Original Paper

Chinese Pathologists' Adoption of Large Language Models: Cross-Sectional Survey on Influencing Factors

Yuting Wang[1*], MPH; Victor Yu Cui[2*], MS; Qinjing Pan[3], MD; Xun Zhang[3], MD; Yao Yang[1], PhD; Peng Xue[1], PhD; Youlin Qiao[1], MD, PhD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]Vanke School of Public Health, Tsinghua University, Beijing, China

[3]Department of Cancer Epidemiology, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Youlin Qiao, MD, PhD
School of Population Medicine and Public Health
Chinese Academy of Medical Sciences and Peking Union Medical College
No 9, Dongdan Third Street, Dongcheng District
Beijing
China
Phone: 86 18104897753
Email: qiaoy@cicams.ac.cn

Abstract

Background: The growing demand for pathology services and the shortage of pathologists in China have created an urgent need for innovative tools to improve diagnostic efficiency and support professional practice. Large language models (LLMs) may assist pathologists in clinical practice, scientific research, and medical education, but their implementation depends on many factors, such as pathologists' awareness and willingness to adopt these technologies.

Objective: This study aimed to assess Chinese pathologists' awareness, practical experience, perceptions, ethical concerns regarding LLMs; their willingness to adopt LLMs; and factors associated with willingness to use LLMs in pathology.

Methods: A cross-sectional questionnaire survey was conducted among certified pathologists practicing in mainland China. The survey assessed participants' demographic and professional characteristics; LLM awareness, use, and perceived performance in clinical practice, scientific research, and medical education; ethical concerns; and suggestions for future improvement. Questionnaire reliability and validity were assessed using Cronbach α, Guttman Lambda 6, and the Kaiser-Meyer-Olkin measure. Group differences in perceived performance were examined using Mann-Whitney U tests with Benjamini-Hochberg false discovery rate (FDR) adjustment. Factors associated with willingness to use LLMs in pathology were measured as the summed score of three 5-point Likert items and analyzed using ordinal logistic regression.

Results: Of 339 eligible pathologists, 248 (73.2%) had heard of LLMs, and 64 of these (25.8%) had used them. ChatGPT was the most recognized and most commonly used model. Compared with nonusers, LLM users rated clinical applications more cautiously, particularly for analyzing patient data, organizing medical history, and supporting the understanding of disease mechanisms and treatments; these differences were significant after domain-specific FDR adjustment (adjusted *P*=.02) but borderline after overall correction across all performance items (adjusted *P*=.05). Perceptions of research and educational applications did not differ significantly between users and nonusers. A higher education level, attending physicians (adjusted odds ratio [aOR] 2.39, 95% CI 1.09-5.22; *P*=.029), chief physicians (aOR 4.64, 95% CI 1.57-13.69; *P*=.005), and more than 43 weekly working hours (aOR 2.33, 95% CI 1.40-3.89; *P*=.001) were associated with greater willingness to use LLMs, whereas working in fourth- or lower-tier cities (aOR 0.42, 95% CI 0.22-0.79; *P*=.007) and reading 1001-3000 slides annually (aOR 0.43, 95% CI 0.20-0.94; *P*=.034) were associated with less willingness. Major ethical concerns included model transparency, sensitive data security, and academic misuse.

Conclusions: Chinese pathologists show high awareness of LLMs and generally favorable attitudes toward their potential application in pathology, but actual use is limited. The gap between awareness and use highlights the need for improved accessibility, targeted training, pathology-specific validation, workflow integration, and ethical governance. Future implementation should

XSL•FO
RenderX

emphasize locally adapted models, data security, and equitable access, particularly for pathologists in lower-tier cities and resource-limited settings.

(JMIR Med Inform 2026;14:e88275) doi: 10.2196/88275

KEYWORDS

pathologists; large language models; ChatGPT; artificial intelligence; AI; technology adoption; cross-sectional survey

Introduction

With the aging population growing and the public awareness of cancer prevention rising in China, an increasing number of residents are actively participating in physical examinations and national screening programs [1,2]. Pathological diagnosis plays a crucial role in these initiatives, as timely and accurate pathology services are essential for early detection, appropriate treatment decisions, and improved patient outcomes. Consequently, the expanding participation in screening programs has significantly intensified the demand for pathology services across the country.

The field of pathology in China faces two critical interrelated challenges. First is a severe shortage of pathology professionals. According to the 2022 *China Health and Health Statistics Yearbook*, there are only ~18,700 registered pathologists and assistant physicians in China, substantially fewer than the estimated demand of over 100,000 professionals, resulting in an alarming gap of 80,000 professionals [3]. This shortage severely impacts patient care, as it leads to delays in diagnosis and increased risk of diagnostic errors, particularly in rural and remote areas where access to high-quality pathology services is extremely limited. Notably, the number of pathology departments in provinces such as Gansu, Qinghai, Ningxia, Inner Mongolia, Tibet, and Xinjiang account for only 11.7% nationwide in China [4].

The second major challenge is the overwhelming workload experienced by Chinese pathologists. According to the "China Pathology Department Survey Report," the average annual workload of a Chinese pathologist is ~4910 cases. This figure exceeds several reported workload estimates from developed health care systems, including a Canadian anatomic pathology study that reported an optimal workload of 2659 surgical pathology cases and 11,323 slides per full-time pathologist and historical UK guidance suggesting ~4000 biopsies per year for pathologists in nontraining hospital settings [5,6]. Chinese pathologists' substantial workload, coupled with the severe shortage and uneven geographic distribution of resources, directly restricts their professional development and significantly limits patients' equitable access to high-quality diagnostic services.

Given these persistent and complex challenges, conventional approaches alone appear insufficient, highlighting the urgent need for innovative solutions to improve the overall quality and responsiveness of pathology services in China. To address these issues, artificial intelligence (AI), particularly large language models (LLMs), presents a promising solution. LLMs possess the capability to process and analyze vast amounts of unstructured clinical data, automate routine pathology tasks, and offer real-time diagnostic assistance, thereby directly addressing the critical challenges identified before. Specifically, LLMs can significantly alleviate pathologists' workload by automating pathology report generation, preliminary screening of digital pathology images, and real-time identification of potential diagnostic errors [7-9]. A practical example is PathChat, an AI-powered tool developed by Modella AI in collaboration with Harvard Medical School. This tool demonstrates promising accuracy in identifying diseases from digital pathology images, offering real-time diagnostic support and significantly reducing diagnostic turnaround times [10]. Tools such as PathChat can be feasibly integrated into Chinese pathology workflows to enhance diagnostic efficiency, improve accuracy, and expand equitable access to high-quality pathology services, particularly in rural and resource-limited regions. However, successful implementation of such advanced AI tools depends heavily on pathologists' willingness to adopt and integrate these technologies into their workflows.

This study aimed to (1) systematically evaluate the perceptions and willingness of Chinese pathologists to adopt LLMs, such as ChatGPT, and (2) identify factors influencing the acceptance of LLMs. Understanding these factors is essential for overcoming potential barriers to technological adoption, facilitating effective integration of LLMs in pathology practice, and, ultimately, addressing the critical workforce and quality challenges in pathology services in China. Insights from this study can also provide valuable guidance for other middle-income countries facing analogous health care service challenges.

Methods

Ethical Considerations

This study was approved by the Institutional Review Board of the Chinese Academy of Medical Sciences and Peking Union Medical College (approval #CAMS and PUMC-IEC-2022-022) and adhered to the relevant principles of the Declaration of Helsinki, ensuring that the rights and confidentiality of all participants are fully protected. Before accessing the questionnaire, participants viewed an electronic information page describing the study purpose, eligibility criteria, survey content, voluntary nature of participation, confidentiality protections, data-handling procedures, and the right to withdraw before submission. Electronic informed consent was obtained from all participants through the Wenjuanxing platform before survey completion. By completing the survey, participants provided their consent to participate in the study. The requirement for written informed consent was waived by the ethics committee because electronic consent was deemed sufficient for this study as the survey was anonymous and involved minimal risk. All participants also provided informed

XSL•FO
RenderX

consent for the publication of the results. No personally identifying information was collected, and all data collected from the participants, including survey responses and any other relevant information, were anonymized to ensure that no individual can be identified and were used solely for academic research purposes. A nonmonetary online pathology learning resource was provided after survey completion as a token of appreciation.

Questionnaire Survey

The questionnaire was developed using Wenjuanxing, an online questionnaire platform widely used in China. Wenjuanxing was selected because it supports QR code distribution, mobile access, anonymous responses, structured questionnaire design, and data export, making it suitable for conference-based web surveys.

The questionnaire was based on a literature review and expert consultation. A presurvey was conducted to assess clarity, relevance, reliability, and validity. The final questionnaire included five sections: demographic and professional characteristics, awareness and prior use of LLMs, perceived utility and willingness to use LLMs, ethical concerns, and future improvement priorities. A total of 21 statements were finalized, including 12 application-related items and 9 ethical concern items. The complete questionnaire is provided in Multimedia Appendix 1.

Eligible participants were certified pathologists practicing in mainland China; pathology medical students, nonphysician industry professionals, and staff from pathology-related companies were excluded. The survey was initially distributed on-site during the 2024 National Pathology Academic Annual Conference held in Xi'an, China, from April 26 to 27, 2024. During the conference, trained research team members invited eligible participants to complete the anonymous online questionnaire by scanning a QR code. Participants were also allowed to share the survey link or QR code with eligible pathologist colleagues. Because of this colleague referral component, the questionnaire remained open until May 15, 2024, and the overall data collection period was from April 26 to May 15, 2024. Because the survey was distributed in an open conference setting and could be shared among eligible colleagues, the number of individuals who viewed the invitation but did not start or submit the questionnaire could not be determined; therefore, a conventional response rate could not be calculated. We, instead, reported the number of submitted responses, exclusions, and eligible responses included in the analysis.

Participants reported the city in which they practiced. The reported work cities were standardized and mapped to the 2024 Ranking of Cities' Business Attractiveness released by the China Business Network New First-Tier Cities Research Institute on May 30, 2024. This ranking is based on the commercial resource concentration, urban connectivity, urban population activity, new economy competitiveness, and future potential. For analysis, city tiers were collapsed into three groups: first- or new first-tier cities, second- and third-tier cities, and fourth- or lower-tier cities. City tier was used as a proxy for a broader urban development and resource context, not as a direct measure of hospital-level pathology capacity. The mapping of unique participant work cities is provided in Multimedia Appendix 2.

Participants were first asked whether they had heard of LLMs before the survey. Those who had heard of LLMs were further asked about information sources, prior use, reasons for nonuse, language used, and subscription to paid LLM services. The perceived utility of LLMs was assessed across clinical practice, scientific research, and medical education using nine items rated on a 5-point Likert scale from 1 to 5, with higher scores indicating stronger agreement; agreement was defined as a rating of 4 or 5. The primary outcome was willingness to use LLMs, measured using three items rated on a 5-point Likert scale covering clinical, research, and educational applications. The summed willingness score ranged from 3 to 15, with higher scores indicating greater willingness. No dichotomization of the willingness score was performed. Ethical concerns were assessed using nine predefined items and summarized descriptively. The three willingness items were treated as a single scale, whereas items related to perceived utility and ethical concern were analyzed as stand-alone statements.

The reliability and validity of the questionnaire were assessed using Cronbach α, Guttman Lambda 6, and the Kaiser-Meyer-Olkin measure.

Statistical Analysis

Categorical variables were summarized as frequencies and percentages, and continuous variables were summarized as means (SDs). Analyses of LLM awareness were conducted among all eligible participants. Analyses of LLM use, information sources, reasons for nonuse, language used, and paid service subscription were conducted among participants who had heard of LLMs, with item-specific denominators reported.

Chi-square tests were performed to compare LLM awareness across demographic and professional characteristics; Fisher exact tests were conducted when expected cell counts were small; and Mann-Whitney *U* tests were performed to compare perceived utility scores between prior LLM users and nonusers. To control for multiple item-level comparisons, *P* values were adjusted using the Benjamini-Hochberg false discovery rate (FDR) procedure, with both domain-specific and overall adjusted *P* values reported.

Factors associated with willingness to use LLMs were examined using ordinal logistic regression with a logit link. The summed willingness score was analyzed as an ordered outcome. Crude odds ratios (cORs) were estimated using univariable models, and adjusted odds ratios (aORs) were estimated using a multivariable model including sex, age group, educational level, city level, professional title, work experience, weekly working hours, annual slide-reading workload, attitude toward LLMs replacing pathologists, research work, teaching work, and prior LLM use. All covariates were entered simultaneously. Multicollinearity was assessed using generalized variance inflation factors (GVIFs), and sensitivity analysis using a collapsed three-level willingness outcome was conducted to examine model robustness. The regression analysis included participants with complete willingness scores and covariate

XSL•FO
RenderX

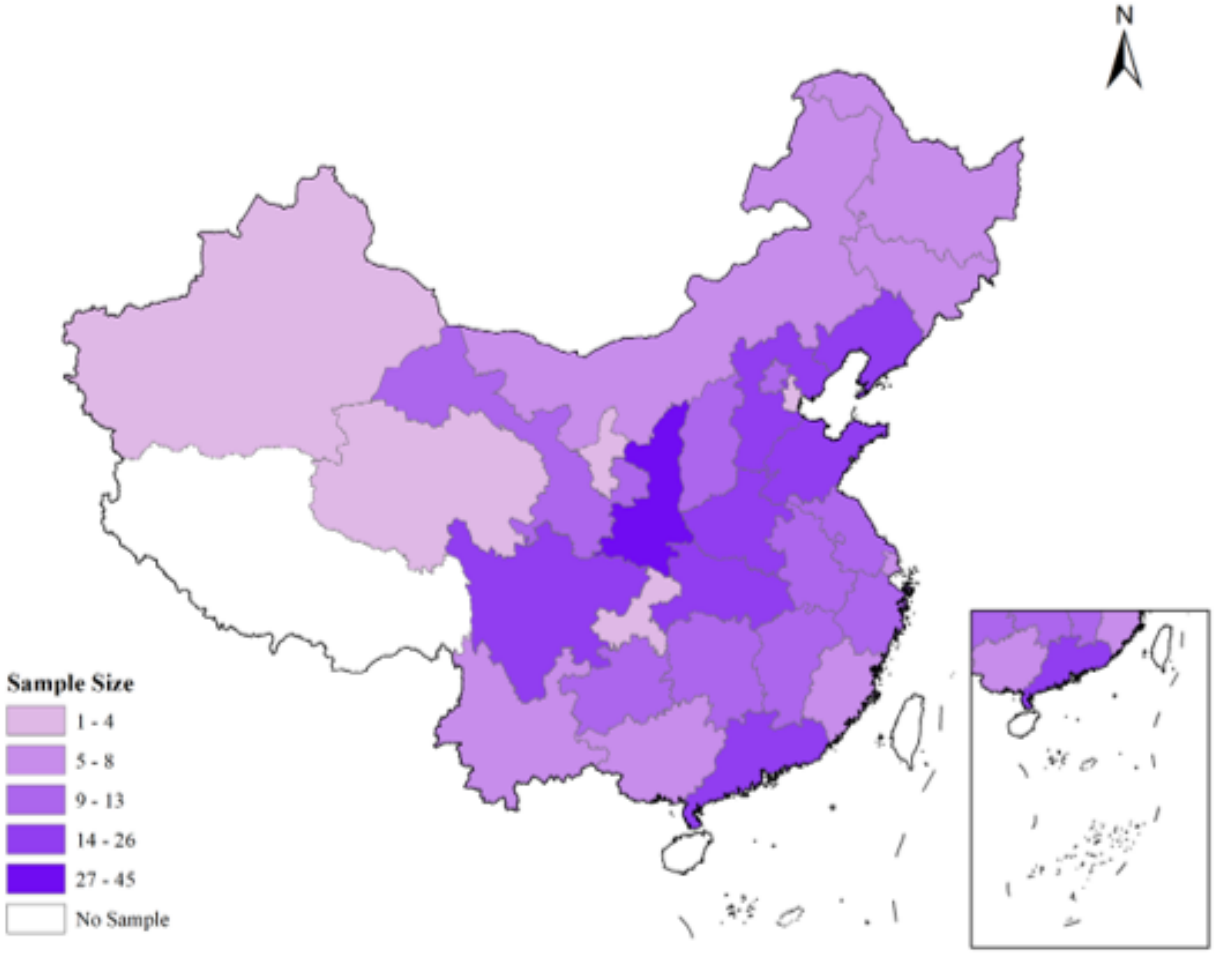
data; no imputation was performed. All analyses were conducted using R version 4.4.1 (R Foundation for Statistical Computing). Statistical tests were two-sided, and $P<.05$ was considered statistically significant.

Results

Basic Characteristics of Participants

Of the 376 individuals who responded to the invitation, 37 (9.8%; n=11, 29.7%, medical students and n=26, 70.3%, staff of pathology-related companies) were excluded from statistical analyses. The questionnaire demonstrated high internal consistency across all domains (Cronbach α=0.951-0.980, Guttman Lambda 6=0.929-0.962, and Kaiser-Meyer-Olkin measure=0.81). The 339 (90.2%) eligible participants included 236 (69.6%) females and 103 (30.4%) males, predominantly within the 30-40–year age group, representing 48.4% (n=164) of the study population. Furthermore, 63.1% (n=214) participants held a bachelor's degree, 46.6% (n=158) were attending physicians, 32.7% (n=111) had a work experience of more than 15 years, 25.7% (n=87) were involved in research work, and 34.5% (n=117) were involved in teaching work. Participants worked 43.01 hours weekly on average, with 34.2% (n=116) of them reading more than 5000 slides annually. Their geographic distribution was broad, with the survey reaching 29 provinces, municipalities, and autonomous regions in mainland China (Figure 1). In the analytic tier grouping used in this study, 122 (36.0%) participants worked in first- or new first-tier cities, 114 (33.6%) in second- or third-tier cities, and 103 (30.4%) in fourth- or lower-tier cities. The detailed mapping of unique participant work cities to analytic city tier categories is provided in Multimedia Appendix 2. The detailed demographic information of participants is shown in Table 1.

Figure 1. Geographical distribution of survey participants (N=339) across mainland China. The map illustrates the number of participants from each province, with darker shades representing higher numbers of respondents. Provinces with no survey participants are marked in white.



XSL•FO
RenderX

Table 1. Demographic and job-related characteristics of study participants (N=339).

Characteristics	Total participants, n (%)	Participants who had heard of LLMs[a] (n=248), n (%)
Gender (*P*=.02)[b]		
Male	103 (30.4)	84 (33.9)
Female	236 (69.6)	164 (66.1)
Age (years;*P*=.59)[b]		
<30	26 (7.7)	21 (8.5)
31-40	164 (48.4)	115 (46.4)
41-50	107 (31.6)	81 (32.7)
>50	42 (12.3)	31 (12.4)
City level (*P*=.01)[b]		
First or new first tier	122 (36.0)	101 (40.7)
Second or third tier	114 (33.6)	76 (30.7)
Fourth or lower tier	103 (30.4)	71 (28.6)
Education (^{P}=.59)[b]		
College degree	13 (3.8)	11 (4.4)
Bachelor's degree	214 (63.1)	157 (63.3)
Master's degree or above	112 (33.1)	80 (32.3)
Professional title (P=.02)[b]		
Medical intern	70 (20.6)	41 (16.5)
Attending doctor	158 (46.6)	123 (49.6)
Associate chief physician	79 (23.3)	59 (23.8)
Chief physician	32 (9.5)	25 (10.1)
Work experience (years; *P*=.71)[b]		
≤5	60 (17.7)	42 (16.9)
6-10	91 (26.8)	64 (25.8)
11-15	77 (22.7)	57 (23.0)
>15	111 (32.8)	85 (34.3)
Weekly working hours (mean 43.01, SD 9.66; *P*=.48)[b]		
≤43	217 (64.0)	156 (62.9)
>43	122 (36.0)	92 (37.1)
Annual slide-reading workload (*P*=.004)[b]		
≤1000	67 (19.8)	46 (18.5)
1001-3000	80 (23.6)	48 (19.4)
3001-5000	76 (22.4)	58 (23.4)
>5000	116 (34.2)	96 (38.7)
Research work (*P*=.46)[b]		
Yes	87 (25.7)	75 (30.2)
No	252 (74.3)	173 (69.8)
Teaching work (*P*=.007)[b]		
Yes	117 (34.5)	96 (38.7)
No	222 (65.5)	152 (61.3)

XSL•FO
RenderX

[a]LLM: large language model.

[b]The chi-square test was performed to analyze demographic differences among participants who had heard of LLMs (n=248, 73.2%).

LLM Awareness and Use and Perceived Barriers to LLM Adoption

Table 2 summarizes participants' awareness of LLMs, sources of information, prior use, reasons for nonuse, language used, and subscription to paid services. Of the 339 pathologists, 248 (73.2%) had heard of LLMs before the survey, whereas 91 (26.8%) had not. Among the 248 participants, social media was the most common source of information (n=176, 71.0%), followed by friends or colleagues (n=31, 12.5%), academic conferences (n=26, 10.5%), and academic papers (n=15, 6.0%).

Table 2. Participants' awareness and use of LLMs[a] and perceived barriers to LLM adoption.

Variables	Participants, n (%)
Heard of LLMs before the survey	
Yes	248/339 (73.2)
No	91/339 (26.8)
Source of information about LLMs	
Friends or colleagues	31/339 (12.5)
Social media	176/339 (71.0)
Academic paper	15/339 (6.0)
Academic conference	26/339 (10.5)
Used LLMs before the survey	
Yes	64/248 (25.8)
No	184/248 (74.2)
Reason for not using LLMs	
I have no need to use it.	55/184 (29.9)
I do not have an account that I can access.	18/184 (9.8)
I do not know where to access.	90/184 (48.9)
I am concerned that the answers provided may not be accurate.	11/184 (6.0)
I am concerned about privacy or data breaches.	7/184 (3.8)
Other.	3/184 (1.6)
Language used among LLM users	
English	8/64 (12.5)
Chinese	56/64 (87.5)
Subscribed to paid LLM services (among LLM users)	
Yes	11/64 (17.2)
No	53/64 (82.8)

[a]LLM: large language model.

Of the 248 (73.2%) participants, 64 (25.8%) had used LLMs before the survey, whereas 184 (74.2%) had not. Of the 64 LLM users, 56 (87.5%) reported using LLMs in Chinese, whereas 8 (13%) reported using them in English. In addition, 11 (17.2%) users had subscribed to paid LLM services, whereas 53 (82.8%) had not.

Among the 184 nonusers, the most frequently reported reason for not using LLMs was not knowing where to access them (n=90, 48.9%), followed by having no need to use LLMs (n=55, 29.9%), not having an accessible account (n=18, 9.8%), concerns about answer accuracy (n=11, 6.0%), concerns about privacy or data breaches (n=7, 3.8%), and other reasons (n=3, 1.6%).

Perceived Utility of LLMs

Only participants who had heard of LLMs before the survey (n=248, 73.2%) were asked to evaluate the potential utility of LLMs. Table 3 compares prior LLM users (n=64, 25.8%) with nonusers (n=184, 74.2%) within this awareness subgroup. In this analysis, nonusers referred to participants who had heard of LLMs but had not previously used them.

XSL•FO
RenderX

Table 3. Perceived utility of LLMs[a] by prior use among participants who had heard of LLMs.

Domain and items	Agreement[b], mean (SD); n/N (%)	Agreement, mean (SD); n/N (%)	*P* value[c]	FDR[d]-adjusted domain-specific/overall *P* values
	LLM users	LLM nonusers		
Clinical practice (n=248)				
Patient data analysis	3.59 (0.99); 36/64 (56.3)	3.91 (0.93); 131/184 (71.2)	.02	.02/.05
Medical-history organization	3.62 (1.05); 39/64 (60.0)	3.98 (0.93); 143/184 (77.7)	.01	.02/.05
Disease mechanism and treatment	3.64 (1.00); 40/64 (62.5)	3.96 (0.93); 137/184 (74.5)	.02	.02/.05
Scientific research (n=75)				
Experimental steps and methods	3.93 (1.18); 20/28 (71.4)	3.91 (1.02); 35/47 (74.5)	.74	.89/.92
Data visualization	3.89 (1.20); 19/28 (67.9)	3.91 (1.04); 35/47 (74.5)	.89	.89/.92
Code writing	3.86 (1.18); 19/28 (67.9)	3.81 (1.10); 33/47 (70.2)	.76	.89/.92
Medical education (n=96)				
Teaching methods	3.94 (0.90); 24/33 (72.7)	4.08 (0.79); 51/63 (81.0)	.48	.92/.92
Explanation of complex concepts	4.03 (0.88); 26/33 (78.8)	4.03 (0.80); 51/63 (81.0)	.92	.92/.92
Teaching efficiency	4.00 (0.94); 24/33 (72.7)	4.08 (0.77); 52/63 (82.5)	.79	.92/.92

[a]LLM: large language model.

[b]Agreement was defined as a rating of 4 or 5 on a 5-point Likert scale.

[c]*P* values were calculated using the Mann-Whitney *U* test.

[d]FDR: false discovery rate. Domain-specific FDR adjustment was performed within each predefined application domain, and overall FDR adjustment was performed across all nine item-level comparisons using the Benjamini-Hochberg procedure.

Clinical Practice Domain

Evaluations included all 248 (73.2%) participants who had heard of LLMs. LLM users rated the items more cautiously than nonusers. Regarding LLM utility for patient data analysis, agreement was reported by 36 (56.3%) of 64 users compared with 131 (71.2%) of 184 nonusers (mean 3.59, SD 0.99, vs mean 3.91, SD 0.93; *P*=.02). Regarding LLM utility for medical history organization, agreement was reported by 39 (60.1%) users compared with 143 (77.7%) nonusers (mean 3.62, SD 1.05, vs mean 3.98, SD 0.93; *P*=.01). Regarding LLM utility for understanding disease mechanisms and treatment, agreement was reported by 40 (62.5%) users compared with 137 (74.5%) nonusers (mean 3.64, SD 1.00, vs mean 3.96, SD 0.93; *P*=.02). These differences were significant after domain-specific FDR adjustment, with adjusted *P*=.02 for all three clinical items, but were borderline after overall FDR adjustment across all nine performance items, with adjusted *P*=.05.

Scientific Research Domain

Evaluations included 75 (30.2%) of 248 participants who had heard of LLMs, including 28 (37.3%) users and 47 (62.7%) nonusers. Ratings were similar between users and nonusers. Regarding LLM utility to organize experimental steps and methods, agreement was reported by 20 (71.4%) users and 35 (74.5%) nonusers (mean 3.93, SD 1.18, vs mean 3.91, SD 1.02; *P*=.74). Regarding LLM utility for data visualization, agreement was reported by 19 (67.9%) users and 35 (74.5%) nonusers (mean 3.89, SD 1.20, vs mean 3.91, SD 1.04; *P*=.89). Regarding LLM utility for code writing, agreement was reported by 19 (67.9%) users and 33 (70.2%) nonusers (mean 3.86, SD 1.18, vs mean 3.81, SD 1.10; *P*=.76). None of these items showed significant between-group differences after FDR adjustment.

Medical Education Domain

Evaluations included 96 (38.7%) of 248 participants who had heard of LLMs, including 33 (34.4%) users and 63 (65.6%) nonusers. Ratings were comparable between users and nonusers. Regarding LLM utility to change teaching methods, agreement was reported by 24 (72.7%) users and 51 (81.0%) nonusers (mean 3.94, SD 0.90, vs mean 4.08, SD 0.79; *P*=.48). Regarding LLM utility to explain complex concepts, agreement was reported by 26 (78.8%) users and 51 (81.0%) nonusers (mean 4.03, SD 0.88, vs mean 4.03, SD 0.80; *P*=.92). Regarding LLM utility to improve teaching efficiency, agreement was reported by 24 (72.7%) users and 52 (82.5%) nonusers (mean 4.00, SD 0.94, vs mean 4.08, SD 0.77; *P*=.79). None of these items showed significant between-group differences after FDR adjustment.

Overall, among participants who had heard of LLMs, prior LLM users expressed more cautious views than nonusers regarding clinical applications, whereas perceptions of scientific research and medical education applications were broadly similar between the two groups.

Factors Associated With Willingness to Use LLMs

Ordinal logistic regression analyses were conducted to examine factors associated with willingness to use LLMs among participants with complete willingness scores and covariate data (n=284, 83.8%; Table 4). Multicollinearity among predictors was low, with $GVIF^{(1/(2\times df))}$ ranging from 1.04 to 1.37, indicating no substantial multicollinearity. Sensitivity analysis using a

XSL•FO
RenderX

collapsed three-level willingness outcome supported the robustness of the ordinal modeling approach.

In the multivariable model, compared with participants with an associate degree or below, those with a bachelor's degree (aOR 6.10, 95% CI 1.45-25.69; *P*=.01) and those with a master's degree or above (aOR 9.69, 95% CI 2.21-42.53; *P*=.003) reported greater willingness to use LLMs. Compared with medical interns, attending physicians (aOR 2.39, 95% CI 1.09-5.22; *P*=.03) and chief physicians (aOR 4.64, 95% CI 1.57-13.69; *P*=.005) also reported greater willingness to use LLMs. Furthermore, working more than 43 hours/week was associated with greater willingness to use LLMs (aOR 2.33, 95% CI 1.40-3.89; *P*=.001). In contrast, working in fourth- or lower-tier cities was associated with less willingness to use LLMs than working in first-tier cities (aOR 0.42, 95% CI 0.22-0.79; *P*=.007). Reading 1001-3000 slides annually was also associated with less willingness to use LLMs than reading 1000 slides or fewer annually (aOR 0.43, 95% CI 0.20-0.94; *P*=.03). Sex, age group, work experience, attitude toward LLMs replacing pathologists, research work, teaching work, and prior LLM use were not significantly associated with willingness to use LLMs postadjustment.

XSL•FO
RenderX

Table 4. Factors associated with willingness to use LLMs[a] in ordinal logistic regression models.

Characteristics	Univariable ordinal logistic regression		Multivariable ordinal logistic regression	
	cOR[b] (95% CI)	*P* value	aOR[c] (95% CI)	*P* value
Gender				
Male	Reference	—[d]	Reference	—
Female	0.70 (0.44-1.12)	.14	0.65 (0.39-1.09)	.10
Age (years)				
<30	Reference	—	Reference	—
30-40	1.71 (0.76-3.82)	.19	2.22 (0.75-6.57)	.15
41-50	1.66 (0.72-3.80)	.23	1.95 (0.55-6.85)	.30
≥51	1.23 (0.47-3.17)	.68	1.51 (0.36-6.36)	.58
Education				
Associate degree and below	Reference	—	Reference	—
Bachelor’s degree	7.37 (2.01-27.01)	.003	6.10 (1.45-25.69)	.01
Master’s degree and above	11.00 (2.93-41.30)	<.001	9.69 (2.21-42.53)	.003
City level				
First or new first tier	Reference	—	Reference	—
Second or third tier	0.78 (0.46-1.33)	.37	0.69 (0.38-1.23)	.21
Fourth or lower tier	0.36 (0.21-0.64)	<.001	0.42 (0.22-0.79)	.007
Professional title				
Medical intern	Reference	—	Reference	—
Attending physician	1.51 (0.76-2.97)	.24	2.39 (1.09-5.22)	.03
Associate chief physician	1.86 (0.88-3.93)	.11	2.40 (0.98-5.86)	.06
Chief physician	3.20 (1.27-8.03)	.01	4.64 (1.57-13.69)	.005
Work experience (years)				
≤5	Reference	—	Reference	—
6-10	0.98 (0.47-2.03)	.95	0.39 (0.14-1.06)	.06
11-15	1.01 (0.48-2.14)	.98	0.40 (0.14-1.16)	.09
>15	1.02 (0.51-2.05)	.95	0.49 (0.15-1.58)	.23
Weekly working hours				
≤43	Reference	—	Reference	—
>43	2.42 (1.51-3.89)	<.001	2.33 (1.40-3.89)	.001
Annual slide-reading workload				
≤1000	Reference	—	Reference	—
1001-3000	0.52 (0.25-1.06)	.07	0.43 (0.20-0.94)	.03
3001-5000	1.17 (0.59-2.29)	.66	0.85 0.40-1.82)	.68
>5000	1.03 (0.55-1.92)	.93	0.68 (0.34-1.37)	.28
Attitude toward LLMs replacing pathologists				
No	Reference	—	Reference	—
Neutral	1.06 (0.63-1.78)	.83	1.23 (0.71-2.11)	.46
Yes	1.32 (0.74-2.35)	.35	1.32 (0.70-2.49)	.39
Research work				
No	Reference	—	Reference	—

XSL•FO
RenderX

Characteristics	Univariable ordinal logistic regression		Multivariable ordinal logistic regression	
	cOR[b] (95% CI)	*P* value	aOR[c] (95% CI)	*P* value
Yes	1.45 (0.89-2.38)	.14	1.03 (0.56-1.88)	.93
Teaching work				
No	Reference	—	Reference	—
Yes	1.00 (0.64-1.58)	.98	0.79 (0.46-1.35)	.39
Prior LLM use				
No	Reference	—	Reference	—
Yes	1.00 (0.61-1.66)	.99	0.82 (0.47-1.41)	.46

[a]LLM: large language model. Willingness to use LLMs was calculated as the summed score of three 5-point Likert-scale items covering clinical practice, scientific research, and medical education, with scores ranging from 3 to 15. Higher scores indicated greater willingness.

[b]cOR: crude odds ratio

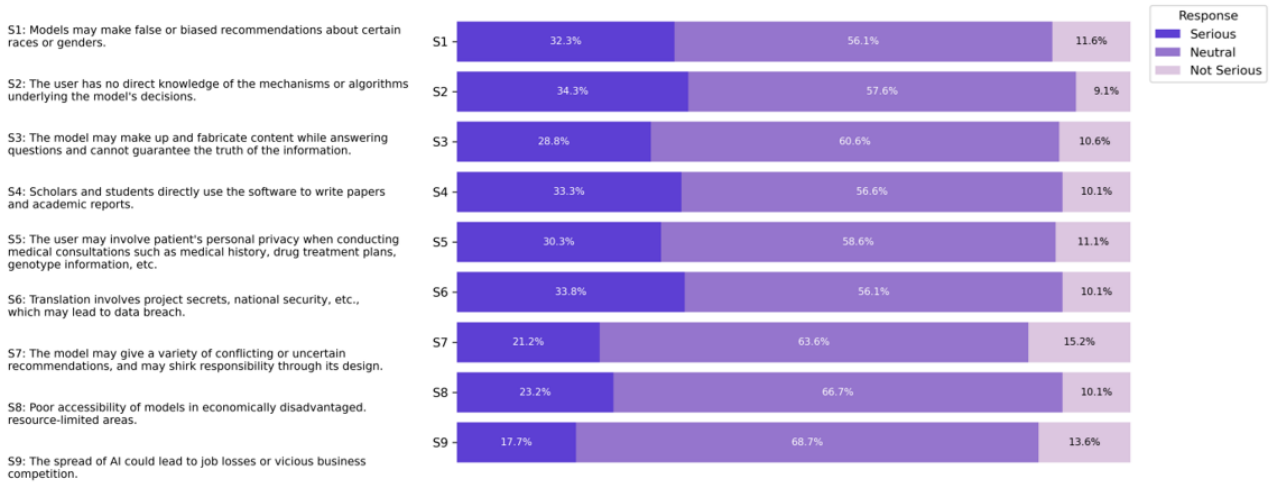
[c]aOR: adjusted odds ratio

[d]Not applicable.

Ethical Concerns and Future Improvement Priorities

Of the 248 (73.2%) participants who had heard of LLMs, 198 (79.8%) acknowledged potential ethical issues related to LLM use in pathology. As shown in Figure 2, among these 198 participants, the concerns most frequently rated as serious were lack of transparency in model mechanisms or algorithms (S2; n=68, 34.3%), potential data breaches involving project secrets or national security information during translation or processing of sensitive content (S6; n=67, 33.8%), and the risk that scholars or students may directly use software to write papers or academic reports (S4; n=66, 33.3%). Other concerns rated as serious included false or biased recommendations about certain races or genders (S1; n=64, 32.3%), involvement of patient privacy during medical consultations (S5; n=60, 30.3%), fabrication of content or inability to guarantee truthfulness of information (S3; n=57, 28.8%), poor accessibility of models in economically disadvantaged or resource-limited areas (S8; n=46, 23.2%), conflicting or uncertain recommendations and unclear responsibility attribution (S7; n=42, 21.2%), and potential job losses or unfair business competition caused by the spread of AI (S9; n=35, 17.7%).

Figure 2. Percentage distribution of assessment for the severity of ethical considerations. Stacked bars show the percentage of participants who rated each ethical concern as "Serious," "Neutral," or "Not serious." The coded items were as follows: S1, LLMs may make false or biased recommendations about certain races or genders; S2, users may have no direct knowledge of the mechanisms or algorithms underlying LLM decisions; S3, LLMs may fabricate content when answering questions and may not guarantee the truthfulness of information; S4, scholars and students may directly use the software to write papers and academic reports; S5, LLM use during medical consultations may involve patient privacy, such as medical history, treatment plans, and genotype information; S6, translation or processing of sensitive content may involve project secrets or national security information and may lead to data breaches; S7, LLMs may provide conflicting or uncertain recommendations and may obscure responsibility through their design; S8, poor LLM accessibility in economically disadvantaged or resource-limited areas may exacerbate inequities; and S9, the spread of AI may lead to job losses or unfair business competition. AI: artificial intelligence; LLM: large language model.

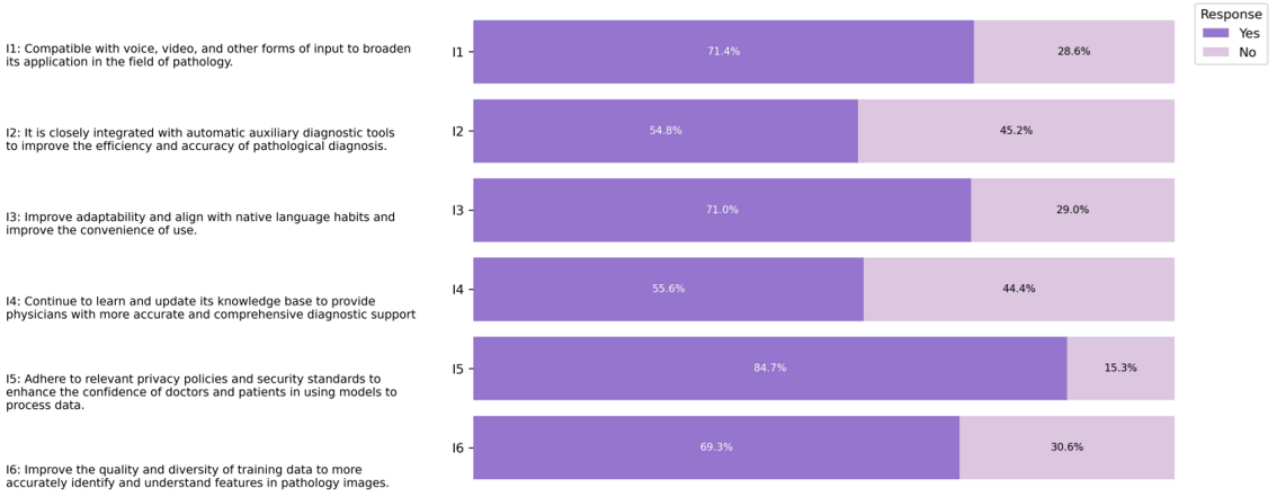


Regarding future improvement priorities, among the 248 participants, the most frequently supported direction was adherence to relevant privacy policies and security standards to enhance doctors' and patients' confidence in using models to process data (I5; n=210, 84.7%), as shown in Figure 3. Other highly supported directions included compatibility with voice, video, and other input formats to broaden LLM applications in pathology (I1; n=177, 71.4%), improved adaptability to native

XSL•FO
RenderX

language habits and greater convenience of use (I3; n=176, 71.0%), and improved quality and diversity of training data to more accurately identify and understand features in pathology images (I6; n=172, 69.3%). In addition, 138 (55.6%) participants supported continuous learning and knowledge base updates to provide more accurate and comprehensive diagnostic support (I4), and 136 (54.8%) participants supported closer integration with automated auxiliary diagnostic tools to improve the efficiency and accuracy of pathological diagnosis (I2).

Figure 3. Percentage distribution of future improvement direction of LLMs. Stacked bars show the percentage of participants who supported or did not support each proposed direction for future improvement. The coded items were as follows: I1, compatibility with voice, video, and other input formats to broaden LLM applications in pathology; I2, integration with automated auxiliary diagnostic tools to improve the efficiency and accuracy of pathological diagnosis; I3, improved adaptability to native language habits and greater convenience of use; I4, continuous learning and updating of the knowledge base to provide physicians with more accurate and comprehensive diagnostic support; I5, adherence to relevant privacy policies and security standards to increase doctors' and patients' confidence in using LLMs to process data; and I6, improvement in the quality and diversity of training data to more accurately identify and understand features in pathology images. LLM: large language model.



Discussion

Principal Findings

This study is the first systematic survey conducted on the perception of and willingness to use LLMs among Chinese pathologists. Although previous studies have explored similar issues, they did not include Chinese pathologists, had small sample sizes, and did not investigate factors influencing the adoption of LLMs [11]. This study filled this gap and provided more broadly applicable and reliable evidence.

The study covered 29 provinces, municipalities, and autonomous regions across China, ensuring geographically diversity. By classifying regional economic levels based on city tiers rather than traditional geographical divisions, the findings provide a clearer understanding of regional disparities. This approach considers political status, economic strength, development level, and urban influence, providing a scientific basis for more effective regional development strategies and medical resource allocation.

In this study, most Chinese pathologists (73.2%) had heard of LLMs and generally held positive views toward their potential use in clinical practice, scientific research, and medical education. However, actual use remained limited, with only 25.8% of pathologists who had heard of LLMs reporting prior experience This is consistent with the findings of Thiyaphat et al [11]. This gap between awareness and use should be interpreted in relation to the LLM landscape at the time of the survey. In early 2024, GPT-based systems, particularly ChatGPT, were among the most visible and widely discussed LLM applications in medicine and pathology, and this trend was reflected in our data: ChatGPT was the most recognized and most frequently used model among respondents, followed by ERNIE Bot and mirror websites [7,9,12]. Therefore, the findings reflect Chinese pathologists' perceptions of mainstream early LLM applications, with GPT-based tools serving as the dominant exemplar. At the same time, LLMs are heterogeneous systems, and differences in model architecture, training data, access route, and local clinical integration may shape users' perceptions of reliability, privacy, and practical usefulness.

Accessibility may partly explain why high awareness did not translate into widespread use. Nearly half of the nonusers reported that they did not know where to access LLMs, and a smaller proportion reported lacking an accessible account. For Chinese pathologists, access to different LLMs may depend on whether the models are officially available, whether domestic alternatives are used, and whether third-party or mirror access routes are involved [11,13]. Such access pathways can influence not only whether pathologists use LLMs but also how they evaluate model stability, privacy protection, and trustworthiness. Thus, limited adoption should not be interpreted solely as lack of interest; it may also reflect the practical accessibility constraints of the Chinese LLM ecosystem.

In contrast to our initial expectation, pathologists working in fourth- or lower-tier cities exhibited less willingness to use LLMs than those working in first- or new first-tier cities after adjustment. This finding suggests that although lower-tier cities may have greater unmet needs for pathology support, practical barriers may constrain willingness to adopt LLMs. These

XSL•FO
RenderX

barriers may include less exposure to AI tools, fewer opportunities for AI-related training, limited institutional support, and uncertainty about data security or workflow integration. Therefore, promoting LLM adoption in lower-tier cities should not rely only on emphasizing potential benefits. Implementation strategies should also improve accessibility, provide targeted training, support Chinese-language and pathology-specific tools, and establish clear governance mechanisms for safe use.

Language is another important dimension of local usability. In our survey, most participants who had used LLMs (n=56, 87.5%) reported Chinese-language use, indicating that adoption barriers were not simply a matter of English proficiency. However, Chinese-language interaction does not eliminate language-related limitations. Many medical LLMs and benchmark datasets have historically emphasized English-language corpora, and model performance may vary when handling Chinese medical terminology, local clinical guidelines, pathology-specific diagnostic nomenclature, and institution-specific reporting conventions. Future implementation should evaluate locally accessible models using Chinese prompts, Chinese pathology terminology, and real-world reporting workflows [14].

Interestingly, despite the widespread positive feedback for LLMs in the field, pathologists who had used these models (group A) rated them lower in clinical and scientific research than those who had only heard of these models (group B). This discrepancy may be due to several factors. For group A, direct experience may have exposed the practical limitations of currently accessible LLMs, particularly when general-purpose conversational models are applied to complex clinical or pathology-specific tasks. For example, although LLMs can summarize large volumes of text and support information retrieval, clinical application in pathology requires accurate interpretation of nuanced diagnostic language, patient history, histomorphological findings, ancillary tests, and local reporting standards. These limitations may lead to a more critical evaluation by group A. In addition, their high expectations prior to using the models may not have been met in practice, leading to more cautious assessments. In contrast, group B, having never used these models, may have been influenced more by external information and theoretical expectations, resulting in a more favorable view of the models. This experience-based caution suggests that adoption should be understood as a sociotechnical process in which perceived usefulness must be weighed against reliability, accountability, privacy protection, and compatibility with routine pathology workflows [15]. A similar pattern was observed in scientific research. Although group A rated the models lower than group B in tasks such as data visualization and code writing, these differences were not statistically significant. This suggests that the application of LLMs in research is still in its early stages, with limited practical impact. The differences in ratings highlight the gap between theoretical expectations and actual use, indicating that the potential of LLMs in complex research tasks has yet to be fully realized [16].

Further analysis showed that pathologists handling more than 5000 slides annually exhibited greater acceptance of LLMs, likely recognizing that LLMs can alleviate workload. LLMs designed specifically for pathology can perform preliminary automated analyses of slide images, identifying common pathological features, thus allowing pathologists to focus on complex cases requiring expert judgment. In addition, LLMs can streamline the process of issuing pathology reports, which are crucial for clinical decision-making. By extracting and structuring information from unformatted text in electronic health records (EHRs), LLMs can assist in generating comprehensive reports that include diagnostic insights and molecular markers.

Truhn et al [9] demonstrated that LLMs, such as GPT-4, can accurately parse unstructured pathology reports and convert them into structured data, integrating medical reasoning to enhance clinical interpretation. This capability not only reduces the writing burden on pathologists but also supports the formulation of optimal patient care strategies.

Some participants engaged in research or teaching also demonstrated heightened interest in LLMs, recognizing their utility in these domains. In research, LLMs excel at automating tasks such as classifying and analyzing extensive textual data, thereby accelerating the academic process. They facilitate rapid compilation of research reports by extracting and summarizing information from diverse sources [17]. In education, LLMs offer customized resources that cater to the complex medical curriculum, enhancing students' learning efficiency through vivid teaching methods and focused content [18]. They also serve as interactive study companions, capable of creating mock exams and providing immediate feedback, enriching the learning experience [19,20]. For pathologists, particularly those with limited access to quality training materials, LLMs are a vital source for medical knowledge acquisition. By applying these models, pathologists can access a wider array of information and reduce training expenses, ultimately improving their professional competencies and the level of pathological diagnosis in underdeveloped regions.

Despite the tremendous potential of LLMs, participants expressed concerns regarding ethical issues, particularly the exacerbation of health inequalities, which refers to the higher systemic health risks faced by certain populations compared to others due to social factors, such as poverty or race. These risk disparities are traced back to Social Determinants of Health (SDoH) that influence health outcomes throughout an individual's life [21]. SDoH are pivotal for patient prognosis, enabling pathologists to tailor diagnostic and treatment plans by understanding underlying health factors and connecting patients to necessary social services. However, capturing SDoH in EHRs remains a challenge due to incomplete structured data [22]. Promisingly, research at Harvard Medical School has demonstrated that fine-tuned models, such as ChatGPT, can effectively identify SDoH indicators in EHRs, outperforming traditional *International Classification of Diseases* (ICD) coding [21]. This capability suggests that with proper training, LLMs could play a crucial role in addressing health inequities. However, it must be noted that the high maintenance costs of these models pose a threat to their widespread application, particularly in low- and middle-income countries [23]. Balancing the widespread adoption of such technology with considerations

XSL•FO
RenderX

of fairness and accessibility remains a key challenge. In addition to equity, our finding that one-third of respondents rated data leakage as a serious concern underscores the need for privacy-preserving deployment, secure data governance, and clear institutional accountability when clinical or report data are processed by LLMs. Because inaccurate or biased outputs, automation bias, and cybersecurity threats have been identified as important risks for health-related LLMs, pathologists should retain final responsibility and use model outputs as assistive information rather than definitive diagnostic conclusions. Balancing the widespread adoption of such technology with considerations of fairness, privacy, and accessibility remains a key challenge.

Regarding future improvements, participants placed the strongest emphasis on privacy and security governance. A total of 210 (84.7%) of 248 participants supported adherence to relevant privacy policies and security standards to enhance doctors' and patients' confidence in using LLMs to process data. This finding is consistent with participants' ethical concerns about data breaches and underscores that responsible implementation requires secure data governance, institutional accountability, and clear rules for handling clinical and report data. Participants also supported improving multimodal input compatibility, adapting models to native language habits, and improving the quality and diversity of training data, suggesting that future pathology-oriented LLMs should be accessible, locally usable, privacy preserving, and supported by pathology-specific datasets. Similar integrations have made progress in the field of dentistry, which could aid in the identification of heteromorphic cells or tissues for image-based diagnoses [24]. Enhancing the quality and diversity of training data to improve the models' ability to recognize pathological features was a priority for 71.4% of participants. Studies, such as that by Haemmerli et al [25] on diagnosing central nervous system tumors, have shown that although LLMs can propose treatment plans, they sometimes struggle with precise tumor classifications. This suggests the need for more multicenter datasets to be trained or validated to improve the generalization ability of models or even to train more models specific to the diagnosis of pathology, such as PathChat [26,27]. In addition, improving the identification of pathological features and the development of multicenter datasets are key challenges that need to be addressed in future LLM development (71%). Addressing the models' limitations in accessing the latest research and data is crucial to avoid providing outdated information [28].

Limitations

This study has several limitations. First, it was conducted in April 2024, prior to the release and widespread use of DeepSeek—a domestically developed LLM that has since gained significant attention in China. As such, our findings provide a valuable baseline of Chinese pathologists' perceptions before the emergence of prominent national LLMs. Although subsequent developments like DeepSeek may shape future acceptance patterns, the key factors assessed (e.g., perceived usefulness, trust, ease of integration, and data security) are fundamental and likely to remain influential across evolving platforms. Although the questionnaire recorded which models participants recognized or used, performance ratings were not stratified by specific model, access route, or interaction language. Future research should investigate how localized models and improved technical performance impact clinicians' attitudes and adoption behaviors over time. Second, participants were primarily recruited from a national academic conference, potentially skewing the sample toward more academically active or technologically engaged pathologists. The colleague referral component may also have amplified network homophily. This may underrepresent those with limited access to or experience with digital tools. Therefore, the proportions of awareness, use, and willingness observed in this study should not be interpreted as population-level estimates for all Chinese pathologists. To address this, future studies should broaden participant diversity, especially by including pathologists from underresourced regions or with minimal digital training, whose perspectives are essential for informing inclusive and effective LLM deployment strategies.

Conclusion

The study revealed a predominantly open and positive stance toward LLMs among Chinese pathologists, with minimal concern about these technologies' potential to replace their professional expertise. Although current use remains limited, pathologists' attitudes suggest cautious readiness to explore LLMs as assistive tools. Given the evolving LLM ecosystem in China, future implementation should prioritize models that are accessible, compatible with the Chinese language, clinically validated, privacy preserving, and well integrated into routine pathology workflows. A key takeaway is the importance of fostering the responsible adoption of LLMs to help support workload management and reduce regional disparities. These efforts have critical implications for advancing the quality and equity of medical services across regions.

Acknowledgments

We sincerely thank all the pathologists who participated in this survey. Their valuable time and insights have been crucial to the formation of this study.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Funding

This work was supported by CAMS Innovation Fund for Medical Sciences (CIFMS; project #2026-I2M-SXZX10-04), the Gates Foundation (project #INV-031449), and the Chongqing Tencent Sustainable Development Foundation through the "AI-Empowered

XSL•FO
RenderX

Demonstration Project for Cervical and Breast Cancer Prevention and Control in Women" (project #SD20260711154745), with strong support from the Tencent Inclusive Health Lab.

Authors' Contributions

Study concept and design: YW, VYC, YY, PX, and YQ
Questionnaire collection: YW, VYC, and YY
Data analysis and manuscript drafting: YW and VYC
Pathological consultation: QP and XZ
Manuscript revision: YW, VYC, QP, XZ, YY, PX, and YQ
Guarantor: YQ

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire on adoption of LLMs in pathology. LLM: large language model.
[DOCX File , 23 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Classification of Chinese cities and example cities.
[DOCX File , 13 KB-Multimedia Appendix 2]

References

1. Han B, Zheng R, Zeng H, Wang S, Sun K, Chen R, et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent. Mar 2024;4(1):47-53. [FREE Full text] [doi: 10.1016/j.jncc.2024.01.006] [Medline: 39036382]
2. Burns KEA, Duffett M, Kho ME, Meade MO, Adhikari NKJ, Sinuff T, et al. ACCADEMY Group. A guide for the design and conduct of self-administered surveys of clinicians. CMAJ. Jul 29, 2008;179(3):245-252. [FREE Full text] [doi: 10.1503/cmaj.080372] [Medline: 18663204]
3. National Health Commission of the People's Republic of China. China Health Statistical Yearbook 2022. Beijing, China. Peking Union Medical College Press; Sep 2022.
4. Chinese Society of Pathology. [Investigation and consideration on the status of pathology departments in 3 831 hospitals of 31 provinces, municipalities and autonomous regions]. Zhonghua Bing Li Xue Za Zhi. Dec 08, 2020;49(12):1217-1220. [FREE Full text] [doi: 10.3760/cma.j.cn112151-20200904-00695] [Medline: 33287503]
5. Maung RT. What is the best indicator to determine anatomic pathology workload? Canadian experience. Am J Clin Pathol. Jan 2005;123(1):45-55. [FREE Full text] [doi: 10.1309/23nygnb2hfnnw4v8] [Medline: 15762279]
6. Usubütün A, Uner S, Harorlu F, Ozer E, Tuzlali S, Ruacan A, et al. [Pathology laboratories staff workload evaluation in Turkey: a survey study]. Turk Patoloji Derg. May 2011;27(2):98-105. [FREE Full text] [doi: 10.5146/tjpath.2011.01056] [Medline: 21630194]
7. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. Aug 2023;29(8):1930-1940. [FREE Full text] [doi: 10.1038/s41591-023-02448-8] [Medline: 37460753]
8. Tayebi Arasteh S, Han T, Lotfinia M, Kuhl C, Kather JN, Truhn D, et al. Large language models streamline automated machine learning for clinical studies. Nat Commun. Mar 21, 2024;15(1):1603. [FREE Full text] [doi: 10.1038/s41467-024-45879-8] [Medline: 38383555]
9. Truhn D, Loeffler CM, Müller-Franzes G, Nebelung S, Hewitt KJ, Brandner S, et al. Extracting structured information from unstructured histopathology reports using generative pre-trained transformer 4 (GPT-4). J Pathol. Mar 14, 2024;262(3):310-319. [FREE Full text] [doi: 10.1002/path.6232] [Medline: 38098169]
10. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, et al. A multimodal generative AI copilot for human pathology. Nature. Oct 12, 2024;634(8033):466-473. [FREE Full text] [doi: 10.1038/s41586-024-07618-3] [Medline: 38866050]
11. Laohawetwanit T, Pinto DG, Bychkov A. A survey analysis of the adoption of large language models among pathologists. Am J Clin Pathol. Jan 28, 2025;163(1):52-59. [FREE Full text] [doi: 10.1093/ajcp/aqae093] [Medline: 39076014]
12. Omar M, Ullanat V, Loda M, Marchionni L, Umeton R. ChatGPT for digital pathology research. Lancet Digit Health. Aug 2024;6(8):e595-e600. [FREE Full text] [doi: 10.1016/s2589-7500(24)00114-6]
13. Lin A, Zeng M, Gan W, Jiang A, Liu Y, Qi C, et al. Medicine digital transformation: evidence from Chinese physicians on generative artificial intelligence implementation and challenges. J Transl Med. Mar 09, 2026;24(1):509. [FREE Full text] [doi: 10.1186/s12967-026-07912-w] [Medline: 41796330]

XSL•FO
RenderX

14. Qiu P, Wu C, Zhang X, Lin W, Wang H, Zhang Y, et al. Towards building multilingual language model for medicine. Nat Commun. Sep 27, 2024;15(1):8384. [FREE Full text] [doi: 10.1038/s41467-024-52417-z] [Medline: 39333468]
15. Salwei ME, Carayon P. A sociotechnical systems framework for the application of artificial intelligence in health care delivery. J Cogn Eng Decis Mak. Dec 11, 2022;16(4):194-206. [FREE Full text] [doi: 10.1177/15553434221097357] [Medline: 36704421]
16. Haimes J, Smith R, Zhang Y. Evaluating the performance gap between theoretical expectations and practical use of large language models. J Artif Intell Res. 2024;32(4):123-135.
17. Dergaa I, Chamari K, Zmijewski P, Ben Saad H. From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. Biol Sport. Apr 2023;40(2):615-622. [FREE Full text] [doi: 10.5114/biolsport.2023.125623] [Medline: 37077800]
18. Holderried F, Stegemann-Philipps C, Herrmann-Werner A, Festl-Wietek T, Holderried M, Eickhoff C, et al. A language model-powered simulated patient with automated feedback for history taking: prospective study. JMIR Med Educ. Aug 16, 2024;10:e59213. [FREE Full text] [doi: 10.2196/59213] [Medline: 39150749]
19. Frommeyer TC, Fursmidt RM, Gilbert MM, Bett ES. The desire of medical students to integrate artificial intelligence into medical education: an opinion article. Front Digit Health. May 13, 2022;4:831123. [FREE Full text] [doi: 10.3389/fdgth.2022.831123] [Medline: 35633734]
20. Fang Q, Reynaldi R, Araminta AS, Kamal I, Saini P, Afshari FS, et al. Artificial intelligence (AI)-driven dental education: exploring the role of chatbots in a clinical learning environment. J Prosthet Dent. Oct 2025;134(4):1296-1303. [FREE Full text] [doi: 10.1016/j.prosdent.2024.03.038] [Medline: 38644064]
21. Guevara M, Chen S, Thomas S, Chaunzwa TL, Franco I, Kann BH, et al. Large language models to identify social determinants of health in electronic health records. NPJ Digit Med. Jan 11, 2024;7(1):6. [FREE Full text] [doi: 10.1038/s41746-023-00970-0] [Medline: 38200151]
22. Truong HP, Luke AA, Hammond G, Wadhera RK, Reidhead M, Maddox KEJ. Utilization of social determinants of health ICD-10 Z-codes among hospitalized patients in the United States, 2016-2017. Med Care. 2020;58(12):1037-1043. [FREE Full text] [doi: 10.1097/mlr.0000000000001418]
23. Denecke K, May R, LLMHealthGroup, Rivera Romero O. Potential of large language models in health care: Delphi Study. J Med Internet Res. May 13, 2024;26:e52399. [FREE Full text] [doi: 10.2196/52399] [Medline: 38739445]
24. Huang H, Zheng O, Wang D, Yin J, Wang Z, Ding S, et al. ChatGPT for shaping the future of dentistry: the potential of multi-modal large language model. Int J Oral Sci. Jul 28, 2023;15(1):29. [FREE Full text] [doi: 10.1038/s41368-023-00239-y] [Medline: 37507396]
25. Haemmerli J, Sveikata L, Nouri A, May A, Egervari K, Freyschlag C, et al. ChatGPT in glioma adjuvant therapy decision making: ready to assume the role of a doctor in the tumour board? BMJ Health Care Inform. Jun 30, 2023;30(1):e100775. [FREE Full text] [doi: 10.1136/bmjhci-2023-100775] [Medline: 37399360]
26. The Lancet Digital Health. Pathology in the era of generative AI. Lancet Digit Health. Aug 2024;6(8):e536. [FREE Full text] [doi: 10.1016/s2589-7500(24)00157-2]
27. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. JAMA. Sep 05, 2023;330(9):866-869. [FREE Full text] [doi: 10.1001/jama.2023.14217] [Medline: 37548965]
28. Kwong JCC, Wang SCY, Nickel GC, Cacciamani GE, Kvedar JC. The long but necessary road to responsible use of large language models in healthcare research. NPJ Digit Med. Jul 04, 2024;7(1):177. [FREE Full text] [doi: 10.1038/s41746-024-01180-y] [Medline: 38965411]

Abbreviations

AI: artificial intelligence
aOR: adjusted odds ratio
cOR: crude odds ratio
EHR: electronic health record
FDR: false discovery rate
GVIF: generalized variance inflation factor
LLM: large language model
SDoH: Social Determinants of Health

XSL•FO
RenderX

Edited by A Benis; submitted 22.Nov.2025; peer-reviewed by C Xu, H Deng; comments to author 31.Mar.2026; revised version received 23.May.2026; accepted 18.Jun.2026; published 20.Nov.2026

Please cite as:
Wang Y, Cui VY, Pan Q, Zhang X, Yang Y, Xue P, Qiao Y
Chinese Pathologists' Adoption of Large Language Models: Cross-Sectional Survey on Influencing Factors
JMIR Med Inform 2026;14:e88275
URL: https://medinform.jmir.org/2026/1/e88275
doi: 10.2196/88275
PMID:

©Yuting Wang, Victor Yu Cui, Qinjing Pan, Xun Zhang, Yao Yang, Peng Xue, Youlin Qiao. Originally published in JMIR Medical Informatics (https://medinform.jmir.org), 20.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on https://medinform.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX