human–AI collaborative decision-making. Effective training programmes must move beyond basic system operation to encompass critical evaluation skills and ethical decision-making frameworks.

## Critical Perspectives Missing from Current Analysis

This study's focus on healthcare provider perspectives, while necessary for understanding adoption barriers, omits several critical stakeholder viewpoints that are essential for comprehensive AI implementation planning. Patient perspectives on AI-assisted surgical decision-making represent a fundamental gap, as patient trust and acceptance directly influence the viability of AI integration in clinical practice. Research in other healthcare AI applications suggests significant patient concerns about algorithmic bias, reduced physician attention during AI-assisted consultations and uncertainty about AI involvement in treatment decisions. Without understanding patient preferences for AI disclosure, their comfort with algorithm-guided recommendations, as well as their expectations concerning human oversight, implementation strategies may fail despite provider readiness.

Organisational and leadership perspectives constitute another critical omission. Hospital administrators, department chairpersons and technology implementation teams face distinct challenges, including budget allocation for AI systems, staff training coordination, legal liability management and integration with existing electronic health records. Leadership

support significantly influences adoption success, yet this study provides no insight into institutional readiness factors or change management strategies.

Additionally, the absence of AI system developers' perspectives limits an understanding of technical constraints and optimisation possibilities. Developer insights could help to illuminate the feasibility of suggested improvements, timeline requirements for system modifications, as well as cost implications concerning enhanced functionality. This gap prevents the realistic assessment of whether the identified barriers can be practically addressed through system design changes.

## The Vast Majority of Surgical Medical Staff Have Not Received DeepSeek Training, Serving as a Major Barrier to Its Use

This study found that 94.8% of the surgical medical staff had not been trained to use the DeepSeek system, but 93.4% of the respondents showed a strong willingness to learn. This finding is consistent with the common challenges faced by medical AI training worldwide.<sup>5,10,12,22,27</sup> Specifically, 53% of medical staff lacked basic AI knowledge, only 10–15% had practical experience and relevant knowledge reserves, and it was generally believed that the current training resources provided by hospitals failed to meet the actual needs. The study further revealed that the training needs of surgical medical staff were significantly concentrated, including mainly clinical case analyses and practical operation guidance (73.8%), questions and feedback skills in human–computer interaction (71.2%), as well as technical principles and application scope (65.6%). Systematic training can help medical staff to skilfully use AI technology, improve digital literacy and, accordingly, play a key role in establishing technical trust.<sup>28</sup>

Therefore, this study concludes that a standardised training system must be formulated. First, a stratified training plan should be developed. According to the post-survey characteristics and technical needs of surgical medical staff, a stratified course including basic theory, operational skills and case analysis should be designed. Second, practical guidance should focus on helping medical staff transform theoretical knowledge into practical operational ability through simulation procedures, case analysis and practical exercises. At the same time, a feedback mechanism must be established to set up a question and feedback link during the training process, solve the questions of medical staff in a timely manner and optimise the training content according to feedback. In addition, continuing education should be promoted; DeepSeek training should be included in the continuing education system for medical staff, and the training content should be updated regularly to ensure that knowledge and maintenance abilities remain current. These measures can help to effectively improve technical mastery and trust level, promote the standardised application of DeepSeek in surgical clinical practice and, accordingly, ensure the safety of patients.

Successful AI implementation also requires strong organisational support. Hospital leadership commitment, clear AI governance policies, and an adequate technical infrastructure are critical factors beyond individual acceptance. Research indicates that AI projects with high-level support have success rates 3.2 times higher than those lacking leadership commitment. The hierarchical structure and collective decision-making patterns in China's healthcare environment mean that senior physicians' technology acceptance may disproportionately influence junior colleagues' adoption behaviours. Future research should include organisational readiness assessments and leadership interviews.

## Surgical Medical Staff Provided Significant Clinical Guidance Regarding the Technical Optimisation Requirements for DeepSeek

This study found that the technical optimisation needs among surgical medical staff concerning DeepSeek had a significant clinical orientation, and 69.3% of the respondents mainly focused on the improvement of data accuracy, which is consistent with a study conducted by Bang et al.<sup>29</sup> Large AI models are inherently probabilistic models, where even small input differences can lead to significant output changes. In the medical and healthcare fields, the accuracy of current applications is still insufficient, and the instability and alignment deviation of models may lead to hallucinations and produce harmful or biased information, triggering the risk of misuse.<sup>29</sup> Therefore, it is recommended that medical and healthcare authorities strengthen medical data governance, gradually improve data quality and promote the establishment of high-quality and open medical standard data sets to promote the open sharing of data resources.<sup>30</sup> In addition, 60.6% of the respondents emphasised the need to simplify the system operation process, and 55.7% focused on the

optimisation of the response speed, which was consistent with the research results of Elhaddad et al.<sup>31</sup> Interruptions in workflow may directly affect work efficiency and patient safety. It is suggested that the human–computer interaction interface be optimised using user-centered design principles and a change management system covering system training, process reengineering and effectiveness evaluation be established to realise the deep integration of AI tools and clinical pathways.<sup>32,33</sup> Through these measures, the clinical application effect of DeepSeek can be effectively improved, and patient safety can be ensured.

Based on the findings of this study, we recommend a phased implementation strategy as follows. (1) Pilot phase: conduct 6-month pilots in departments with higher technology acceptance (operating rooms, neurosurgery). (2) Training phase: implement case analysis training targeting the 73.8% survey-based demand and human–computer interaction skills training addressing the 71.2% needs gap, based on the survey. (3) Expansion phase: gradually extend from low-risk applications (teaching and research) to clinical decision-making support. (4) Optimisation phase: continuously improve data accuracy based on the 69.3% survey user feedback and streamline the operational processes to address the 60.6% survey-based user concerns.

### **Surgical Medical Staff’s Understanding of DeepSeek is Affected by Many Factors Medical Staff Whose Departments Involved the Operating Room and Neurosurgery Department, and Who Used Scenarios Focused on Life Services, Had a Higher Understanding of DeepSeek**

This study showed that operating room and neurosurgery medical staff were more knowledgeable about DeepSeek than staff in other departments. As a key place for surgical operations and emergency and critical care rescue, nurses in operating rooms have high-precision and high-efficiency clinical needs and, as such, may be more strongly inclined to learn how to use AI technology. Neurointerventional surgery is accelerating the integration of generative AI technologies, for example, using generative adversarial networks to improve the resolution of medical images and NLP models to optimise doctor–patient communication. These technologies must meet strict requirements for accuracy and efficiency in neurosurgery and, in this way, can help encourage medical staff to pay attention to and accept AI tools. In addition, this study found that healthcare workers who focused their use case scenario on life services knew more about DeepSeek, which is consistent with the findings of a study conducted by Marinovich et al.<sup>14</sup> Life service scenarios (such as health consultation and chronic disease management) are high-frequency and low-structured, and medical staff must quickly respond to diverse needs in a fragmented scenario. The natural language interaction and instant information retrieval functions of DeepSeek complement the flexibility of such a scenario, reduce the threshold for technology use, enhance user engagement and improve cognitive depth. It is suggested that a specialised AI training system be constructed, an image-assisted decision-making module for the operating room be created, and the intelligent question-and-answer knowledge base for the chronic disease management department be optimised to enhance the clinical transformation value of technology application.

### **Surgical Medical Staff with Low Frequency Usage Have a Low Understanding of DeepSeek**

This study found that medical staff who never/occasionally used DeepSeek had a limited understanding of the technology, which is similar to the results of a study conducted by Qurashi et al.<sup>9</sup> The core reason for this lies in the lack of practical experience and limited technical understanding. Low-frequency use leads medical staff to only form a shallow understanding of basic functions (such as information retrieval) but limits their understanding of core capabilities (such as data analysis and decision-making suggestions) due to a lack of interaction. It is suggested that surgical medical staff embed DeepSeek into their daily diagnosis and treatment processes (eg preoperative planning, intraoperative navigation and postoperative follow-up) to help gradually deepen their understanding of DeepSeek through a “use–feedback–optimisation” cycle.

Patient acceptance of AI-assisted surgical decision-making represents an important limitation of this study. Previous research<sup>23</sup> suggests that patients generally prefer to be explicitly informed when AI is involved in their treatment decisions, and many patients tend to trust human expert opinions more when AI recommendations conflict with physician judgment. However, when AI systems provide transparent reasoning pathways, patient acceptance increased from 42% to 71%. Traditional Chinese patients’ trust in a physician in a position of authority may influence AI acceptance, indicating

the need for specialised patient education and transparency protocols. Future research should investigate patient attitudes toward AI-assisted surgical planning and its impact on treatment adherence.

## Adaptation Strategies for Resource-Constrained Healthcare Settings

While our study focused on tertiary hospitals with advanced infrastructure, the adaptation of DeepSeek for low-resource hospitals with limited digital infrastructure represents a critical implementation challenge. Based on our findings regarding technical optimization needs (69.3% emphasizing data accuracy improvement, 60.6% requesting operational simplification), we propose several adaptation strategies for resource-constrained environments: (1) Lightweight system architecture: Development of streamlined DeepSeek versions with reduced computational requirements that can operate effectively on basic hardware configurations commonly found in secondary and primary care facilities; (2) Offline/semi-offline capabilities: Implementation of local processing modules that can function with intermittent internet connectivity, storing essential medical knowledge databases locally while synchronizing with central systems when connectivity permits; (3) Simplified user interfaces: Design of intuitive, touch-based interfaces requiring minimal training, addressing the 51.9% barrier of insufficient training identified in our study; (4) Modular implementation approach: Development of scalable modules allowing hospitals to adopt specific AI functions (diagnostic assistance, treatment recommendations, or documentation support) based on their infrastructure capacity and clinical priorities; (5) Cost-effective deployment models: Implementation of shared-service models where multiple smaller hospitals can access centralized AI processing through cloud-based solutions, reducing individual infrastructure costs while maintaining functionality.

## Summary

This study reveals significant implementation gaps for DeepSeek in surgical settings. While 67.0% of medical staff understood its basic functions, only 9.7% achieved proficient usage levels, with practical applications primarily confined to low-risk scenarios. Key barriers include inadequate training (94.8%), data privacy concerns (57.5%), and over-reliance fears (58.5%).

To accelerate meaningful AI adoption, healthcare institutions should: (1) implement evidence-based stratified training programmes targeting the identified knowledge gaps; (2) establish clear governance frameworks addressing ethical and legal responsibilities; (3) develop staged implementation protocols progressing from low-risk to high-stakes clinical scenarios. Regulatory bodies should provide explicit guidance on AI accountability in surgical decision-making, while technology developers must prioritise transparent performance reporting and robust validation studies to build justified confidence among surgical practitioners. Additionally, multi-centre prospective studies are needed to validate the actual impact of AI-assisted surgical planning on patient outcomes, providing evidence for evidence-based implementation.

## Limitations

First, the scope of the survey was limited to Class III Grade A hospitals, which represent the highest-level medical institutions in China. This greatly restricts the generalizability of the research results to lower-level hospitals (which cover a wider population) facing different resource constraints and infrastructure limitations. At the same time, it may overestimate the technical readiness for artificial intelligence applications and the acceptance of staff in resource-limited medical environments. Second, the convenience sampling method used in this study may lead to potential biases in the research results. On the one hand, medical staff who are interested in or hold a positive attitude towards artificial intelligence technology are more likely to participate in this survey. This may result in an overestimation of the overall acceptance level, and the proportion of relevant groups in the sample may even be over-represented. On the other hand, the under-representation of certain professional groups (eg anesthesiologists only account for 4.7%) limits our ability to draw meaningful conclusions about departmental differences in artificial intelligence application models. Third, the cross-sectional design used in this paper can only provide a snapshot of the current attitudes and usage patterns. It cannot evaluate how cognition evolves over time, nor can it assess how the initial adoption experience affects long-term usage behavior. In addition, it must be acknowledged that there are some methodological limitations: the survey did not conduct a formal test-retest reliability assessment, which may affect the stability of the measurement results over time; the questionnaire did not include a validated technology acceptance or change readiness scale, which could have provided

a more detailed understanding of the determinants of application; and the study completely relied on self-reported usage data without objective verification through system logs or direct observation, which may introduce social desirability bias, that is, respondents may exaggerate the use of artificial intelligence or hide their concerns to show their technical competence. Finally, this study did not conduct in-depth exploration of the personal experiences of surgical medical staff through qualitative methods.

Future research should consider including hospitals at different levels to improve the generalizability and representativeness of the research results. Combining longitudinal design and qualitative research methods, future research should dynamically track changes in the usage behavior of surgical medical staff, and conduct in-depth analysis of their usage experiences and the impact on patients' clinical outcomes, so as to provide a more comprehensive basis for evaluating the quality and safety of DeepSeek.

## Data Sharing Statement

All data generated or analyzed during this study are included in the article.

## Ethics Approval and Consent to Participate

This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Hunan Provincial People's Hospital.

## Funding

Application of goal-directed hypothermia technology based on multimodal data fusion in brain protection of patients with type A aortic dissection (2023JJ60313).

## Disclosure

None of the authors have any personal, financial, commercial, or academic conflicts of interest.

## References

1. Lv HR, Jiang R, Zhang XG, et al. DeepSeek and medical big language model: technology innovation and medical service model reconstruction. *J Med Intell.* 2025;46(02):1–7, 13. doi:10.3969/j.issn.1673-6036.2025.02.001
2. Elsevier. Clinician of the Future Report 2023. 2025. Available from: <https://www.elsevier.com/resources/clinician-of-the-future-2023>. Accessed November 14, 2025.
3. National Health Commission. Notice of the General Office of the National Health Commission on issuing reference Guidelines for the application scenarios of artificial intelligence in the health industry. Beijing: People's Medical Publishing House; 2024.
4. Nelson CA, Pachauri S, Balk R, et al. Dermatologists' perspectives on artificial intelligence and augmented intelligence - a cross-sectional survey. *JAMA Dermatol.* 2021;157(7):871–874. doi:10.1001/jamadermatol.2021.1685
5. Chen M, Zhang B, Cai Z, et al. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. *Front Med.* 2022;9:990604. doi:10.3389/fmed.2022.990604
6. Sun ZQ, Xu YY. *Medical Statistics*. 4th ed. Beijing: People's Medical Publishing House; 2014.
7. Li M, Li YX, Dai L, et al. Analysis of doctors' willingness to use medical artificial intelligence and its influencing factors. *J Med Intell.* 2024;45(5):26–31. doi:10.3969/j.issn.1673-6036.2024.05.005
8. Ba HJ, Chen JR, Hu H, et al. Resident on the application of the artificial intelligence cognition and attitude survey. *Chin J Med Educ.* 2025;45(3):194–197. doi:10.3760/cma.j.cn115259-20241126-01215
9. Qurashi AA, Alanazi RK, Alhazmi YM, et al. Saudi radiology personnel's perceptions of artificial intelligence implementation: a cross-sectional study. *J Multidiscip Healthc.* 2021;14:3225–3231. doi:10.2147/JMDH.S340786
10. Scheetz J, Rothschild P, McGuinness M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. *Sci Rep.* 2021;11(1):5193. doi:10.1038/s41598-021-84698-5
11. Leenhardt R, Fernandez-Urien Sainz I, Rondonotti E, et al. PEACE: perception and expectations toward artificial intelligence in capsule endoscopy. *J Clin Med.* 2021;10(23):5708. doi:10.3390/jcm10235708
12. Zheng B, Wu MN, Zhu SJ, et al. Attitudes of medical workers in China toward artificial intelligence in ophthalmology: a comparative survey. *BMC Health Serv Res.* 2021;21(1):1067. doi:10.1186/s12913-021-07044-5
13. Reeder K, Lee H. Impact of artificial intelligence on US medical students' choice of radiology. *Clin Imaging.* 2022;81:67–71. doi:10.1016/j.clinimag.2021.09.018
14. Marinovich ML, Wylie E, Lotter W, et al. Artificial intelligence (AI) to enhance breast cancer screening: protocol for population-based cohort study of cancer detection. *BMJ Open.* 2022;12(1):e054005. doi:10.1136/bmjopen-2021-054005
15. Xiang Q, Yang W. Integration and innovation of ChatGPT in clinical medical education: opportunities, challenges and strategies. *Health Vocat Educ.* 2025;43(01):1–4. doi:10.20037/j.issn.1671-1246.2025.01.01

16. Wadhwa V, Alagappan M, Gonzalez A, et al. Physician sentiment toward artificial intelligence (AI) in colonoscopic practice: a survey of US gastroenterologists. *Endosc Int Open*. 2020;8(10):E1379–E1384. doi:10.1055/a-1223-1926
17. Sarwar S, Dent A, Faust K, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ Digit Med*. 2019;2:28. doi:10.1038/s41746-019-0106-0
18. Huisman M, Ranschaert E, Parker W, et al. An international survey on AI in radiology in 1041 radiologists and radiology residents part 2: expectations, hurdles to implementation, and education. *Eur Radiol*. 2021;31(11):8797–8806. doi:10.1007/s00330-021-07782-4
19. Lim SS, Phan TD, Law M, et al. Non-radiologist perception of the use of artificial intelligence (AI) in diagnostic medical imaging reports. *J Med Imaging Radiat Oncol*. 2022;66(8):1029–1034. doi:10.1111/1754-9485.13388
20. Ray TR, Kellogg RT, Fargen KM, et al. The perils and promises of generative artificial intelligence in neurointerventional surgery. *J Neurointerv Surg*. 2023;16(1):4–7. doi:10.1136/jnis-2023-020353
21. Zhu RS, Feng ZW, Chen Q, et al. International security challenges and governance of artificial intelligence. *Forum Sci Technol China*. 2023;2023(3):160–167, 179.
22. Shelmerdine SC, Rosendahl K, Arthurs OJ. Artificial intelligence in paediatric radiology: international survey of health care professionals' opinions. *Pediatr Radiol*. 2022;52(1):30–41. doi:10.1007/s00247-021-05195-5
23. Nong P, Platt J. Patients' trust in health systems to use artificial intelligence. *JAMA Network Open*. 2025;8(2):e2460628. doi:10.1001/jamanetworkopen.2024.60628
24. Jungmann F, Jorg T, Hahn F, et al. Attitudes toward artificial intelligence among radiologists, IT specialists, and industry. *Acad Radiol*. 2021;28(6):834–840. doi:10.1016/j.acra.2020.04.011
25. Hagos DH, Battle R, Rawat DB. Recent advances in generative AI and large language models: current status, challenges, and perspectives. *IEEE Trans Artif Intell*. 2024;5(12):5873–5893. doi:10.1109/TAI.2024.3444742
26. Li T, Zhang J, Li YZ, et al. Innovative applications, risks, challenges and governance countermeasures of artificial intelligence in healthcare. *J Med Intell*. 2020;46(01):2–8, 16. doi:10.3969/j.issn.1673-6036.2025.01.001
27. Abuzaid MM, Elshami W, Tekin H, et al. Assessment of the willingness of radiologists and radiographers to accept the integration of artificial intelligence into radiology practice. *Acad Radiol*. 2022;29(1):87–94. doi:10.1016/j.acra.2020.09.014
28. Banerjee M, Chiew D, Patel KT, et al. The impact of artificial intelligence on clinical education: perceptions of postgraduate trainee doctors in London (UK) and recommendations for trainers. *BMC Med Educ*. 2021;21(1):429. doi:10.1186/s12909-021-02870-x
29. Bang Y, Cahyawijaya S, Lee N, et al. A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity. 2023. doi:10.48550/arXiv.2302.04023.
30. Ma A, Ge XL. Research on the application of artificial intelligence large model in medical and health field. *China Mod Dr*. 2024;62(33):89–95. doi:10.3969/j.issn.1673-9701.2024.33.018
31. Elhaddad M, Hamam S. AI-driven clinical decision support systems: an ongoing pursuit of potential. *Cureus*. 2024;16(4):e57728. doi:10.7759/cureus.57728
32. Golden G, Popescu C, Israel S, et al. Applying artificial intelligence to clinical decision support in mental health: what have we learned? *Health Policy Technol*. 2024;13(2):100844. doi:10.1016/j.hlpt.2024.100844
33. Shamszare H, Choudhury A. Clinicians' perceptions of artificial intelligence: focus on workload, risk, trust, clinical decision making, and clinical integration. *Healthcare*. 2023;11(16):2308. doi:10.3390/healthcare11162308

Journal of Multidisciplinary Healthcare

Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or healthcare processes in general. The journal covers a very wide range of areas and welcomes submissions from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-multidisciplinary-healthcare-journal>

**Dovepress**  
Taylor & Francis Group