Digital information for patient education, volume II

Edited by
Xiaofei Zhang, Feng Guo and Paul H. Lee

Published in
Frontiers in Public Health

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Digital information for patient education, volume II

Topic editors

Xiaofei Zhang — Nankai University, China

Feng Guo — Tianjin University, China

Paul H. Lee — University of Southampton, United Kingdom

Table of
contents

Check for updates

Editorial: Digital information for patient education, volume II

Xiaofei Zhang[1]*, Feng Guo[2]* and Paul Lee[3]

[1]School of Management, Harbin Institute of Technology, Harbin, China, [2]College of Management and Economics, Tianjin University, Tianjin, China, [3]Southampton Clinical Trials Unit, University of Southampton, Southampton, United Kingdom

Editorial on the Research Topic
Digital information for patient education, volume II

1 Introduction

Recent advances in digital technologies—such as online platforms, big data, live streaming, and artificial intelligence (AI)—are transforming how health and medical information is generated, shared, and managed, giving rise to new models of medical information management (1–3). Through digital health information systems, patients can access medical knowledge remotely, which has been shown to improve their wellbeing (4, 5). Within this broader shift, online health communities (OHCs) have become an increasingly important channel for patient care and education (6, 7). By offering fee-based online consultations, OHCs help patients overcome time and location constraints and obtain professional medical advice without the need for in-person visits (8, 9). Recent studies further show that the design of asynchronous interactions—such as communication depth, information intensity, and relationship duration—plays a key role in shaping patient satisfaction on these platforms, providing useful guidance for improving digital health information delivery (10, 11).

Nevertheless, existing research points to a persistent gap between the availability of digital health information and its effectiveness in supporting patient education (12). For example, studies of mobile health monitoring services show that affective factors—such as emotional attachment and affective attitudes—play a central role in sustained use, with user satisfaction and confirmation perceptions jointly shaping long-term engagement (13). Despite these insights, there remains limited guidance on how digital health information can be systematically designed and delivered to improve patient understanding and learning outcomes. This challenge is compounded by evidence that digital health innovations may unintentionally reinforce existing inequalities. High-quality online medical resources tend to be concentrated in economically developed regions, which can widen gaps in digital health literacy among vulnerable populations (14). Because effective health management requires patients to fully comprehend medical information, the ability to understand online health content is critical to the broader digital transformation of healthcare services (15). Yet much medical knowledge—particularly technical terminology—remains difficult for patients to interpret. Digital technologies offer promising ways to address this problem by enabling richer information types, formats, and presentation modes. These tools can enhance patients'

understanding of medical knowledge and support more informed health-related decision making. One illustrative example is the emergence of online medical teams, which rely on multi-physician collaboration to deliver more comprehensive and authoritative advice (16). Recent evidence shows that such team-based digital models significantly increase patients' adoption of medical recommendations (17), highlighting a concrete pathway through which digital innovation can improve information comprehension and downstream behavioral outcomes.

This Research Topic aims to enhance patients' understanding of online health information by examining how digital health content is organized and presented, while highlighting the distinctive roles that mobile communication, big data, and AI play in contemporary healthcare. In this editorial, we outline the thematic focus of the Research Topic and highlight the key insights offered by the papers included in this Research Topic.

2 Papers in this research topic

This Research Topic brings together 13 articles that have successfully completed the standard peer-review process of Frontiers in Public Health. Although the studies differ in research focus, methodological approaches, and theoretical perspectives, they are united by a shared emphasis on digital information for patient education and address one or more of the predefined research themes. To provide a structured overview, we synthesize the contributions and organize them into six thematic areas: (1) medical and data-enabled technologies in information management for patient education; (2) the role of mobile health information and communication technologies (ICTs) in information sharing; (3) the provision of health information based on big data; (4) applications of AI in patient education; (5) the effectiveness of digital technologies in enhancing patient education outcomes; and (6) emerging digital innovations aimed at improving patients' understanding of health information.

Two studies address the theme of med-tech and data-enabled technologies in information management for patient education, examining how data-driven tools can enhance the organization and dissemination of health information. Huang, Guo et al. draw on motivation theory to examine how physicians' professional motivation shapes online knowledge sharing for patient education, with particular attention to the moderating roles of online experience and offline expertise. Using panel data from 11,839 physicians on a leading Chinese online health platform, the authors find that professional motivation positively influences both the quantity and quality of online knowledge sharing. Moreover, online experience strengthens this positive relationship, whereas offline expertise weakens the effect on sharing quantity. This study enriches the literature on motivation theory and online knowledge sharing and offers practical implications for physicians and platform managers. Liu et al. adopt a dual-factor model to investigate how discharged patients' individual characteristics affect their online health information-seeking behaviors, while considering the moderating role of living with children. Based on survey data from 292 discharged patients and analyzed using structural equation modeling, the results show that perceived stress

encourages information-seeking behavior, whereas resistance to change and learned helplessness inhibit it. In addition, cohabitation with children moderates the negative effect of resistance to change. This study deepens our understanding of discharged patients' information-seeking behaviors and provides actionable insights for online health information providers.

Two articles focus on the role of mobile health ICTs in supporting information sharing for patient education. Using repeated cross-sectional surveys, Sakamoto et al. examine changes in parents' knowledge and concerns about pediatric fever between 2017 and 2024, as well as the association with the smartphone application "Oshiete! Doctor." Their findings show that while parents' knowledge about fever improved substantially over time, anxiety related to potential brain damage and seizures increased. Moreover, the dissemination of the app alone was insufficient to counter broader societal anxiety. This study highlights the complex and evolving impact of mobile health applications on parental health literacy. Ma et al. investigate the dynamic evolution of users' continuance intention toward mHealth services by integrating the expectation–confirmation model with a latent growth modeling approach. Drawing on three waves of longitudinal survey data from 236 respondents, they find that continuance intention declines over time. Expectations accelerate this decline, whereas confirmation mitigates it, with the effect of expectation being more pronounced among users without chronic diseases. By extending the expectation–confirmation framework into a dynamic context, this study advances understanding of long-term mHealth service usage.

Two studies address the provision of health information driven by big data. Deng et al. conduct a scientometric and bibliometric analysis of medication literacy research published between 2003 and 2024, drawing on 1,968 articles from the Web of Science Core Collection and using CiteSpace and VOSviewer. Their big data analysis identifies three major research streams—medication literacy and adherence, the development of assessment tools, and psychosocial factors—and highlights key contributors, including the United States and Northwestern University. This study offers a comprehensive mapping of the field and helps clarify future research directions. Luo et al. analyze 331 atrial fibrillation–related scientific videos from YouTube, Bilibili, TikTok, and Douyin using big data techniques such as principal component analysis and normalization. The results show that videos produced by creators with medical backgrounds tend to attract greater popularity, whereas non-medical creators often provide higher-quality basic educational content. The study also points to limited content regulation across platforms. Together, these findings introduce novel approaches for large-sample, cross-platform analyses of digital health information.

Two articles examine the application of AI in patient education. Huang, Wang et al. assess the performance of five large language models—ChatGPT-4o, DeepSeek-V3, Doubao, Wenxin Yiyan 4.0 Turbo, and Qwen—in responding to 31 frequently asked ophthalmology consultation questions. Using a five-point Likert scale combined with textual analysis, the authors find that ChatGPT-4o and DeepSeek-V3 achieve the highest overall performance and are more suitable for lay users, whereas Doubao and Qwen are better aligned with the needs of medically

trained audiences. This study offers practical guidance for the clinical deployment of large language models in ophthalmology. Bhattacharjee and Bhattacharya apply AI-driven nudge theory to improve hand hygiene compliance by introducing smart dispensers and wearable devices that deliver real-time visual and auditory reminders. Their findings show that these interventions can increase compliance by up to 30%, with additional gains when combined with gamification elements. This work highlights the potential of AI-enabled nudging technologies to support infection control and enhance patient education.

Three studies examine the effectiveness of digital technologies in patient education. Jiang et al. assess the quality of 200 Mandarin TikTok videos on diabetic retinopathy using the DISCERN and PEMAT-A/V instruments. Their analysis reveals substantial variation in content quality across uploader types, with videos produced by non-profit organizations scoring highest. Moreover, engagement metrics are strongly correlated with content quality, underscoring the importance of professional involvement and platform-level regulation in improving patient education outcomes. Impito and Azevedo evaluate the effectiveness of video-based health education (VbHE) in 33 least-developed African countries between 2020 and 2024, synthesizing evidence from 15 empirical studies. The findings indicate that VbHE can significantly enhance knowledge retention and health outcomes, particularly in areas such as cancer, HIV, and maternity care. At the same time, the authors highlight persistent challenges related to internet connectivity and cultural adaptation, emphasizing both the promise and the limitations of VbHE for underserved populations. Pavuloori et al. conduct a systematic review of 41 studies examining tools used to evaluate YouTube videos on surgical procedures. Their review shows that DISCERN, the Global Quality Score (GQS), and JAMA criteria are the most frequently applied assessment tools. However, they also find that approximately 88% of the videos fail to provide adequate patient education, while content from professional sources consistently offers higher educational value. This study calls for the development of standardized evaluation frameworks and improvements in the quality of online surgical education videos.

Two additional studies examine innovative digital technologies that enhance patients' understanding of health information. Zhou et al. conduct a cross-platform analysis of 271 pulmonary nodule–related videos published on YouTube, Bilibili, and TikTok, assessing both content quality and user engagement. Their findings indicate that long-form platforms such as YouTube and Bilibili tend to provide higher-quality medical information, whereas short-form platforms like TikTok generate greater user engagement. The study also shows that non-profit organizations and physicians are the primary content contributors. Based on these results, the authors offer practical recommendations for content creators, platform managers, and viewers. Wang et al. extend the theory of planned behavior to explore how sex-biased HPV vaccination information influences vaccination intentions among male college students. Using survey data from 240 respondents and structural equation modeling, they find that male-targeted information increases vaccination willingness by shaping attitudes, subjective norms, and perceived behavioral control. However, price sensitivity constrains the extent to which favorable attitudes translate into intention.

This study provides valuable insights for designing gender-sensitive health communication strategies.

3 Conclusion

Research on digital information for patient education has grown rapidly, yet it remains fragmented across technological applications, information quality, and practical outcomes. This Research Topic addresses this gap by bringing together thirteen studies that examine key dimensions of digital patient education—ranging from med-tech, AI, and mobile health to big data–driven information provision and ethical considerations. Together, these articles offer both theoretical and empirical insights into how digital technologies shape patient education. The findings yield actionable implications for patients, clinicians, and platform designers, highlighting ways to improve digital health engagement and the delivery of medical information. At the same time, the Research Topic underscores several limitations, including regional specificity and the relative lack of evidence on long-term effects. These gaps point to promising avenues for future research, such as cross-platform comparisons, longitudinal analyses, and the integration of multiple digital technologies. Overall, this Research Topic provides a foundation for advancing digital patient education and encourages continued inquiry in this evolving field.

Author contributions

XZ: Conceptualization, Writing – original draft. FG: Writing – original draft. PL: Writing – review & editing.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Meng F, Zhang X, Liu L, Ren C. Converting readers to patients? From free to paid knowledge-sharing in online health communities. *Inform Process Manag.* (2021) 58:102490. doi: 10.1016/j.ipm.2021.102490

2. Stephanie L, Sharma RS. Digital health eco-systems: an epochal review of practice-oriented research. *Int J Inf Manage.* (2020) 53:102032. doi: 10.1016/j.ijinfomgt.2019.10.017

3. Satti FA, Hussain M, Ali SI, Saleem M, Ali H, Chung TC, et al. semantic sequence similarity based approach for extracting medical entities from clinical conversations. *Inf Process Manag.* (2023) 60:103213. doi: 10.1016/j.ipm.2022.103213

4. Pian W, Song S, Zhang Y. Consumer health information needs: a systematic review of measures. *Inf Process Manag.* (2020) 57:102077. doi: 10.1016/j.ipm.2019.102077

5. Liu Y, Ren C, Shi D, Li K, Zhang X. Evaluating the social value of online health information for third-party patients: is uncertainty always bad? *Inf Process Manag.* (2020) 57:102259. doi: 10.1016/j.ipm.2020.102259

6. Liu QB, Liu X, Guo X. The effects of participating in a physician-driven online health community in managing chronic disease: evidence from two natural experiments. *MIS Quart.* (2020) 44:391–419. doi: 10.25300/MISQ/2020/15102

7. Agarwal R, Gao G, DesRoches C, Jha AK. Research commentary—the digital transformation of healthcare: current status and the road ahead. *Inform Syst Res.* (2010) 21:796–809. doi: 10.1287/isre.1100.0327

8. Goh JM, Gao G, Agarwal R. The creation of social value: can an online health community reduce rural–urban health disparities? *MIS Quart.* (2016) 40:247–64. doi: 10.25300/MISQ/2016/40.1.11

9. Zhang X, Xie K, Gu B, Guo X. Engaging physicians with introductory incentives: references to online and offline income. *MIS Quart.* (2025) 49:917–52. doi: 10.25300/MISQ/2024/18172

10. Zhang X, Wu Y, Valacich JS, Jenkins JL. Transforming patient-physician interaction through asynchronous online health interaction: a relational communication perspective. *J Assoc Inform Syst.* (2025) 26:1681–702. doi: 10.17705/1jais.00952

11. Zhang X, Li K, Wu Y, Liang S, Yu M. Transforming e-commerce with AI: navigating innovation, personalization, and ethical challenges. *J Theoret Appl Electr Commerce Res.* (2026) 21:29. doi: 10.3390/jtaer21010029

12. De Freitas J, Uguralp AK, Oguz-Uguralp Z, Puntoni S. Chatbots and mental health: insights into the safety of generative AI. *J Cons Psychol.* (2023) 34:481–91. doi: 10.1002/jcpy.1393

13. Zhang X, Yin C, Guo X, Lai K-H. The role of previous affective responses in the continuance of mhealth monitoring services: a longitudinal investigation. *IEEE Trans Eng Manag.* (2024) 71:6982–94. doi: 10.1109/TEM.2023.3275924

14. Chen S, Lai K-h, Guo X, Zhang X. The influence of digital health technology on the allocation of regional medical resources in China. *Health Policy Technol.* (2025) 14:101013. doi: 10.1016/j.hlpt.2025.101013

15. Nicolas-Rocca TS, Schooley B, Joo S-J. Design and development of a patient-centered e-health system to improve patient understanding at discharge. *Commun Assoc Inform Syst.* (2014) 34:1–24. doi: 10.17705/1CAIS.03424

16. Zhang X, Pu J, Lu Y, Guo F. Team makes you better: evidence from online medical consultation platforms. *Inform Syst Res.* (2025) 36:1738–59. doi: 10.1287/isre.2022.0689

17. Zhang X, Zhou L, Wang S, Fan C, Huang D. Facilitating patient adoption of online medical advice through team-based online consultation. *J Theoret Appl Electr Commerce Res.* (2025) 20:231. doi: 10.3390/jtaer20030231

Leveraging AI-driven nudge theory to enhance hand hygiene compliance: paving the path for future infection control

Samiksha Bhattacharjee[1] and Sudip Bhattacharya[1]*

[1]Department of Clinical Pharmacology, All India Institute of Medical Sciences Deoghar (AIIMS Deoghar), Deoghar, India, [2]Department of Community and Family Medicine, All India Institute of Medical Sciences Deoghar (AIIMS Deoghar), Deoghar, India

Hand hygiene is critical for preventing infections, yet maintaining compliance remains challenging across healthcare, schools, and communities. Despite strong evidence, lapses occur due to cognitive barriers, understaffing, limited resources, and antimicrobial resistance. Behavioral science highlights factors like time constraints and cognitive biases affecting adherence, with compliance rates as low as 40%. Nudge theory, developed by Thaler and Sunstein, offers promising solutions by using subtle interventions, like visual or auditory cues, to encourage hand hygiene without imposing strict regulations. Recent innovations integrate artificial intelligence (AI) with nudges, enhancing compliance through real-time feedback. AI-powered systems, such as smart dispensers and wearable devices, provide reminders using visual or auditory cues at critical moments. For example, dispensers may light up or chime when a healthcare worker enters a patient's room, prompting hand hygiene. Studies show these AI-driven interventions significantly improve compliance, with rates increasing by up to 30% in some cases. AI can also analyze patterns of non-compliance, deploying personalized nudges during high-risk periods. Combining nudge theory with gamification, such as team-based competitions and rewards, further reinforces positive habits. However, implementing AI solutions in countries like India faces challenges, including limited resources, resistance to new technologies, and cultural barriers. Despite hurdles, integrating AI-driven nudges with behavioral strategies has the potential to transform hand hygiene practices. This approach fosters accountability, reduces infection rates, and ensures safer patient care by embedding compliance into daily routines, paving the way for sustainable improvements in infection control.

KEYWORDS

hand hygiene, infection control, artificial intelligence, nudge theory, behavioral science

Introduction

Hand hygiene is a key element in infection prevention, recognized worldwide for reducing pathogen transmission. Despite its effectiveness, adherence remains a challenge in healthcare, schools, and communities. A meta-analysis by Allegranzi and Pittet (1) highlighted that proper hand hygiene among healthcare workers significantly reduces healthcare-associated infections (HAIs), improving patient outcomes and lowering costs.

Despite strong evidence supporting the effectiveness of hand hygiene in preventing infections, maintaining compliance remains a persistent challenge. Infection control rates often fluctuate due to various interrelated factors, including lapses in hygiene practices, inadequate implementation of standardized protocols, and deficiencies in healthcare infrastructure. These challenges are exacerbated in resource-limited settings, where access

to essential supplies such as soap, alcohol-based hand sanitizers, or even clean water may be inconsistent.

Understaffing in healthcare facilities significantly contributes to non-compliance, as healthcare workers often face overwhelming workloads and time constraints that deprioritize proper hand hygiene. Additionally, limited access to hygiene resources within high-pressure environments, such as emergency rooms or rural health centers, further hampers adherence to recommended practices. The situation is further complicated by the growing prevalence of antimicrobial-resistant (AMR) pathogens. AMR pathogens not only increase the risks associated with lapses in hygiene but also escalate the consequences of infection outbreaks. Poor hand hygiene provides a pathway for the transmission of these resistant strains, making the task of infection control even more critical and complex. These multifaceted challenges underscore the importance of not only promoting education and awareness around hand hygiene but also ensuring the availability of adequate resources, reinforcing protocols, and addressing systemic issues in healthcare infrastructure. Innovative solutions, such as the integration of technology and behavioral strategies, are increasingly being explored to overcome these barriers and ensure sustainable compliance with hand hygiene standards (1).

Are we looking in the wrong places? AI in healthcare could help tackle human behavior challenges. Despite education, healthcare workers face cognitive barriers to hand hygiene. AI and behavioral science, like nudge theory, offer solutions through real-time feedback, enhancing compliance and infection control.

The role of human behavior in hand hygiene compliance

Healthcare workers (HCWs) navigate numerous cognitive demands daily, and hand hygiene frequently takes a backseat to more immediate tasks. Behavioral science identifies key factors that contribute to lapses in hygiene, including time constraints, risk perceptions, and cognitive biases (2). A systematic review by Eramus et al. points out the hand hygiene guideline compliance among 96 empirical studies were pretty low (40%) and a minority of those included studies only looked for behavioral determinants of hand hygiene stated previously. Nudge theory, developed by Richard Thaler and Cass Sunstein, suggests that small, subtle interventions, or "nudges," can significantly influence behavior without limiting choices as humans are driven by "choice architecture" (3). Nudges are particularly useful in healthcare because they guide professionals toward the desired action— hand hygiene—without imposing stringent regulations (3).

A common example is the "default bias," where individuals are more likely to follow preset guidelines if they do not require much effort to deviate from. It is a psychological tendency where individuals are more likely to stick with a preselected option or follow preset guidelines, primarily because it requires minimal effort to opt out or make an alternative choice. This bias leverages human preference for convenience and simplicity, capitalizing on the tendency to avoid additional cognitive or physical effort.

In the context of hand hygiene, default bias can be harnessed to encourage compliance by designing environments and systems that make the desired behavior—such as washing or sanitizing hands—the easiest or most natural action. For example, strategically placing hand sanitizer dispensers at prominent, unavoidable locations like the entrance of patient rooms or near medical equipment ensures that healthcare workers encounter these reminders as part of their routine workflow. The physical proximity and accessibility of these dispensers reduce the effort needed to perform hand hygiene, aligning with default bias.

Another application involves visual or auditory cues, such as dispensers that light up or emit a soft chime when approached. These cues serve as gentle nudges, reinforcing the preselected guideline of sanitizing hands before or after patient interaction. By incorporating these prompts into the environment, compliance becomes the path of least resistance, minimizing the likelihood of neglect. Default bias is particularly effective because it works subtly, without requiring conscious decision-making. By embedding the desired action into the default setting, healthcare institutions can promote consistent hand hygiene practices while reducing the cognitive burden on busy staff. This approach not only improves compliance but also fosters habitual behavior, contributing to better infection control outcomes. In hand hygiene, visual or auditory cues serve as nudges that subtly remind HCWs to wash their hands, even in low-risk settings where compliance is typically lower (4). By tapping into human psychology, nudges can trigger automatic compliance behaviors that would otherwise be neglected (4).

AI-driven nudges can enhance hand hygiene practices

Recent innovations in AI are reshaping how healthcare settings implement nudges. Automated systems such as AI-driven smart dispensers and wearable devices and other automated electronic motoring system can monitor hand hygiene and offer real-time feedback based on individual behaviors (5). AI can go beyond passive monitoring and employ nudges like visual cues (lights on dispensers) or auditory prompts (alerts or chimes) to remind healthcare workers to wash their hands at the most opportune moments (5). For instance, when a healthcare worker approaches a patient room, the hand sanitizer dispenser could glow or emit a gentle chime as a nudge, reminding them to sanitize before entering. On these lights, a systematic review by Wang C et al. have found that, these can be a useful tool while keeping in mind the challenges that these bring along, such as issues of accuracy, usability, associated costs etc. (5).

Research shows that these AI-driven nudges have been highly effective in boosting compliance. A study involving AI-powered badges that provide auditory feedback every time a hand hygiene opportunity was missed showed a 30% increase in compliance rates (6). These interventions combine real-time feedback with gentle nudges, reinforcing habitual behavior without imposing direct control. It was found to be as effective as human observation in one of its kind trial done by Singh et al. (7).

Furthermore, AI-driven predictive analytics can identify patterns of non-compliance, allowing systems to deploy nudges at the most critical times. For example, AI can detect if a healthcare worker has a history of skipping hand hygiene during high-stress

periods or after prolonged patient interaction. In such cases, personalized nudges, such as a voice prompt from the dispenser or a vibrating wearable, could pre-emptively remind the worker to wash their hands.

Integrating nudge theory and AI for sustainable behavioral compliance

While technology plays a critical role, it is the combination of AI-driven nudges and behavioral feedback that can sustain long-term compliance. Nudge theory posits that consistent, subtle cues will eventually create habits that persist even when the prompts are no longer present (3–5). Healthcare settings that integrate nudges into their hand hygiene protocols use behavioral models, such as the cue-action-reward loop, to reinforce the desired action.

In hospitals, visual cues like color-changing lights and auditory signals such as chimes help reinforce hand hygiene, gradually becoming part of healthcare workers' routines and reducing the effort needed to remember hygiene practices. Studies have shown that combining these nudges with gamification—another behavioral strategy—can enhance adherence (6). This technology can be leveraged here as when AI systems track team performance and offer rewards for high compliance, healthcare workers are more likely to sustain positive habits. One study at Netherlands involving nine public hospitals introduced team-based competitions with public recognition for high hand hygiene scores, leading to a significant (9%) improvement in compliance among nursing officers. Several institutions have successfully implemented AI-powered nudge systems to enhance hand hygiene compliance. At, New York City Health + Hospitals/South Brooklyn Health AHHR (automated hand hygiene reminder system) combined with real-time feedback increased hand hygiene compliance from around 86 to 89% in MSU wards over a 2 week period (8). The system was successfully implemented in that hospital and stands as a solid example to this initiative. Talking about India's experiments, A Delhi based design studio "Quicksan" in partnership with Jhon's Hopkin's University launched a project called "Safehands" in labor rooms and district hospitals and community health centers across India aiming to find a solution to hand hygiene compliance. The device named was "HAIgeine," consisting a camera to capture the hand movements, an AI based processing unit to give feedback. Although there was probably less flaw in the novel tool, rather the acceptance was something to be concerned for. There has been report that at some centers the physician's particularly avoided the washbasin with this system enabled (9).

However, implementing AI-enabled nudge theory for hand hygiene in India faces challenges due to the country's diverse healthcare system. In rural areas, inconsistent access to clean water and overstretched staff makes AI solutions difficult. Cultural shifts have complicated traditional hygiene practices, and resistance to AI systems is common due to unfamiliarity. High costs further limit adoption, especially in rural hospitals, widening healthcare inequalities. AI alone cannot address deeper behavioral issues like hierarchy, work pressure, and overconfidence.

The future of hand hygiene will combine AI with personalized nudges and wearable feedback, maintaining high compliance even during busy times (4, 5) AI can use hand hygiene data to deliver tailored nudges, such as personalized cues for specific situations like sanitizing after glove removal. This precision can significantly enhance compliance, lower infection rates, and improve patient outcomes.

Conclusion

Combining AI-driven nudges with behavioral feedback effectively addresses hand hygiene compliance. This approach subtly influences behavior, boosts compliance, and fosters a culture of accountability, leading to safer patient care and better infection control.

Recommendations

- To improve hand hygiene compliance, healthcare facilities should leverage behavioral insights like default bias and integrate them with technology-driven solutions. This can be achieved by designing workflows and environments where hand hygiene becomes the easiest and most natural action. For example, hand sanitizer dispensers should be strategically placed at critical points such as patient room entrances, ensuring their accessibility aligns with staff movements.
- Facilities should also adopt AI-enabled systems that provide real-time reminders through visual or auditory cues, such as lights or chimes, to prompt compliance. These systems should be tailored to individual behaviors, using predictive analytics to deploy nudges at high-risk times, such as during busy shifts or after prolonged patient interactions. Combining these interventions with gamification strategies, like team-based competitions and performance recognition, can further motivate healthcare workers to adhere to hygiene protocols.
- Additionally, efforts should focus on addressing systemic barriers, such as understaffing and resource limitations, to sustain compliance. This includes ensuring consistent availability of hygiene supplies, improving staff training, and fostering a culture of accountability through leadership support and ongoing feedback.
- For resource-limited settings, scalable and cost-effective solutions, like low-cost visual cues or simplified AI systems, should be prioritized. Implementing these measures will not only enhance hand hygiene practices but also contribute to reducing healthcare-associated infections and combating antimicrobial resistance, ultimately improving patient outcomes and safety.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SaB: Writing – original draft, Writing – review & editing. SuB: Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that Generative AI was used in the creation of this manuscript. We acknowledge the use of generative AI tools in the preparation of this manuscript. These tools were employed solely for checking grammatical mistakes and language errors, ensuring clarity and precision in the text. The authors affirm that they have critically reviewed and refined the manuscript and take full responsibility for its content, including the ideas, interpretations, and conclusions presented herein.

Publisher's note

References

1. Allegranzi B, Pittet D. Role of hand hygiene in healthcare-associated infection prevention. *J Hosp Infect*. (2009) 73:305–15. doi: 10.1016/j.jhin.2009.04.019

2. Erasmus V, Daha TJ, Brug H, Richardus JH, Behrendt MD, Vos MC, et al. Systematic review of studies on compliance with hand hygiene guidelines in hospital care. *Infect Control Hosp Epidemiol*. (2010) 31:283–94. doi: 10.1086/650451

3. Thaler RH, Sunstein CR. Nudge: Improving decisions about health, wealth, and happiness. (2008). Yale University Press.

4. Pittet D. Hand hygiene: it's all about when and how. *Infect Control Hosp Epidemiol*. (2008) 29:957–9. doi: 10.1086/592218

5. Wang C, Jiang W, Yang K, Yu D, Newn J, Sarsenbayeva Z, et al. Electronic monitoring systems for hand hygiene: systematic review of technology. *J Med Internet Res*. (2021) 23:e27880. doi: 10.2196/27880

6. Fisher DA, Seetoh T, Oh May-Lin H, Viswanathan S, Toh Y, Yin WC, et al. Automated measures of hand hygiene compliance among healthcare workers using ultrasound: validation and a randomized controlled trial. *Infect Control Hosp Epidemiol*. (2013) 34:919–28. doi: 10.1086/671738

7. Singh A, Haque A, Alahi A, Yeung S, Guo M, Glassman JR, et al. Automatic detection of hand hygiene using computer vision technology. *J Am Med Inform Assoc*. (2020) 27:1316–20. doi: 10.1093/jamia/ocaa115

8. Seferi A, Parginos K, Jean W, Calero C, Fogel J, Modeste S, et al. Hand hygiene behavior change: a review and pilot study of an automated hand hygiene reminder system implementation in a public hospital. *Antimicrob Steward Healthc Epidemiol*. (2023) 3:e122. doi: 10.1017/ash.2023.195

9. Chen Y, Clayton EW, Novak LL, Anders S, Malin B. Human-Centered Design to Address Biases in Artificial Intelligence. *J Med Internet Res*. (2023) 25:e43251. doi: 10.2196/43251

Exploring the impact of discharged patients' characteristics on online health information-seeking behaviors: insights from patients' dilemmas

Fei Liu[1], Xiangyin Kong[1], Tian Xia[2] and Huijing Guo[3]*

[1]School of Economics and Management, Harbin Engineering University, Harbin, China, [2]School of Management, Harbin Institute of Technology, Harbin, China, [3]School of Economics and Management, China University of Mining and Technology, Xuzhou, China

Introduction: Aligned with the dual-factor model, this study aims to investigate why discharged patients seek online health information, considering the complexity of patients' dilemmas. Additionally, we account for specific cultural context factor in China and seek to examine the role of living with children in mitigating the dilemmas faced by discharged patients in their pursuit of additional online information.

Methods: We empirically tested the research model using data collected from 292 discharged patients. The data was examined through structural equation modeling, employing Smart PLS.

Results: The findings suggest that perceived stress facilitates discharged patients' engagement in seeking online health information, whereas resistance to change and learned helplessness impede such behaviors. Furthermore, our analysis reveals that cohabiting with children moderates the effects of resistance to change on online health information-seeking behavior.

Discussion: In conclusion, this paper extends the literature by examining the role of discharged patients' characteristics on online health information-seeking behaviors. Following the practices in China, this study involves living arrangements (with children) as an essential factor in the research model. This paper offer suggestions to online providers to make health-related information more suitable for discharged patients.

KEYWORDS

dilemmas, discharged patients, dual-factor model, living with children, online health information-seeking behaviors

1 Introduction

With the popularization of the Internet and the promotion of online health services, online health information services are becoming increasingly popular among individuals with health information demands (Cline and Haynes, 2001; Gulec et al., 2022). Studies have reported that, on the demand side, approximately 74% of young Americans (Basch et al., 2018) and 65% of Australians (Nikoloudakis et al., 2018) have sought health information online, including conditions, symptoms, and treatment options advice. In China, patients are more accustomed to seeking health resources from traditional offline channels. However, due to the large

population in China, it is inconvenient to acquire health information exclusively through offline channels (Matsumoto et al., 2012), leading to patients being unable to satisfy their quest for health information. According to Xinhua Net, the news website of the Chinese government, online health information-seeking behaviors rank as the second main category of people's information-seeking behaviors. With the development of the Chinese Internet in the healthcare sector, more and more people are seeking web-based health information and services (Cao et al., 2016; Zhang et al., 2017).

Health information-seeking behaviors involve acquiring health-related information to assist individual medical decision-making and self-management (Lambert and Loiselle, 2007). Online health information services can benefit healthy people by preventing disease (Zhao et al., 2021). It can also help patients manage their health conditions. Patients who suffer from diseases can search for health-related information to find ways to minimize their suffering related to health conditions, such as recovery problems. Online information could also provide an alternative explanation of doctors' suggestions for patients (Chang and Huang, 2020).

Patients usually have health information demands during the complex self-management period (Sedrak et al., 2020). Although online health information services provide a more convenient and low-cost way for patients to meet their information needs, patients still cannot avoid the dilemmas discussed herein. In China, discharged patients lack basic health-related information, and they cannot acquire well-matched information or distinguish it from less relevant information (Meng et al., 2019). In addition to technical factors, online-seeking behaviors are influenced by the characteristics of discharged patients themselves. Although searching for health information online has become common among patients, existing research primarily focuses on the health information seek for the general patient population (Kim et al., 2023). It is important to note that discharged patients, particularly in China, are often middle-aged or elderly individuals who have unique experiences and may face additional challenges when seeking health information online.

In this context, studies are needed to understand the real dilemmas discharged patients face when seeking online health information. Such dilemmas include dependence on clinical opinions (Eibergen et al., 2018), vulnerability (Walter et al., 2019), extensive negative emotions (Yan et al., 2021), and stress (Musarezaie et al., 2019) after a long clinical experience. They usually experience a high level of perceived stress while dealing with their own health conditions (Liu et al., 2021) and are unwilling to change their original information acquisition routes (Shahbaz et al., 2020). At the same time, they usually feel a helplessness when seeking health information online due to the conflicting information they find (Fassett-Carman et al., 2019; Wanke and Schwabe, 2020).

Resistance is often manifested as the failure of users to transition from existing business models to new ones (Polites and Karahanna, 2012). Online seeking health information provide a new source of information that is accessible 7 days a week and without spatial limitations. Moreover, this method has indeed supplanted traditional offline means of obtaining health information to some extent. Unlike typical patients, discharged patients may encounter greater resistance when shifting to online sources for health information, as they have become accustomed to receiving it from professional healthcare providers in person. This resistance to seeking online health information among discharged patients may be related to their

hospitalization experiences. However, existing studies on the impact of patient characteristics on online health information-seeking behaviors have primarily focused on facilitating factors, such as the roles of health-related stress (de Looper et al., 2021) and health literacy (Nangsangna and Vroom, 2019). There has been limited attention to the influence of discharged patients' characteristics on their online health information-seeking behaviors, particularly concerning inhibitory factors. Understanding these factors not only enables health information providers to develop effective dissemination strategies but also enhances our comprehension of facilitating factors from a new perspective (Tsai et al., 2019). Drawing on the dilemmas discharged patients face, this study adopted the dual-factor model to understand the role of discharged patients' characteristics in their information-seeking processes. This study analyzed how facilitating and inhibitory factors, especially the latter, affect online health information-seeking behaviors. Thus, the first research question is as follows:

> RQ 1. Which factors, especially inhibitors, can impact patients seeking health information online based on their dilemmas after discharge?

At the same time, considering the traditional Chinese culture, the role the family plays in health information acquisition and decision-making cannot be negligible (Zhang et al., 2020). To assist with care delivery and avoid accidents, discharged patients usually live with their children, if they have any, making living arrangements essential for patients' self-management (Irani et al., 2019). Thus, the second research question is as follows:

> RQ 2. How do living arrangements affect the relationship between discharged patients' characteristics and online health information-seeking behaviors?

To narrow this research gap, we adopted the dual-factor model and proposed a research model to investigate the enablers and inhibitors of patients' online health information-seeking behaviors. According to the dilemmas discharged patients face, we regard the enablers (perceived stress) and inhibitors (resistance to change and learned helplessness) as factors affecting online health information-seeking behaviors.

We tested our hypotheses using a survey of 292 respondents. The results supported all the direct influencing routes. Perceived stress was positively associated with online health information-seeking behaviors, while resistance to change and learned helplessness were negatively associated with these behaviors. By comparing patients who live with their children with those who do not, we found that living with children can only significantly weaken the relationship between resistance to change and online health information-seeking behaviors. The result rejected the moderating effect of the other two direct routes.

This paper contributes to the health-related online-seeking literature in several ways. First, this paper extended the literature by examining the positive relationship between the perceived stress of discharged patients and their online health information-seeking behaviors. Second, we enriched the literature on online health information-seeking behaviors by introducing the role of inhibitors following the health and information dilemmas discharged patients face. Third, this study contributes to understanding how living arrangements (i.e., with children) influence discharged patients'

online health information-seeking behaviors. In practical terms, this research starts from the dilemma discharged patients face to better understand why patients do or do not use online information-seeking services. Using the findings, this paper offers suggestions to online health information providers to make online information more relevant to discharged patients. Finally, this paper emphasizes the important role of living arrangements on discharged patients' online health information-seeking behaviors.

The rest of the paper proceeds as follows. In the theoretical background section, we describe prior research on online health information-seeking behaviors, the dual-factor model, perceived stress, resistance to change, and learned helplessness to set the literary context for our study, and the research model incorporates the dual-factor model. Next, the research methodology in the subsequent section, including data collection and analysis, to empirically test our hypothesized models and introduces the statistics and calculation results of the partial least squares. Finally, this study provides the main findings, implications, and limitations.

1.1 Online health information-seeking behaviors

Online health information-seeking behaviors refers to the process by which individuals obtain health information about diseases, treatment plans, and prognoses through search engines (e.g., Google and Baidu), dedicated online health platforms (e.g., Haodaifu), and social media platforms (e.g., Facebook and WeChat) (Chen et al., 2025). As popular internet behaviors, online health information-seeking behaviors may generate multiple health-related outcomes, such as assisting patients' self-diagnosis under specific health conditions (Jiang and Street, 2017; McMullan et al., 2019) and supporting health or medical services distribution (Shamlou et al., 2022). Driven by multiple motives, needs, and environmental factors, people may adopt online health information services (Reifegerste et al., 2019). Related phenomena have led researchers to explore the influencing factors of online health information-seeking behaviors in recent years.

Previous studies usually examined the impact of internal (personal traits) and external (environmental characteristics) factors on online information-seeking behaviors. Such studies found that individual health-related characteristics, such as lacking health literacy (Wong and Cheung, 2019) and suffering from disease symptoms (Berle et al., 2020), were the essential influencing factors of online health information-seeking behaviors. At the same time, social surroundings, such as lacking health resources (Xiao et al., 2014), could also influence online health information-seeking behaviors.

Patients are key consumption populations adopting online health information services (Sedrak et al., 2020). To date, a rich and extensive literature has examined the enablers of patients' online health information-seeking behaviors (de Looper et al., 2021; Nangsangna and Vroom, 2019; Zhao et al., 2020). As a part of healthcare services, online health information could compensate for the shortage of traditional medical services and expand patients' medical resources (Miller and Bell, 2012; Zhao et al., 2020). Then, with the expectation of improving personal health status and self-management levels, patients may seek information online (Kitchens et al., 2014). Yet once patients are discharged from hospitals, they may face more obstacles

when seeking online health information due to their unique characteristics.

1.2 Dual-factor model of online health information-seeking behaviors

The dual-factor model highlights the need to consider both enablers and inhibitors when using information systems (Cenfetelli, 2004). Enablers are those factors that promote system usage whereas inhibitors focus on the factors that predict users' tendency to reject the technology. According to the dual-factor model, it is essential to understand the inhibitors from the perspective of personal characteristics, such as technology anxiety (Guo et al., 2013) and dispositional resistance to change (Hsieh, 2016), when using information systems. Focusing on the elderly, previous studies have investigated the correlation of online health information-seeking behaviors with possible barriers, such as low e-health literacy, conflicting information, and web design factors (Pourrazavi et al., 2022). Due to past disease experiences and uncontrollable health conditions, discharged patients may generate negative emotions and dependence on clinical opinions when they encounter barriers while seeking health information, leading to a negative impact on factors related to use behaviors (Bhattacherjee and Hikmet, 2007). Although the existing literature has studied online health information-seeking behaviors from many aspects, it is rare to treat discharged patients as subjects while exploring the inhibitors of these demand groups. Hence, in this study, we adopted the dual-factor model to understand the inhibitors of discharged patients' online health information-seeking behaviors.

1.3 Dilemmas and characteristics of discharged patients

Patients usually differ from general population in their specific health concerns and stressful situations. Due to the uncertainty and perceived risks of their health status, seeking online health information is more purposeful for patients (Rains and Tukachinsky, 2015), such as seeking disease treatment information, than for general populations. Researchers have found that primarily personal factors, such as stress and anxiety generated from complex health conditions (McMullan et al., 2019), drive patients to seek health information online. Due to long-term clinical experience, discharged patients tend to be dependent on clinical opinions (Eibergen et al., 2018), vulnerable (Walter et al., 2019), full of negative emotions (Yan et al., 2021), and stressed (Musarezaie et al., 2019). The dilemma discharged patients face, caused by their complicated conditions, recovery problems, distance from offline medical resources, and special needs for disease-related information, will affect their online health information-seeking behaviors. Thus, representative personal characteristics of discharged patients were identified based on several specific dilemmas.

Unlike general populations who are remote from health resources, discharged patients have usually adopted several clinical resources to deal with their health problems during hospitalization. Although patients receive clinical guidance and advice, limited health resources support them in going through the uncertainty and unprovability of health conditions after

discharge (Chang and Huang, 2020). These health-related conditions may generate related concerns and contribute to a higher level of perceived stress for discharged patients than general individuals. Perceived stress reflects the subjective evaluation of the stress level individuals experience when they encounter a stressful objective event (Cohen et al., 1983). Patient populations usually have problems with perceived stress (Liu et al., 2021). Related studies have shown that patients experience higher levels of perceived stress due to health threats (Liu et al., 2021) and emotional distress (Yan et al., 2021) than healthy people. One previous study demonstrated that patients with stressful health conditions and concerns have a greater need for health information, leading to frequent health information-seeking behaviors (Musarezaie et al., 2019). When they are discharged from the hospitals, perceived stress becomes a significant characteristic of discharged patients in health information acquisition to be explored.

Previous studies have indicated that, after discharge, patients tend to search for information using health queries to reduce uncertainty (Yan et al., 2021). As an effective strategy to relieve health-related stress, online health information-seeking behaviors, which are low-cost ways to acquire health-related information, can comfort stressed discharged patients. Thus, we hypothesize that:

> *Hypothesis 1*. Discharged patients' perceived stress increases the number of their online health information-seeking behaviors.

As health-related problems have accompanied discharged patients for some time, they are usually dependent on clinical opinions and a resistance to changing their information behaviors. Resistance to change refers to the unwillingness to take action to adapt to the pressure to change the status quo (Shahbaz et al., 2020). Previous studies have found that resistance to change could negatively impact an individual's intention to change health behaviors (Deng et al., 2014; Laumer et al., 2016). Resistance to change also weakens the positive relationship between intention to use and actual use (Shahbaz et al., 2020). Discharged patients with long-term illness experience have become accustomed to seeking clinical support from offline physicians (Sedrak et al., 2020), causing them to distrust alternative health-related resources. As a result, discharged patients may resist changing their health behaviors due to the influence of the external environment. Once discharged patients who are used to offline diagnosis and treatment opinions will exhibit resistance to change as a key characteristic of their online health information seeking.

Online health information services could provide an accessible way for discharged patients to find health-related information (Soroya et al., 2021). Discharged patients' efforts to seek health-related information online could reshape their ways of handling health issues (Guo et al., 2013). However, discharged patients are mainly accustomed to adopting information acquired from face-to-face hospital medical services rather than using online information technology, which could cause them to resist changing their previous health habits (Deng et al., 2014) related to acquiring health information. In sum, we believe that previous health information-acquiring habits may generate a resistant attitude that leads to decreased online health information-seeking behaviors of patients after discharge. Thus, we hypothesize that:

> *Hypothesis 2*. Discharged patients' resistance to change decreases the number of their online health information-seeking behaviors.

Online health information-seeking behaviors may generate unpredictability and uncertainty in the search results (Soroya et al., 2021), which discharged patients who lack basic health-related knowledge may not be able to discern (Meng et al., 2019). And individual discharged patients have complicated health issues (Musial et al., 2020), while online health information services usually lack personalization (Meng et al., 2019), resulting in difficulties in acquiring well-matched information online. This dilemma may result in helplessness for discharged patients seeking information online (Fassett-Carman et al., 2019; Wanke and Schwabe, 2020). Learned helplessness is grounded in cognitive learning theory (Maier, 1980) and indicates that people think, feel, or act passively and helplessly when they cannot control their surroundings (Trindade et al., 2020). The root of learned helplessness is the belief that the individual has lost control of the situation and is incapable of making changes (Baum et al., 1986). Because of the uncertainty of clinical treatment (Djulbegovic et al., 2011), discharged patients usually feel incapable of dealing with their health conditions through personal efforts, resulting in learned helplessness. Previous studies have shown that learned helplessness harms mental health and behavioral changes (López Steinmetz et al., 2021; Xue et al., 2023). If patients' online efforts to seek health information are unable to improve their health outcomes, the characteristics of learned helplessness will become more pronounced.

If individuals believe that their efforts may not produce improved health conditions and may even worsen these conditions, they may stop their attempts to find health information (Xue et al., 2023; Ying et al., 2021) due to their helplessness. Feelings of helplessness when repeatedly seeking health information online without changing the patient's health conditions lead to the emergence of learned helplessness (Benko et al., 2021), resulting in a negative attitude toward seeking online health information. Thus, we hypothesize that:

> *Hypothesis 3*. Discharged patients' learned helplessness decreases the number of their online health information-seeking behaviors.

1.4 Living arrangement

People who live with their children tend to have improved health cognition (Yu et al., 2022) and more appropriate health management behaviors. Understanding the role of this living arrangement in discharged patients' online health information-seeking behaviors is essential.

Although seeking information online is an essential strategy for releasing perceived stress, discharged patients living with their children have other strategies for coping with stress issues, such as face-to-face communication with their children (Yan et al., 2021). We can conclude that, under the same level of perceived stress, discharged patients who live with their children may find more alternative strategies to relieve perceived stress than those who do not live with their children. Therefore, we hypothesize that:

> *Hypothesis 4*. Under the same level of perceived stress, discharged patients who live with their children seek less online health

information than discharged patients who do not live with their children.

Related studies have demonstrated that social support (such as the companionship of children) is associated with attitude change (Stroebe and Diehl, 1988). Living with their children may reduce discharged patients' online health information-seeking behaviors caused by a resistant attitude and may enhance their self-management behaviors (Irani et al., 2019). Likewise, discharged patients who live with their children could also reduce online health information-seeking behaviors caused by resistance to change through social support provided by their children. Then, we hypothesize that:

Hypothesis 5. Under the same level of resistance to change, discharged patients who live with their children seek more online health information than discharged patients who do not live with their children.

People who live with their children usually have less leisure time and are busy dealing with various domestic affairs (Walter and Haun, 2021), making them more vulnerable to managing situations with little outside resources. As a result, people who live with their children may experience weakened positive thinking (Walter and Haun, 2021). For discharged patients with feelings of learned helplessness, the negative thinking may be enhanced by living with their children. Therefore, we hypothesize that:

Hypothesis 6. Under the same level of learned helplessness, discharged patients who live with their children seek less online health information than discharged patients who do not live with their children.

1.5 Research model

Following the dual-factor model, we developed a research model to understand the enabler (perceived stress) and inhibitors (resistance to change and learned helplessness) in discharged patients' online health information-seeking behaviors (see Figure 1). We built this research model from the insights of discharged patients' dilemmas related to their health conditions and online information service engagement. Furthermore, considering the essential role of living arrangements in discharged patients' daily life, we proposed that living with children is a moderator that affects the relationship among our three direct affect routes.

We also included several personal characteristics in the model as control variables to alleviate the influence due to covariance issues. Previous literature has suggested that individuals' age, gender (Weber et al., 2020), levels of education (Berle et al., 2020), and disease experience (Sedrak et al., 2020) can influence their health-seeking behaviors on the internet. Based on our research context, we adopted the number of hospitalization experiences instead of the disease experience. We thus controlled the effects of these factors on online health information-seeking behaviors.

2 Methodology

2.1 Participants

To pretest the questionnaire, we initially collected data from a small sample of hospitalized patients in offline hospitals. However, as we began to collect the questionnaires, the spread of COVID-19 resulted in the loss of opportunities for face-to-face interactions with
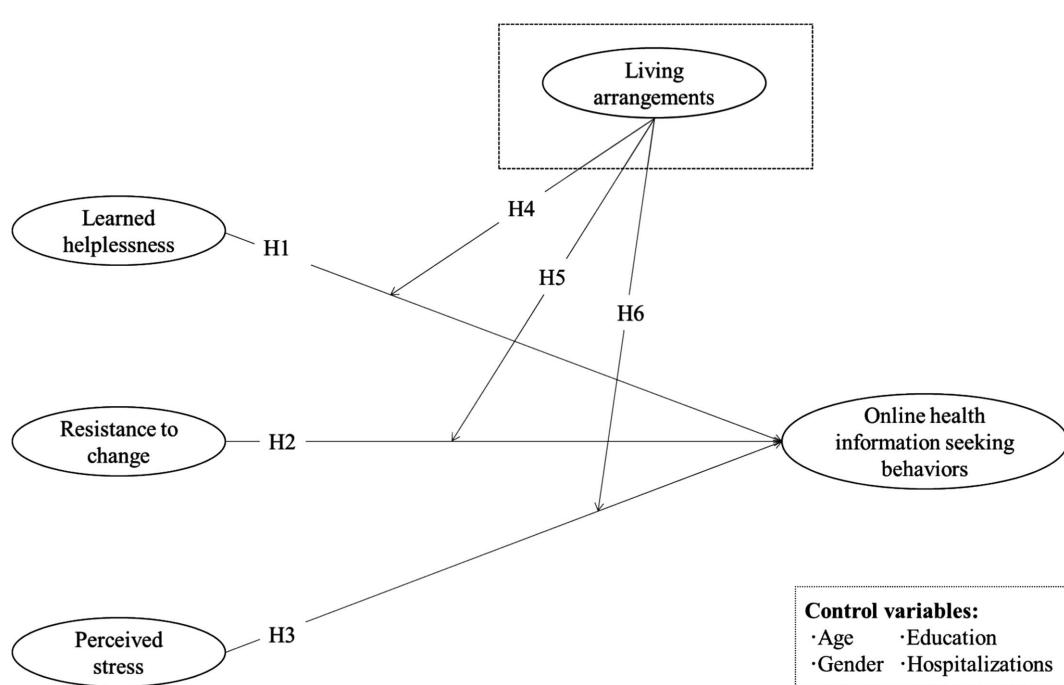
FIGURE 1
Research model.

patients. Therefore, we utilized the data collection service of China's leading online survey platform, Wenjuanxing, to manage the survey. There were two primary reasons for conducting the survey online. First, it relates to our research topic, online health information seeking behaviors, which is inherently digital and primarily targets internet users. Second, it has been reported that the internet has become the preferred source of information for patients, surpassing healthcare professionals (Chen et al., 2025). Thus, the online survey can be considered representative to some extent. The questionnaire was divided into two parts. In the first part, respondents were informed about the purpose of data collection and provided their consent. Subsequently, they answered questions regarding their hospitalization experience and health management after discharge. In the second part, they were asked about their perspectives on online health information seeking behaviors during their self-management period post-discharge. To ensure data quality and reduce social desirability bias, we have taken several measures when designing questionnaires and cleaning up the collected data. First, we added attention traps and reverse coding questions to reduce single-method bias. Second, we ruled out all cases where the problem is missing a value or a similar value. The questionnaires were distributed online from September 1 through September 30, 2021.

The website initially collected 472 questionnaires, however, due to concerns about patient privacy and the absence of effective incentives, 158 responses were incomplete. Of the 314 pieces of data the website provided, the authors excluded another 22 pieces of data with more than eight consecutive measurement items of the same value (indicating that the respondents did not answer seriously). Ultimately, 292 valid questionnaires were used to conduct the quantitative analysis.

Table 1 presents the demographic characteristics. According to the data, the proportion of women in the survey was a bit higher than that of men, reaching 54.46%, and the age was mainly 25–34 and 35–44 years old. Most of the survey group had a master's degree or above, and they had been hospitalized once or twice. Meanwhile, most of them did not live alone; 59.25% lived with their children.

2.2 Measures

To test the variables in this study, we adopted and developed the measurement items from previous studies. We first adopted the measurement of online health information-seeking behaviors (Bell et al., 2011) and then developed the measurement items based on the practices in China. We summed the five answers up to measure the number of online health information-seeking behaviors. We also adopted measurements of learned helplessness (Quinless and Nelson, 1988), resistance to change (Bhattacherjee and Hikmet, 2007), and perceived stress (Cohen et al., 1983) from previous research on a 7-point Likert-type scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"). We examined living with children by directly asking the participants whether they lived with their children (regardless of whether they have a child or not). Based on the above support and the background of this study, we translated the questionnaire into Chinese and made it more suitable for answering.

TABLE 1 Demographic characteristics.

	Frequency	Percentage (%)
Gender		
Male	133	45.54
Female	159	54.46
Age		
18–24	17	5.82
25–34	102	34.93
35–44	89	30.48
45–54	63	21.58
55–64	13	4.45
≥65	8	2.74
Education		
Middle school and below	9	3.08
Technical secondary school and high school	29	9.93
Junior college	38	13.01
Bachelor's degree or above	216	73.98
Hospitalization frequency		
0	18	7.00
1	161	54.14
2	83	28.98
≥3	30	9.87
Live alone		
Yes	18	6.16
No	274	93.84
Live with children		
Yes	173	59.25
No	119	40.75

2.3 Data analysis

Partial least squares (PLS), a form of structural equation modeling (SEM), provides value for behavioral research fields (Lowry and Gaskin, 2014). PLS offers some data analysis advantages, and we used this technique to test the research model for the following reasons. First, PLS has low requirements on the measurement scale and requires a relatively small sample size (Hair et al., 2011). Moreover, SEM allows the simultaneous operation of multiple related equations, and a model is measured multiple times. PLS can be used for both the confirmation of existing model parameters and exploratory verification (Chin, 1998).

We used the Smart PLS 4.0 software as the statistical tool to examine the research model. We analyzed our data in three stages. The first stage was testing the measurement model, which evaluated the reliability and validity of the construct to ensure its appropriateness. In the second stage, we tested the structural model and the hypotheses. Finally, we used a multigroup analysis to compare whether the respondents live with their children or not.

3 Results

3.1 Measurement model testing

We evaluated model reliability by examining composite reliability (CR) and average variance extracted (AVE) (Fornell and Larcker, 1981; Hsu and Lin, 2008). The results are shown in Table 2. The CR ranged was from 0.864 to 0.937 (higher than 0.7), and the AVE ranged from 0.672 to 0.789 (higher than 0.5). Comparing these two sets of results with the critical values of 0.7 and 0.5, respectively (Fornell and Larcker, 1981; Hsu and Lin, 2008; Chin, 1998) indicated that the structure had very good reliability. Discriminant validity could be measured by the square root of AVE (Fornell and Larcker, 1981). As shown in Table 2, the square root of AVE ranged from 0.820 to 0.888, which was greater than variable correlations. Thus, it could be concluded that the discriminant validity was acceptable.

We evaluated the convergent validity by measuring the items' loadings. The results in Table 3 show that the loadings of all items ranged from 0.782 to 1.000, all of which were greater than 0.70 (Anderson and Gerbing, 1988). In addition, the loadings of these items were greater than their cross-loadings with other items (Gefen and Straub, 2005), indicating that the structure had good convergent validity.

3.2 Structural model testing

Based on our research hypotheses, the measurement of the structural model was tested in stage 1. The PLS results of the structural model are shown in Table 4. We tested the effects of learned helplessness, perceived stress, resistance to change, and control variables (age, hospitalization frequency, education, and gender) on online health information-seeking behaviors.

The results showed that learned helplessness and resistance to change were negatively associated with online search behavior (b1 = −0.144, t = 2.635; b2 = −0.195, t = 2.798) while perceived stress was positively associated with online search behavior (b3 = 0.273, t = 4.300), thereby lending support to H1, H2, and H3. Finally, the control variables had insignificant effects on online health information-seeking behaviors (b4 = 0.079, t = 1.216; b5 = 0.103, t = 1.694; b6 = −0.041, t = 0.672; b7 = 0.105, t = 1.060).

Our low R-squared values can be explained by three aspects. First, due to the difficulty in predicting individual behavior, low values are common in social science research that predicts human behavior (online health information seeking) and widely exist in previous studies (Agarwal and Prasad, 1999). Second, our study starts from the dilemma discharged patients face in order to investigate their online health information-seeking behaviors, which is inherently complex, so a lower R-squared value is also acceptable. Finally, there are

significant differences in the factors that affect the online health information-seeking behaviors of discharged patients, and we did not consider them based solely on individuals' predicament.

3.3 Multigroup analysis

Multigroup analysis (MGA) is an efficient approach to evaluate moderation using partial least squares path modeling (PLSPM), as recent research has shown (Cheah et al., 2020). We tested the hypotheses by comparing whether individuals lived with their children. Following Keil et al. (2000), we statistically compared the corresponding path coefficients and computed the p-value (see Table 5). The p-value suggests that H5 is supported, but H4 and H6 are not.

4 Discussion

The overall goal of this study was to explore the factors influencing online health information-seeking behaviors, especially the inhibitors, rooted in the dilemmas of discharged patients. We presented a theoretical model based on the dual-factor model, testing the positive (perceived stress) and negative (learned helplessness and resistance to change) impact on online health information-seeking behaviors. As living with children can alter discharged patients' health-management behaviors (Yu et al., 2022), such as information-related behaviors (Zhou et al., 2019), we examined living with children as a moderating variable of the direct routes.

4.1 Discussion of findings

The results supported all the directed influencing relationships on online health information-seeking behaviors. However, only the moderating role of living with children in the relationship between resistance to change and online health information-seeking behaviors was supported; the other two moderating hypotheses were rejected by the results.

First, the results demonstrate that perceived stress significantly positively affects the number of online health information-seeking behaviors. When facing the same dilemma, discharged patients with a higher level of perceived stress seek more online health information than patients with a lower level of perceived stress. As an effective and low-cost strategy to relieve health-related stress, online health information-seeking behaviors can comfort stressed discharged patients.

Second, resistance to change and learned helplessness significantly negatively impact the number of online health information-seeking behaviors. As expected, the relationship between resistance to change and online health information-seeking behaviors is significantly

TABLE 2 Reliability and discriminant validity.

Constructs	Cronbach's α	CR	AVE	LH	PS	RTC
LH	0.706	0.864	0.762	**0.873**		
PS	0.837	0.891	0.672	0.233	**0.820**	
RTC	0.915	0.937	0.789	0.178	0.293	**0.888**

The bold diagonal data refer to the square roots of AVE.

TABLE 3 Convergent validity.

	Item loading	OHIB	LH	RTC	PS
OHIB	1.000	**1.000**	−0.127	−0.136	0.220
LH1	0.937	−0.102	**0.937**	0.143	0.261
LH2	0.804	−0.132	**0.804**	0.184	0.115
RTC1	0.832	−0.062	0.157	**0.832**	0.340
RTC2	0.901	−0.109	0.155	**0.901**	0.291
RTC3	0.904	−0.122	0.206	**0.904**	0.273
RTC4	0.914	−0.155	0.129	**0.914**	0.205
PS1	0.834	0.213	0.156	0.241	**0.834**
PS2	0.836	0.160	0.192	0.326	**0.836**
PS3	0.782	0.161	0.204	0.169	**0.782**
PS4	0.826	0.185	0.211	0.230	**0.826**

The bold data refer to the factor loadings of each construct.

TABLE 4 Partial least squares results.

Path	Coefficient (bi)	t-statistics	R Square
LH → OHIB	−0.144	2.635***	0.141
RTC → OHIB	−0.195	2.798***	
PS → OHIB	0.273	4.300***	
AGE → OHIB	0.079	1.216	
HF → OHIB	0.103	1.694	
EDU → OHIB	−0.041	0.672	
GEN → OHIB	0.105	1.060	

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

TABLE 5 Multigroup analysis results.

Path	Path coefficients-difference	p value
LH → OHIB	0.004	0.975
RTC → OHIB	0.284	0.017**
PS → OHIB	0.047	0.689

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

negative, suggesting that discharged patients with resistance to change may rely on existing information-accessing habits instead of seeking information online. Meanwhile, with a negative attitude of learned helplessness toward online health information, discharged patients decrease the number of online health information-seeking behaviors.

Third, the results confirmed that living with children moderates the relationship between resistance to change and online health information-seeking behaviors. Discharged patients living with children demonstrate fewer online health information-seeking behaviors caused by resistance to change than patients who do not live with their children. The other two moderation effects of living with children on direct relationships are insignificant. The following considerations may explain these phenomena. Living with children is an essential social determinant of health-related behaviors, with a high dependence on age (Rattay and von der Lippe, 2020). In this study, our data sample mainly consisted of patients in the age groups of 25–34 and 35–44, which may have resulted in the insignificant moderating

effect of living with children. In addition, although people who live with their children may weaken their positive thinking (Walter and Haun, 2021), a related paper also concluded that parents tend to show positive attitudes toward media (Vittrup et al., 2016), resulting in a less clear role of living with children in online health-information seeking behaviors.

4.2 Theoretical and practical implications

Following dual-factor model, we conduct a theoretical model of influencing online health information-seeking behaviors, especially inhibitors. This study contributed to the existing literature in several ways.

First, we extended the literature by examining the positive relationship between discharged patients' perceived stress and their online health information-seeking behaviors. For the general population, research has already provided support for the claim that perceived stress has an effect on search behaviors (Kugbey et al., 2019; Liu et al., 2021). In our study, we extended the influence to discharged patients based on their health dilemmas.

Second, to give full insights into discharged patients' online information-seeking behaviors, we explored the inhibitors of online health information-seeking behaviors. Following the dilemma of discharged patients' contradiction in previous health information-adoption habits and online-seeking approach, we empirically proved that resistance to change could be an inhibitor of online information-seeking behaviors. Furthermore, as discharged patients experience feelings of helplessness about online seeking results, we found learned helplessness could decrease the number of online health information-seeking behaviors.

Third, this study contributes to understanding how living with children influences discharged patients' online health information-seeking behaviors. This study provides a new viewpoint on the moderating effect between resistance to change and online health information-seeking behaviors.

This research also offers practical implications in assisting discharged patients in coping with dilemmas for three groups—namely, discharged patients themselves, online health information service providers, and family members.

First, our research analyzed the real dilemma of discharged patients, offering insights into understanding the reasons why they cannot escape these dilemmas using the available low-cost online health information services. Although discharged patients have limited health resources from traditional medical systems, they might not search for alternative online health-related information services due to their previous information adoption habits and feelings of helplessness related to online seeking results.

Second, this study's results lead to suggestions for online health information service providers to adjust the online content and information delivery ways to achieve available and appropriate services to assist discharged patients. These providers could link online services with traditional hospitals to reduce patients' resistance to online health information service adoption. At the same time, the provider of online health information should enhance the credibility and personality of such information to reduce feelings of learned helplessness. By following these suggestions, some of the inhibitors of online health information-seeking behaviors can be removed.

Third, living arrangements could provide an essential role for discharged patients to seek health information online. This study found that discharged patients who live with their children might have a weakened relationship between discharged patients' characteristics and online health information-seeking behaviors, illustrating the critical role of living with children in changing the online-seeking behaviors generated by a resistant attitude.

4.3 Limitations and potential for future study

Despite this study's theoretical and practical implications, we should also address several limitations of the study. First, our study used online survey data to test the facilitators and inhibitors of online health information-seeking behaviors for discharged patients. The reason for collecting data online was that the cooperating hospital could not support survey distribution during the COVID-19 pandemic in China. The research sample was mainly between the ages of 25–34 and 35–44. However, older adults are an essential part of discharged patients. In our future research, we will explore the inhibitors of online health information-seeking behaviors among older adult patients after discharge.

Second, the research sample of this study was patients with multiple health conditions who had been hospitalized and discharged. However, different diseases may result in different health conditions and self-management behaviors for discharged patients. In our future research, we will analyze and examine discharged patients with specific diseases, such as chronic diseases, to achieve more detailed insights into these disease characteristics.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

FL: Writing – review & editing. XK: Writing – original draft. TX: Writing – original draft. HG: Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agarwal, R., and Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decis. Sci.* 30, 361–391. doi: 10.1111/j.1540-5915.1999.tb01614.x

Anderson, J. C., and Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychol. Bull.* 103, 411–423. doi: 10.1037/0033-2909.103.3.411

Basch, C. H., Mac Lean, S. A., Romero, R.-A., and Ethan, D. (2018). Health information seeking behavior among college students. *J. Community Health* 43, 1094–1099. doi: 10.1007/s10900-018-0526-9

Baum, A., Fleming, R., and Reddy, D. M. (1986). Unemployment stress: loss of control, reactance and learned helplessness. *Soc. Sci. Med.* 22, 509–516. doi: 10.1016/0277-9536(86)90016-x

Bell, R. A., Hu, X., Orrange, S. E., and Kravitz, R. L. (2011). Lingering questions and doubts: online information-seeking of support forum members following their medical visits. *Patient Educ. Couns.* 85, 525–528. doi: 10.1016/j.pec.2011.01.015

Benko, E., Babnik, K., and von Humboldt, S. (2021). "Learned helplessness" in Encyclopedia of gerontology and population aging. eds. D. Gu and M. E. Dupre (Cham: Springer).

Berle, D., Starcevic, V., Khazaal, Y., Viswasam, K., Hede, V., and McMullan, R. D. (2020). Relationships between online health information seeking and psychopathology. *Gen. Hosp. Psychiatry* 62, 96–97. doi: 10.1016/j.genhosppsych.2019.04.006

Bhattacherjee, A., and Hikmet, N. (2007). Physicians' resistance toward healthcare information technology: a theoretical model and empirical test. *Eur. J. Inf. Syst.* 16, 725–737. doi: 10.1057/palgrave.ejis.3000717

Cao, W., Zhang, X., Xu, K., and Wang, Y. (2016). Modeling online health information-seeking behavior in China: the roles of source characteristics, reward assessment, and internet self-efficacy. *Health Commun.* 31, 1105–1114. doi: 10.1080/10410236.2015.1045236

Cenfetelli, R. T. (2004). Inhibitors and enablers as dual factor concepts in technology usage. *J. Assoc. Inf. Syst.* 5, 472–492. doi: 10.17705/1jais.00059

Chang, C. C., and Huang, M. H. (2020). Antecedents predicting health information seeking: a systematic review and meta-analysis. *Int. J. Inf. Manag.* 54:102115. doi: 10.1016/j.ijinfomgt.2020.102115

Cheah, J. H., Thurasamy, R., Memon, M. A., Chuah, F., and Ting, H. (2020). Multigroup analysis using smart PLS: step-by-step guidelines for business research. *Asian. J. Bus. Res.* 10:I-XIX. doi: 10.14707/ajbr.200087

Chen, J., Duan, Y., Xia, H., Xiao, R., Cai, T., and Yuan, C. (2025). Online health information seeking behavior among breast cancer patients and survivors: a scoping review. *BMC Womens Health* 25:1. doi: 10.1186/s12905-024-03509-x

Chin, W. W. (1998). Commentary: issues and opinion on structural equation modeling. *MIS Q.* 22, vii–xvi. doi: 10.2307/249674

Cline, R. J., and Haynes, K. M. (2001). Consumer health information seeking on the internet: the state of the art. *Health Educ. Res.* 16, 671–692. doi: 10.1093/her/16.6.671

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *J. Health Soc. Behav.* 24, 385–396. doi: 10.2307/2136404

De Looper, M., van Weert, J. C., Schouten, B. C., Bolle, S., Belgers, E. H., Eddes, E. H., et al. (2021). The influence of online health information seeking before a consultation on anxiety, satisfaction, and information recall, mediated by patient participation: field study. *J. Med. Internet Res.* 23:e23670. doi: 10.2196/23670

Deng, Z., Mo, X., and Liu, S. (2014). Comparison of the middle-aged and older users' adoption of mobile health services in China. *Int. J. Med. Inform.* 83, 210–224. doi: 10.1016/j.ijmedinf.2013.12.002

Djulbegovic, B., Hozo, I., and Greenland, S. (2011). "Uncertainty in clinical medicine" in Philosophy of medicine, vol. 16, 299–356.

Eibergen, L., Janssen, M. J., Blom, L., and Karapinar-Çarkit, F. (2018). Informational needs and recall of in-hospital medication changes of recently discharged patients. *Res. Soc. Adm. Pharm.* 14, 146–152. doi: 10.1016/j.sapharm.2017.01.006

Fassett-Carman, A., Hankin, B. L., and Snyder, H. R. (2019). Appraisals of dependent stressor controllability and severity are associated with depression and anxiety symptoms in youth. *Anxiety Stress Coping* 32, 32–49. doi: 10.1080/10615806.2018.1532504

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.1177/002224378101800104

Gefen, D., and Straub, D. (2005). A practical guide to factorial validity using PLS-graph: tutorial and annotated example. *Commun. Assoc. Inf. Syst.* 16:5. doi: 10.17705/1CAIS.01605

Gulec, H., Kvardova, N., and Smahel, D. (2022). Adolescents' disease-and fitness-related online health information seeking behaviors: the roles of perceived trust in online health information, eHealth literacy, and parental factors. *Comput. Hum. Behav.* 134:107318. doi: 10.1016/j.chb.2022.107318

Guo, X., Sun, Y., Wang, N., Peng, Z., and Yan, Z. (2013). The dark side of elderly acceptance of preventive mobile health services in China. *Electron. Mark.* 23, 49–61. doi: 10.1007/s12525-012-0112-4

Hair, J. F., Ringle, C. M., and Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *J. Mark. Theory Pract.* 19, 139–152. doi: 10.2753/MTP1069-6679190202

Hsieh, P. J. (2016). An empirical investigation of patients' acceptance and resistance toward the health cloud: the dual factor perspective. *Comput. Hum. Behav.* 63, 959–969. doi: 10.1016/j.chb.2016.06.029

Hsu, C. L., and Lin, J. C. C. (2008). Acceptance of blog usage: the roles of technology acceptance, social influence and knowledge sharing motivation. *Inf. Manag.* 45, 65–74. doi: 10.1016/j.im.2007.11.001

Irani, E., Moore, S. E., Hickman, R. L., Dolansky, M. A., Josephson, R. A., and Hughes, J. W. (2019). The contribution of living arrangements, social support, and self-efficacy to self-management behaviors among individuals with heart failure: a path analysis. *J. Cardiovasc. Nurs.* 34, 319–326. doi: 10.1097/JCN.0000000000000581

Jiang, S., and Street, R. L. (2017). Pathway linking internet health information seeking to better health: a moderated mediation study. *Health Commun.* 32, 1024–1031. doi: 10.1080/10410236.2016.1196514

Keil, M., Tan, B. C., Wei, K. K., Saarinen, T., Tuunainen, V., and Wassenaar, A. (2000). A cross-cultural study on escalation of commitment behavior in software projects. *MIS Q.* 24:299. doi: 10.2307/3250940

Kim, K., Shin, S., Kim, S., and Lee, E. (2023). The relation between eHealth literacy and health-related behaviors: systematic review and meta-analysis. *J. Med. Internet Res.* 25:e40778. doi: 10.2196/40778

Kitchens, B., Harle, C. A., and Li, S. (2014). Quality of health-related online search results. *Decis. Support. Syst.* 57, 454–462. doi: 10.1016/j.dss.2012.10.050

Kugbey, N., Meyer-Weitz, A., and Asante, K. O. (2019). Access to health information, health literacy and health-related quality of life among women living with breast cancer: depression and anxiety as mediators. *Patient Educ. Couns.* 102, 1357–1363. doi: 10.1016/j.pec.2019.02.014

Lambert, S. D., and Loiselle, C. G. (2007). Health information—seeking behavior. *Qual. Health Res.* 17, 1006–1019. doi: 10.1177/1049732307305199

Laumer, S., Maier, C., Eckhardt, A., and Weitzel, T. (2016). Work routines as an object of resistance during information systems implementations: theoretical foundation and empirical evidence. *Eur. J. Inf. Syst.* 25, 317–343. doi: 10.1057/ejis.2016.1

Liu, S., Lithopoulos, A., Zhang, C. Q., Garcia-Barrera, M. A., and Rhodes, R. E. (2021). Personality and perceived stress during COVID-19 pandemic: testing the mediating role of perceived threat and efficacy. *Personal. Individ. Differ.* 168:110351. doi: 10.1016/j.paid.2020.110351

López Steinmetz, L. C., Leyes, C. A., Dutto Florio, M. A., Fong, S. B., López Steinmetz, R. L., and Godoy, J. C. (2021). Mental health impacts in Argentinean college students during COVID-19 quarantine. *Front. Psych.* 12:557880. doi: 10.3389/fpsyt.2021.557880

Lowry, P. B., and Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: when to choose it and how to use it. *IEEE Trans. Prof. Commun.* 57, 123–146. doi: 10.1109/TPC.2014.2312452

Maier, S. F. (1980). Learned helplessness and the schedule-shift hypotheses. *J. Res. Pers.* 14, 170–186. doi: 10.1016/0092-6566(80)90025-2

Matsumoto, M., Inoue, K., Kashima, S., and Takeuchi, K. (2012). Does the insufficient supply of physicians worsen their urban-rural distribution? A Hiroshima-Nagasaki comparison. *Rural Remote Health* 12, 228–236. doi: 10.3316/informit.346237890546324

McMullan, R. D., Berle, D., Arnáez, S., and Starcevic, V. (2019). The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. *J. Affect. Disord.* 245, 270–278. doi: 10.1016/j.jad.2018.11.037

Meng, F., Guo, X., Peng, Z., Zhang, X., and Vogel, D. (2019). The routine use of mobile health services in the presence of health consciousness. *Electron. Commer. Res. Appl.* 35:100847. doi: 10.1016/j.elerap.2019.100847

Miller, L. M. S., and Bell, R. A. (2012). Online health information seeking: the influence of age, information trustworthiness, and search challenges. *J. Aging Health* 24, 525–541. doi: 10.1177/0898264311428167

Musarezaie, N., Samouei, R., Shahrzadi, L., and Ashrafi-Rizi, H. (2019). Prediction of health information-seeking behavior components based on health anxiety among users of public libraries. *J. Educ. Health Promot.* 8:227. doi: 10.4103/jehp.jehp_262_19

Musial, A., Butts, B., Loechtenfeldt, A., Herrmann, L. E., Schmidlin, E., Kelley, J., et al. (2020). Challenges following hospital discharge for children with medical complexity. *Hosp. Pediatr.* 10, 531–536. doi: 10.1542/hpeds.2019-0306

Nangsangna, R. D., and Vroom, F. D. C. (2019). Factors influencing online health information seeking behaviour among patients in Kwahu west municipal, Nkawkaw, Ghana. *Online J. Public Health Inform.* 11:e13. doi: 10.5210/ojphi.v11i2.10141

Nikoloudakis, I. A., Vandelanotte, C., Rebar, A. L., Schoeppe, S., Alley, S., Duncan, M. J., et al. (2018). Examining the correlates of online health information–seeking behavior among men compared with women. *Am. J. Mens Health* 12, 1358–1367. doi: 10.1177/1557988316650625

Polites, G. L., and Karahanna, E. (2012). Shackled to the status quo: the inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. *MIS Q.* 36, 21–42. doi: 10.25300/MISQ/2012/36.1.02

Pourrazavi, S., Kouzekanani, K., Asghari Jafarabadi, M., Bazargan-Hejazi, S., Hashemiparast, M., and Allahverdipour, H. (2022). Correlates of older adults' E-health information-seeking behaviors. *Gerontology* 68, 935–942. doi: 10.1159/000521251

Quinless, F. W., and Nelson, M. A. M. (1988). Development of a measure of learned helplessness. *Nurs. Res.* 37, 11–15. doi: 10.1097/00006199-198801000-00003

Rains, S. A., and Tukachinsky, R. (2015). An examination of the relationships among uncertainty, appraisal, and information-seeking behavior proposed in uncertainty management theory. *Health Commun.* 30, 339–349. doi: 10.1080/10410236.2013.858285

Rattay, P., and von der Lippe, E. (2020). Association between living with children and the health and health behavior of women and men. Are there differences by age? Results of the "German health update" (GEDA) study. *Int. J. Environ. Res. Public Health* 17:3180. doi: 10.3390/ijerph17093180

Reifegerste, D., Czerwinski, F., Rosset, M., Baumann, E., Kludt, E., and Weg-Remers, S. (2019). Demographic and cancer-related differences between self-seeking patients and supported patients: analysis of cancer information–service data. *Psycho-Oncology* 28, 759–766. doi: 10.1002/pon.5016

Sedrak, M. S., Soto-Perez-De-Celis, E., Nelson, R. A., Liu, J., Waring, M. E., Lane, D. S., et al. (2020). Online health information–seeking among older women with chronic illness: analysis of the women's health initiative. *J. Med. Internet Res.* 22:e15906. doi: 10.2196/15906

Shahbaz, M., Gao, C., Zhai, L., Shahzad, F., and Arshad, M. R. (2020). Moderating effects of gender and resistance to change on the adoption of big data analytics in healthcare. *Complexity* 2020, 1–13. doi: 10.1155/2020/2173765

Shamlou, Z., Saberi, M. K., and Amiri, M. R. (2022). Application of theory of planned behavior in identifying factors affecting online health information seeking intention and behavior of women. *Aslib J. Inf. Manag.* 74, 727–744. doi: 10.1108/AJIM-07-2021-0209

Soroya, S. H., Farooq, A., Mahmood, K., Isoaho, J., and Zara, S. E. (2021). From information seeking to information avoidance: understanding the health information behavior during a global health crisis. *Inf. Process. Manag.* 58:102440. doi: 10.1016/j.ipm.2020.102440

Stroebe, W., and Diehl, M. (1988). When social support fails: supporter characteristics in compliance-induced attitude change. *Personal. Soc. Psychol. Bull.* 14, 136–144. doi: 10.1177/0146167288141014

Trindade, I. A., Mendes, A. L., and Ferreira, N. B. (2020). The moderating effect of psychological flexibility on the link between learned helplessness and depression symptomatology: a preliminary study. *J. Contextual Behav. Sci.* 15, 68–72. doi: 10.1016/j.jcbs.2019.12.001

Tsai, J. M., Cheng, M. J., Tsai, H. H., Hung, S. W., and Chen, Y. L. (2019). Acceptance and resistance of telehealth: the perspective of dual-factor concepts in technology adoption. *Int. J. Inf. Manag.* 49, 34–44. doi: 10.1016/j.ijinfomgt.2019.03.003

Vittrup, B., Snider, S., Rose, K. K., and Rippy, J. (2016). Parental perceptions of the role of media and technology in their young children's lives. *J. Early Child. Res.* 14, 43–54. doi: 10.1177/1476718X14523749

Walter, F., Carr, M. J., Mok, P. L., Antonsen, S., Pedersen, C. B., Appleby, L., et al. (2019). Multiple adverse outcomes following first discharge from inpatient psychiatric care: a national cohort study. *Lancet Psychiatry* 6, 582–589. doi: 10.1016/S2215-0366(19) 30180-4

Walter, J., and Haun, V. C. (2021). Positive and negative work reflection, engagement and exhaustion in dual-earner couples: exploring living with children and work-linkage as moderators. *German J. Hum. Resour. Manag.* 35, 249–273. doi: 10.1177/2397002220 9649

Wanke, N., and Schwabe, L. (2020). Subjective uncontrollability over aversive events reduces working memory performance and related large-scale network interactions. *Cereb. Cortex* 30, 3116–3129. doi: 10.1093/cercor/bhz298

Weber, W., Reinhardt, A., and Rossmann, C. (2020). Lifestyle segmentation to explain the online health information–seeking behavior of older adults: representative telephone survey. *J. Med. Internet Res.* 22:e15099. doi: 10.2196/15099

Wong, D. K. K., and Cheung, M. K. (2019). Online health information seeking and eHealth literacy among patients attending a primary care clinic in Hong Kong: a cross-sectional survey. *J. Med. Internet Res.* 21:e10831. doi: 10.2196/10831

Xiao, N., Sharman, R., Rao, H. R., and Upadhyaya, S. (2014). Factors influencing online health information search: an empirical analysis of a national cancer-related survey. *Decis. Support. Syst.* 57, 417–427. doi: 10.1016/j.dss.2012.10.047

Xue, S., Gu, Q., Zhu, K., and Jiang, J. (2023). Self-compassion buffers the impact of learned helplessness on adverse mental health during COVID-19 lockdown. *J. Affect. Disord.* 327, 285–291. doi: 10.1016/j.jad.2023.01.099

Yan, L., Gan, Y., Ding, X., Wu, J., and Duan, H. (2021). The relationship between perceived stress and emotional distress during the COVID-19 outbreak: effects of boredom proneness and coping style. *J. Anxiety Disord.* 77:102328. doi: 10.1016/j. janxdis.2020.102328

Ying, T., Tang, J., Wen, J., Ye, S., Zhou, Y., and Li, F. S. (2021). Traveling with pets: constraints, negotiation, and learned helplessness. *Tour. Manag.* 82:104183. doi: 10.1016/j.tourman.2020.104183

Yu, Y., Lv, J., Liu, J., Chen, Y., Chen, K., and Yang, Y. (2022). Association between living arrangements and cognitive decline in older adults: a nationally representative longitudinal study in China. *BMC Geriatr.* 22:843. doi: 10.1186/s12877-022-03473-x

Zhang, L., Qin, Y., and Li, P. (2020). Media complementarity and health information acquisition: a cross-sectional analysis of the 2017 HINTS-China survey. *J. Health Commun.* 25, 291–300. doi: 10.1080/10810730.2020.1746868

Zhang, X., Wen, D., Liang, J., and Lei, J. (2017). How the public uses social media wechat to obtain health information in China: a survey study. *BMC Med. Inform. Decis. Mak.* 17, 66–79. doi: 10.1186/s12911-017-0470-0

Zhao, X., Fan, J., Basnyat, I., and Hu, B. (2020). Online health information seeking using "# COVID-19 patient seeking help" on Weibo in Wuhan, China: descriptive study. *J. Med. Internet Res.* 22:e22910. doi: 10.2196/22910

Zhao, J., Han, H., Zhong, B., Xie, W., Chen, Y., and Zhi, M. (2021). Health information on social media helps mitigate Crohn's disease symptoms and improves patients' clinical course. *Comput. Hum. Behav.* 115:106588. doi: 10.1016/j.chb.2020.106588

Zhou, Z., Cai, L., Zhuang, M., Hong, Y. A., and Fang, Y. (2019). A longitudinal analysis of the association between the living arrangements and psychological well-being of older Chinese adults: the role of income sources. *BMC Geriatr.* 19, 347–349. doi: 10.1186/ s12877-019-1371-0

*CORRESPONDENCE
Weitao Liang
✉ 1023444366@qq.com
Zhong Wu
✉ wuzhong71@scu.edu.cn

†These authors have contributed equally to
this work and share first authorship

Cross-platform analysis of atrial fibrillation scientific videos: using composite index and a basic assessment scale

Chong Luo[1†], Xiaoli Qin[1†], Xiaoyu Xie[2†], Jie Gao[3], Yuwei Wu[4], Weitao Liang[1]* and Zhong Wu[1]*

[1]Department of Cardiovascular Surgery, West China Hospital, Sichuan University, Chengdu, Sichuan, China, [2]Department of Anesthesia, West China Hospital, Sichuan University, Chengdu, Sichuan, China, [3]West China School of Medicine, Sichuan University, Chengdu, Sichuan, China, [4]West China School of Public Health, Sichuan University, Chengdu, Sichuan, China

Background: Currently, video platforms were filled with many low-quality, uncensored scientific videos, and patients who utilize the Internet to gain knowledge about specific diseases are vulnerable to being misled and possibly delaying treatment as a result. Therefore, a large sample survey on the content quality and popularity of online scientific videos was of great significance for future targeted reforms.

Objective: This study utilized normalization data analysis methods and a basic assessment scale, providing a new aspect for future research across multiple platforms with large sample sizes and for the development of video content quality assessment scales.

Methods: This cross-sectional study analyzed a sample of 331 videos retrieved from YouTube, BiliBili, TikTok, and Douyin on June 13, 2024. In the analysis of atrial fibrillation scientific videos across four social media platforms, comprehensive metrics and a basic scoring scale revealed associations between platforms, creators, and the popularity and content quality of the videos. Data analysis employed principal component analysis, normalization data processing, non-parametric tests, paired t-tests, and negative binomial regression.

Results: Analysis of the user engagement data using a composite index revealed a significant difference in the popularity of videos from publishers with a medical background ($z = -4.285$, $p < 0.001$), no aforementioned findings were found among video platforms, however, except for the Bilibili platform. As for content quality, while the difference in the total number of videos between the two groups was almost 2-fold (229:102), the difference in qualified videos was only 1.47-fold (47:32), a ratio that was even more unbalanced among the top 30% of videos with the most popularity. Notably, the overall content quality of videos from publishers without a medical background was also significantly higher ($z = -2.299$, $p = 0.02$).

Conclusion: This analysis of atrial fibrillation information on multiple social media platforms found that people prefer videos from publishers with a medical background. However, it appeared that these publishers did not sufficiently create high-quality, suitable videos for the public, and the platforms seemed to lack a rigorous censorship system and policy support for high-quality content. Moreover, the normalized data processing method and the basic assessment scale that we attempted to use in this study provided new ideas for future large-sample surveys and content quality review.

1 Introduction

Atrial fibrillation (AF) is the most common clinical arrhythmia, affecting approximately 59 million individuals worldwide (1). Patients with AF face a 2-fold increased risk of myocardial infarction (2) and a 5-fold increased risk of heart failure (3). Additionally, they have a 5-fold higher risk of ischemic stroke, which rises to 20-fold in the presence of mitral stenosis (4, 5). According to Lin et al., 28.6% of AF patients are asymptomatic and often hospitalized for other primary conditions (6). Therefore, educating patients about the health risks associated with AF may improve adherence to treatment more effectively than focusing solely on medication dosages and routine follow-ups.

Traditional medical education methods, such as posters and manuals, are often limited in their reach due to accessibility and geographic constraints (7). This restricts the dissemination of disease-related information to both chronic disease patients and the public (6, 8, 9). Today, with the widespread use of smartphones and personal computers, individuals can access health information anytime and anywhere. This shift has made online video platforms a powerful tool for distributing medical education and raising awareness (10–12). Many healthcare institutions on platforms like YouTube have obtained certification and are actively publishing videos (13). Lee et al. found that 40.8% of U.S. adults use YouTube to search for health-related information, leading to a 30% increase in physical activity (14). However, some meta-analysis of educational videos on online platforms revealed that there is still considerable room for improvement, particularly in terms of content accuracy and regulatory oversight on platforms like YouTube (15–17). Platforms demonstrate inadequate content moderation for user-uploaded videos, enabling virtually any user to disseminate disease-related articles or videos without rigorous vetting. Coupled with social media's intrinsic real-time nature that facilitates viral propagation speeds, misinformation, unverified claims, and pseudoscientific content become freely accessible through platform search algorithms (15, 18, 19), The insufficient creator-audience interaction diminishes the signature dialogic nature of social media platforms (20), effectively rendering them functionally equivalent to traditional didactic health broadcasts.

Current research on the quality of educational videos predominantly consists of single-platform, single-disease studies with small sample sizes and video quality evaluations rely on horizontal comparisons of isolated unidimensional metrics (e.g., view counts, likes), without in-depth exploration of inter-indicator correlations. There is a notable lack of cross-platform quality comparative analyses and disease-specific quality assessment tools (depending on generic scales like DISCERN or JAMA). In this study, AF educational videos were retrieved from four platforms: YouTube, BiliBili, TikTok, and Douyin. We experimentally applied a normalization method to unify interaction data from different platforms for statistical analysis and used a 'basic and essential' professional rating scale to assess the content quality of sampled videos.

This study provided a new perspective and method for future large-scale data analysis and content quality evaluation across different platforms.

2 Methods

2.1 Study design

This study was a cross-sectional analysis. On June 13, 2024, we performed keyword searches for "Atrial Fibrillation," "AF," and "AF + Management" across YouTube, BiliBili, TikTok, and Douyin platforms. For each video identified, we recorded the URL, the number of views, likes, and the number of comments and replies on the same day to mitigate potential changes over time. The content quality of the videos was then evaluated over the subsequent month.

2.2 Measures

The keyword searches were independently conducted by two reviewers using a web browser with a cleared cache. The top 50 results for each search term were selected. Exclusion criteria included duplicate videos, audio-only content, videos with titles that did not match the content, and videos in languages other than English or Chinese. In cases where the two reviewers disagreed, a third reviewer was consulted to make the final decision. After applying these criteria, a total of 331 videos were included in the analysis.

2.3 Data collection

2.3.1 Interaction metrics

The interaction metrics collected for each video included: view count, number of likes, comments, and replies. To avoid potential bias, we matched 1:1 after data collection in rows stratified by a reply/comment ratio of 3:1. Metrics between different platforms were examined using Kruskal-Wallis tests (Table 1A). Principal Component Analysis (PCA) and weighted scoring were employed to preprocess interaction metrics data [Kaiser-Meyer-Olkin (KMO) = 0.741; and Bartlett's test of sphericity < 0.001]. Based on the PC1 loading, the Heatscore, a composite index, was obtained using a normalization method (Figure 1; Table 1B).

2.3.2 Specialized metrics

The DISCERN and JAMA scales have been utilized to evaluate the accuracy of information sources. However, the aforementioned scales are capable of assessing the reliability of the information sources present in the video but offer a limited capacity for the evaluation of content richness on a professional area. This observation is consistent with the findings of previous studies that have expressed reservations regarding the application of the DISCERN and JAMA scales (21–23), Therefore, the classification criteria developed by clinicians were used in this study. A scoring system, the Atrial Fibrillation Specific Score (AFSS), was

TABLE 1A Interaction metrics categorized by platforms.

Platform	Views	Likes	Comments	Replies
Bilibili (n = 100)	1557.0 (456.5,8927.5)	17.5 (7.0,172.5)	2.0 (0.0,15.8)	0.0 (0.0,0.0)
Tiktok (n = 79)	41200.0 (16600.0,124300.0)	1369.0 (344.0,3469.0)	44.0 (19.0,165.0)	1.0 (0.0,4.0)
YouTube (n = 49)	40825.0 (9285.5,173419.5)	485.0 (92.5,1910.5)	11.0 (1.0,81.5)	0.0 (0.0,0.0)
Douyin (n = 103)	27822.3 (8464.5,62007.8)	1628.0 (517.0,3623.0)	117.0 (39.0,266.0)	0.0 (0.0,0.0)
p value	0.000**	0.000**	0.000**	0.000**

*$p < 0.05$; **$p < 0.01$; Kruskal-Wallis test.
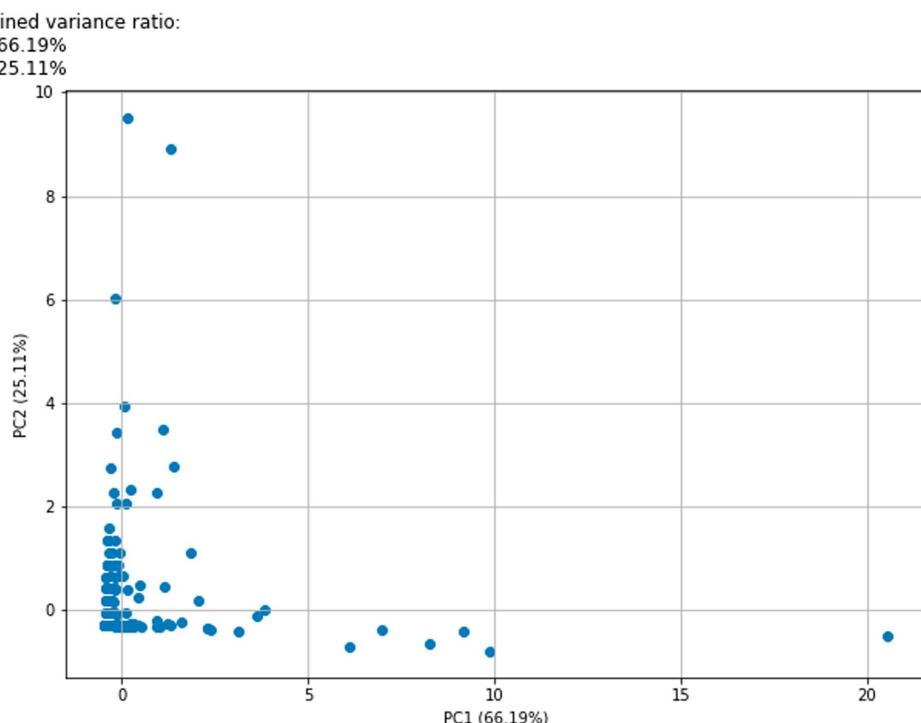
FIGURE 1
Interaction metrics PCA 2D scatter plot. The horizontal and vertical coordinate are the principal component (PC1,2) axis and principal component contributions. PC1 explains 66.19% of the variance in the dataset. Specific data for each principal component loading can be found in Table 1B. Blue dots represent each sample included in the analysis.

TABLE 1B PCA loadings.

	PC1	PC2	PC3	PC4
Views	0.57	0.05	0.79	0.22
Likes	0.59	0.03	0.20	0.78
Comments	0.58	0.06	0.57	0.58
Replies	0.02	1.00	0.08	0.00

designed based on existing literature and expert guidelines (24, 25). However, using such complex and precise scales to guide the creation of educational videos may result in content that is too advanced for viewers to comprehend, limiting their ability to engage in shared decision-making (26). In the course of routine treatment, it has been observed that patients tend to demonstrate a limited interest in the pathogenesis, preferring instead to inquire about the treatment options and the distinctions between them. However, during outpatient consultations, patients frequently encounter challenges in comprehending and assimilating a substantial volume of disease-related information within

a limited timeframe. This observation underscores the imperative for the development of a scoring scale that quantifies the quality of popularized video content, with the objective of enhancing public comprehension of medical information. On the one hand, it can effectively promote the platform's content review, and on the other hand, we hope that video authors can produce videos based on this scale to provide a good way for the public to understand and absorb disease-related knowledge before and after the treatment. To address this, we developed the Essential Score (Escore) as a simplified version of the AFSS framework. We have attempted to remove from the AFSS framework about diagnosis and management of the disease that may require a medical background to understand, and to focus on knowledge that is of more interest to the general public, such as the course of the disease, its associated dangers, methods for self-diagnosis, and available treatment options. Five specialized cardiac surgeons from our hospital were consulted to evaluate the content validity of both the AFSS and Escore rating scales. After multiple rounds of adjustments and assessments, the final criteria are presented in Table 2. Each video was scored for content quality independently by three authors, or by the corresponding author if the

TABLE 2 Specialized metrics scales.

Diagnose	Treatment	Management
1. Atrail Fibrillation Defination (**Essential**)	1. Anticoagulant/Avoid stroke (**Essential**)	1. Discussion of the indications/contradictions/possible complications of various treatments
2. Risk factors for AF	2. Surgical options such as AF ablation procedures(catheter/surgical)/left atrial appendage (**Essential**)	2. Notifications of postoperative or drug follow-up care
3. Common clinical manifestation (**Essential**)	3. Antiarrhythmic drug therapy (**Essential**)	3. Mentions of AF is a disease that requires long- term treatment and needs patients' own participation in the management of this disease (**Essential**)
4. Findings of a physical examination		4. Attentions to the prevention and control of risk factors of AF
5. Transthoracic echocardiography as diagnostic technique		
6. The consequences of allowing AF to be left untreated (**Essential**)		

Concordance between the three authors: Fleiss Kappa = 0.765. Escore reliability: Cronbach's a = 0.715; Validity: Kaiser-Meyer-Olkin (KMO) = 0.746; and Bartlett's test of sphericity <0.001; Content validity: S-CVI/UA = 0.857, S-CVI/Ave = 0.971.

variation between them was more than three. The reliability, validity, and content validity indexes confirmed that these scales are appropriate for evaluating the content quality of the videos analyzed in this study (Table 2).

In this study, the content quality of the sample videos was evaluated using the Escore with one point awarded for fulfilling one item in the Diagnose and Management section, but slightly different in the Treatment section. With a clear diagnosis of a chronic disease, there is mostly an emphasis on long-term treatment and regular follow-up (27–29), and it is important for patients to have a thorough understanding of the treatments for their condition (30, 31), so a missing item in the Treatment section of the Escore was worth zero points. In video quality grading, treatment videos scoring below 3 were classified as disqualification. Scores of 5–7 and 3–4 were classified as adequacy and eligibility.

2.4 Statistical analysis

All videos were categorized based on publishing platform and author type. Categorical variables were analyzed using frequency and relative frequency, whereas continuous variables were summarized using median values. Python was utilized to perform PCA and calculate the Heatscore index. Pandas and NumPy were employed for data processing and numerical computations, PCA and the calculation of the Heatscore. IBM SPSS Statistics version 26.0 was applied for statistical analysis and the computation of the metrics. According to the Kolmogorov–Smirnov test, the data in this study exhibited non-normal distribution, necessitating the use of non-parametric tests to evaluate differences between groups.

3 Results

Out of the initial pool of 600 videos, 269 were excluded due to duplication, lack of audio, or irrelevance. This left 331 videos for analysis, with 31% from non-medical backgrounds (NMB), such as self-media and public accounts, and 69% from medical backgrounds (MB), including doctors and hospital media.

In the first phase, we analyzed interaction metrics (views, likes, comments, and replies) and specialized metrics (AFSS and Escore) based on video platform and author type. In the interaction metrics section, after processing through PCA and weighted scoring, we found that the Heatscore still showed significant differences among various video platforms. A subsequent Nemenyi post-hoc test revealed that the source of the difference was solely the video data from the Bilibili platform (Figure 2a). The MB group had higher Heatscore than the NMB group (z = −4.285, p < 0.001; Figure 2b).

In terms of specialized metrics, there was no significant difference in AFSS scores between MB and NMB groups, but the Escore was notably higher in the NMB group (z = −2.299, p = 0.02; Figures 3a,b). Further paired t-tests and regression analysis showed that AFSS scores were significantly higher than Escore in two groups (t-statistic = 17.051/ 10.814, p < 0.001/ < 0.001; Table 3A). Using AFSS scores as the independent variable and Escore as the dependent variable in a negative binomial regression analysis, it was found that AFSS scores had a significant positive effect on Escore (Table 3B).

In the second part, it was observed that most adequacy scientific videos (50/57, 88%) originate from Bilibili and YouTube. Concurrently, all video platforms were inundated with a substantial number of disqualified videos (252/331, 76%). When categorized by author type, MB creators produced nearly twice as many videos as NMB creators (229:102). Yet, the proportion of qualified videos (Escore ≥ 3) was only 1.47:1 (47:32), while non-compliant videos were 2.6 times higher in the MB group (182:70; Figures 4a,b). Among high-popularity videos (top 30% in Heatscore), the MB group again had more videos (78:21) but with ratio of disqualified to qualified videos, approximately 2:1.

4 Discussion

4.1 Interaction metrics

In previous similar reports, the data from different platforms were analyzed by directly profiling the number of plays, likes, and replies of the sample video (32–34). However, we recognize that different platforms encompass diverse user bases and content strategies. This variability
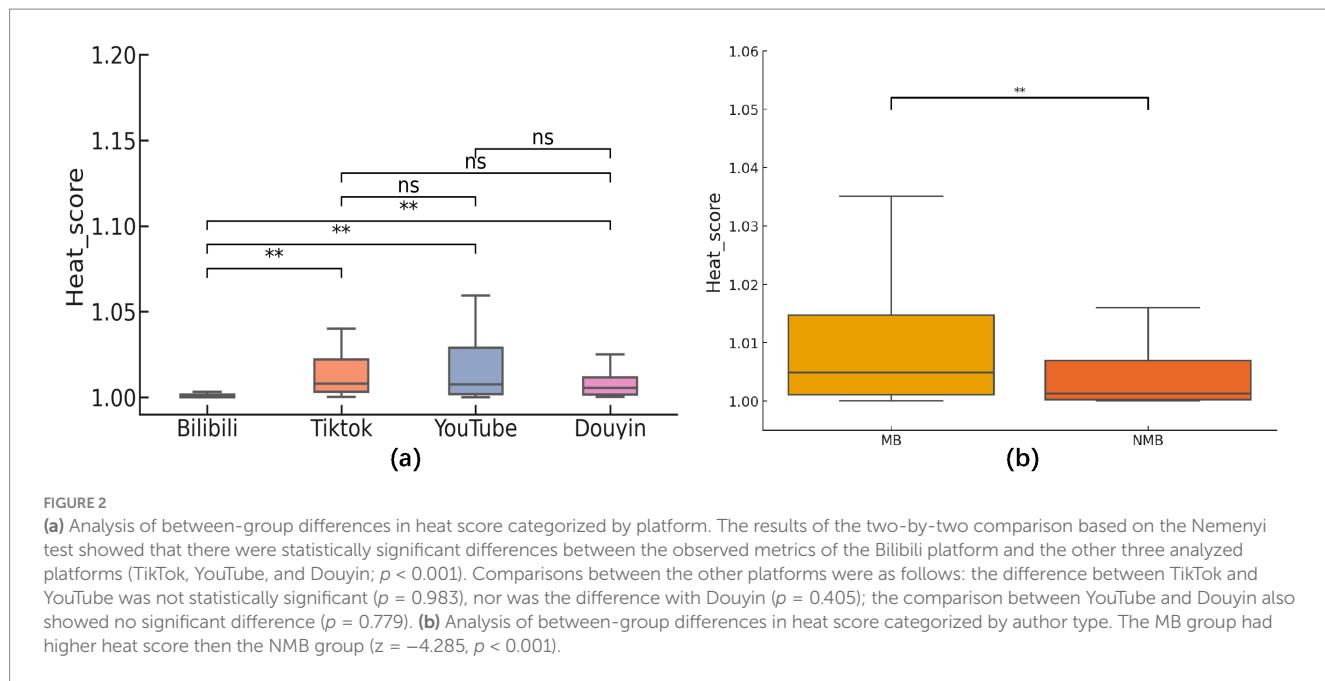
FIGURE 2
(a) Analysis of between-group differences in heat score categorized by platform. The results of the two-by-two comparison based on the Nemenyi test showed that there were statistically significant differences between the observed metrics of the Bilibili platform and the other three analyzed platforms (TikTok, YouTube, and Douyin; $p < 0.001$). Comparisons between the other platforms were as follows: the difference between TikTok and YouTube was not statistically significant ($p = 0.983$), nor was the difference with Douyin ($p = 0.405$); the comparison between YouTube and Douyin also showed no significant difference ($p = 0.779$). **(b)** Analysis of between-group differences in heat score categorized by author type. The MB group had higher heat score then the NMB group ($z = -4.285$, $p < 0.001$).
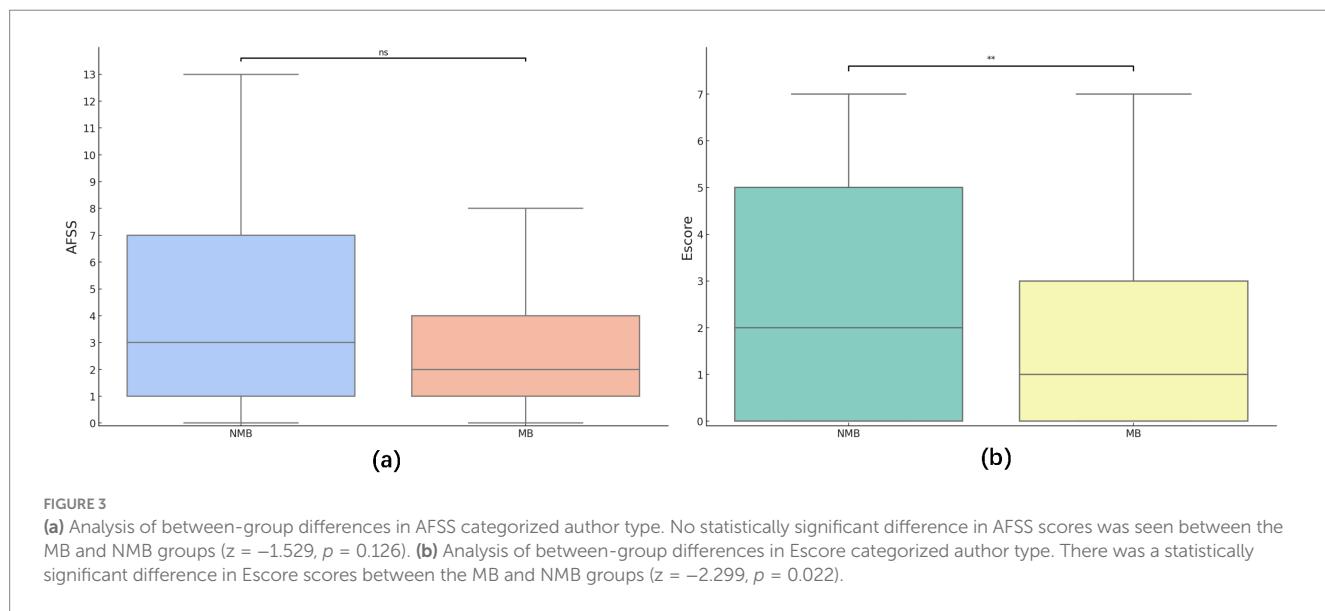
FIGURE 3
(a) Analysis of between-group differences in AFSS categorized author type. No statistically significant difference in AFSS scores was seen between the MB and NMB groups ($z = -1.529$, $p = 0.126$). **(b)** Analysis of between-group differences in Escore categorized author type. There was a statistically significant difference in Escore scores between the MB and NMB groups ($z = -2.299$, $p = 0.022$).

TABLE 3A AFSS and Escore paired *t* test.

Items	Paired (M ± SD)		Mean difference	p	Cohen's d
	AFSS	Escore			
AFSS paired Escore (MB)	2.97 ± 2.28	1.59 ± 1.80	1.38	0.000**	1.024
AFSS paired Escore (NMB)	3.99 ± 3.38	2.44 ± 2.39	1.55	0.000**	1.127

Within the MB and NMB groups, respectively, the variability of the scores of the two rating scales was tested using paired t-tests, and both the AFSS and Escore showed statistically significant differences ($p < 0.01$), with the mean of the AFSS (2.97/3.99) being higher than the mean of the Escore (1.69/2.44). Cohen's d value represents the effect size magnitude, where larger values indicate greater differences between groups. An effect size exceeding the threshold of 0.8 (considered large) demonstrates both statistically and clinically significant divergence, reflecting a substantial discrepancy between the two comparison groups. *$p < 0.05$; **$p < 0.01$.

complicates the interpretation of results, given the disparity in data across platforms (Table 1A). Lookingbil et al. took a randomized permutation test approach to statistically analyze user engagement (i.e., the number of views, likes, comments, and shares) of videos on a single video platform and revealed a significant correlation between the number of publishers' followers, video likes, and the level of user engagement through linear regression, an approach that should be viewed as an attempt at integrative analysis (23). We believe that direct analysis of data

TABLE 3B Negative binomial regression analysis.

Items	Coefficient	Std. Error	z value	p	OR	OR 95% CI	McFadden's R²
MB							
Intercept	−1.093	0.164	−6.649	0.000**	0.335	0.243–0.462	
AFSS	0.398	0.036	11.04	0.000**	1.489	1.387–1.597	0.135
NMB							
Intercept	−0.771	0.233	−3.303	0.000**	0.463	0.293–0.731	
AFSS	0.303	0.037	8.084	0.000**	1.354	1.258–1.457	0.147

Dependent variable: Escore. AFSS in both the MB and NMB groups showed a significant positive influence relationship on Escore scores (Coefficient = 0.398/ 0.303, $p < 0.01$), as well as a 1.489/1.354-fold increase in the change in Escore with a one-point increase in AFSS. OR(MB) = 1.489, 95% CI = 1.387–1.597, OR(NMB) = 1.354, 95% CI = 1.258–1.457. McFadden's R²: 0.135(MB), 0.147(NMB) suggests that the effect of AFSS is statistically significant, but the overall explanatory power of the model is not at the desired level (>0.2), and a larger sample needs to be included in the future for a more in-depth exploration. *$p < 0.05$; **$p < 0.01$.
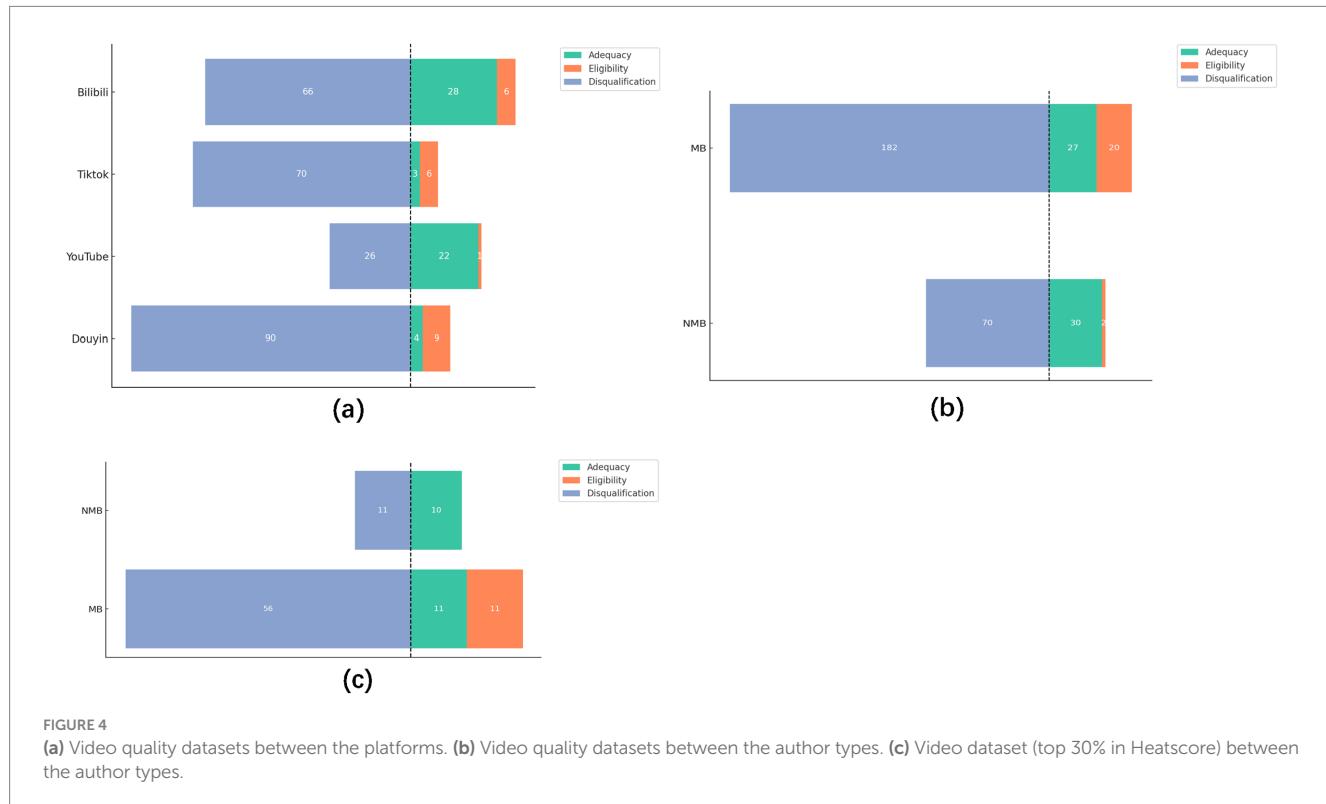
FIGURE 4
(a) Video quality datasets between the platforms. **(b)** Video quality datasets between the author types. **(c)** Video dataset (top 30% in Heatscore) between the author types.

from different platforms can well explain the differences in audience population and policy support of different platforms, but a comprehensive indicator may be needed to make a judgment on the popularity of the video itself. Therefore, our study introduces methods like PCA and normalization techniques for further processing interactive metrics (Table 1B). Importantly, the outcomes of these two statistical approaches yielded notably divergent results (Table 1A; Figure 2a).

The above analysis showed that although YouTube and TikTok outperform Douyin and Bilibili on paper, there is little difference in the Heatscore (popularity) among YouTube, TikTok and Douyin platforms. Bilibili consistently trails behind these platforms in terms of engagement. These findings align closely with similar studies in literature, where videos posted by individuals with medical backgrounds (MB groups) tend to garner higher popularity among viewers (35–37) (Figure 2b).

However, the findings are not cause for celebration. PCA analysis revealed that among the variables examined, the number of comments

and replies accounted for 58 and 2% of the total variance in the raw data, respectively. Notably, the median number of replies across all platforms was zero, indicating a widespread lack of publisher responsiveness to viewer inquiries. Moreover, despite TikTok's highest average view count of 41,200, this figure pales in comparison to the nearly 59 million individuals worldwide affected by atrial fibrillation (AF) (1). This disparity underscores the challenge in effectively reaching and engaging with the expansive AF community through current social media platforms.

4.2 Specialized metrics

Several previous studies have evaluated content quality using established scales such as DISCERN, JAMA, PEMAT-A/V, and GQS (32, 33, 38, 39) alongside self-constructed scales based on disease-specific guidelines (35, 40–43). The former has the advantages that the

scale has been extensively validated, and the reliability of the conclusions is greater, but the applicability is poorer, and the content of the evaluation lacks the specialization of the disease in which it is related, as opposed to the latter, which is the same in reverse. The situation was such that video platforms were overwhelmed with low-quality and misleading content (41, 44, 45). Hence, specialized scales might be more advantageous in the evaluation of the quality of video content for related diseases, and their disadvantages could be ameliorated through multiple validations. In this study, the AFSS scale was created to cover diagnosis, treatment, and management in accordance with the European and US guidelines for the diagnosis and management of AF. Subsequently, we introduced the Escore scale, designed to prioritize essential and fundamental aspects, reflecting a more necessity approach. In this section, this study had two purposes, firstly to investigate the potential use of a "more basic necessity" rating scale. Secondly, to evaluate whether the content quality of the sample videos was satisfactory.

In developing the AFSS scale, it was found that the detailed scoring criteria, while professional enough, could overwhelm viewers seeking concise information from scientific videos. Hence, the Escore scale was crafted to distill essential elements more comprehensibly.

To assess the Escore scale's efficacy, we conducted paired t-tests on groups with MB and NMB. Results showed that Escore were significantly lower than the AFSS, it was able to obtain fewer scores than the AFSS, suggesting that the Escore was not a parallel scale to the AFSS (Table 3A). Moreover, negative binomial regression analyses indicated a strong positive correlation between AFSS and Escore scores (Table 3B). The above results indicated that the "basic necessity" Escore scale maintained a high level of consistency and validity with the "complexity and comprehensiveness" AFSS scale. The similar OR valves suggested that the Escore performance is consistent across the two datasets, with no significant variations, supporting the stability of the Escore. As such, the Escore was qualified and potentially useful as a "more basic necessity" scale.

4.3 Satisfaction with content quality

In recent years, it has become a common way for the public to learn about diseases through online video sites or software (46–48). Tan et al. found that videos featuring medical professionals, highlighted by titles or attire, tend to attract more engagement (49). However, the findings of several studies are not positive about the quality of video content from professional authors (50–52).

Both publishers and platforms were involved in ensuring the quality of the content of the videos uploaded on platforms. This study assessed the publishers' by rating the content quality of the sample videos using the AFSS/ Escore scale. It was noted with surprise that the median AFSS scores were only 2.0 and 3.0 out of a total of 13, and that there was no statistically significant difference in the scores between the two groups (Figure 3a). It indicated that the overall content quality of current streaming videos is low, and that the MB group has a higher Heatscore (popularity), but does not have the favorable conditions to create higher quality videos. On the one hand, the MB group may be more likely to be believed by the general public due to their platform accreditation or with the title of medical practitioners, but the quality of their existing videos is not sufficient to provide adequate information to the general public. On the other hand, regardless of whether or not the publishers have a medical background, the overall quality of the video content is substandard, and the incomplete introduction of the

treatment content related to atrial fibrillation may result in patients having a biased understanding of the treatment, regardless of the medical background of the publisher. This may lead to compliance problems in long-term treatment. These shortcomings were even more prominent in the Escore scores, where it appeared that publishers in the NMB group, who did not have a medical background, were better able to create "more basic and necessary" scientific videos ($z = -2.299$, $p = 0.005$; Figure 3b). It is possible that NMB creators may encounter challenges in comprehending the intricate ECG manifestations, electrophysiological changes and disease management of AF. Consequently, these creators may tend to the production of video content that focuses on more readily comprehensible aspects of the disease, such as its manifestations and therapeutic treatments which are more aligned with the Escore scoring. Conversely, for creators in the MB group, the traditional medical education they received—which may dedicate several hours to comprehensively teaching a single disease—forces them to focus only on partial aspects of a condition when producing videos. This limitation stems from three critical constraints: their ingrained educational paradigms, video duration limits, and personal energy reserves. These factors collectively contribute to their challenges in crafting Escore scale based high-quality videos. However, of greater concern to us was the fact that, like the findings from the interactive metrics, overall, the quality of content in both groups remained low.

Next, we categorized the content quality into Adequacy, Eligibility, and Disqualification according to the Escore score range, aiming to figure out whether the platforms have reviewed the quality of the content. The results are consistent with the concerns expressed in related studies (41, 44, 45). Although platforms have already introduced medical certification measures (53–55). Still, video platforms were overflowing with disqualification videos (252/331, 76%), and there were only 14 more eligible videos in the MB group (1.47-fold) but 2.6-fold more disqualification videos in the NMB group when there were twice as many sample videos as in the NMB group (Figures 4a,b). We analyzed the top 30% of videos in the Heatscore to figure out whether platforms are grading and supporting high-quality videos. It was found that even among the MB group videos with high popularity, there is still a large proportion disqualification video (56/78, 72%). The MB groups had 3.7-fold videos than the NMB group but had 5.1-fold disqualification video while having only 2-fold qualified video (Figure 4c). Combined with the previous finding that the Heatscore of videos in the MB group was significantly higher than that of the NMB (Figure 2b), we assume that the platform did not have a well re-assessment process of the content quality while providing support to publishers with medical backgrounds, resulting in a situation where videos in the MB group were currently high in popularity, but a large number of disqualification videos still existed.

In the context of Internet globalization, social media platforms have been shown to be faster, more convenient and possess unique social attributes in comparison to traditional means of publicity. This provides a vast fertile ground for the widespread dissemination of disease-related knowledge, but it also breeds 'bacteria'. It is incumbent upon platforms to provide creators with guidelines for uploading videos, establish a comprehensive background and content review mechanism to eliminate defective and shoddy works, provide views support for well-produced videos with reliable and detailed content, and dynamically monitor the view data of popular science videos so that searchers can find newer and better-quality videos. For the creators, it is essential to adhere to rigorous standards in medical science video

production, ensuring that treatments are not selectively or utilitarian introduced. Secondly, the content should be presented in a manner that is more accessible to the general public, with a reduction in text and the adoption of simpler forms of expression such as images or animations.

5 Limitation

We only included videos that ranked in the top 50 search results, and it is possible that this strategy does not fully include search terms that may be used by the public. Second, this study attempted to use the Heatscore and Escore, although they were developed based on previous literature and authoritative guidelines and literature, their reliability needs to be explored in subsequent studies.

6 Conclusion

In this study assessing the quality of scientific videos on AF knowledge across different video platforms, with four platforms, it appeared that there was insufficient support for high-quality videos and a lack of a rigorous process for reviewing the quality of the content. Despite having medical backgrounds, creators in the MB group did not consistently produce higher-quality videos. Furthermore, this study introduced a normalization method to analyze data, which revealed significant differences between groups, yielding insights distinct from those obtained through raw data analysis. This methodological innovation presented a new way for future studies with larger sample sizes and across multiple platforms. In terms of content quality evaluation, this study pioneered the validation of a "basic and essential" scoring system, designed to better suit public consumption. This innovative approach offered a fresh perspective for future content reviews on video platforms.

The prevailing tendency among the general public to seek information regarding diseases from online sources has become increasingly pervasive. It is incumbent upon platforms to develop vetting standards, optimize recommendation algorithms and establish dynamic monitoring. Creators should consider forming interdisciplinary teams that integrate physicians (to ensure content authority), media scholars (to refine narrative structure), and visual designers (to achieve cognitive load reduction).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Ethics statement

Ethical approval was not required for the study involving human data in accordance with the local legislation and institutional requirements. Written informed consent was not required, for either participation in the study or for the publication of potentially/indirectly identifying information, in accordance with the local legislation and institutional requirements. The social media data was accessed and analyzed in accordance with the

platform's terms of use and all relevant institutional/national regulations.

Author contributions

CL: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. XQ: Data curation, Formal analysis, Writing – original draft. XX: Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft. JG: Data curation, Investigation, Writing – original draft. YW: Data curation, Formal analysis, Investigation, Writing – original draft. WL: Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing. ZW: Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2025.1507776/full#supplementary-material

References

1. Roth Gregory A, Mensah George A, Johnson Catherine O. Global burden of cardiovascular diseases and risk factors, 1990–2019. *J Am Coll Cardiol*. (2020) 76:2982–3021. doi: 10.1016/j.jacc.2020.11.010

2. Frederiksen TC, Dahm CC, Preis SR, Lin H, Trinquart L, Benjamin EJ, et al. The bidirectional association between atrial fibrillation and myocardial infarction. *Nat Rev Cardiol*. (2023) 20:631–44. doi: 10.1038/s41569-023-00857-3

3. Santhanakrishnan R, Wang N, Larson MG, Magnani JW, McManus DD, Lubitz SA, et al. Atrial fibrillation begets heart failure and vice versa. *Circulation*. (2016) 133:484–92. doi: 10.1161/CIRCULATIONAHA.115.018614

4. Marini C, De Santis F, Sacco S. Contribution of atrial fibrillation to incidence and outcome of ischemic stroke. *Stroke*. (2005) 36:1115–9. doi: 10.1161/01.STR.0000166053.83476.4a

5. Rogers PA, Bernard ML, Madias C, Thihalolipavan S, Mark Estes NA, Morin DP. Current evidence-based understanding of the epidemiology, prevention, and treatment of atrial fibrillation. *Curr Probl Cardiol*. (2018) 43:241–83. doi: 10.1016/j.cpcardiol.2017.06.001

6. Lin J, Wu X-Y, Long D-Y, Jiang CX, Sang CH, Tang RB, et al. Asymptomatic atrial fibrillation among hospitalized patients: clinical correlates and in-hospital outcomes in improving Care for Cardiovascular Disease in China-atrial fibrillation. *EP Europace*. (2023) 25:euad272. doi: 10.1093/europace/euad272

7. Kingsley K, Galbraith GM, Herring M, Stowers E, Stewart T, Kingsley KV. Why not just Google it? An assessment of information literacy skills in a biomedical science curriculum. *BMC Med Educ*. (2011) 11:17. doi: 10.1186/1472-6920-11-17

8. Sharma D, Kanneganti TD. The cell biology of inflammasomes: mechanisms of inflammasome activation and regulation. *J Cell Biol*. (2016) 213:617–29. doi: 10.1083/jcb.201602089

9. Newman D, Tong M, Levine E, Kishore S. Prevalence of multiple chronic conditions by U.S. state and territory, 2017. *PLoS One*. (2020) 15:e0232346. doi: 10.1371/journal.pone.0232346

10. Dragseth MR. Building student engagement through social media. *J Polit Sci Educ*. (2020) 16:243–56. doi: 10.1080/15512169.2018.1550421

11. Bik HM, Goldstein MC. An introduction to social media for scientists. *PLoS Biol*. (2013) 11:e1001535. doi: 10.1371/journal.pbio.1001535

12. Fischhoff B. The sciences of science communication. *Proc Natl Acad Sci*. (2013) 110:14033–9. doi: 10.1073/pnas.1213273110

13. Knight E, Intzandt B, MacDougall A, Saunders TJ. Information seeking in social media: a review of YouTube for sedentary behavior content. *Interact J Med Res*. (2015) 4:e3. doi: 10.2196/ijmr.3835

14. Lee J, Turner K, Xie Z, Kadhim B, Hong Y-R. Association between health information–seeking behavior on YouTube and physical activity among U.S. adults: results from health information trends survey 2020. *AJPM Focus*. (2022) 1:100035. doi: 10.1016/j.focus.2022.100035

15. Yeung AWK, Tosevska A, Klager E, Eibensteiner F, Tsagkaris C, Parvanov ED, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. *J Med Internet Res*. (2022) 24:e28152. doi: 10.2196/28152

16. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. *J Med Internet Res*. (2022) 24:e37367. doi: 10.2196/37367

17. Javidan A, Nelms MW, Li A, Lee Y, Kayssi A, Naji F. Evaluating YouTube as a source of education for patients undergoing surgery: a systematic review. *Ann Surg*. (2023) 278:e712–e718. doi: 10.1097/SLA.0000000000005892

18. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med*. (2019) 240:112552. doi: 10.1016/j.socscimed.2019.112552

19. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res*. (2021) 23:e17187. doi: 10.2196/17187

20. Liu Z, Chen Y, Lin Y. YouTube/ Bilibili/ TikTok videos as sources of medical information on laryngeal carcinoma: cross-sectional content analysis study. *BMC Public Health*. (2024) 24:1594. doi: 10.1186/s12889-024-19077-6

21. Azer SA. Are DISCERN and JAMA suitable instruments for assessing YouTube videos on thyroid Cancer? Methodological concerns. *J Cancer Educ*. (2020) 35:1267–77. doi: 10.1007/s13187-020-01763-9

22. Tam J, Porter EK, Lee UJ. Examination of information and misinformation about urinary tract infections on TikTok and YouTube. *Urology*. (2022) 168:35–40. doi: 10.1016/j.urology.2022.06.030

23. Lookingbill V, Mohammadi E, Cai Y. Assessment of accuracy, user engagement, and themes of eating disorder content in social media short videos. *JAMA Netw Open*. (2023) 6:e238897. doi: 10.1001/jamanetworkopen.2023.8897

24. Joglar JA, Chung MK, Armbruster AL, Benjamin EJ, Chyou JY, Cronin EM, et al. Peer review committee members 2023 ACC/AHA/ACCP/HRS guideline for the diagnosis and Management of Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines. *Circulation*. (2024) 149:e1–e156. doi: 10.1161/CIR.0000000000001193

25. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, et al. 2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): the task force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) developed with the special contribution of the European heart rhythm Association (EHRA) of the ESC. *Eur Heart J*. (2021) 42:373–498. doi: 10.1093/eurheartj/ehaa612

26. Barton JL, Trupin L, Tonner C, Imboden J, Katz P, Schillinger D, et al. English language proficiency, health literacy, and trust in physician are associated with shared decision making in rheumatoid arthritis. *J Rheumatol*. (2014) 41:1290–7. doi: 10.3899/jrheum.131350

27. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. International Society of Hypertension Global Hypertension Practice Guidelines. *Hypertension*. (2020) 75:1334–57. doi: 10.1161/HYPERTENSIONAHA.120.15026

28. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. Introduction and methodology: standards of Care in Diabetes—2023. *Diabetes Care*. (2022) 46:S1–4. doi: 10.2337/dc23-Sint

29. KDIGO. 2024 clinical practice guideline for the evaluation and Management of Chronic Kidney Disease. *Kidney Int*. (2024) 105:S117–s314. doi: 10.1016/j.kint.2023.10.018

30. Barbarot S, Stalder JF. Therapeutic patient education in atopic eczema. *Br J Dermatol*. (2014) 170:44–8. doi: 10.1111/bjd.12932

31. Walker D, Adebajo A, Bukhari M. The benefits and challenges of providing patient education digitally. *Rheumatology*. (2020) 59:3591–2. doi: 10.1093/rheumatology/keaa642

32. Wang M, Yao N, Wang J, Chen W, Ouyang Y, Xie C. Bilibili, TikTok, and YouTube as sources of information on gastric cancer: assessment and analysis of the content and quality. *BMC Public Health*. (2024) 24:57. doi: 10.1186/s12889-023-17323-x

33. Suresh CH, Leng K, Washnik NJ, Parida S. The portrayal of hearing loss information in online mandarin videos. *J Otolaryngol*. (2023) 18:152–9. doi: 10.1016/j.joto.2023.05.007

34. Chen Z, Pan S, Zuo S. TikTok and YouTube as sources of information on anal fissure: a comparative analysis. *Front Public Health*. (2022) 10:1000338. doi: 10.3389/fpubh.2022.1000338

35. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ*. (2022) 22:382. doi: 10.1186/s12909-022-03446-z

36. Beautemps J, Bresges A. What comprises a successful educational science YouTube video? A five-thousand user survey on viewing behaviors and self-perceived importance of various variables controlled by content creators. Original research. *Front Commun*. (2021) 5. doi: 10.3389/fcomm.2020.600595

37. Liu C, Shi Z-Q, Zhuo J-W, Chen H, Chiou W-K. Research on short video health communication in medical and health institutions under the background of "healthy 2030". Springer Nature; Switzerland. (2023):476–485.

38. Yeung A, Ng E, Abi-Jaoude E. TikTok and attention-deficit/hyperactivity disorder: a cross-sectional study of social media content quality. *Can J Psychiatr*. (2022) 67:899–906. doi: 10.1177/07067437221082854

39. Ferhatoglu SY, Kudsioglu T. Evaluation of the reliability, utility, and quality of the information in cardiopulmonary resuscitation videos shared on open access video sharing platform YouTube. *Australasian Emergency Care*. (2020) 23:211–6. doi: 10.1016/j.auec.2020.05.005

40. Brooks F, Lawrence H, Jones A, McCarthy M. YouTube™ as a source of patient information for lumbar discectomy. *Annals Royal College of Surgeons of England*. (2014) 96:144–6. doi: 10.1308/003588414X13814021676396

41. Aragon-Guevara D, Castle G, Sheridan E, Vivanti G. The reach and accuracy of information on autism on TikTok. *J Autism Dev Disord*. (2023). doi: 10.1007/s10803-023-06084-6

42. Radonjic A, Fat Hing NN, Harlock J, Naji F. YouTube as a source of patient information for abdominal aortic aneurysms. *J Vasc Surg*. (2020) 71:637–44. doi: 10.1016/j.jvs.2019.08.230

43. Camm CF, Russell E, Ji Xu A, Rajappan K. Does YouTube provide high-quality resources for patient education on atrial fibrillation ablation? *Int J Cardiol*. (2018) 272:189–93. doi: 10.1016/j.ijcard.2018.08.066

44. Burki T. Vaccine misinformation and social media. *Lancet Digital Health*. (2019) 1:e258–9. doi: 10.1016/S2589-7500(19)30136-0

45. Benecke O, DeYoung SE. Anti-vaccine decision-making and measles resurgence in the United States. *Glob Pediatr Health*. (2019) 6:2333794x19862949. doi: 10.1177/2333794x19862949

46. Parabhoi L, Sahu RR, Dewey RS, Verma MK, Kumar Seth A, Parabhoi D. YouTube as a source of information during the Covid-19 pandemic: a content analysis of YouTube videos published during January to march 2020. *BMC Med Inform Decis Mak*. (2021) 21:255. doi: 10.1186/s12911-021-01613-8

47. Utunen H, Tokar A, Dancante M, Piroux C. Online learning for WHO priority diseases with pandemic potential: evidence from existing courses and preparing for disease X. *Arch Public Health*. (2023) 81:61. doi: 10.1186/s13690-023-01080-9

48. Chatterjee A, Strong G, Meinert E, Milne-Ives M, Halkes M, Wyatt-Haines E. The use of video for patient information and education: a scoping review of the variability and effectiveness of interventions. *Patient Educ Couns*. (2021) 104:2189–99. doi: 10.1016/j.pec.2021.02.009

49. Tan Y, Geng S, Chen L, Wu L. How doctor image features engage health science short video viewers? Investigating the age and gender bias. *Ind Manag Data Syst*. (2023) 123:2319–48. doi: 10.1108/IMDS-08-2022-0510

50. Kunze KN, Krivicich LM, Verma NN, Chahla J. Quality of online video resources concerning patient education for the Meniscus: a YouTube-based quality-control study. *Arthroscopy*. (2020) 36:233–8. doi: 10.1016/j.arthro.2019.07.033

51. Bora K, Pagdhune A, Patgiri SJ, Barman B, Das D, Borah P. Does social media provide adequate health education for prevention of COVID-19? A case study of YouTube videos on social distancing and hand-washing. *Health Educ Res*. (2022) 36:398–411. doi: 10.1093/her/cyab028

52. Kunze KN, Alter KH, Cohn MR, Vadhera AS, Verma NN, Yanke AB, et al. YouTube videos provide low-quality educational content about rotator cuff disease. *Clin Shoulder Elb*. (2022) 25:217–23. doi: 10.5397/cise.2022.00927

53. Bilibili Certification System. (2024). Available online at: https://account.bilibili.com/account/official/home?spm_id_from=333.999.0.0 (Accessed June 29, 2024).

54. Expanding equitable access to health information on YouTube. (2023). Available online at: https://blog.youtube/news-and-events/expanding-equitable-access-to-health-information-on-youtube/ (Accessed June 29, 2024).

55. Doiyin Certification System. Available online at: https://lf3-cdn-tos.draftstatic.com/obj/ies-hotsoon-draft/douyin/cd0175f4-34b4-4e0f-8197-72089c9d7709.html (Accessed June 29, 2024).

Examining the impact of sex-biased information on health behaviors: a study of HPV vaccination among male college students based on the extended theory of planned behavior

Tong-Chen Lucas Wang[1,2], Mei-Juan Zhang[1] and
Hualin Zhang[3]*

[1]School of Media and Communication, Shenzhen University, Shenzhen, China, [2]Institute of Global
Communication, Shenzhen University, Shenzhen, China, [3]Division of Arts, Shenzhen University,
Shenzhen, China

Introduction: Human papillomavirus (HPV) is an exceptionally transmissible virus with a widespread presence that represents a considerable public health concern. Notably, the existing information landscape regarding HPV vaccination tends to favor female perspectives, which may inadvertently neglect the risks associated with HPV infection in men-particularly among male college students, who constitute a vulnerable demographic. To address this gap, effective health communication strategies are essential to encourage vaccination behaviors. This study seeks to extend the Theory of Planned Behavior (TPB) to investigate how sex-biased information influences the HPV vaccination intentions of male college students.

Methods: A survey involving 240 participants was conducted to analyze the correlation between the HPV vaccine information environment and vaccination intentions. Utilizing Structural Equation Modeling (SEM), we assessed both direct and indirect effects that shape this relationship.

Results: Findings reveal that sex-biased HPV vaccination information can have both direct and indirect effects on the vaccination behaviors of male university students. Specifically, attitudes towards the vaccine positively influence the willingness to receive free vaccinations, while subjective norms and perceived behavioral control positively impact the intention to obtain both paid and free vaccinations.

Discussion: This study highlights that the female-oriented focus of HPV vaccination information, which underscores feminine norms, may limit the effectiveness of HPV health education for male college students. Additionally, price considerations have a limiting impact on the favorable attitudes of male university students towards vaccination. Conversely, male-targeted informational campaigns that heighten perceptions of HPV-related risks can diminish price sensitivity regarding vaccines, ultimately fostering increased willingness to vaccinate. The paper concludes by suggesting avenues for future research aimed at developing health communication strategies to enhance HPV vaccination rates among men. Implications and limitations are discussed.

KEYWORDS

theory of planned behavior (TPB), male HPV vaccination, sex bias, HPV information management, public health campaign

1 Introduction

Human Papillomavirus (HPV) infection, primarily transmitted through sexual contact, mother-to-child transmission, and contact with skin and mucous membranes, represents a highly infectious category of viruses. This infection impacts individuals regardless of gender and is associated with various diseases, thereby continuing to present a substantial public health challenge worldwide (1). The epidemiology of HPV in females has been well studied and documented. In contrast, the epidemiological landscape of HPV in males remains less understood. A systematic meta-analysis published in *The Lancet Global Health* underscores the significant prevalence of HPV among the male population. As of 2023, it is estimated that approximately one-third of males worldwide have been infected with at least one variant of genital HPV (2). Furthermore, around one-fifth of males are reported to be infected with one or more high-risk HPV types associated with cancer (2). Cervical cancer, as emphasized by Okunade et al. (3), is one of the malignancies induced by HPV, alongside anal cancer (4) and oral squamous cell carcinoma (5). These HPV-related cancers impose a substantial burden on both the physical and mental well-being of affected individuals. Research indicates that cervical cancer screening data in China reveals an overall HPV infection prevalence rate of 19.1% (6). This prevalence presents a considerable challenge to China's public health infrastructure, particularly considering its large population, and represents a critical issue regarding the reproductive health of its citizens. In response, China has implemented comprehensive strategies for HPV prevention, which include the National Health Commission's promotion of free HPV vaccinations for eligible girls, as well as the dissemination of health education and awareness campaigns about HPV through various mediums, notably social media. Nonetheless, the focus of these health initiatives and vaccination campaigns has predominantly been on the female demographic, with insufficient attention directed towards males, who are comparatively disadvantaged in terms of access to vaccine information.

Several studies have shown that men serve as a primary transmission group for HPV (7) and exhibit higher rates of infection (7–10). However, vaccination rates among men remain notably low (11–13). China aims to enhance HPV vaccination efforts among males, and numerous scholars have underscored the importance of this initiative, especially for specific populations such as men who have sex with men (MSM) (11, 13, 14). However, existing literature does not provide clear insights into current vaccination trends and the determinants influencing vaccination among male populations, which constitutes the primary focus of this study.

The observed phenomenon of a high rate of HPV infection coupled with a low vaccination rate among males can be attributed to three primary factors. First, at the individual level, men are less likely to encounter relevant information regarding HPV vaccines within the current information ecosystem, particularly on algorithm-driven social media platforms. Here, discussions surrounding HPV vaccination are frequently linked to 'cervical cancer,' a condition specific to the female reproductive system. Even when men do encounter information on these platforms, it is often tailored to a female audience and fails to directly address male health concerns (15). Consequently, the prevailing health information environment emphasizes content and strategies that primarily influence female behavioral intentions, neglecting to engage men as key targets for

health promotion. This oversight exacerbates the existing information gap regarding HPV vaccination among males and further diminishes awareness of the implications of sex-biased health information.

Second, at the societal level, the stigma surrounding HPV infection is deeply rooted in prevailing social norms. In the Chinese context, sexually transmitted diseases are often viewed as indicators of a lack of self-respect and self-care, with traditional beliefs disproportionately attributing the responsibility for the occurrence and transmission of these diseases to women (16–18). As a result, women in China are more likely to conform to societal expectations by engaging in preventive measures against sexually transmitted infections, such as HPV vaccination, which has gained acceptance as a mainstream health intervention in recent years. In contrast, societal norms do not similarly impose responsibilities on men, leading to a lower uptake of preventive behaviors among this demographic, including HPV vaccination (112, 113). This dynamic not only reinforces gender norms and stereotypes within sexual health discourse, but also overlooks the significant sexual health risks faced by China's male population, which numbers approximately 720 million[1].

Third, from a public health perspective, existing HPV-related health policies in China largely disregard the male population, resulting in a notable gap in media coverage related to men's health issues. Furthermore, the current public health framework provides various policy incentives for women to receive HPV vaccinations, including free vaccination options for eligible women. In recent years, Chinese public health authorities have been proactively advocating for policies aimed at encouraging men to receive HPV vaccinations. The topic of HPV vaccine pricing for men in China has persistently garnered significant attention and discourse, especially within the dichotomy of out-of-pocket expenses versus free vaccination options. This study also explores the impact of vaccine pricing on the vaccination intentions of males, serving as a preliminary investigation to inform future policy considerations aimed at expanding HPV vaccination access to men.

Considering the previously identified issues, this research enhances the Theory of Planned Behavior (TPB) by integrating sex bias as a precursor variable, while also broadening the scope of HPV vaccination intentions to encompass both self-funded and free vaccination options. This study investigates the mechanism through which sex-biased information affects men's intentions regarding self-paying for or receiving free HPV vaccinations. The findings aim to offer valuable insights for the future utilization and promotion of HPV vaccines among male populations.

2 Literature review

2.1 Gendered narratives and systemic gaps in HPV research

HPV vaccination research exhibits a female bias, with a scarcity of male study participants. In studies with a female-oriented focus,

---

1 Data source: National Bureau of Statistics of China https://data.stats.gov.cn/easyquery.htm?cn=C01

scholars have identified social media as a primary source of health information, highlighting that individual engagement with social media content related to vaccines can enhance their awareness of HPV vaccines. A significant positive correlation has been established between the cognitive levels regarding HPV vaccines and the willingness to receive the vaccine among both female populations and parents of young girls (19–21). Additionally, research indicates that vaccination attitudes can influence vaccination behaviors, with parents, as the primary decision-makers regarding their daughters' HPV vaccination, often hesitant due to uncertainties about vaccine safety (22, 23). Among young women, skepticism regarding the side effects of HPV vaccines emerges as the strongest predictor of vaccination behavior, followed by attitudes towards the vaccination process itself (24, 25). Overall, prior research has consistently demonstrated that the degree of awareness and the prevailing attitudes towards HPV vaccines significantly impact the likelihood of vaccination among female populations.

Vaccination necessitates that individuals engage in rational evaluations of their personal capabilities and the surrounding environmental factors, including economic conditions and the availability of vaccines. Research suggests that the uptake of HPV vaccination among female populations is affected by factors associated with perceived behavioral control. This concept has been shown to be a positive predictor of vaccination behavior among female college students, as well as the vaccination decisions made by parents for their daughters (26, 27). Moreover, self-efficacy, which is closely linked to perceived behavioral control, also plays a crucial role in influencing individuals' willingness to be vaccinated (28). Environmental challenges, such as shortages of vaccines, pose significant barriers to HPV vaccination, particularly impacting women from low- and middle-income backgrounds (15, 29, 30). Thus, the administration of the HPV vaccine represents a logical decision-making process for individuals, as women consistently assess the effectiveness of their vaccination.

The promotion and administration of HPV vaccines are also shaped by a variety of social and cultural norms, which can exert both positive and negative influences, particularly within Asian societies characterized by intricate family and ethnic dynamics and stringent expectations regarding women's behavior. Young women from diverse ethnic backgrounds often harbor stigmatizing perceptions of HPV vaccination, including the concern that receiving the vaccine may be construed as an endorsement of promiscuity (31–33). In several Asian nations, discussions surrounding sexual health are frequently suppressed due to prevailing social norms, leading to the perception that HPV vaccination may inadvertently endorse excessive sexual behavior or encourage risky sexual practices (34). Nevertheless, in contemporary Chinese society, certain non-stigmatizing social norms significantly contribute to the encouragement of women's acceptance of HPV vaccination. For example, within a child-centric cultural framework, Chinese parents demonstrate a willingness to embrace scientific approaches in managing their children's health, with the moral imperative of 'scientific parenting' emerging as a significant predictor of their decisions regarding HPV vaccination (35). Furthermore, health advocacy for HPV vaccination on Chinese social media predominantly targets women (36, 37), while the "feminization" of HPV creates societal expectations that encourage Chinese women to pursue vaccination (38). However, in practice, males, who are notable carriers of sexually transmitted infections, may neglect their own HPV prevention efforts. This negligence can be attributed to the lack of traditional social norms and insufficient media attention regarding HPV vaccination for men. Consequently, they may emerge as new focal points for HPV transmission. This issue represents a significant aspect that has been overlooked in prior research, particularly in studies on sexually transmitted infections conducted in Asia.

The uptake of HPV vaccination is generally shaped by individual factors, particularly among females. These factors include the levels of knowledge, educational background, attitudes towards the vaccine, and personal capabilities. Additionally, social influences, such as cultural norms and the availability of the vaccine, play a significant role. Nonetheless, most existing research has predominantly emphasized the female viewpoint, thereby neglecting male perspectives. This oversight has resulted in a notable gender bias within academic discourse. Consequently, there has been a relative lack of scholarly attention directed towards understanding the disparities in HPV vaccination rates within a sex-biased informational context, particularly regarding the perceptions and health information needs of male populations.

2.2 Sex-biased health information environment

Sex bias is characterized by the unequal representation of information and perspectives between genders, typically favoring one gender over the other (39). Sex bias in HPV vaccination information is defined by the unequal distribution of vaccine-related content across male and female populations, whereby the health needs of one gender are prioritized at the expense of the other. Recent studies focusing on the feminization of HPV vaccination (15, 40, 41) have identified three key dimensions in which female-biased HPV vaccination information manifests. Firstly, in the context of information presentation, there is a pronounced emphasis on the significance of vaccination for women, while the infection risks, disease implications, and vaccination advantages pertinent to men are either minimized or overlooked. Secondly, regarding the targeted populations, promotional initiatives predominantly emphasize the protection of women's health, particularly in relation to the prevention of cervical cancer. In contrast, there is a noticeable lack of attention given to health issues that are relevant to men, such as the prevention of anal cancer, oropharyngeal cancer, and genital warts. Lastly, in framing responsibility for vaccination, there is a tendency to characterize it as a "female obligation," rather than acknowledging it as a collective public health responsibility that necessitates engagement from individuals of all genders. These three dimensions underscore the notable disparities present in messaging that favors female perspectives over male considerations in HPV vaccination discourse.

Researchers contend that the information landscape of social media—including the dissemination of scientific knowledge, online public discourse, and digital imagery—demonstrates significant sex bias (42–47). Existing studies on sex bias primarily concentrate on descriptive analyses and the identification of influencing factors, aiming to clarify the underlying mechanisms that manifest societal forms of sex bias. Nonetheless, the academic discourse has largely overlooked the implications of sex bias for the promotion of individual health behaviors through information dissemination. The

aforementioned study utilizes a feminist perspective to analyze the impacts of male-oriented information environments on women situated within patriarchal structures.

A sex-biased information environment can significantly influence individuals' attitudes and behaviors, thereby shaping their perceptions of gender role differences. Gender-biased information environments can significantly shape behavioral attitudes by reinforcing existing stereotypes and promoting the internalization of specific beliefs. For instance, the pervasive presence of gender-biased information can influence individuals' evaluations of certain behaviors through the repeated exposure to concepts such as gender discrimination and stereotypical divisions of labor (48). When these biased messages gain widespread acceptance within society, individuals tend to assimilate them into their personal belief systems. Research supporting this assertion demonstrates that the utilization of sex-biased language plays a significant role in the continuation of gender discrimination and influences individuals' perceptions and attitudes regarding gender-related concepts (49–51). Furthermore, sex bias is associated with the health outcomes of both men and women (52). In the realm of vaccination, most textual information environments are free from sex bias; for instance, communications regarding flu and hepatitis B vaccines generally concentrate on the diseases themselves rather than on gender distinctions. However, when the information environment pertaining to vaccination decisions is characterized by sex bias, as observed with HPV vaccines, this bias—whether manifested through gender stereotyping or stigmatization—can influence attitudes towards vaccination, thereby impacting both intentions and actual vaccination behaviors. Currently, the information environment surrounding HPV vaccination in China frequently links HPV to female identity, which may lead the public to perceive a greater societal expectation for women, as opposed to men, to receive the HPV vaccine. Social media platforms generate a substantial volume of information daily, varying in quality, which can contribute to a pronounced sex-biased information environment.

The gender-biased information environment influences the subjective norms regarding HPV vaccination among early-adult male college students through gender norms. Firstly, gender roles can be considered a form of social norm (114). Secondly, collectivist cultures emphasize group harmony and adherence to social norms, resulting in individual behaviors being significantly influenced by group opinions (115). Male college students, being in the early stages of adulthood, are particularly sensitive to the perceptions of their peers and experience a strong sense of social identity (116). Within a collectivist context, these male students may be more attuned to socially constructed norms shaped by gender roles and may fear being labeled as non-conforming to these roles. In a gender-biased social media information environment, where HPV vaccination is predominantly viewed as a concern for women, male students may perceive the act of receiving the HPV vaccine as inconsistent with traditional masculine roles, potentially leading to ridicule or social exclusion. As a result, they might avoid vaccination despite being aware of its health benefits. A study investigating the willingness to vaccinate against HPV among Chinese college students studying in the United States revealed that participants more strongly influenced by collectivist culture were less likely to receive the HPV vaccine (53). Therefore, it can be posited that information environments characterized by gender bias have a substantial impact on individuals'

attitudes and subjective norms, especially among individuals embedded in collectivist cultures.

Behavioral attitudes and subjective norms pertain to an individual's assessment of subjective elements, whereas perceived behavioral control encompasses a consideration of objective factors (54). For individuals, sex-biased information represents an objective reality. The information environment characterized by gender bias can initially diminish individual self-efficacy by influencing perceptions of ability. For instance, the sustained emphasis on male advancement in the fields of science and technology may cause women to undervalue their own competencies within these domains, fostering a belief that "I am not proficient in science and engineering" (55). Furthermore, when the information environment implies that certain resources are preferentially allocated to one gender, individuals may perceive a constricted ability to exercise control over their actions. Specifically, women might show reluctance to pursue executive roles due to prevalent discussions surrounding the "glass ceiling" phenomenon (56). Specifically, information favoring males regarding HPV vaccination and information favoring females influence men's cognition, particularly among male college students, thereby impacting their perceived behavioral control. Consequently, this study centers on the relationships between sex-biased informational contexts and various factors, including attitudes, subjective norms, perceived behavioral control, and actual behaviors.

Gender roles represent the socialized outcomes of individual behaviors, which are shaped by biological inheritance and influenced by various political, economic, and cultural factors within society. These roles contribute to societal perceptions of gender, with socially prescribed or implicit gender roles guiding individuals' behavioral patterns through mechanisms such as information dissemination (57). Informed by TPB, some researchers have sought to incorporate gender as a new variable to enhance the TPB framework, investigating its effect on behavioral development within this model. For instance, Kyrrestad et al. (58) integrated gender differences in TPB were detected, Intention to increase the frequency of drinking is predicted by subjective norms (SN) and attitude for girls and by subjective norms (SN) for boys. Similarly, Xu et al. (59) analyzed the differences in learning behaviors between male and female engineering students under peer pressure, and found that gender differences have a positive impact on subjective norms, and high-performing female students can motivate male classmates and improve their learning intentions and behaviors. Under the guidance of TPB theory, Xie et al. (60) tried to explore the influence of gender differences on healthy office building strategies, and the experimental results showed that there was a positive relationship between female employees' health behavior intentions and architectural design strategies.

These studies primarily focus on gender as a precursor variable within the TPB framework, highlighting its significant impact on individual behavioral patterns. Nonetheless, the mechanisms through which sex-biased information contributes to behavioral differences warrant further investigation. Research suggests that gender can function as a precursor variable influencing individual attitudes, social norms, and perceived behavioral control in the initial stage, ultimately shaping behavioral variations. Furthermore, sex-biased information also impacts societal perceptions of gender, which in turn reinforces attitudes, social norms, and perceived behavioral control, culminating in differentiated behavioral patterns. This study aims to examine whether sex-biased information plays a pivotal role in shaping

individuals' behavioral patterns, with a particular focus on HPV vaccination behavior. This exploration seeks to understand how cognitive differences may lead to varied behavioral orientations.

2.3 Extended theory of planned behavior and HPV vaccination

The Theory of Planned Behavior (TPB) is a framework within social psychology that seeks to explain and forecast individual behavioral actions. According to this theory, individuals make decisions regarding their behavior in a rational manner, where three key components—Attitude toward the Behavior (AB), Subjective Norms (SN), and Perceived Behavioral Control (PBC)—interact to shape behavioral intentions. These intentions subsequently influence the likelihood of actual behavior occurring (54). As a well-established theoretical model, the TPB has been widely applied in health communication research, particularly in studies related to HPV vaccination in recent years. Researches have made various modifications and extensions to the theory to better suit different research contexts. Notably, the TPB has demonstrated strong predictive capabilities, providing valuable insights into the processes underlying the formation of new behaviors in specific scenarios, including detailed investigations of HPV vaccination behaviors (27, 61–67).

In the context of deciding whether to receive the HPV vaccination, individuals are required to engage in rational decision-making processes. This involves evaluating their personal health status, considering the economic costs associated with vaccination, and conducting a comprehensive analysis of the potential health benefits. This rational analysis ultimately informs their decision-making process regarding vaccination. Notably, the cost of vaccines has emerged as a significant factor contributing to vaccine hesitancy within the Chinese population (68, 69). In mainland China, HPV vaccines are classified into two categories: free and self-funded. The free vaccines are primarily distributed through the national vaccination program and are currently accessible in 15 provinces. In contrast, the self-funded vaccines are predominantly imported, with an average price exceeding $160, a substantial financial burden relative to the average income of residents in China. Research indicates that the willingness of Chinese women to receive the HPV vaccine is significantly influenced by the cost of the vaccine, with higher prices negatively impacting vaccination rates (70–72). This study specifically aims to investigate the relationship between the willingness to receive the HPV vaccine and its price among males, with the objective of enhancing the understanding of the decision-making processes within this demographic.

Attitude toward the behavior encompass a range of beliefs regarding the likelihood of outcomes associated with specific behaviors (54). Generally, a more favorable attitude towards a behavior correlates with a stronger intention to engage in that behavior, thereby increasing the likelihood of its actual occurrence. Research by Fazio and Zanna (73) indicates that individuals often form positive or negative expectations about behavioral outcomes based on direct experiences, with attitudes grounded in such experiences exerting a greater influence on behavior than those stemming from indirect experiences. In previous vaccination campaigns, including those for the H7N9 virus, hepatitis B, and

COVID-19, there were no significant gender differences observed in the information environments fostered on social media platforms. Attitudes towards these vaccines primarily derived from the accumulation of direct past experiences, which were not markedly affected by sex-biased information. In contrast, the current dynamics surrounding HPV vaccination reveal distinct characteristics: attitudes towards HPV vaccination are considerably shaped by a sex-biased information environment, diverging from the mechanisms influencing other vaccines. Consequently, sex bias has not been recognized as a primary factor in previous investigations into vaccine attitudes and vaccination behaviors (74–77).

Within the Chinese context, HPV vaccines are classified into two categories: one that is provided at no cost under the national immunization program and another that requires out-of-pocket payment. The disparity in vaccine pricing significantly influences individuals' attitudes towards vaccination and their actual vaccination behaviors. For instance, research conducted in Asian nations has shown that public willingness to vaccinate often declines markedly when confronted with higher-priced private vaccines (78). In summary, investigations into HPV vaccine uptake among women indicate that attitudes toward the vaccine can effectively predict vaccination behavior, with more positive attitudes leading to a greater willingness to vaccinate, although this relationship is moderated by the vaccine's price (38, 79–81). Therefore, a positive attitude towards HPV vaccination should also be regarded as a critical factor in forecasting HPV vaccination rates among males in China. Therefore, this study proposes:

> H1: Attitudes toward HPV vaccination behavior positively influence (a) the willingness of males to pay for HPV vaccination and (b) the willingness to receive free HPV vaccination.

Subjective Norms are defined as an individual's perception of societal expectations that either endorse or discourage a particular behavior (54). Cultural contexts significantly shape health perceptions, resulting in varying interpretations and narratives regarding the same illness in different countries, regions, and even within the same nation. From a cultural perspective, China is a collectivist society, where social values are centered on collective interests. In collectivist societies, when making decisions about specific behaviors, individuals place greater emphasis on the support and opposition of others and groups in society, and are more inclined to conform to social expectations (82). In China, the vaccination behaviors of individuals are intricately associated with their sense of social responsibility. Studies suggest that individuals who possess a heightened sense of social responsibility are more likely to engage in vaccination, particularly evident during the COVID-19 vaccination campaign (83, 84). Furthermore, existing research indicates that subjective norms play a crucial role in influencing HPV vaccination rates among women (27, 35, 38, 66). Therefore, it is reasonable to propose that subjective norms, including the concept of social responsibility, also impact HPV vaccination behaviors among male populations in China. Therefore, the present study posits:

> H2: The subjective norms regarding HPV vaccination have a positive influence on (a) the willingness of male populations to pay for HPV vaccination and (b) their willingness to receive free HPV vaccination.

Perceived behavioral control refers to an individual's belief in their ability to manage and perform a specific behavior, regardless of its complexity (54). Economic capability stands as a significant determinant for individuals when evaluating their decision-making capacities. In research targeting female populations, economic factors—including family income and socioeconomic status—have been found to have a strong correlation with HPV vaccination rates (85, 86). Most of these studies indicate that perceived behavioral control serves as a positive predictor of HPV vaccination among women (23, 38, 66). Therefore, it is plausible to assert that perceived behavioral control regarding HPV vaccination among Chinese men will similarly serve as a positive predictor of their intention to receive the HPV vaccine (66, 70, 87). Therefore, we propose:

H3: Perceived behavioral control of HPV vaccination has a positive impact on (a) the willingness of men to pay for HPV vaccination and (b) the willingness of men to receive free HPV vaccination.

This research also examines the influence of sex-biased information on the willingness of male populations to receive the HPV vaccination. Consequently, drawing from prior research and the theoretical framework outlined earlier, sex-biased information is introduced as a precursor variable within the TPB model, thereby developing an extended TPB model that incorporates sex-biased information. Thus, the current study aims to explore the following research questions:

RQ1: Does sex-biased HPV vaccination information have a direct impact on the HPV vaccination intentions of male individuals?

RQ2: Does sex-biased HPV vaccination information exert an indirect influence on the HPV vaccination intentions of male

individuals through (a) attitudes toward HPV vaccination behavior, (b) subjective norms regarding HPV vaccination, and (c) perceived behavioral control over HPV vaccination?

Overall, this study, which adopts a male viewpoint and employs TPB, develops an enhanced TPB model that incorporates the concept of information sex bias. The research investigates the impact of sex-biased information on the intention of men to receive the HPV vaccination, while also examining the mechanisms that underlie this relationship. The proposed extended TPB model, informed by sex-biased information, is illustrated in Figure 1.

3 Methods

3.1 Data and sample

The methodology employed in this study was a questionnaire survey. The questionnaires were disseminated via the WJX.cn platform on Chinese social media channels. Our team distributed questionnaires online in mainland China during September to October 2024. Participants provided informed consent and received a random reward ranging from RMB 0.5 to 3 upon completing the questionnaire. Given that all items in the questionnaire were mandatory, there were no missing values in the data. However, we established criteria for identifying and removing outliers. First, attention check questions were included in the survey; responses indicating "no" to the item stating "Please select 'yes'" were excluded from the dataset. Second, responses that were completed in less than 100 s were also removed from analysis.

Given that the peak age for male patients positive for HPV across various provincial regions in China is between 20 and 39 years (88),
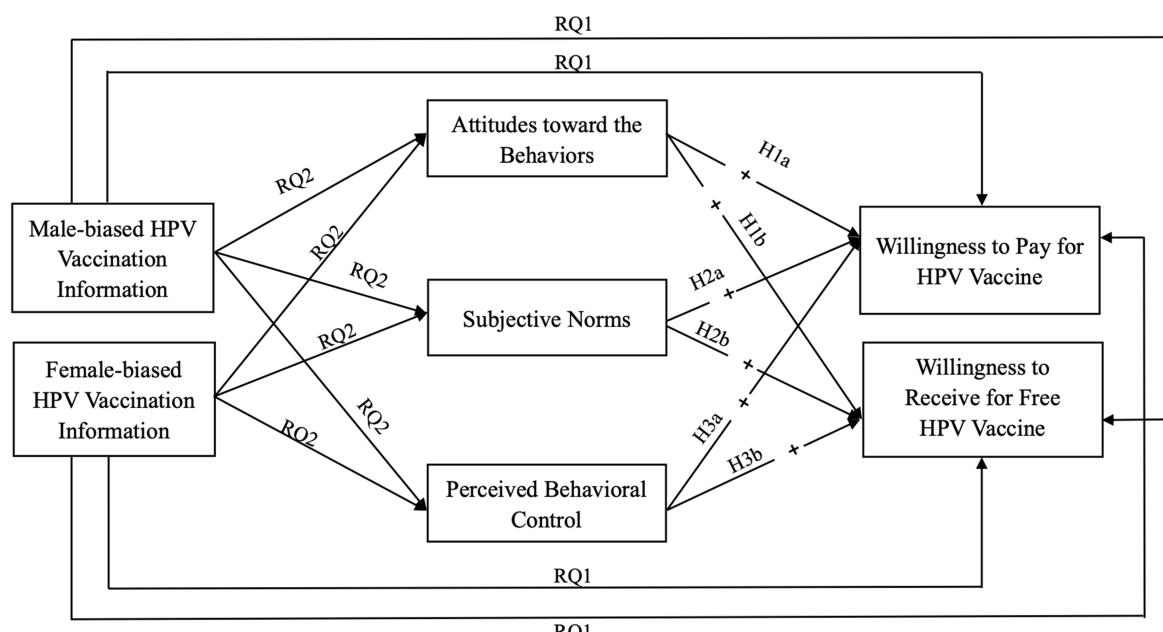

FIGURE 1
Theoretical framework.

which coincides with the age demographic of students in higher education institutions, the participants in this research were male students from colleges and universities. A convenience sampling approach was utilized to survey individuals aged 18 years and older within this target group. Regarding the sample size of this study, as we ultimately construct a structural equation model for statistical analysis, the widely accepted and cited assertion that "at least 100–200" samples are required is supported by references such as Principles and Practice of Structural Equation Modeling [(89), p. 16]. The sample size should be at least 5–10 times the maximum path length in the model. MacCallum et al. (90) extensively discussed various factors influencing SEM sample size and proposed the 5–10 times guideline as a general principle. In our model, the longest path comprises three variables, indicating a required sample size of approximately 150 to 300 participants. Accordingly, we set our target at the midpoint of this range, which is 225. After collecting a total of 264 questionnaires and discarding invalid responses, we achieved a final valid sample size of 240, thereby meeting the established criteria, also resulting in an effective recovery rate of 90.23%. However, it is important to note that due to constraints associated with the research topic, many potential respondents were reluctant to engage in surveys concerning male HPV issues. This limitation may regrettably hinder our ability to secure a larger sample size.

The sample comprised individuals aged between 19 and 30 years (M = 22.2, SD = 1.98), and included 16 junior college students (6.7%), 187 undergraduate students (77.9%), 36 master's students (15.0%), and 1 doctoral student (0.4%). Regarding sexual orientation, 220 participants identified as heterosexual (91.7%), while 20 participants identified as homosexual, bisexual, or other orientations (8.3%). The findings indicated that 107 male college students (44.6%) reported having prior sexual experience, of whom 17 (15.9% of those with sexual experience) indicated having engaged in sexual relations with individuals of the same sex. Furthermore, a total of 65 male college students (27.1%) within the sample had received the HPV vaccination (see Table 1).

3.2 Measurement

3.2.1 Demographic variables

Demographic factors included age, education level (0 = college diploma, 1 = bachelor's degree, 2 = master's degree, 3 = doctoral degree), sexual orientation (0 = heterosexual, 1 = homosexual/bisexual/other), self-reported sexual behavior such as sexual history (0 = yes, 1 = no), homosexual history (0 = yes, 1 = no), and HPV vaccination history (0 = yes, 1 = no).

3.2.2 Sex-biased HPV vaccination information

Sex-biased HPV vaccination information was measured as male-biased HPV vaccination information (HPV-MI) and female-biased HPV vaccination information (HPV-WI). The questionnaire was adapted based on prior research (91). The dimensions assessed included self-publishing, friend publishing, and media publishing. Self-published information is generally rooted in individual experiences, opinions, and emotions, and is shaped by external inputs from friends and media. Thus, self-publishing was integrated into the analysis. Conversely, media-published information is predominantly factual, data-driven, and centered on public events, encompassing

TABLE 1 Sample characteristics (N = 240).

Variable	M (SD) or N (%)
Demographic characteristics	
Age	22.2 (1.98)
Education level	
Junior college	16 (6.7%)
Bachelor degree	187 (77.9%)
University master	36 (15.0%)
University doctor	1 (0.4%)
Sexual orientation	
Heterosexuality	220 (91.7%)
Homosexual/Bisexual/Other	20 (8.3%)
Self-reported sexual behavior	
Sexual history	
Yes	107 (44.6%)
No	133 (55.4%)
Homosexual history	
Yes	17 (15.9%)
No	90 (84.1%)
HPV vaccination history	
Yes	65 (27.1%)
No	175 (72.9%)

dimensions related to social phenomena, public affairs, and cultural dynamics. This type of information is typically more objective, reflecting societal norms and shared characteristics. The frequency of information shared by friends falls between these two extremes. Both variables are measured using three items (1 = I never, 2 = Once a month, 3 = Every few weeks; 4 = Once a week; 5 = Every day), including: (1) "How frequently have you posted or mentioned information related to HPV vaccination for males/females on social media platforms in the past 6 months"; (2) "How frequently have your friends posted or mentioned information related to HPV vaccination for males/females on social media platforms in the past 6 months"; (3) "How frequently have you seen information related to HPV vaccination for males/females on social media platforms in the past 6 months" (where "Male-biased HPV vaccination information" items: M = 1.731, SD = 0.062, Cronbach's α = 0.852; "Female-biased HPV vaccination information" items: M = 2.022, SD = 0.112, Cronbach's α = 0.744).

3.2.3 Attitudes towards HPV vaccination behaviors

The concept of attitudes towards HPV vaccination behaviors (HPV-AB) encompasses individuals' subjective assessments pertaining to human papillomavirus (HPV) vaccination, including the preventive effect of HPV vaccine, its maintenance effect, the gains and losses of not getting vaccinated, and the benefits of vaccination. The questionnaire employs a Likert 5-point scale (1 = strongly disagree, 5 = strongly agree) and is borrowed from previous research (92). There are four items: (1) "Getting the HPV vaccine can prevent human papillomavirus infection"; (2) "The effect of getting the HPV vaccine

is good"; (3) "Not getting the HPV vaccine makes it easier to be infected with human papillomavirus"; (4) "Getting the HPV vaccine can reduce the expenses due to human papillomavirus infection" ($M = 4.01$, $SD = 0.614$, Cronbach's $\alpha = 0.704$).

3.2.4 HPV vaccination subjective norms

The HPV Vaccination Subjective Norm (HPV-SN) is a construct that assesses the perceived social pressure individuals experience when considering the HPV vaccination. This social pressure may arise from various sources, including family members, friends, governmental entities, and media representations. The measurement of HPV-SN consists of five specific items and utilizes a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) (92): (1) "My relatives and friends support me getting the HPV vaccine"; (2) "I would get the HPV vaccine because my relatives and friends recommend it"; (3) "I would get the HPV vaccine because the government provides adequate HPV vaccines"; (4) "I would get the HPV vaccine because the government provides free HPV vaccines"; (5) "I get the HPV vaccine because it is promoted for use" ($M = 3.88$, $SD = 0.931$, Cronbach's $\alpha = 0.832$).

3.2.5 Perceived behavioral control of HPV vaccination

The Perceived Behavioral Control of HPV Vaccination (HPV-PBC) assesses individuals' perceived ease of self-administration of the HPV vaccine. This construct encompasses several factors, including the complexity of personal decision-making, the availability of time for vaccination, the accessibility of vaccination services, and financial considerations associated with vaccination. The measure consists of four statements in a Likert scale (92) (1 = strongly disagree, 5 = strongly agree): (1) "I can decide for myself whether to get the HPV vaccine"; (2) "I have time to get the HPV vaccine"; (3) "It is convenient to get the HPV vaccine"; (4) "The cost of getting the HPV vaccine is not significant" ($M = 3.8$, $SD = 0.920$, Cronbach's $\alpha = 0.704$).

3.2.6 Willingness to HPV vaccination

The dependent variable was defined by participants' self-reported willingness to receive the HPV vaccination either at their own expense (HPV-P) or at no cost (HPV-F). The item was "I am willing to (pay for/receive for free) the HPV vaccine" ($M_{pay} = 3.34$, $SD_{pay} = 1.003$; $M_{free} = 4.35$, $SD_{free} = 0.620$).

All variables have been restructured based on previously established scales. A detailed account of the constructs, including Sex-biased HPV Vaccination Information, Attitudes Towards HPV Vaccination Behaviors, Subjective Norms Regarding HPV Vaccination, Perceived Behavioral Control over HPV Vaccination, and Willingness to Receive HPV Vaccination, is presented in Table 2.

3.3 Data analysis

This research employed SPSS 26.0 and Amos 28.0 software for the analysis of data results. Initially, SPSS 26.0 was used to evaluate the reliability and validity of the questionnaire. The findings from the reliability assessment indicated that the Cronbach's alpha coefficients for each primary variable ranged from 0.704 to 0.852, exceeding the threshold of 0.7, which suggests that the questionnaire demonstrates adequate reliability. Furthermore, the examination of structural validity revealed the following: (1) The KMO statistic and

Bartlett's test of sphericity for each variable yielded a KMO value of 0.808, which is above the 0.7 threshold, indicating that the intercorrelations among the variables are not significantly different, thus rendering the data highly appropriate for factor analysis (93); (2) The p-values obtained from Bartlett's test of sphericity were consistently 0.000, which is below the 0.001 level, leading to the rejection of the sphericity hypothesis and confirming the presence of correlations among the original variables, thereby affirming the suitability of the data for factor analysis.

Subsequently, a confirmatory factor analysis was performed utilizing AMOS 28.0 software to assess a measurement model comprising five latent variables: HPV-MI (HPV vaccination information favoring males), HPV-WI (HPV vaccination information favoring females), HPV-AB (behavioral attitudes toward HPV vaccination), HPV-SN (subjective norms pertaining to HPV vaccination), HPV-PBC (perceived behavioral control over HPV vaccination), as well as two observed variables, HPV-P (self-funded HPV vaccination) and HPV-F (complimentary HPV vaccination). The evaluation metrics employed were convergent validity and discriminant validity. The findings demonstrated that the overall fit indices of the measurement model were satisfactory, with the following values recorded: $\chi^2 = 353.196$, $\chi^2/df = 2.102$, RMSEA = 0.068, CFI = 0.909, and IFI = 0.911. The composite reliability (CR) values exceeded 0.7, and the average variance extracted (AVE) values were all greater than 0.35, confirming the high convergent validity of the measurement scale (see Table 3).

Before implementing the specific verification model, we conducted a correlation analysis among the various variables, with detailed results presented in Table 4. Additionally, we performed a multicollinearity assessment on the variables included in the model. The findings indicated that there were no multicollinearity issues among the variables incorporated into the model. Specifically, the minimum variance inflation factor (VIF) value was 1.267, while the maximum VIF value reached 1.778, both of which met the critical threshold of VIF < 10 (117: 102).

4 Results

The results of the path analysis for the extended TPB model were derived using structural equation modeling, with the corresponding path coefficients presented in Table 5.

Information regarding HPV vaccination that is biased towards males has a substantial impact on various factors related to the willingness to pay for the HPV vaccine. Specifically, the analysis reveals that such information significantly affects self-reported willingness to pay for vaccination ($\beta = 0.215$, $SE = 0.089$, $p < 0.001$), attitudes towards HPV vaccination behaviors ($\beta = 0.187$, $SE = 0.076$, $p = 0.040$), subjective norms associated with HPV vaccination ($\beta = 0.215$, $SE = 0.079$, $p = 0.011$), and perceived behavioral control over HPV vaccination ($\beta = 0.247$, $SE = 0.065$, $p = 0.025$). Overall, the findings indicate that male-biased HPV vaccination information exerts a positive influence on individuals' behavioral attitudes, subjective norms, perceived control, and their willingness to invest in vaccination.

The analysis revealed no significant differences in the willingness to pay for or accept free HPV vaccinations, as well as in attitudes towards HPV vaccination behavior, subjective norms, and

TABLE 2 Details of questionnaire adaptation.

Previous research inventory	This study adapted the scale
Sex-biased HPV vaccination information	
In the past 6 months, how often did you post or mention e-cigarettes on (insert platform)?	How frequently have you posted or mentioned information related to HPV vaccination for males/females on social media platforms in the past 6 months?
In the past 6 months, how often did a friend post or mention e-cigarettes on (insert platform)?	How frequently have your friends posted or mentioned information related to HPV vaccination for males/females on social media platforms in the past 6 months?
In the past 6 months, how often did you see advertisements about e-cigs on (insert platform)?	How frequently have you seen information related to HPV vaccination for males/females on social media platforms in the past 6 months?
Attitudes towards HPV vaccination behaviors	
Getting the EV71 vaccine can prevent Hand, foot and mouth disease.	Getting the HPV vaccine can prevent human papillomavirus infection.
The effect of getting the EV71 vaccine is good.	The effect of getting the HPV vaccine is good.
Not getting the EV71 vaccine makes it easier to be infected with Hand, foot and mouth disease.	Not getting the HPV vaccine makes it easier to be infected with human papillomavirus.
Getting the EV71 vaccine can reduce the expenses due to Hand, foot and mouth disease.	Getting the HPV vaccine can reduce the expenses due to human papillomavirus infection.
HPV vaccination subjective norms	
My relatives and friends support me getting the EV71 vaccine for my children.	My relatives and friends support me getting the HPV vaccine.
I would get the EV71 vaccine for my children because my relatives and friends recommend it	I would get the HPV vaccine because my relatives and friends recommend it.
I would get the EV71 vaccine for my children because the government provides HPV vaccines.	I would get the HPV vaccine because the government provides adequate HPV vaccines.
	I would get the HPV vaccine because the government provides free HPV vaccines.
I get the EV71 vaccine for my children because it is promoted for use.	I get the HPV vaccine because it is promoted for use.
Perceived behavioral control of HPV vaccination	
I can decide for my children whether to get the EV71 vaccine.	I can decide for myself whether to get the HPV vaccine.
I have time to get the EV71 vaccine for my children.	I have time to get the HPV vaccine.
It is convenient to get the EV71 vaccine for my children.	It is convenient to get the HPV vaccine.
The cost of getting the EV71 vaccine is not significant.	The cost of getting the HPV vaccine is not significant.
Willingness to HPV vaccination	
I am willing to pay for the EV71 vaccine for my children.	I am willing to pay for the HPV vaccine.
I am willing to get the EV71 vaccine for my children.	I am willing to receive for free the HPV vaccine.

perceived behavioral control among individuals exposed to sex-biased HPV vaccination information. This finding suggests that such information exerts neither a positive nor a negative influence on the aforementioned variables. Therefore, RQ1 and RQ2 have been addressed.

A significant difference was identified in the willingness to accept free HPV vaccination, contingent upon individuals' attitudes towards HPV vaccination behaviors ($\beta = 0.208$, $SE = 0.106$, $p = 0.015$). This finding suggests that positive attitudes towards HPV vaccination behaviors significantly enhance the likelihood of accepting free HPV vaccination. Therefore, H1b is supported, while H1a is not.

The subjective norms associated with HPV vaccination had a significant impact on individuals' willingness to pay for the vaccine ($\beta = 0.429$, $SE = 0.156$, $p < 0.001$) as well as their willingness to accept a free HPV vaccination ($\beta = 0.312$, $SE = 0.109$, $p = 0.002$). These findings indicate that subjective norms exert a positive influence on the willingness to either pay for or receive a complimentary HPV vaccination. Therefore, H2a and H2b are supported.

The perceived behavioral control regarding HPV vaccination plays a significant role in influencing individuals' willingness to pay for the vaccine ($\beta = 0.265$, $SE = 0.249$, $p = 0.017$) as well as their willingness to accept free HPV vaccination ($\beta = 0.327$, $SE = 0.204$, $p = 0.005$). These findings suggest that individuals who perceive a greater degree of control over HPV vaccination are more likely to express a willingness to both pay for and receive the vaccination at no cost. Therefore, H3a and H3b are supported. The overall framework of the extended TPB model, which incorporates considerations of gender bias, is presented in Figure 2.

In summary, sex-biased information regarding HPV vaccination can have both direct and indirect effects on the vaccination behaviors of male college students. More specifically, male-biased information regarding HPV vaccination is associated with an enhancement in attitudes toward vaccination behavior, subjective norms, and perceived behavioral control. This, in turn, directly influences the willingness to pay for HPV vaccination and indirectly impacts both the willingness to pay and the willingness to accept free HPV vaccination. Conversely,

TABLE 3 Potential dimension reliability analysis.

Dimension	Latent variable	Standardized regression weights	Regression weights	S.E.	T-value	P	SMC	CR	AVE
HPV-MI	HPV-MI_3	0.745	1.000				0.555	0.866	0.685
	HPV-MI_2	0.890	1.061	0.077	13.792	***	0.792		
	HPV-MI_1	0.841	0.832	0.060	13.848	***	0.707		
HPV-AB	HPV-AB_4	0.712	1.000				0.507	0.737	0.417
	HPV-AB_3	0.467	0.826	0.131	6.294	***	0.218		
	HPV-AB_2	0.678	0.802	0.095	8.479	***	0.460		
	HPV-AB_1	0.696	0.739	0.085	8.660	***	0.484		
HPV-SN	HPV-SN_5	0.746	1.000				0.557	0.838	0.510
	HPV-SN_4	0.642	0.794	0.085	9.377	***	0.412		
	HPV-SN_3	0.770	1.059	0.091	11.698	***	0.593		
	HPV-SN_2	0.759	0.978	0.088	11.139	***	0.576		
	HPV-SN_1	0.641	0.847	0.092	9.228	***	0.411		
HPV-PBC	HPV-PBC_4	0.414	1.000				0.171	0.689	0.365
	HPV-PBC_3	0.595	1.334	0.219	6.102	***	0.354		
	HPV-PBC_2	0.736	1.468	0.287	5.115	***	0.542		
	HPV-PBC_1	0.628	1.184	0.248	4.778	***	0.394		
HPV-WI	HPV-WI_3	0.619	1.000				0.383	0.784	0.557
	HPV-WI_2	0.939	1.395	0.154	9.037	***	0.882		
	HPV-WI_1	0.637	0.905	0.101	8.945	***	0.406		

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 4 Correlation analysis.

Variables	M	SD	1	2	3	4	5	6
1 HPV-AB	15.940	2.666						
2 HPV-SN	19.323	3.677	0.537**					
3 HPV-PBC	15.149	2.854	0.433**	0.575**				
4 HPV-MI	12.746	2.435	−0.077	−0.129*	−0.187**			
5 HPV-WI	11.907	2.402	−0.050	−0.091	−0.005	0.442**		
6 HPV-P	3.330	1.000	0.278**	0.532**	0.516**	−0.279**	−0.082	
7 HPV-F	4.320	0.810	0.509**	0.646**	0.515**	−0.047	−0.082	0.326**

*$p < 0.05$, **$p < 0.01$.

female-biased information does not exert any significant positive or negative influence on these variables.

5 Discussion

Existing research on the determinants of HPV vaccination behavior predominantly concentrates on female populations. This study expands upon this foundation by examining the effects of sex-biased information on HPV vaccination behaviors among male populations. The results indicate that sex-biased information can exert both direct and indirect influences on the health behaviors of male individuals. This paper posits the following conclusions: (1) the relationship between such information and an individual's gender identity is notable; (2) sex-biased information activates distinct gender concepts, thereby

prompting individuals to employ varying methods of information processing; (3) the financial cost of the vaccine emerges as a significant factor contributing to hesitancy toward HPV vaccination.

5.1 Promotion of vaccination intentions through a male-biased information environment

The current research indicates that targeted information positively influences male college students' attitudes, subjective norms, and perceived control concerning HPV vaccination, ultimately enhancing their willingness to either pay for or accept free vaccination. When male college students are exposed to information specifically addressing HPV vaccination for males,

TABLE 5 TPB model path coefficient.

Pathway	Estimate	P	95% CI		S.E.	C.R.
			Lower	Upper		
HPV-MI→HPV-P	0.215***	***	0.160	0.374	0.089	3.192
HPV-WI→HPV-P	0.029	0.657	−0.005	0.211	0.100	0.444
HPV-MI→HPV-F	−0.098	0.060	−0.018	0.155	0.063	−1.622
HPV-WI→HPV-F	0.057	0.328	−0.043	0.150	0.072	0.978
HPV-MI→HPV-AB	0.187*	0.040	0.043	0.199	0.076	2.057
HPV-WI→HPV-AB	0.096	0.280	−0.022	0.167	0.087	1.080
HPV-AB→HPV-P	−0.182	0.060	−0.074	0.323	0.152	−1.884
HPV-AB→HPV-F	0.208*	0.015	0.185	0.435	0.106	2.436
HPV-MI→HPV-SN	0.215*	0.011	0.043	0.210	0.079	2.549
HPV-WI→HPV-SN	0.044	0.600	−0.022	0.160	0.092	0.525
HPV-SN→HPV-P	0.429***	***	0.260	0.536	0.156	3.891
HPV-SN→HPV-F	0.312**	0.002	0.243	0.517	0.109	3.174
HPV-MI→HPV-PBC	0.247*	0.025	0.002	0.214	0.065	2.235
HPV-WI→HPV-PBC	−0.043	0.634	−0.041	0.082	0.063	−0.477
HPV-PBC→HPV-P	0.265*	0.017	0.089	0.416	0.249	2.388
HPV-PBC→HPV-F	0.327**	0.005	0.124	0.345	0.204	2.835

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

FIGURE 2
Model results.
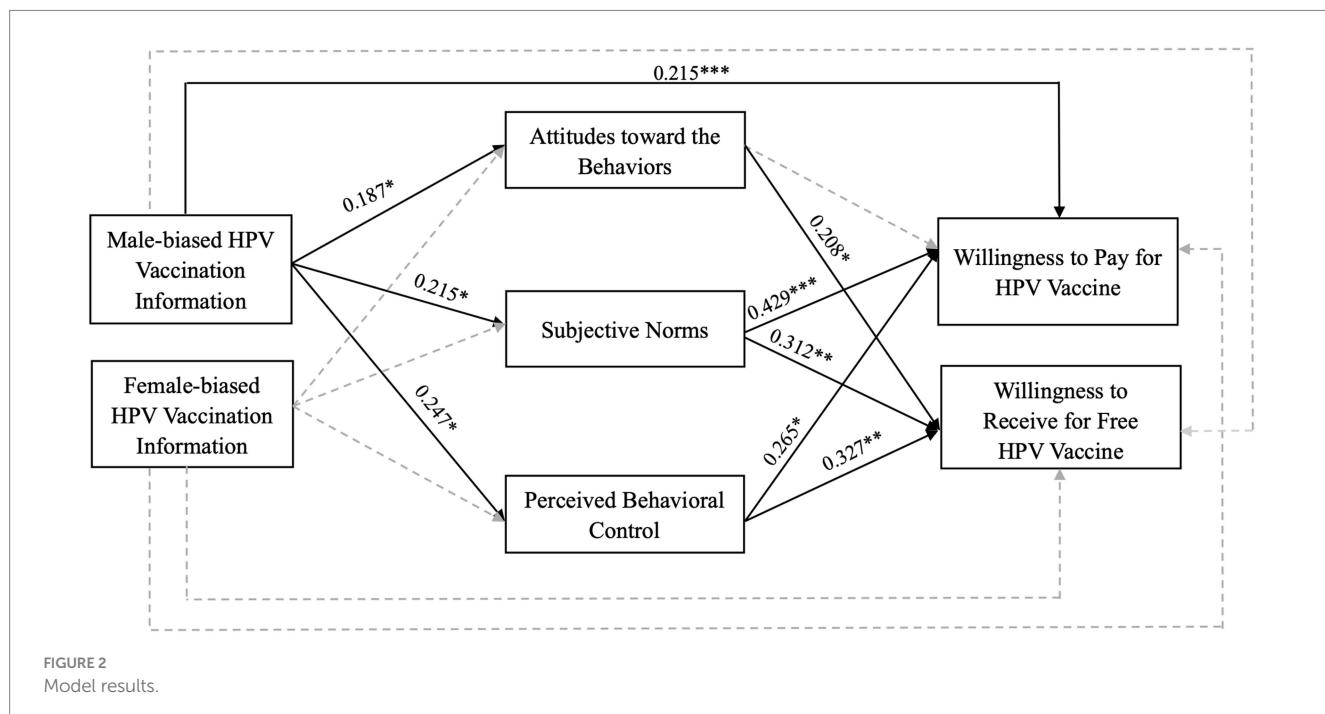
they receive affirmative messages regarding the vaccine's safety, efficacy, and its role in preventing diseases associated with persistent HPV infection. Such exposure is likely to foster more favorable attitudes towards vaccination, thus increasing the likelihood of participation. Furthermore, the prevalence of HPV-related information aimed at males, once it reaches a critical threshold within the information environment, can generate social pressure among male college students. This pressure arises

from a societal consensus that views HPV vaccination as a responsible action. Recommendations from family members, friends, and healthcare professionals can further amplify this social pressure, making individuals more inclined to pursue vaccination.

Perceived behavioral control, which is shaped by perceptions of difficulty and control (118), plays a crucial role in predicting behavioral intentions. Notably, perceived difficulty tends to exert a

stronger influence than perceived control, and it is significantly affected by key message components. When male individuals are exposed to essential information—such as guidelines on how and when to receive the HPV vaccine—they experience a reduction in perceived difficulty, leading to an enhancement in perceived behavioral control (94). This, in turn, influences their behavioral intentions. Thus, the dissemination of core informational elements, whether through publications, discussions with peers, or social media platforms, captures male attention and stimulates cognitive engagement. This cognitive engagement diminishes the perceived difficulty associated with obtaining the HPV vaccine, prompting males to modify or refine their vaccination strategies in response to new information or circumstances. As a result, HPV vaccine information specifically designed for males can more effectively reshape their perceptions and assessments regarding vaccination, significantly increasing their willingness to receive the HPV vaccine, whether at their own expense or at no cost. To improve the rate of HPV vaccination among males, it is essential to implement an information dissemination strategy that is specifically tailored to this demographic.

5.2 Challenges within a female-biased information environment

In our constructed model, certain pathways exhibit statistical insignificance. The primary group of non-significant pathways relates to the direct and indirect impacts of female-biased information on men's willingness to receive the HPV vaccine, whether provided through free vaccination programs or paid services. We suggest that entrenched gender roles may contribute to this phenomenon. Men may perceive HPV risks as predominantly associated with females, thus developing a psychological defense mechanism that renders them less receptive to health promotion messages (95). Furthermore, the framing of information appears insufficient; existing female-biased narratives surrounding the HPV vaccine, particularly in advertisements, tend to emphasize cervical cancer prevention while neglecting male-specific risks such as anal and oropharyngeal cancers. This oversight fails to address critical health concerns pertinent to men (96), ultimately hindering the ability of female-biased information to influence men's willingness to receive the HPV vaccine, irrespective of cost considerations.

From a cognitive theoretical perspective, we propose a potential explanation for this phenomenon. According to (119) Gender Schema Theory, men may often process information through the lens of socially constructed gender schemas, which leads them to prioritize information that aligns with traditional masculine roles—such as competition and goal-oriented behavior—while overlooking or minimizing information associated with femininity, such as emotional support and caregiving. This cognitive framework influences how attention is allocated. Consequently, when men encounter information regarding HPV vaccination, they tend to categorize it based on gender labels. Through these socially developed gender schemas, men may automatically classify 'HPV vaccination' as relevant primarily to the 'female domain.' When presented with the term 'female HPV,' men may activate cognitive filtering mechanisms due to a perceived incongruence with their gender role expectations, resulting in either neglect or minimization of the information. Conversely, if HPV

vaccines are framed as something that "men can also receive," this reframing can integrate the topic into acceptable male gender schemas by removing the "feminine" label—such as highlighting that men are also at risk for contracting HPV—thereby attracting their attention. Several studies examining men's cognition and attitudes towards HPV and its vaccine have identified low vaccination rates among men as being linked to misunderstandings about HPV and a belief that it is not pertinent to them (97–99). This may lead to a negligible effect on health education regarding HPV among male university students.

Additionally, several mediating variables have not demonstrated their mediating effects under the influence of female-biased information. On one hand, such information lacks content specifically targeting male audiences, which diminishes its potential to foster positive perceptions regarding the HPV vaccine among men (100). On the other hand, there is currently a deficit of male opinion leaders advocating for HPV vaccination (101). The absence of gender-specific service guidance in female-biased messaging—such as designated appointment channels for male vaccinations—further inhibits the activation of perceived behavioral control (102).

The non-significant relationship observed in the second group indicates that male-biased information regarding HPV vaccination does not have a meaningful direct impact on the willingness to accept free HPV vaccination. This lack of a direct effect may be explained by the cost sensitivity paradox. Researchers suggest that when a service is offered for free, it can lead to a "zero-price effect," whereby individuals believe that free products decrease their production costs (103). Considering this phenomenon, the current study posits that the provision of free HPV vaccinations could unintentionally heighten men's doubts about the quality of the vaccine. Unlike information targeted at females, which may encounter less skepticism due to the availability of free HPV vaccination programs for women in certain regions of China, men are likely to prioritize vaccine quality and question the effectiveness of complimentary vaccines. Consequently, their willingness to pay for vaccination can be shaped by subjective norms—such as messages from parents, friends, or healthcare professionals promoting male HPV vaccination (104)—and perceived behavioral control, which includes guidance from healthcare providers regarding the practicality and benefits of receiving vaccinations (105).

Similarly, the lack of a significant path in the third group suggests that people's attitudes towards vaccination do not significantly affect their willingness to pay for it. This is mainly due to economic considerations that diminish the impact of positive attitudes. Studies have shown that the costs associated with vaccines can reduce the positive feelings people have about vaccination and act as a major obstacle to getting vaccinated (13, 106–108). As a result, individuals' attitudes cannot reliably predict men's willingness to pay for vaccinations. We will also explore the mechanisms through which price influence vaccination behavior in subsequent discussions.

5.3 Understanding the psychological logic of male HPV vaccination behavior through effect size

In accordance with the effect size criteria established by Kline (89), where a β-value of 0.10 signifies a small effect, 0.30 indicates a medium effect, and 0.50 represents a large effect, the independent variable "male-biased HPV vaccination information" demonstrates

notable effects on three variables outlined in the TPB: attitudes toward the behavior (β = 0.187), subjective norms (β = 0.215), and perceived behavioral control (β = 0.247). All these effects are classified as small, with perceived behavioral control approaching the threshold for a medium effect. This finding implies that gender-biased information is particularly effective in influencing perceived behavioral control. According to the theoretical framework of Ajzen (54), perceived behavioral control refers to an individual's personal assessment of their capability to successfully engage in a specific behavior, which encompasses their understanding of potential challenges they may face during its execution. Male-biased information typically manifests in the form of prescriptive texts that instruct men on when, how, and where to receive vaccinations. Such information plays a vital role in shaping individuals' perceptions regarding the practical challenges associated with HPV vaccination.

The effects of the variables from TPB on individuals' willingness to vaccinate differ across various contexts. In the scenario of free vaccination, both subjective norms (β = 0.312) and perceived behavioral control (β = 0.327) exhibit moderate effect sizes, whereas attitudes toward vaccination behaviors (β = 0.208) are characterized by a small effect size. This suggests that the inclination to accept free vaccination is more significantly influenced by the interaction between social norms and individuals' self-assessed capabilities, highlighting the importance of personal beliefs regarding the necessity and method of receiving the HPV vaccine. Conversely, in scenarios involving paid vaccination, subjective norms (β = 0.429) exert a notably stronger influence compared to their impact on willingness for free vaccination. This implies that when vaccination incurs a cost, individuals' decisions are largely guided by the preferences expressed by significant others concerning their vaccination status. The effect of perceived behavioral control in this context is relatively small (β = 0.265), which may indicate that financial considerations act as a considerable impediment to engaging in vaccination behavior. For men specifically, concerns regarding economic barriers may diminish their perception of control over these barriers, leading to a lesser impact overall. Furthermore, the direct effect of "male-targeted HPV vaccination information" on "willingness to pay for HPV vaccine" is also classified as having a small effect size (β = 0.215). This underscores that in order to enhance outcomes related to vaccination willingness and behavior, it may be advantageous to leverage TPB mediating pathways for a more comprehensive influence.

5.4 The mechanism of Price influence on HPV vaccine uptake among males

The economic implications associated with HPV vaccines may negatively affect the attitudes of male college students towards vaccination. Our research demonstrates that favorable attitudes towards vaccination among male college students enhance their willingness to accept free HPV vaccinations; however, no significant difference is observed in their propensity to pay for these vaccinations. Additionally, subjective norms have a positive effect on both their willingness to pay for and to receive free HPV vaccinations, while perceived behavioral control also positively influences both aspects of vaccination willingness. Information regarding HPV vaccination targeted at men predominantly emphasizes the vaccine's protective benefits against HPV, which can foster positive perceptions and

attitudes towards vaccination. Nevertheless, in scenarios where individuals must finance their own vaccinations, economic costs emerge as a critical variable influencing decision-making, thereby diminishing the role of attitudes in vaccination choices. Consequently, attitudes towards vaccination exert varying levels of influence depending on the economic context.

In the framework of the Theory of Planned Behavior, economic factors primarily shape behavioral attitudes, rather than subjective norms or perceived behavioral control. This differentiation arises from the theory's classification of distinct psychological constructs, each of which corresponds to unique psychological processes and social influences. Behavioral attitudes inherently encompass considerations of economic factors, with economic costs forming a substantial component of overall behavioral costs. Conversely, subjective norms focus on the social pressures individuals encounter, while perceived behavioral control pertains to individuals' internal assessments of their capabilities. The latter two constructs involve behavioral costs that are largely unassociated with economic factors. Perceived behavioral control and subjective norms have a significant impact on intentions to receive both free and paid HPV vaccinations. Subjective norms pertain to individuals' perceptions of social expectations from family, friends, and society regarding male HPV vaccination. On the other hand, perceived behavioral control relates to beliefs about one's capability to carry out the vaccination behavior, including factors like access to vaccination services, affordability, and logistical considerations. These elements emerged as important predictors of vaccination intentions. Regardless of vaccination cost, male college students integrated social norms (i.e., what others expect of them) and self-assessed capabilities (i.e., whether they could practically obtain the vaccine) into their decision-making. This aligns with theoretical distinctions between subjective norms (social pressure to conform) and perceived behavioral control (perceived ease or difficulty of action), highlighting their unique roles in shaping health behaviors (120).

Information regarding HPV vaccination that is specifically targeted towards males appears to mitigate price sensitivity towards HPV vaccines by amplifying the perceived threat of HPV. Such informational cues play a significant role in shaping individuals' perceptions of disease threats, encompassing both susceptibility and severity, thereby promoting proactive health behaviors (109). The findings of this study suggest that male-targeted HPV vaccination information positively influences men's willingness to pay for HPV vaccines, while there is no notable difference in their willingness to accept free HPV vaccines. Researchers contend that the pricing of vaccines is a crucial determinant of vaccine hesitancy among the public (68), which subsequently impacts individuals' choices regarding the timing of health behaviors, including HPV vaccination. The informational landscape surrounding HPV vaccination for men, influenced by personal experiences, peer discussions, and social media, encompasses essential knowledge such as the appropriate age for vaccination and the procedural aspects, as well as more nuanced medical insights regarding the significance of male HPV vaccination, including prevalence, incidence, and transmission pathways. This context facilitates a shift in male college students' perceptions of the health threats posed by HPV, bolstering their awareness of both susceptibility and severity of the virus, which in turn affects their perceptions of the cost-effectiveness of vaccination and enhances their willingness to invest in out-of-pocket vaccinations.

5.5 Practical implications

The findings of this study suggest that tailored information regarding HPV vaccination aimed at men is likely to increase their willingness to receive the vaccine. Therefore, it is essential to enhance the dissemination of various forms of male-oriented HPV vaccine promotion, particularly through media channels, with a focus on social media platforms. By emphasizing the male perspective, specifically highlighting men's vulnerability to HPV-related diseases, we can facilitate a greater understanding among men regarding the highly transmissible nature of HPV, which affects individuals regardless of gender. This approach also serves to underscore the effectiveness of the HPV vaccine, thereby enhancing men's awareness and acceptance of vaccination. Moreover, implementing a range of engaging and entertaining promotional strategies could elevate the significance of HPV awareness while mitigating perceived risks associated with vaccination. For instance, in China, health campaigns such as the "Playful Immunization Promotion" initiative (110), which utilized platforms like TikTok, have demonstrated effectiveness in capturing public interest and increasing vaccination rates. Similar initiatives could be strategically developed to specifically target male audiences.

Cost sensitivity has emerged as a significant barrier to HPV vaccination among men, as identified in this study. The introduction of economic stimuli like financial subsidies for vaccine administration could serve as one of the essential strategies for increasing vaccination uptake. For instance, several regions in China, including Guangdong and Zhejiang provinces, have enacted policies that provide free HPV vaccinations to certain age groups of women (111). This provision of free vaccines is closely associated with female identity, and the messaging surrounding "free HPV vaccinations" does not directly engage men; rather, it indirectly motivates men to consider accepting free HPV vaccinations through the influence of behavioral attitudes, subjective norms, and perceived behavioral control. The sex-biased information regarding HPV vaccination significantly shapes the health behaviors of male college students, and the underlying structural factors must not be overlooked. The implementation of free vaccination initiatives, including those for HPV, is contingent upon various factors such as local fiscal allocations, public health policies, and vaccine availability. Presently, free HPV vaccination policies primarily target specific adolescent age groups rather than a wider demographic. To enhance the willingness of males, particularly male college students, to pay for HPV vaccination, it is essential to provide them with repeated exposure to HPV vaccination information that is specifically designed for a male audience.

6 Conclusion and limitations

This research employed a questionnaire survey approach to explore the sex-biased information environment related to HPV vaccination, delving into whether such sex-specific information has a notable effect on male individuals' health behavior decisions, particularly their intention to receive the HPV vaccine. The study also extended the Theory of Planned Behavior model to analyze the underlying mechanisms. The results indicate that information favoring males exerts a greater influence on the vaccination intentions of male college students compared to information favoring females.

In particular, the cost of the vaccine was identified as a significant barrier for male college students seeking vaccination. Additionally, consistently delivering HPV vaccination information that is tailored towards males effectively diminishes their sensitivity to price, thereby increasing their willingness to receive the vaccine.

In recent years, marketing strategies for HPV vaccines have increasingly associated the virus predominantly with women. However, it is important to recognize that HPV infection is not inherently sex-specific. To effectively combat cervical cancer, widespread vaccination among women is vital; nevertheless, the participation of men in vaccination efforts is equally important. Given that men exhibit a significant prevalence of HPV infection, it is essential to actively promote HPV vaccination initiatives targeting male populations. To enhance men's willingness to receive the vaccine and facilitate their participation, it is crucial to improve the dissemination of information tailored specifically for men within the predominantly female-focused HPV vaccination discourse. This approach will contribute to safeguarding the sexual health of sexually active men, including those who engage in sexual activities with other men, within China's substantial male demographic of 720 million individuals.

This study remains in its initial phases, and further investigation will enhance the understanding of the research questions posed. Currently, the article presents several limitations. First, the dataset utilized is relatively small, leading to concerns regarding its representativeness. Also, the study employs convenience sampling, which raises concerns regarding the representativeness of the sample. Specifically, web users may not accurately reflect the broader population, potentially resulting in selection bias. Additionally, the reliance on voluntary participation often leads to the selection of participants based on their interests rather than randomization, further introducing bias into the findings. Moreover, this study is based on self-reported data from participants, which may be susceptible to recall bias and the influence of social desirability. To more accurately assess the impact of sex-biased information while minimizing subjective interference, subsequent studies could implement experimental methodologies. To improve the generalizability of the findings, future research should prioritize the expansion of both the scope and scale of data collection.

Second, the current study employs a single-item measure to assess the willingness of the sample to accept both paid and free vaccines. Although this approach has been utilized in previous research (e.g., 121), the use of a single-item assessment carries inherent risks, particularly concerning reliability issues. This methodology may consequently introduce biases into the findings of this research. Furthermore, this study did not provide illustrative examples of male-biased and female-biased HPV vaccination-related information on social media to aid participants in completing the questionnaire. This omission may have hindered their understanding of the concept of "information related to male/female HPV vaccination." Consequently, this lack of clarity could have led to a certain degree of bias in the participants' evaluations of the items presented.

Moreover, this study is limited to a sample exclusively composed of males, thus lacking a comparative analysis that includes different gender perspectives. To achieve a more holistic understanding of the effects of sex-biased information, future research should include both male and female participants, comparing sex-specific reactions to

such information to uncover potential differential impacts. Research design employing the aforementioned approach will contribute to a deeper comprehension of the role of sex-biased information in health communication and provide a scientific foundation for developing more effective health intervention strategies. In addition, our research did not assess the perceived relative cost of the HPV vaccine. Exploring this dimension may yield additional noteworthy and significant insights.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

T-CW: Formal analysis, Project administration, Supervision, Writing – original draft, Writing – review & editing. M-JZ: Data curation, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing – original draft, Writing – review & editing. HZ: Conceptualization, Project administration,

Writing – original draft, Writing – review & editing, Supervision, Validation, Funding acquisition.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Reference

1. Baptista AD, Simão CX, Santos VCGD, Melgaço JG, Cavalcanti SMB, Fonseca SC, et al. Knowledge of human papillomavirus and pap test among Brazilian university students. *Rev Assoc Med Bras*. (2019) 65:625–32. doi: 10.1590/1806-9282.65.5.625

2. Bruni L, Albero G, Rowley J, Alemany L, Arbyn M, Giuliano AR, et al. Global and regional estimates of genital human papillomavirus prevalence among men: a systematic review and meta-analysis. *Lancet Glob Health*. (2023) 11:e1345–62. doi: 10.1016/S2214-109X(23)00305-4

3. Okunade KS. Human papillomavirus and cervical cancer. *J Obstet Gynaecol*. (2020) 40:602–8. doi: 10.1080/01443615.2019.1634030

4. Serrano-Villar S, Tincati C, Raju SC, Sáenz JS, Moreno E, Bargiela R, et al. Microbiome-derived cobalamin and succinyl-CoA as biomarkers for improved screening of anal cancer. *Nat Med*. (2023) 29:1738–49. doi: 10.1038/s41591-023-02407-3

5. Melo BADC, Vilar LG, Oliveira NRD, Lima POD, Pinheiro MDB, Domingueti CP, et al. Human papillomavirus infection and oral squamous cell carcinoma—a systematic review. *Braz J Otorhinolaryngol*. (2021) 87:346–52. doi: 10.1016/j.bjorl.2020.10.017

6. Zeng Z, Austin RM, Wang L, Guo X, Zeng Q, Zheng B, et al. Nationwide prevalence and genotype distribution of high-risk human papillomavirus infection in China: analysis of 2,458,227 test results from 29 Chinese provinces. *Am J Clin Pathol*. (2022) 157:718–23. doi: 10.1093/ajcp/aqab181

7. Bustamante G, Reiter PL, McRee AL. Cervical cancer screening among sexual minority women: findings from a national survey. *Cancer Causes Control*. (2021) 32:911–7. doi: 10.1007/s10552-021-01442-0

8. Dykens JA, Peterson CE, Holt HK, Harper DM. Gender neutral HPV vaccination programs: reconsidering policies to expand cancer prevention globally. *Front Public Health*. (2023) 11:1067299. doi: 10.3389/fpubh.2023.1067299

9. Goldfarb JA, Comber JD. Human papillomavirus (HPV) infection and vaccination: a cross-sectional study of college students' knowledge, awareness, and attitudes in Villanova. *PA Vacc. X*. (2022) 10:100141. doi: 10.1016/j.jvacx.2022.100141

10. Lu Y, Li P, Luo G, Liu D, Zou H. Cancer attributable to human papillomavirus infection in China: burden and trends. *Cancer*. (2020) 126:3719–32. doi: 10.1002/cncr.32986

11. Grandahl M, Nevéus T. Barriers towards HPV vaccinations for boys and young men: a narrative review. *Viruses*. (2021) 13:1644. doi: 10.3390/v13081644

12. Ran H, Chen Y, Gao J, Guo H, Peng S. Low awareness of HPV infection and willingness of HPV vaccination among Chinese male college students in the east of China. *Front Public Health*. (2022) 10:971707. doi: 10.3389/fpubh.2022.971707

13. Spînu AD, Anghel RF, Marcu DR, Iorga DL, Cherciu A, Mischianu DLD. HPV vaccine for men: where to? *Exp Ther Med*. (2021) 22:1266–6. doi: 10.3892/etm.2021.10701

14. Zhao Y, Xin X, Deng H, Xu J, Weng W, Zhang M, et al. Improving the acceptability of human papillomavirus vaccines among men who have sex with men according to the associated factors: a systematic review and meta-analysis. *Front Pharmacol*. (2021) 12:600273. doi: 10.3389/fphar.2021.600273

15. Li Y, Guo Y, Lin H. Feminized virus: a content analysis of social Media's representation of HPV vaccine. *Soc Media*. (2022) 8:1. doi: 10.1177/20563051221104232

16. Jin J, Han Y. Understanding cultural perceptions of sexuality in China and their influence on human papillomavirus vaccine hesitancy. *Front Public Health*. (2025) 12:1462722. doi: 10.3389/fpubh.2024.1462722

17. Liu Z, Zou W. Breaking taboos in women's reproductive health: the communication strategies used by top OB/GYN influencers in Chinese social media. *Health Commun*. (2024) 39:685–96. doi: 10.1080/10410236.2023.2181677

18. Xiao X, Su Y. Still a "female problem": a framing analysis of the human papillomavirus (HPV) vaccine in Chinese online news. *Chin J Commun*. (2020) 13:275–92. doi: 10.1080/17544750.2020.1714683

19. Alhusayn KO, Alkhenizan A, Abdulkarim A, Sultana H, Alsulaiman T, Alendijani Y. Attitude and hesitancy of human papillomavirus vaccine among Saudi parents. *J Family Med Prim Care*. (2022) 11:2909–16. doi: 10.4103/jfmpc.jfmpc_2377_21

20. Sallam M, Al-Mahzoum K, Eid H, Assaf AM, Abdaljaleel M, Al-Abbadi M, et al. Attitude towards HPV vaccination and the intention to get vaccinated among female university students in health schools in Jordan. *Vaccine*. (2021) 9:1432. doi: 10.3390/vaccines9121432

21. Smolarczyk K, Duszewska A, Drozd S, Majewski S. Parents' knowledge and attitude towards HPV and HPV vaccination in Poland. *Vaccine*. (2022) 10:228. doi: 10.3390/vaccines10020228

22. Bogka E, Naoum P, Pavi E, Athanasakis K. What influences parents on their decision to vaccinate their daughters against HPV? *J Pediatr Adolesc Gynecol*. (2024) 37:396–401. doi: 10.1016/j.jpag.2024.02.001

23. Wang H, Wang X, Chen P, Xu H, Liu Y, Kang R, et al. Factors influencing Chinese female college students' willingness to receive human papillomavirus vaccine: a cross-sectional study based on information-motivation-behavioral skills model. *Hum Vaccin Immunother*. (2022) 18:2140550. doi: 10.1080/21645515.2022.2140550

24. Gauna F, Verger P, Fressard L, Jardin M, Ward JK, Peretti-Watel P. Vaccine hesitancy about the HPV vaccine among French young women and their parents: a telephone survey. *BMC Public Health*. (2023) 23:628. doi: 10.1186/s12889-023-15334-2

25. Kim HW, Lee EJ, Lee YJ, Kim SY, Kim Y, Lee JL. Knowledge, attitudes, and perceptions associated with HPV vaccination among female Korean and Chinese university students. *BMC Womens Health*. (2022) 22:51. doi: 10.1186/s12905-022-01624-1

26. Kang EH, Lee EH. Updated theory of planned behavior in predicting parents' intentions to vaccinate their sons in elementary school against human papillomavirus. *Res Commun Public Health Nurs*. (2021) 32:195–204. doi: 10.12799/jkachn.2021.32.2.195

27. Yi Y, Xiu S, Shi N, Huang Y, Zhang S, Wang Q, et al. Perceptions and acceptability of HPV vaccination among parents of female adolescents 9–14 in China: a cross-sectional survey based on the theory of planned behavior. *Hum Vaccin Immunother*. (2023) 19:2225994. doi: 10.1080/21645515.2023.2225994

28. Khoa BT. The role of self-efficacy and firm size in the online advertising services continuous adoption intention: theory of planned behavior approach. *J Open Innov: Technol Mark Complex*. (2023) 9:100025. doi: 10.1016/j.joitmc.2023.100025

29. Colzani E, Johansen K, Johnson H, Celentano LP. Human papillomavirus vaccination in the European Union/European economic area and globally: a moral dilemma. *Eur Secur*. (2021) 26:2001659. doi: 10.2807/1560-7917.ES.2021.26.50.2001659

30. Malvolti S, Soble A, Bloem P, LaMontagne DS, Aggarwal R, Pitisuttithum P, et al. The global demand and supply balance of the human papillomavirus vaccine: implications for the global strategy for the elimination of cervical Cancer. *Vaccine*. (2023) 12:4. doi: 10.3390/vaccines12010004

31. Foster S, Carvallo M, Lee J, Fisher R, Traxler H. An implication of impurity: the impact of feminine honor on human papillomavirus (HPV) screenings and the decision to authorize daughter's HPV vaccinations. *Stigma Health*. (2021) 6:216–27. doi: 10.1037/sah0000230

32. Ginjupalli R, Mundaden R, Choi Y, Herfel E, Oketch SY, Watt MH, et al. Developing a framework to describe stigma related to cervical cancer and HPV in western Kenya. *BMC Womens Health*. (2022) 22:39. doi: 10.1186/s12905-022-01619-y

33. Morse RM, Brown J, Gage JC, Prieto BA, Jurczuk M, Matos A, et al. "Easy women get it": pre-existing stigma associated with HPV and cervical cancer in a low-resource setting prior to implementation of an HPV screen-and-treat program. *BMC Public Health*. (2023) 23:2396. doi: 10.1186/s12889-023-17324-w

34. Wong LP, Wong PF, Megat Hashim MMAA, Han L, Lin Y, Hu Z, et al. Multidimensional social and cultural norms influencing HPV vaccine hesitancy in Asia. *Hum Vaccin Immunother*. (2020) 16:1611–22. doi: 10.1080/21645515.2020.1756670

35. Zhang L, Yang J, Cao Y, Kang W. Sociocultural–psychological predictors influencing parents' decision-making regarding HPV vaccination for their adolescent daughters in mainland China: an extended TPB model. *Front Public Health*. (2023) 10:1035658. doi: 10.3389/fpubh.2022.1035658

36. Hu J, Whyke TW, Lopez-Mugica J. Investigating media coverage and public perceptions of the HPV vaccine in China–a content analysis of Weibo posts. *Sex Cult*. (2023) 27:363–88. doi: 10.1007/s12119-022-10017-3

37. Su X. Content analysis of HPV vaccine messages on Chinese social media. *J Messenger*. (2020) 12:63–73. doi: 10.26623/themessenger.v12i1.1814

38. Li AJ, Kyesi F, Mwengee W, Mphuru A, Giattas MR, Shayo B, et al. Impact of the human papillomavirus (HPV) vaccine supply shortage on Tanzania's national HPV vaccine introduction. *Vaccine*. (2022) 40:A26–9. doi: 10.1016/j.vaccine.2021.01.036

39. Zhang H, Sun Y, Shi S eds. Modern English-Chinese psychological vocabulary. Beijing: China Light Industry Press (2016). 230 p.

40. Mehta A, Markman B, Rodriguez-Cintron W. Don't jump down my throat: gender gap in HPV vaccinations risk long term cancer threats. *Expert Rev Vaccines*. (2022) 21:1045–53. doi: 10.1080/14760584.2022.2075850

41. Sobierajski T, Małecka I, Augustynowicz E. Feminized vaccine? Parents' attitudes toward HPV vaccination of adolescents in Poland: a representative study. *Hum Vaccin Immunother*. (2023) 19:2186105. doi: 10.1080/21645515.2023.2186105

42. Friedman S. E., Schmer-Galunder S., Chen A., Goldman R. P., Ausman M. (2020). Gender gaps correlate with gender Bias in social media word Embeddings. In CogSci. Available online at: https://escholarship.org/uc/item/45d4b391

43. Hale M.-L. (2023). -Gender based biases and discrimination—new developments in social media contexts- [doctoral thesis, Unilu—University of Luxembourg]. ORBilu-University of Luxembourg. Available online at: https://orbilu.uni.lu/handle/10993/58871

44. Horvát EÁ, González-Bailón S. Quantifying gender disparities and bias online: editors' introduction to "gender gaps in digital spaces" special issue. Journal of computer-mediated. *Communication*. (2024) 29:zmad054. doi: 10.1093/jcmc/zmad054

45. Nascimento FR, Cavalcanti GD, Da Costa-Abreu M. Unintended bias evaluation: an analysis of hate speech detection and gender bias mitigation on social media using ensemble learning. *Expert Syst Appl*. (2022) 201:117032. doi: 10.1016/j.eswa.2022.117032

46. Shen K, Ding L, Kong L, Liu X. From physical space to cyberspace: recessive gender biases in social media mirror the real world. *Cities*. (2024) 152:105149. doi: 10.1016/j.cities.2024.105149

47. Ulloa R, Richter AC, Makhortykh M, Urman A, Kacperski CS. Representativeness and face-ism: gender bias in image search. *New Media Soc*. (2024) 26:3541–67. doi: 10.1177/14614448221100699

48. Lamer SA, Dvorak P, Biddle AM, Pauker K, Weisbuch M. The transmission of gender stereotypes through televised patterns of nonverbal bias. *J Pers Soc Psychol*. (2022) 123:1315–35. doi: 10.1037/pspi0000390

49. Azmy K, Rahman F, Muntasir M. Discourse in gender studies: how language shapes gender narratives? *Saree Res Gender Stud*. (2024) 6:25–38. doi: 10.47766/saree.v6i1.2553

50. Montano MRF, Opeña ARC, Miranda MMS. Language as an agent of change: promoting gender fairness. *Technium Soc Sci J*. (2024) 53:336–54. doi: 10.47577/tssj.v53i1.10313

51. Šumiga J. (2024). Linguistic gender-related biases (doctoral dissertation, University of Rijeka. Faculty of Humanities and Social Sciences. Department of English Language and Literature). Available online at: https://urn.nsk.hr/urn:nbn:hr:186:706872

52. Homan P. Structural sexism and health in the United States: a new perspective on health inequality and the gender system. *Am Sociol Rev*. (2019) 84:486–516. doi: 10.1177/0003122419848723

53. Tung IL, Machalek DA, Garland SM. Attitudes, knowledge and factors associated with human papillomavirus (HPV) vaccine uptake in adolescent girls and young women in Victoria, Australia. *PLoS one*. (2016) 11:e0161846. doi: 10.1371/journal.pone.0161846

54. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process*. (1991) 50:179–211. doi: 10.1016/0749-5978(91)90020-T

55. James R, Fisher JR, Carlos-Grotjahn C, Boylan MS, Dembereldash B, Demissie MZ, et al. Gender bias and inequity holds women back in their conservation careers. *Front Environ Sci*. (2023) 10:1056751. doi: 10.3389/fenvs.2022.1056751

56. Yates J, Plagnol AC. Female computer science students: a qualitative exploration of women's experiences studying computer science at university in the UK. *Educ Inf Technol*. (2022) 27:3079–105. doi: 10.1080/14635240.2021.1995776

57. Yang Z, Hao X. The dictionary of psychology = the dictionary of psychology. Shanghai: Shanghai Dictionaries Publishing House (2016). 516 p.

58. Kyrrestad H, Mabille G, Adolfsen F, Koposov R, Martinussen M. Gender differences in alcohol onset and drinking frequency in adolescents: an application of the theory of planned behavior. *Drugs Educ Prev Policy*. (2022) 29:21–31. doi: 10.1080/09687637.2020.1865271

59. Xu L, Zhang J, Ding Y, Zheng J, Sun G, Zhang W, et al. Understanding the role of peer pressure on engineering students' learning behavior: a TPB perspective. *Front Public Health*. (2023) 10:1069384. doi: 10.3389/fpubh.2022.1069384

60. Xie H, Zhu HY, Jiang NJ, Yin YN. Awareness of HPV and HPV vaccines, acceptance to vaccination and its influence factors among parents of adolescents 9 to 18 years of age in China: a cross-sectional study. *J Pediatr Nurs*. (2023) 71:73–8. doi: 10.1016/j.pedn.2023.03.007

61. Orji C, Brown CM, Barner J, Moczygemba L, Morales-Campos D. Determinants of human papillomavirus (HPV) vaccination intentions among young adult college students using the theory of planned behavior. *J Am Coll Heal*. (2024) 1-14:1–14. doi: 10.1080/07448481.2024.2325935

62. Degarege A, Watanabe-Galloway S, Mansilla K, Sileshi RM, Peters ES. Evaluation of theoretical frameworks to detect correlates of HPV vaccination in the Midwest, US. *Using Struct Equation Model Vacc*. (2023) 11:1856. doi: 10.3390/vaccines11121856

63. Harper K, Short MB, Bistricky S, Kusters IS. 1-2-3! Catch-up for HPV: a theoretically informed pilot intervention to increase HPV vaccine uptake among young adults. *Am J Health Educ*. (2023) 54:119–34. doi: 10.1080/19325037.2022.2163005

64. Li L, Li J. Factors affecting young Chinese women's intentions to uptake human papillomavirus vaccination: an extension of the theory of planned behavior model. *Hum Vaccin Immunother*. (2020) 16:3123–30. doi: 10.1080/21645515.2020.1779518

65. Rafeie L, Vizeshfar F, Nick N. The effect of education based on planned behavior theory on women's knowledge and attitudes about human papillomavirus. *Sci Rep*. (2024) 14:18581. doi: 10.1038/s41598-024-69340-4

66. Wang H, Xu Y, Zhang H, Chen N. Determinants of HPV vaccine uptake intentions in Chinese clinical interns: an extended theory of planned behavior approach. *Front Public Health*. (2024) 12:1345530. doi: 10.3389/fpubh.2024.1345530

67. Yarmohammadi S, Mousavi S, Ajori L, Mohammadzade F, Marashi T. Impact of attitude, subjective norms and perceived behavioral control on health behaviors among women with low-risk HPV, and the mediating role of behavioral intention: an interventional study. *Int J Health Promot Educ*. (2023) 61:70–82. doi: 10.1080/14635240.2021.1995776

68. Wang J, Ji Q, Dong S, Zhao S, Li X, Zhu Q, et al. Factors influencing vaccine hesitancy in China: a qualitative study. *Vaccine*. (2021) 9:1291. doi: 10.3390/vaccines9111291

69. Zhang P, Zhang Q, Guan H, Fan K, Bi X, Huang Y, et al. Who is more likely to hesitate to accept COVID-19 vaccine: a cross-sectional survey in China. *Expert Rev Vaccines*. (2022) 21:397–406. doi: 10.1080/14760584.2022.201981

70. Lu X, Ji M, Wagner AL, Huang W, Shao X, Zhou W, et al. Willingness to pay for HPV vaccine among female health care workers in a Chinese nationwide survey. *BMC Health Serv Res*. (2022) 22:1324. doi: 10.1186/s12913-022-08716-6

71. Zhou L, Gu B, Xu X, Li Y, Cheng P, Huo Y, et al. On imported and domestic human papillomavirus vaccines: cognition, attitude, and willingness to pay in Chinese medical students. *Front Public Health*. (2022) 10:863748. doi: 10.3389/fpubh.2022.863748

72. Zou M, Liu H, Liu H, Wang M, Zou Z, Zhang L. Vaccinating women previously treated for human papillomavirus-related cervical precancerous lesions is highly cost-effective in China. *Front Immunol*. (2023) 14:1119566. doi: 10.3389/fimmu.2023.1119566

73. Fazio RH, Zanna M. On the predictive validity of attitudes: the roles of direct experience and confidence. *J Pers*. (1978) 46:228–43. doi: 10.1111/j.1467-6494.1978.tb00177.x

74. Li S, Hao J, Su Y, Zhan H, Zhou N, Qiu Y, et al. COVID-19 vaccine hesitancy and influencing factors among Chinese hospital staff: a cross-sectional study. *Sci Rep*. (2024) 14:4369. doi: 10.1038/s41598-024-55001-z

75. Liang Y, Bai X, Liu X, Zhang Z, Pang X, Nie L, et al. Hepatitis B vaccination coverage rates and associated factors: a community-based, cross-sectional study conducted in Beijing, 2019–2020. *Vaccine*. (2021) 9:1070. doi: 10.3390/vaccines9101070

76. Liu C, Nicholas S, Wang J. The association between protection motivation and hepatitis b vaccination intention among migrant workers in Tianjin, China: a cross-sectional study. *BMC Public Health*. (2020) 20:1219. doi: 10.1186/s12889-020-09292-2

77. Liu PL, Ao SH, Zhao X, Zhang L. Associations between COVID-19 information acquisition and vaccination intention: the roles of anticipated regret and collective responsibility. *Health Commun*. (2023) 38:2198–209. doi: 10.1080/10410236.2022.2059801

78. Arabyat RM, Nusair MB, Al-Azzam SI, Amawi HA, El-Hajji FD. Willingness to pay for COVID-19 vaccines: applying the health belief model. *Res Soc Adm Pharm*. (2023) 19:95–101. doi: 10.1016/j.sapharm.2022.09.003

79. Kasting ML, Christy SM, Stout ME, Zimet GD, Mosher CE. Attitudinal correlates of HPV vaccination in college women. *Clin Nurs Res*. (2022) 31:826–35. doi: 10.1177/10547738211045227

80. Samrawit T. (2023). Mother's intention towards human Pappiloma virus vaccination among their 9–14 years of daughters in Debretabor town administration, north West Ethiopia 2022, using theory of planned behavior (doctoral dissertation).

81. Zomordi G, Moradi M, Hasanzadeh M, Ghavami V. The effect of education based on the theory of planned behavior on the intention of vaccination against human papillomavirus in female students: a controlled educational trial. *J Educ Health Prom*. (2022) 11:237. doi: 10.4103/jehp.jehp_1145_21

82. Earley PC. Social loafing and collectivism: a comparison of the United States and the People's republic of China. *Adm Sci Q*. (1989) 34:565. doi: 10.2307/2393567

83. Liu Y, Kuang K. Predictors of COVID-19 vaccination hesitancy in China: a meta-analysis. *Public Health*. (2023) 220:135–41. doi: 10.1016/j.puhe.2023.05.009

84. Su Y, Li S, Huang F, Xue J, Zhu T. Exploring the influencing factors of COVID-19 vaccination willingness among young adults in China. *Int J Environ Res Public Health*. (2023) 20:3960. doi: 10.3390/ijerph20053960

85. Chang J, Zhu S, Zhang Y, Carvalho N, Xu S, Lu Y, et al. Determinants of parental demand of human papillomavirus vaccination for adolescent daughters in China: contingent valuation survey. *Int J Health Plann Manag*. (2024) 39:1456–81. doi: 10.1002/hpm.3818

86. Gong X, Xu J, He Y, Zou G, Liu J. Socioeconomic inequalities in human papillomavirus knowledge and vaccine uptake: evidence from a cross-sectional study in China. *Front Public Health*. (2024) 12:1399192. doi: 10.3389/fpubh.2024.1399192

87. Wang Q, Zhou F, Zhang W, Tang C. A study of parental decision-making over the vaccination of girls, based on the protection motivation theory and the elaboration likelihood model. *Front Public Health*. (2022) 10:1024399. doi: 10.3389/fpubh.2022.1024399

88. Xiang J, Han L, Fan Y, Feng B, Wu H, Hu C, et al. Prevalence and genotype distribution of human papillomavirus among attendees at a sexually transmitted diseases Clinic in Urban Tianjin, China. *Int J Gen Med*. (2021) 14:1983–90. doi: 10.2147/IJGM.S308215

89. Kline RB. *Principles and practice of structural equation modeling*. New York: The Guilford Press (2023).

90. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychol Methods*. (1996) 4:84–99. doi: 10.1037/1082-989X.4.1.84

91. Sawdey MD, Hancock L, Messner M, Prom-Wormley EC. Assessing the association between E-cigarette use and exposure to social Media in College Students: a cross-sectional study. *Subst Use Misuse*. (2017) 52:1910–7. doi: 10.1080/10826084.2017.1319390

92. Xu B, Lu F, Sun Y, Wu Y, Wang D. Analysis of influencing factors of EV71 vaccination intention of parents of nursery children based on planned behavior theory. *Modern Prev Med*. (2020) 47:3518–3521+3540. (in Chinese). doi: 10.20043/j.cnki.mpm.2020.19.017

93. Kaiser HF. An index of factorial simplicity. *Psychometrika*. (1974) 39:31–6. doi: 10.1007/BF02291575

94. Cheung S. (2000). Perceived behavioral control.

95. Horio F, Ikeda T, Zaitsu M, Takebe D, Tabata A, Matsukura M, et al. Knowledge and awareness of human papillomavirus vaccination and cervical cancer among men and women in Japan: a questionnaire survey. *Asian Pacific J Cancer Prev*. (2023) 24:1063–71. doi: 10.31557/APJCP.2023.24.3.1063

96. Fitzgerald E. Gendered responsibility: a critique of HPV vaccine advertisements, 2006–2016. J White-Farnham, B Siegel Finer and C Molloy editors. In Women's health advocacy. New York: Routledge (2019). 102–110.

97. Jia S, Pan B, Hong D, Zhang Q, Jiang H, Hong Y, et al. A survey of potential acceptance of 9-valent HPV vaccine among Chinese male college students. *Hum Vaccin Immunother*. (2023) 19:2272533. doi: 10.1080/21645515.2023.2272533

98. Lindsay AC, Delgado D, Valdez MJ, Restrepo E, Guzman YM. "I don't think he needs the HPV vaccine cause boys can't have cervical cancer": a qualitative study of Latina mothers'(mis) understandings about human papillomavirus transmission, associated cancers, and the vaccine. *J Cancer Educ*. (2022) 37:370–8. doi: 10.1007/s13187-020-01824-z

99. Nkwonta CA, Dawson RM, Adegboyega A. "I don't think I have a chance to get it": international university student HPV knowledge and preventive behaviors. *J Am Coll Heal*. (2022) 70:240–7. doi: 10.1080/07448481.2020.1740232

100. Laserson AK, Oliffe JL, Krist J, Kelly MT. HPV vaccine and college-age men: a scoping review. *Am J Mens Health*. (2020) 14:1557988320973826. doi: 10.1177/1557988320973826

101. Janette Maria P, Ikhsano A, Sutjipto AM. Accelerating diffusion innovation using Celebritites as opinion leader Millenials: HPV vaccines in Indonesia. *Malaysian J Commun*. (2020) 36:32–48. doi: 10.17576/JKMJC-2020-3603-03

102. Amantea C, Foschi N, Gavi F, Borrelli I, Rossi MF, Spuntarelli V, et al. HPV vaccination adherence in working-age men: a systematic review and meta-analysis. *Vaccine*. (2023) 11:443. doi: 10.3390/vaccines11020443

103. Ariely D, Shampan'er K. How small is zero price?: the true value of free products. *Federal Reserve Bank of Boston*. (2006) 6:1–44. doi: 10.2139/ssrn.951742

104. Stout ME, Christy SM, Winger JG, Vadaparampil ST, Mosher CE. Self-efficacy and HPV vaccine attitudes mediate the relationship between social norms and intentions to receive the HPV vaccine among college students. *J Community Health*. (2020) 45:1187–95. doi: 10.1007/s10900-020-00837-5

105. Koskan A, Stecher C, Helitzer D. College males' behaviors, intentions, and influencing factors related to vaccinating against HPV. *Hum Vaccin Immunother*. (2021) 17:1044–51. doi: 10.1080/21645515.2020.1819101

106. Olusanya OA, Tomar A, Thomas J, Alonge K, Wigfall LT. Application of the theoretical domains framework to identify factors influencing catch-up HPV vaccinations among male college students in the United States: a review of evidence and recommendations. *Vaccine*. (2023) 41:3564–76. doi: 10.1016/j.vaccine.2023.04.071

107. Tao Y, Shao H, Zhang T, Pu J, Tang C. Factors influencing men's attitudes toward HPV vaccination in males included in the Chinese national immunization program. *Vaccine*. (2022) 10:1054. doi: 10.3390/vaccines10071054

108. Xie X, Wang R, Gou Z, Chen S. A comparative study of gender differences in healthy office building strategies. *Front Psychol*. (2023) 14:1146260. doi: 10.3389/fpsyg.2023.1146260

109. Ranjit YS, Shin H, First JM, Houston JB. COVID-19 protective model: the role of threat perceptions and informational cues in influencing behavior. *J Risk Res*. (2021) 24:449–65. doi: 10.1080/13669877.2021.1887328

110. Wu J, Yang R. Campaign governance and playfulness: unraveling Chinese HPV immunization promotion efforts on Douyin. *Healthcare*. (2024) 12:1657. doi: 10.3390/healthcare12161657

111. Huang Y, Zhang D, Yin L, Zhao J, Li Z, Lu J, et al. Modeling the health impact and cost-effectiveness of a combined schoolgirl HPV vaccination and cervical Cancer screening program in Guangdong Province, China. *Children*. (2024) 11:103. doi: 10.3390/children11010103

112. Rujumba, J, Akugizibwe, M, Basta, NE, Banura, C. (2021). Why don't adolescent girls in a rural Uganda district initiate or complete routine 2-dose HPV vaccine series: Perspectives of adolescent girls, their caregivers, healthcare workers, community health workers and teachers. *PloS one*, 16:e0253735. doi: 10.1371/journal.pone.0253735

113. Doshi, RH, Casey, RM, Adrien, N, Ndiaye, A, Brennan, T, Roka, JL, et al (2022). Feasibility and acceptability of nationwide HPV vaccine introduction in Senegal: Findings from community-level cross-sectional surveys, 2020. *PLOS Global Public Health*, 2:e0000130. doi: 10.1371/journal.pgph.0000130

114. Cislaghi, B, Heise, L. (2020). Gender norms and social norms: differences, similarities and why they matter in prevention science. *Soc Health Illness*, 42:407–422. doi: 10.1111/1467-9566.13008

115. Triandis, H. (1988). Collectivism v. individualism: A reconceptualisation of a basic concept in cross-cultural social psychology. In *Cross-cultural studies of personality, attitudes and cognition*. 60–95. London: Palgrave Macmillan UK.

116. Reed, MB, Lange, JE, Ketchie, JM, and Clapp, JD. (2007). The relationship between social identity, normative information, and college student drinking. *Soc Influence*, 2:269–294. doi: 10.1080/15534510701476617

117. James, G, Witten, D, Hastie, T, Tibshirani, R. (2013). Linear Regression. In: G Allen, R De Veaux, R Nugent editors, *An introduction to statistical learning*. 102. New York: springer.

118. Rodgers, WM, Conner, M, Murray, TC. (2008). Distinguishing among perceived control, perceived difficulty, and self-efficacy as determinants of intentions and behaviours. *British J Social Psychol*, 47:607–630. doi: 10.1348/014466607X248903

119. Bem, SL. (1981). The BSRI and gender schema theory: A reply to Spence and Helmreich. *Psychol Rev*, 88:369–371. doi: 10.1037/0033-295X.88.4.369

120. Meijer, MA, Brabers, AE, de Jong, JD. (2023). Social context matters: The role of social support and social norms in support for solidarity in healthcare financing. *Plos One*, 18:e0291530. doi: 10.1371/journal.pone.0291530

121. Alsanafi, M, Salim, NA, Sallam, M. (2023). Willingness to get HPV vaccination among female university students in Kuwait and its relation to vaccine conspiracy beliefs. *Human Vaccines Immunothera*, 19:2194772. doi: 10.1080/21645515.2023.2194772

Knowledge domain and emerging trends in medication literacy research from 2003 to 2024: a scientometric and bibliometric analysis using CiteSpace and VOSviewer

Penghong Deng[1†], Xiaoxia Liu[2†], Caiyun Li[3,4†], Xingping Zhu[5], Junli Cui[5], Ping Hua[5]* and Gang Chen[5]*

[1]Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China, [2]Department of Infectious Diseases, The Affiliated Dongtai Hospital of Nantong University, Dongtai, China, [3]Department of Pharmacy, Suzhou Hospital, Xiyuan Hospital of China Academy of Chinese Medical Sciences (Suzhou TCM Hospital), Suzhou, China, [4]Department of Pharmacy, Suzhou Traditional Chinese Medicine Hospital Affiliated to Nanjing University of Chinese Medicine, Suzhou, China, [5]Department of Pharmacy, The Affiliated Dongtai Hospital of Nantong University, Dongtai, China

Background: Medication literacy (ML) has emerged as a critical global public health concern, garnering growing scholarly attention over the past two decades. To delineate major research domains, identify evolving trends, and inform future research priorities, we conducted a scientometric analysis of the scientific literature on ML.

Methods: A systematic search was performed to retrieve publications on ML from the Web of Science Core Collection, covering the period from 2003 to 2024. Scientometric analyses were executed using CiteSpace and VOSviewer to visualize and evaluate collaborative networks, including co-citation references, co-occurring keywords, and contributions by countries, institutions, authors, and journals.

Results: The analysis incorporated 1,968 eligible publications. A rapidly growing trend in research interest in ML was observed, with an average annual growth rate of 46.1% in publications between 2003 and 2022. Three major research trends were identified: relationship between ML and medication adherence, the development of ML-specific assessment tools, and investigation of psychosocial factors associated with ML. The United States of America, Northwestern University, Davis Tc, and Patient Education and Counseling were identified as the most cited and influential entities within this field, representing the leading country, institution, author, and journal, respectively.

Conclusion: Scientometric analysis provides invaluable insights to clinicians and researchers involved in ML research by identifying leading contributors, intellectual bases and research trends. ML is evolving from unidimensional analysis to multidisciplinary exploration of dynamic mechanisms. Future research on ML is facing significant challenges, including the exploration of adherence mechanisms, validation of digital assessment tools, and the moderating effect model of socio-psychological factors on ML.

KEYWORDS

medication literacy, scientometric, bibliometric, CiteSpace, VOSviewer

1 Introduction

Global demographic aging is intensifying at an accelerated pace, with the proportion of population aged ≥60 years projected to surge from 10% in 2000 to 21% by 2050 (1). This demographic shift is fundamentally altering global disease spectrum, particularly through the escalating burden of non-communicable chronic diseases (NCDs) (2). Medication interventions, serving as the primary therapeutic approach for NCDs, critically influence disease trajectory modulation and health-related quality of life. For instance, patients diagnosed with hypertension or diabetes mellitus require long-term medication adherence to maintain stable blood pressure and glycemic control. However, suboptimal medication practices presented substantial challenges: the World Health Organization (WHO) data indicated that medication-related complications accounted for one-third of annual mortality, with associated economic losses exceeding $42 billion yearly (3). Medication non-adherence, dosing inaccuracies, and inappropriate drug utilization collectively contributed to diminished treatment efficacy, elevated hospital readmissions, therapeutic failures, healthcare system strain, and excess mortality (4–6). Numerous studies have demonstrated that these clinical and economic consequences can be prevented by strengthening medication literacy (ML) (7, 8).

Evolving from the broader concept of health literacy, ML was initially conceptualized in the 2005 UK Medication Safety Report and formally operationalized by Raynor (9) as individuals' capacity to retrieve, interpret, and apply medication-related information for informed decision-making. Numerous studies had proposed conceptually related terms to ML, including prescription literacy (specifically addressing the comprehension of prescription information) (10), pharmacotherapy literacy (emphasizing comprehensive therapeutic decision-making capabilities) (11), and pharmaceutical literacy (focusing on the understanding of specialized pharmacological knowledge) (12). In 2018, Pouliot A. et al. proposed a widely recognized academic definition, stating that ML referred to the extent to which individuals can obtain, comprehend, communicate, calculate, and process patient-specific information regarding their medications to make informed medication and health decisions (7). Research consistently identified this competence as a key determinant of medication adherence, with higher proficiency levels correlating with improved regimen adherence (13–15). Enhanced medication-related knowledge facilitated accurate interpretation of therapeutic instructions, thereby optimizing clinical outcomes and minimizing medication risks. Conversely, deficiencies in this domain were associated with poorer cardiovascular outcomes in coronary artery disease patients and reduced functional capacity among older adult populations (4, 16). These findings collectively underscored ML as a critical mediator of patient safety and therapeutic success.

The Third Global Patient Safety Challenge "Medication without Harm" strategic plan, launched by the World Health Organization (WHO), proposed that instruments and techniques should be employed to improve patients' medication literacy and interventions should be developed to promote patients' knowledge of drug use (17). This initiative highlighted the fact that ML and safety was one of the main research priorities in drug safety worldwide (17). In response to the emerging health goals, ML research has captured considerable interest and attention over the past two decades. Previous studies had found a low level of ML among patients with NCDs worldwide, characterized by poor understanding of medication-related knowledge, low medication adherence, and inadequate healthcare provider engagement (18–20). In addition, numerous scholars have conducted substantial research on definitions, predictive models, current problems, assessment tools, influencing factors and interventions for ML (21–24). Despite the existing publications providing insights into specific aspects of ML, this field lacks systematic integration of cumulative knowledge and research prioritization.

The exponential expansion of scholarly output necessitates advanced analytical methodologies to map this domain's intellectual architecture. Scientometric approaches, combining bibliometric analysis with data visualization, provide robust mechanisms for quantifying research trends and knowledge dissemination patterns—a methodological paradigm distinct from traditional systematic reviews (25, 26). Importantly, this approach establishes a systematic pathway to anticipate emerging paradigms and address complex research challenges by synthesizing interdisciplinary scientific frameworks with advanced methodological tools (27, 28). Such analyses prove particularly valuable for identifying collaborative networks, benchmarking institutional contributions, and detecting disciplinary gaps, though their application remains limited in ML research. This study's principal aim involves conducting a longitudinal scientometric evaluation to delineate the evolution, current frontiers, and emerging directions in ML research since over the past two decades. Secondary objectives focus on characterizing international collaboration dynamics, institutional productivity patterns, and knowledge dissemination channels while identifying critical research voids requiring scholarly attention.

2 Methods

2.1 Data source

The bibliometric dataset was systematically retrieved from the Web of Science Core Collection (WOSCC), a premier research database encompassing scholarly publications across 254 subject categories (29). This database is characterized by rigorous journal selection criteria, a comprehensive citation network, and standardized bibliographic fields optimized for scientometric analysis (29, 30). Compared to other major databases, WOSCC's extensive disciplinary coverage and high-quality data have established it as the preferred source for mainstream bibliometric analysis tools, with native compatibility in software such as CiteSpace and VOSviewer (30, 31). Additionally, this dataset's analytical utility extends beyond conventional bibliographic metadata (authorship, institutional affiliations, geographic distributions) through its integrated citation mapping functionality and multi-layered indexing architecture, establishing it as the benchmark and classical data source for scientometric investigations (29). Within the WOSCC, the Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI) were selected as primary data channels. The SCIE focuses on natural sciences, encompassing fields such as physics, medicine, and engineering, while the SSCI indexes social sciences disciplines including economics, psychology, and education. Both are important components of the WOSCC, but differ in disciplinary scope and citation patterns (29). Recognized as the gold standard for disciplinary coverage, these indices employ stringent journal inclusion criteria encompassing editorial rigor, citation impact metrics, and

international diversity (29), thereby ensuring the methodological validity of our analytical framework. For this study, publications addressing ML from the SCIE and SSCI were specifically extracted.

2.2 Retrieval strategy and data collection

A standardized literature search was executed by a single investigator (P. D.) on September 1, 2024, to control for temporal variability in database content. The selection of core search terms was derived from both MeSH (Medical Subject Headings) terminology and the widely recognized conceptual framework in this research domain. The search strategy combined title (TI) and author keyword (AK) fields using the following Boolean parameters: (TI = (medication literacy OR drug literacy OR pharmaceutical literacy OR medication knowledge OR medication understanding OR prescription understanding OR prescription knowledge OR medication attitude OR healthy medication behavior)) OR (AK = (medication literacy OR drug literacy OR pharmaceutical literacy OR medication knowledge OR medication understanding OR prescription understanding OR prescription knowledge OR medication attitude OR healthy medication behavior)). The temporal scope encompassed January 2003 to September 2024, restricted to the original research and review papers in English. Post-retrieval processing involved implementation of predefined exclusion criteria and cross-database deduplication procedures, with the selection process visually summarized in Figure 1.

2.3 Measures

We employed two techniques to investigate research evolution and trends:

1. Co-citation network of references: The co-citation network is based on the relationship between two documents being cited by a third document at a specific time, representing the intellectual foundation of the third document (32). As the subject evolves, the co-citation network expands from a single network to multiple networks, illustrating the shifts in the intellectual foundation over time (32). These transitions reflect the research tracks and trends in the citing documents. By analyzing the co-citation reference network, an intellectual landscape is constructed using highly cited literature and research frontiers (identified by extracting themes from the citing literature).
2. Co-occurring network of author keywords: Keywords provide insights into the specific research areas and directly address the research hotspots within the field. The co-occurrence network measures the frequency of paired keywords within a collection of documents and captures their associations. The process of co-occurrence analysis involves extracting keywords from the documents, tallying keyword frequencies, and identifying clusters, bursts, and connections among keywords (27).

As for our secondary objectives, we constructed collaborative networks of countries, institutions, authors, and journal co-citation networks. The collaborative network incorporated countries, institutions, and authors with at least one international collaboration,

enabling researchers to identify leading scholars and innovative research groups within the field. Author co-citation analysis was performed to identify highly cited authors, examine their connections, and explore the corresponding intellectual structure within the field. Additionally, journal co-citation networks help identify high-impact journals, reveal connections, and provide insights into the distribution of disciplinary knowledge domains.

2.4 Software and data analysis

This study employed two specialized bibliometric tools: VOSviewer (version 1.6.20) and CiteSpace (version 6.3. R3 Advanced), as documented in seminal works by van Eck & Waltman and Chen et al. (27, 33). VOSviewer, initially developed by Waltman et al. (33), provides an intuitive platform for network construction and visualization, facilitating the analysis of geographical distributions, institutional productivity, collaborative networks, and lexical co-occurrence patterns. We employed VOSviewer to analyze the networks of authors' countries, institutions, co-author collaborations, co-occurring keywords, and density map of keywords and journals. CiteSpace, a Java-based application introduced by Chen et al. (27), specializes in emerging trends detection and knowledge domain mapping through systematic mapping, integrated bibliometric analysis, and data mining algorithms. Systematic mapping offers a comprehensive overview of existing scholarly knowledge, facilitating the identification of research domains that are sufficiently mature for meta-synthesis and those warranting further empirical investigation (27). As a quantitative analytical paradigm rooted in mathematical and statistical principles, bibliometrics enables researchers to elucidate the structural relationships and evidentiary connections within scientific literature (28). By utilizing CiteSpace, we were able to identify intellectual bases, emerging research fronts, temporal trends, and citation dynamics.

In CiteSpace, we configured the analysis with 1-year time slices. The g-index (k = 25) was employed to assess research impact, which effectively accounted for both high-cited publications and less-cited works. Clusters were groups of tightly connected nodes identified by optimizing modularity in the network, and the labeling relied on statistical likelihood to extract representative terms (27). In this study, cluster labels were derived through log-likelihood ratio (LLR) algorithmic processing of keyword corpora ($p < 0.001$). The knowledge networks generated by VOSviewer and CiteSpace comprise two fundamental elements: nodes (representing bibliographic entities including references, keywords, countries, authors, institutions and journals) and edges (denoting relational linkages through collaboration, co-citation, or co-occurrence). Node diameter correlates positively with bibliometric indicators such as citation frequency, occurrence count, or centrality metrics, serving as visual proxies for scholarly influence. Chromatic encoding of nodes and edges conveys information about the year of the corresponding citations, clusters, or occurrences. Highly connected nodes are included between and within clusters, revealing relevant areas and their evolution throughout the years.

The study employed CiteSpace's structural variation analysis and burst detection algorithms to investigate critical factors shaping network topology and identify emerging research trajectories. Structural variation analysis quantifies the boundary-spanning

FIGURE 1
Flow chart of the scientometric study. #1: TI = (medication literacy OR drug literacy OR pharmaceutical literacy OR medication knowledge OR medication understanding OR prescription understanding OR prescription knowledge OR medication attitude OR healthy medication behavior); #2: AK = (medication literacy OR drug literacy OR pharmaceutical literacy OR medication knowledge OR medication understanding OR prescription understanding OR prescription knowledge OR medication attitude OR healthy medication behavior).

potential of scholarly works through novel linkage formation metrics (34). Publications that establish interdisciplinary connections are particularly significant, as they often represent pivotal points of knowledge integration and potential catalysts for paradigm shifts (34). Complementarily, burst detection analysis, implemented through temporal data streaming algorithms, identifies citation and term patterns exhibiting sudden frequency and intensity anomalies (35). These temporal signatures frequently indicate the emergence of novel research fronts or innovations. We conducted burst detection analysis on cited references, keywords, authors, and journals to synthesize and reveal possible future research priorities. Additionally, to illustrate the evolutions and connections among clusters, we utilized timeline analysis, which spatially distributes nodes along temporal axes.

Three critical graph-theoretical indices guided cluster interpretation, following Chen et al.'s methodological framework (27): (1) Betweenness Centrality: This metric quantifies node brokerage potential through shortest-path analysis, identifying intra-cluster core nodes and inter-cluster bridging hubs. Nodes with higher centrality scores indicate their critical role within the research field, such as

highly influential publications or interdisciplinary researchers. (2) Modularity (Q): This metric evaluates the tightness of intra-group connections and the separation between groups to validate the rationality of network clustering structure (Q∈[0, 1]). A Modularity value greater than 0.3 suggests that the network's clustered structure is well-defined and meaningful. (3) Silhouette Coefficient (S): This measure assesses the homogeneity within clusters and the accuracy of node classification (e.g., whether a given publication is correctly assigned to its thematic cluster, S∈[−1, 1]), with S > 0.7 confirming substantial node similarity. Additionally, Centrality Divergence was calculated as the standard deviation of betweenness centrality distributions, serving as an indicator of structural innovation potential in boundary-spanning publications (34).

3 Results

Two different analytical software tools were employed to systematically map the evolution of ML research over the past two

decades. This dual-method approach enabled a comprehensive evaluation of publication trends and the construction of knowledge networks, including co-cited references, author keywords, and contributions across countries, institutions, authors, and journals.

3.1 Analysis of publication outputs and trends

The final analysis comprised 1,968 unique scholarly publications, including 1,852 research articles and 116 review papers, which collectively accumulated 38,669 citations following screening and exclusion protocols. The authorship network encompassed 9,177 contributors, averaging 4.66 authors per publication, representing 6,732 institutions across 571 countries/territories. Temporal analysis revealed substantial growth in scholarly output, with annual publications increasing from 19 in 2003 to 194 in 2022, reflecting a compound annual growth rate of 46.1%. However, the dataset's temporal boundary (September 2024) resulted in an apparent decline in annual publication and citation metrics. Despite relatively modest annual publications, citation impact demonstrated significant enhancement, with the average citations per document (total annual citations divided by annual publications) rising from 0.37 (7/19) in 2003 to 22.65 (4,431/178) in 2023 (Supplementary Figure S1).

3.2 Analysis of co-citation references

3.2.1 Clusters of research

A co-citation network was conducted to identify the influential and representative research in the ML field (Figure 2A). Additionally, the co-citation reference network analysis yielded 11 distinct thematic clusters, exhibiting robust modularity (Q = 0.8259) and exceptional intra-cluster homogeneity (S = 0.9385), confirming both the credibility and distinctiveness of the groups (Figure 2B). The cluster labels were synthetically generated based on representative noun phrases extracted from the keyword lists of cited references within each cluster. More detailed descriptions of each cluster are available in Table 1. Three major research trends were identified based on the largest linkage pathways between clusters. The clusters contributing to these trends are presented with their cluster label, size, silhouette score, average year of publication, and the most representative reference.

The first major research trend focused on the association between ML and medication adherence. This trend emerged in 2002 with the identification of cluster #8 (*"readability"*; 37, S = 0.973, 2002) in our database, alongside a seminal article published by Schillinger et al. in JAMA, which examined the relationship between health literacy and diabetes outcomes (36). This cluster subsequently evolved into cluster #0 (*"health literacy"*; 110, S = 0.908, 2005), which served as a conceptual foundation for ML. Within this cluster, numerous studies explored the relationship between various dimensions of ML, such as literacy levels, the number of prescription medications, the understanding of prescriptions, and appropriate medication use (37, 38). Subsequently, a strong interconnection was observed between cluster 0 (*"health literacy"*) and cluster 1 (*"medication adherence"*; 109, S = 0.907, 2010) (39), suggesting a rapidly evolving phase in the exploration of the relationship between ML and adherence. Moreover, the emerging knowledge linkages between cluster #1

(*"medication adherence"*) and cluster #3 (*"deprescribing"*; 74, S = 0.927, 2014) (40) represented a new research frontier, highlighting the growing academic interest in the association between deprescribing practices and medication adherence.

The second major research trend revolved around the development of ML-specific assessment instruments. This trend began with cluster #6 (*"electronic health records"*; 48, S = 0.942, 2009) and highlighted the impact of medication administration tools and patient-centered labels on ML levels, providing a foundation for the development of subsequent assessment tools (41, 42). Over the past decade, this research field had further enriched and converged into the third largest cluster #2 (*"medication literacy assessment"*; 108, S = 0.908, 2018) (7). Within this clustering, an international definition of ML was recognized through expert consensus, significantly contributing to the standardization of ML assessment tools (7).

The third research trend focused on the investigations of psychosocial factors associated with ML. According to the inter-cluster links, clusters associated with this research topic trend were cluster #4 (*"self-medication"*; 66, S = 0.995, 2017) (43), cluster #7 (*"hypertension"*; 40, S = 0.946, 2015) (24), and cluster #10 (*"medications for opioid use disorder"*; 18, S = 0.996, 2010). Within these clusters, research participants were mainly the older adult and patients with chronic diseases, with influencing factors involving disease status, education level, economic income, and psychological cognition.

The timeline map provided a visual representation of the duration and historical progression of each cluster, effectively capturing the trends mentioned earlier. It also allowed us to pinpoint the temporal placement of landmark publications. Notably, the most recent and dynamically active clusters in the analysis were cluster #2 (*"medication literacy assessment"*), cluster #4 (*"self-medication"*), cluster #5 (*"attitude to health"*), cluster #9 (*"long-term care"*), and cluster #10 (*"medications for opioid use disorder"*), indicating a growing research interest in these areas (Figure 3).

3.2.2 Most cited references and transformative papers

Table 2 presented the top ten most cited references, which played a crucial role in shaping the intellectual foundations of the clustering studies. A comprehensive review of strategies to assess and improve medication adherence/compliance conducted by L. Osterberg et al. in 2005 emerged as the most co-cited paper, with 98 citations within our reference network (44). Notably, a cross-sectional investigation of drug literacy and comprehension of prescription drug labels authored by Davis TC et al. in the Annals of Internal Medicine received 91 co-citations within our network (37). It is noteworthy that these two publications demonstrated substantial citation bursts, with burst strengths of 6.01 and 13.20, respectively, indicating their potential to exert significant influence on ML research (Supplementary Table S1).

Furthermore, a structural variation analysis was conducted to identify transformative papers that derived disciplinary evolution in the research field through interdisciplinary knowledge integration. Utilizing centrality divergence metrics, we identified three paradigm-shifting publications: an illustrated medication schedule developed by Kripalani S et al. for better understanding of prescription drugs (45), an investigation of the relationship between patient literacy level and self-reported HIV medication adherence (46), and a multicenter study conducted by Persell SD et al. of health literacy on medication reconciliation in ambulatory care (47). These papers have made
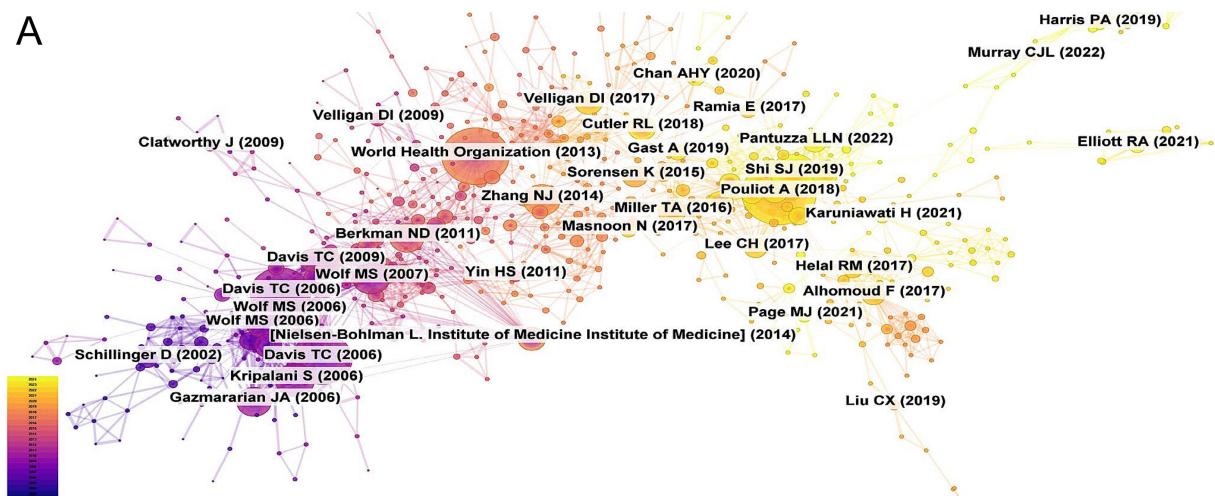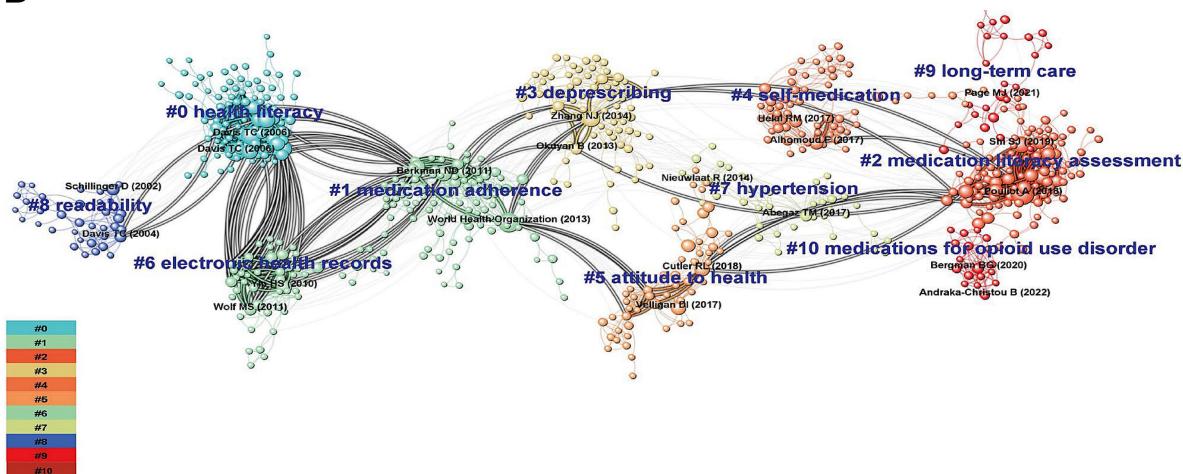
FIGURE 2
Co-citation references network **(A)** and corresponding clustering visualization **(B)** generated by CiteSpace (2003-2024). A node represents a cited reference. Each node represent one highly co-cited article. The size of a node is proportional to the co-citation count. Nodes are organized in different clusters gathered into a network of co-citation. The highlighted lines represent the evolution and connections among different clusters.

significant contributions to the field and have been instrumental in advancing our understanding of ML research.

3.3 Analysis of co-occurring author keywords

Figure 4 depicted a timeline visualization derived from the co-occurrence analysis of author keywords using CiteSpace, illustrating the evolution of thematic clusters in ML research. The keyword clustering exhibited robust validity, supported by high modularity and silhouette scores (Q = 0.3144; S = 0.6923), indicating well-defined and internally coherent groupings. The cluster labels were synthetically generated based on homogeneous, high-frequency keywords extracted from the citing literature. Six major clusters were identified (ranked by size): cluster #0 ("*medication errors*"; 134; S = 0.652; 2012), #1 ("*health literacy*"; 123; S = 0.671; 2011), #2 ("*self-medication*"; 107; S = 0.726; 2011), #3 ("*schizophrenia*"; 100; S = 0.742; 2009), #4 ("*buprenorphine*"; 90; S = 0.646; 2013), #5

("*public health*"; 19; S = 0.884; 2016). We found that cluster #0, #1, #2, #3, and #4 showed extensive temporal spans (>20 years), suggesting their foundational role as core research themes within this domain. The concentration of high-frequency keywords in cluster #0, #1, and #3 during early stages indicated theoretical maturation in medication literacy, errors, and adherence research, with a visible translational shift toward clinical implementation studies in recent years. In contrast, the emergence of high-frequency keywords in cluster #2 and #4 reflected their status as burgeoning research frontiers, as evidenced by intensified scholarly activity recently (Figure 4).

Furthermore, keywords were analyzed for burstiness to identify keywords that exhibited significant temporal fluctuations in academic attention (Supplementary Table S2). The keywords with the highest burst intensity were *quality* (strongest), *adherence*, and *nonadherence*. The most persistent keywords based on when the citation outbreak began were *comprehension*, *adverse drug events*, and *physicians*. Notably, *self-efficacy* and *resistance* emerged as areas of recent academic focus, showing continued prominence

TABLE 1 Summary of the largest clusters identified for co-citation network of references.

Cluster ID	Size	Silhouette	Mean (Year)	Top five extracted terms based on keywords
0	110	0.908	2005	health literacy; labels; safety; prescription; medication literacy
1	109	0.907	2010	medication adherence; cognition; pregnancy; medication knowledge; self-medication
2	108	0.908	2018	medication literacy assessment; relationship; structural equation model; hypertension; adherence
3	74	0.927	2014	deprescribing; polypharmacy; pharmaceutical literacy; diabetes; older adults
4	66	0.995	2017	self-medication; antibiotics; knowledge; health literacy; antibiotic
5	56	0.957	2015	attitude to health; antipsychotics; health behavior; schizophrenia; lifestyle
6	48	0.942	2009	electronic health records; ambulatory care; medication error; drug labeling; crowdsourcing
7	40	0.946	2015	hypertension; medication taking; hypertensive patients; medication history; emergency medical services
8	37	0.973	2002	readability; prescriptions; prescription drug labels; perceived effectiveness; ethnic/racial differences
9	31	0.98	2020	long-term care; medication disposal; medication adherence; pharmaceutical waste; dementia
10	18	0.996	2020	medications for opioid use disorder; acceptability; subjective norms; peer recovery specialist; treatment

FIGURE 3
Timeline visualization of co-citation references network (2003–2024). A node represents a cited reference. The size of a node depends on its betweenness centrality. For each cluster, nodes are organized by their year of publication on horizontal lines. Nodes with large coloured tree rings are those with high betweenness centrality (external purple tree rings) and burst strength (central red tree rings). The color of lines indicate the time of links between nodes or between clusters.

between 2018 and 2024. Additionally, VOSviewer software was to generate temporal overlay visualizations mapped to average publication year. The most frequently cited keywords encapsulated the major research trends-medication adherence, medication knowledge, health literacy, and medication attitudes-which were highly conceptually aligned with the thematic focus of this study (Figure 5A).

3.4 Analysis of collaboration networks across countries and institutions

Figure 5B displays the cooperation networks of countries, while Figure 5C shows the cooperation networks of institutions. In total, 72 countries or territories were captured in the analysis. The *United States of America* (USA) held a central position, with the highest number of

TABLE 2 The top 10 most cited references.

Co-citations	Author	Year	Title	Journal
98	Osterberg L	2005	Adherence to medication	New England Journal of Medicine
91	Davis Tc	2006	Literacy and misunderstanding prescription drug labels	Annals of Internal Medicine
84	Horne R	1999	The Beliefs about Medicines Questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication	Psychology & Health
70	Hogan Tp	1983	A self-report scale predictive of drug compliance in schizophrenics: reliability and discriminative validity	Psychological Medicine
70	Morisky De	1986	Concurrent and predictive validity of a self-reported measure of medication adherence	Medical Care
67	Davis Tc	2006	Low literacy impairs comprehension of prescription drug warning labels	Journal of General Internal Medicine
66	Berkman Nd	2011	Low health literacy and health outcomes: an updated systematic review	Annals of Internal Medicine
66	Davis Tc	1993	Rapid estimate of adult literacy in medicine: a shortened screening instrument	Family medicine
66	Morisky De	2008	Predictive validity of a medication adherence measure in an outpatient setting	The Journal of Clinical Hypertension
62	Horne R	1999	Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness	Journal of Psychosomatic Research

publications ($n = 768$), followed by *China* with 181 publications and *Australia* with 164 publications. In terms of citations, the *USA* was also the most cited country ($n = 20,363$), followed by *England* ($n = 3,281$) and *Canada* ($n = 2,325$). Furthermore, VOSviewer identified 150 institutions from the dataset. *Northwestern University* emerged as both the most published institution ($n = 62$) and the most cited institution ($n = 3,564$). *Emory University* also produced 49 publications, while *Sydney University* had 32 publications. In terms of citations, *Emory University* ranked the second ($n = 3,254$), followed by *Louisiana State University* ($n = 1729$).

3.5 Analysis of co-authorship networks

A network of co-cited authors was established, demonstrating significant modularity and silhouette scores ($Q = 6,128$; $S = 0.8674$) (Supplementary Figure S2). Cluster #0, titled *"health literacy"*, emerged as the predominant thematic cluster, central to the network and encompassing research domains including medication adherence, attitudes, knowledge, errors, and self-management. The top three most cited authors were *Davis TC* ($n = 192$), *World Health Organization* ($n = 188$), and *Wolf MS* ($n = 169$). Analysis of betweenness centrality, metric reflecting authors' roles in bridging network subfields, entified *Hogan TP* (centrality = 0.14), *Bandura A* (0.12), and *Cramer JA* (0.10) as key interdisciplinary connectors. *Kalichman SC* was identified as the highest citation burst intensity, indicating a significant increase in scholarly influence, and the most active contributor during the 2005–2011 period (Supplementary Table S3).

Furthermore, analysis of collaborative author networks revealed prominent collaborative clusters anchored by *Wolf MS*, *Horne R*, *Davis TC*, *Parker RM*, and *Kripalani S* (Figure 5D). These scholars occupied central network positions, driving both collaborative synergies and advancements in the field.

3.6 Analysis of journal occurrence and citations

This study employed VOSviewer to construct journal co-occurrence (Figure 6A) and journal citation density maps (Figure 6B), systematically indicating the distribution characteristics of journals in ML research. *Research in Social & Administrative Pharmacy* (IF = 3.7, Q1) ranked first with 55 published papers, followed by *Patient Preference and Adherence* ($n = 52$, IF = 2.0, Q1) and *Patient Education and Counseling* ($n = 46$, IF = 2.9, Q1), collectively forming the core knowledge dissemination platforms in this domain. From an academic influence perspective, *Patient Education and Counseling* dominated with 1,180 total citations, followed by the *Journal of General Internal Medicine* ($n = 1,151$, IF = 4.3, Q1) and *JAMA-Journal of the American Medical Association* ($n = 928$, IF = 63.5, Q1), highlighting their disciplinary leadership. Dual-map overlay analysis further revealed knowledge flow patterns (Figure 6C). The citing journal cluster (left) and cited journal cluster (right) demonstrated two prominent knowledge transfer pathways: (1) Publications from *"Health/Nursing/Medicine"* journals primarily informed advancements in clinical medicine and health education research; (2) Outputs from *"Psychology/Education/Social"* journals were predominantly inherited by *"Psychology/Education/Health"* journals. This interdisciplinary citation paradigm underscores the dual attributes of ML research as clinical treatment and socio-educational relevance, providing theoretical guidance for journal selection strategies.

4 Discussion

4.1 Summary of the main findings

This study offered a comprehensive scientometric assessment of the global ML research landscape, delineating its intellectual structure,

FIGURE 4
Timeline visualization of co-occurring keywords network (2003–2024). A node represents a keyword. The position of the node corresponds to the year of keyword occurrence. The size of a node is proportional to the frequency of its occurrence. The clusters are labelled in blue on the far right of the timeline map.
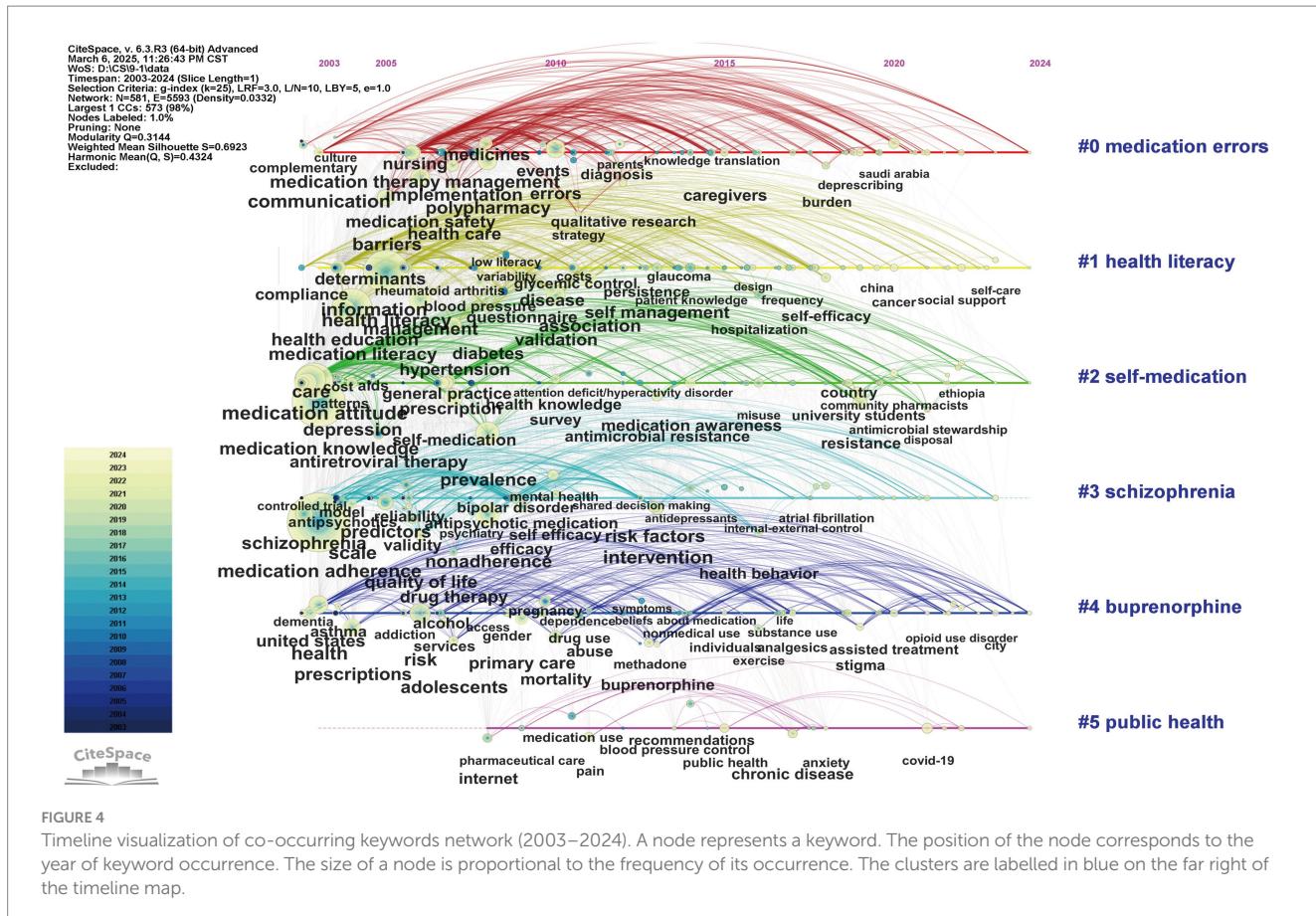
thematic evolution, and collaborative patterns over a 20-year period. While annual publication output remained modest, the field exhibited a marked growth trajectory, reflecting its rising scholarly focus. Analyses of co-cited literature networks and author keyword clusters revealed robust linkages across 11 and 10 thematic clusters, respectively, converged into three research themes: the relationship between ML and medication adherence, development of ML-specific assessment instruments, and investigations of psychosocial factors associated with ML. *The United States* served as the leading contributor nationally, with *Northwestern University* emerging as the most productive institutional entity. *Davis TC*, *Wolf MS*, and *Horne R* were identified as the most frequently cited authors, while *Wolf MS* demonstrated the highest publication productivity. The journals *Patient Education and Counseling*, *Journal of General Internal Medicine*, and *JAMA-Journal of the American Medical Association* ranked as the most influential outlets in the field.

4.2 Identification of research trends

The resulting co-citation reference network and author's keyword analysis extracted three distinct major research trends in ML research from 2003 to 2024, which were also captured by the qualitative analysis of highly cited literature. The first research trend focused on the association between ML and medication adherence. The interplay between ML and medication adherence had emerged as a critical determinant of clinical outcomes. ML referred to a patient's capacity

to acquire, comprehend, and apply medication-related information, encompassing knowledge of drug nomenclature, dosing protocols, administration standards, and risk profiles (7). Medication knowledge represented the fundamental understanding of drug-related information (e.g., dosage, administration), whereas ML emphasized higher-order competencies in acquiring, evaluating, and applying such information in clinical practice (e.g., dose adjustment, adverse reaction identification) (48). Existed Studies had demonstrated that patients with adequate knowledge but insufficient literacy remained at significantly elevated risk of medication errors (48). Medication adherence was manifested as behavioral consistency with the prescribed treatment regimen (49). Previous evidence had demonstrated that ML positively modulated adherence through multilevel synergistic mechanisms (40). The first mechanism was cognitive reinforcement. Enhanced understanding of pharmacological mechanisms and long-term therapeutic necessity reduced self-discontinuation behaviors triggered by symptomatic relief. Patients might benefit from routine medication use reviews (MURs) with their healthcare providers to identify and address potential medication-related problems in advance (50). The second one was skill empowerment. Pharmaceutical care interventions, including regimen simplification (e.g., reduced dosing frequency) and visual medication aids (e.g., dosing calendars), mitigated non-adherence stemming from operational errors (51). Additionally, based on the Health Belief Model, ML increased people's awareness of disease severity and therapeutic benefits, fostering a paradigm shift from passive compliance to active therapeutic engagement (52). This transformation
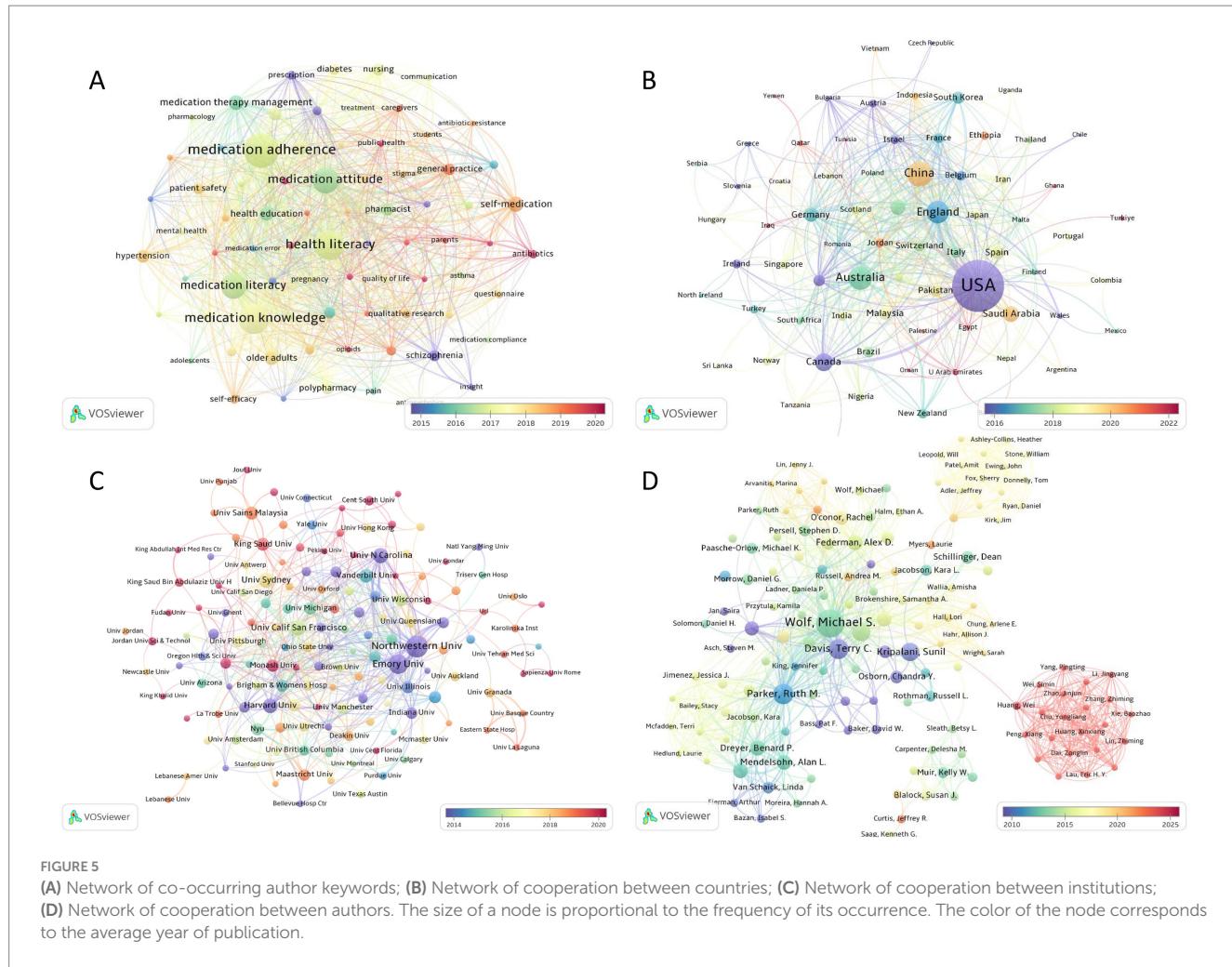
FIGURE 5
(A) Network of co-occurring author keywords; **(B)** Network of cooperation between countries; **(C)** Network of cooperation between institutions; **(D)** Network of cooperation between authors. The size of a node is proportional to the frequency of its occurrence. The color of the node corresponds to the average year of publication.

proved particularly salient in chronic disease and mental health management (53). Attitudes and behaviors were identified as independent predictors of medication adherence (14, 54). While socioeconomic and polypharmacy challenges persisted as adherence inhibitors, integrative strategies combining ML interventions (e.g., personalized education, cost-reduction policy dissemination) and digital health technologies (e.g., medication reminder applications) demonstrated synergistic potential to counterbalance these impediments (55–58).

The second major and influential research trend involved the development of measurement tools for assessing ML. Scientifically validated and contextually appropriate assessment instruments constituted a critical prerequisite for both accurately evaluating individual ML levels and designing evidence-based intervention strategies. Given the multifactorial nature of ML, encompassing cognitive, behavioral, and sociocultural dimensions, its concept and evaluative criteria must be contextually adapted rather than universally standardized across various populations and healthcare contexts (7). As emphasized by Gentizon et al. (59), assessments required tailored calibration to align theoretical constructs of ML with practical measurement paradigms, ensuring congruence with specific demographic profiles and clinical settings. However, the heterogeneous quality of existing ML measurement instruments and the diverse emphases in their assessment scopes posed substantial challenges in the selection of ML assessment tools that were suitable for specific chronic disease patients (12, 19, 59). The unidimensional 14-item MedLitRxSE, as developed by Sauceda et al. (60), was presently the sole instrument formally recommended by the Agency for Healthcare Research and Quality (AHRQ) for evaluating ML among adult care recipients and their informal caregivers. This instrument exhibited reliable values, satisfactory content validity, structural validity, and internal consistency; however, its overall reliability remained uncertain. Additionally, the extensive content coverage of the C-MLSHP (14) and the Pharmacy Consumer Health Literacy Questionnaire (61), along with the satisfactory content validity of the PTHL-SR and MedLit-NSAID (19, 62), provided a robust foundation for the psychometric analysis of subsequent measurement instruments. The content validity of these instruments required further investigation, particularly with a focus on systematic engagement of target populations. Furthermore, previous studies had indicated that the PWMIL (63), the RALPH Interview Guide (64), and the Medication Literacy Questionnaire for Discharged Patients (65) provided uncertain evidence, necessitating further testing for both content and agency. Performance-based ML assessment instruments (e.g., MedLitRxSE) utilized standardized scenario testing to objectively quantify medication management competencies (e.g., dosage calculation, medication label interpretation), offering strong reliability and reproducibility (60).
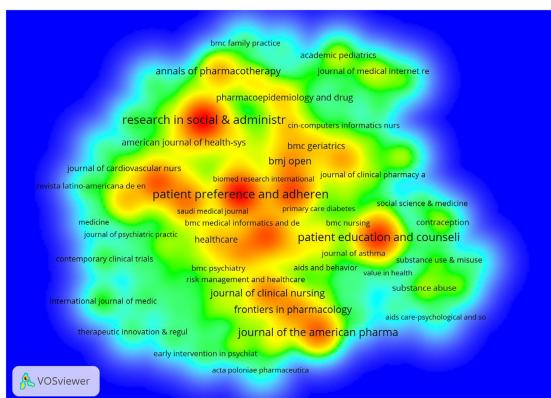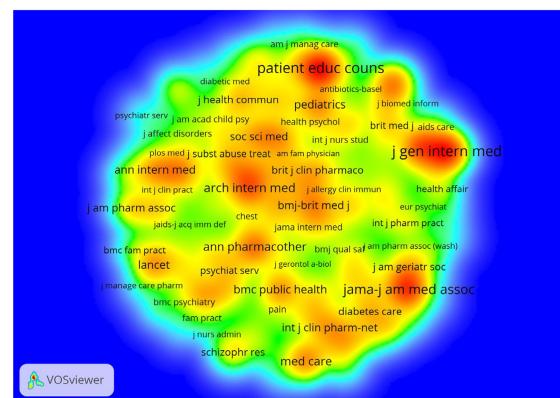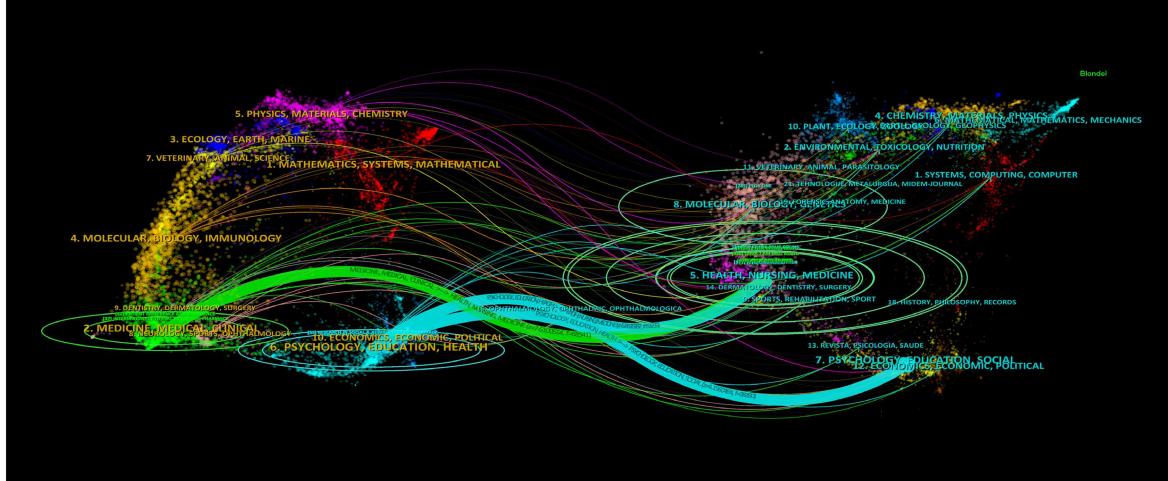
FIGURE 6
(A) Density map of journal publications; (B) Density map of cited journals; (C) Map of citation trajectories of citing and cited journals.

However, these tools might fail to fully capture patients' adaptive capacities in real-world medication use contexts. Conversely, perception-based measures (e.g., the HeLMS questionnaire) assessed self-reported medication-taking confidence and challenges through subjective evaluations, which effectively identify experiential barriers but remain vulnerable to recall bias and social desirability effects (66). Adopting integrated methodologies in future research—combining performance-based tools to identify skill gaps with perception-based measures to uncover behavioral determinants—will be critical for comprehensive evaluation.

The development of ML were shaped by multidimensional psychosocial factors, centered on the dynamic interaction mechanisms between individuals and their environments (24). According to the Social Cognitive Theory (SCT) and Health Belief Model (HBM), patients' medication literacy was not solely determined by cognitive ability, but also closely associated with psychological state, social support and cultural context (67–69). First, patients' ability to interpret and integrate medication-related information was directly affected by education level (18). Due to barriers in understanding technical terminology or deficiencies in logical analysis, populations with a low education level often struggled to accurately assess medication risks and benefits. For example, previous studies revealed that, compared to 78% among highly educated groups, only 32% of chronic disease

patients in low-income communities could correctly interpret dosage adjustment instructions on drug labels (70). Second, family and community support were important external resources for ML practice (71). Medication adherence can be increased by more than 40% through the proactive involvement of family members, including medication reminders, emotional reassurance, and behavioral monitoring (72). Conversely, social isolation or familial conflicts may contribute to medication discontinuation behaviors. For example, depressed patients who were lacking emotional support had a 2.3-fold higher risk of self-reducing antidepressant dosages (72). Third, patients' perceptions of disease severity, treatment benefits, and self-management confidence constituted intrinsic drivers of ML (58, 59). Brod et al. found that psychological insulin resistance was associated with patients' beliefs regarding diabetes and insulin, negative self-perceptions and attitudinal barriers, and fears of adverse outcomes and complications of insulin use, thereby contributing to patients' reluctance to initiate and intensify treatment (73). In addition, the risk of medication use was exacerbated by financial stress and sensitivity to the cost of medication, which prompted some patients to purchase medication through informal sources or to self-adjust their dosage. Approximately half of patients with NCDs in developing countries were forced to reduce their use of prescription medications due to financial constraints (74).

Overall, the identified research trends were interconnected, forming a cyclical "assessment-mechanism-intervention" framework that collectively advanced ML research. The development of standardized ML assessment tools established a methodological foundation, enabling robust validation of the relationship between ML and medication adherence. Identification of the psychosocial factors served dual purposes: explaining individual variations in ML levels and enhancing the cultural relevance of assessment tools—both critical for designing personalized interventions. Importantly, the optimized interventions not only improved adherence but may also have their effectiveness moderated by psychosocial factors. The systematic associations represented the significant scientific value of ML research proceeding from assessment to intervention.

4.3 Outputs and influence networks

The analysis of research outputs and influence networks constitutes a secondary objective of this investigation, aiming to capture geographic distributions, identify gaps, and recognize high-impact countries, research groups, and authors within specific subjects. The presented collaborative networks, co-citation visualizations, and associated bibliometric indices, provide readers, particularly active researchers, with critical insights into the field's epistemic architecture. At the national/institutional level, the USA and Northwestern University emerged as the foremost contributors in both publication volume and citations, which can be attributed to their top researchers and well-established biomedical foundations. Contrasted with the comparative lag of developing regions, the extensive collaboration in ML research in the Western countries and institutions reveals structural inequalities in medication-related health research globally. Substantive support for under-resourced nations and institutions is imperative to advance understanding of how various healthcare systems and sociocultural paradigms influence ML research. Furthermore, our co-cited author network highlights the significant contributions of Davis TC to the field, particularly in the relationship between ML and understanding of prescription drug labels (37, 38). While co-authorship networks provide limited proxy measures of scholarly influence, systematic examination of high-impact citations and transformative literature enables identification of field-shaping contributors. It warrants emphasis that journal rankings derived from WOSCC publication/citation counts constitute imperfect quality proxies. However, the analysis of co-cited journals does identify the most cited journals in a given research area, such as Patient Education and Counseling in our network, which are considered appropriate for specific topics. In addition, an emerging trend of cross-disciplinary research was observed in the ML field. Different domains, including clinical practice, social behavior, social education, and digital medications, were establishing a multidimensional ML-enhancement system. A paradigm shift of research mindset was required for investigators within this field.

4.4 Potential trends for future research

Acting as a pivotal bridge between patient cognition and medication behavior, ML is transitioning from a unidimensional analysis to a multidisciplinary exploration of dynamic mechanisms (75, 76). Future research should prioritize examining the impact of cultural heterogeneity on ML and the potential of digital interventions, such as AI-assisted medication guidance, to enhance adherence (77, 78). Cross-disciplinary collaborations, such as integrating psychology with pharmaceutical care, could optimize adherence strategies, particularly for vulnerable populations like older adults and individuals with limited education (79). Current assessment tools, such as MedLitRxSE, demonstrate limitations in dimensional coverage and technical adaptability within digital health contexts (59). To address these limitations, future efforts should focus on developing comprehensive scales that integrate functional, critical, interactive, and digital literacy dimensions, augmented by natural language processing technologies to enable real-time dynamic evaluations (77). For example, intelligent platforms utilizing computerized adaptive testing could incorporate electronic health record data to generate personalized feedback, thereby enhancing clinical utility (77). Additionally, the mediating and moderating roles of psychosocial factors, such as self-efficacy and social support, in the "ML to adherence" pathway require systematic validation (80). Mixed-methods approaches, combining longitudinal data with qualitative interviews, could elucidate the dynamic mediating effects of self-efficacy on ML-adherence relationships and the buffering role of familial support in low-education populations (24). Future studies must quantify the relative weights of psychosocial determinants and design precision interventions targeting at-risk subgroups, such as individuals with low literacy and ethnic minorities, to establish an ecological support network spanning individual, community, and policy levels. This integrated approach will advance the construction of a robust framework for enhancing ML, ultimately promoting equitable medication management and improved health outcomes.

4.5 Strengths and limitations

Compared to a narrative review, scientometric analysis provides a more systematic and comprehensive approach to mapping research landscapes, offering clinicians and researchers critical insights into emerging trends and intellectual structures. This method contributes to identifying underexplored scientific questions, thereby guiding the direction of future research efforts (25). Furthermore, it enables the identification of influential authors, journals, and institutions within the field of ML, fostering opportunities for collaboration and knowledge exchange across specialized research domains. However, several limitations of this study must be acknowledged. First, while co-citation analysis is an essential component of scientometric methods, it is susceptible to citation biases, including publication bias, self-citation, authorship bias, literature type bias, and journal impact factor bias, which may undermine the objectivity of the findings (32). Second, data collection was restricted to the SCIE and SSCI within the WOSCC, limiting the scope of retrieved publications. Other prominent databases, such as PubMed and Embase, which provide full-text references and citation lists, were excluded (31). Third, the co-citation network analysis focused solely on first authors, potentially overlooking the contributions of co-authors. Additionally, the keyword co-occurrence networks were susceptible to variations in keyword expressions, thus affecting cluster interpretation. Finally, the

co-citation network's ability to capture recent trends was constrained by the limited citation of newly published literature.

5 Conclusion

This first scientometric study provides a comprehensive analysis of the historical trends and research landscape of ML research, revealing sustained growth in scholarly output over two decades, with publication volumes peaking in 2022. The analysis identifies leading contributors, including the most productive countries, institutions, authors, and journals, while mapping thematic priorities such as the relationship between ML and medication adherence, development of ML-specific assessment instruments, and investigations of psychosocial factors associated with ML. The findings underscore the necessity for strengthened cross-institutional collaboration, particularly among European, U.S., and Chinese entities—to leverage the influence of key opinion leaders. By synthesizing current research trends and emerging frontiers, this work provides clinicians and researchers with an empirical foundation to guide future inquiries, while offering funding bodies strategic insights into priority areas.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

PD: Methodology, Writing – original draft, Investigation. XL: Writing – review & editing, Investigation, Conceptualization. CL: Investigation, Conceptualization, Writing – review & editing, Validation. XZ: Methodology, Conceptualization, Data curation, Writing – review & editing, Software. JC: Conceptualization, Writing – review & editing, Data curation, Methodology. PH: Validation, Investigation, Writing – review & editing, Supervision. GC: Writing – review & editing, Validation, Supervision.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2025.1598482/full#supplementary-material

References

1. The United Nations. (2022). World population prospects 2022. The United Nations. Available online at: https://doi.org/10.18356/9789210014380 (accessed April 25, 2025).

2. Murray CJL. The global burden of disease study at 30 years. *Nat Med*. (2022) 28:2019–26. doi: 10.1038/s41591-022-01990-1

3. World Health Organization. (2023). Global burden of preventable medication-related harm in health care: a systematic review. World Health Organization. Available online at: https://iris.who.int/handle/10665/376203 (accessed December 1, 2024).

4. Dobbels F, De Geest S, van Cleemput J, Droogne W, Vanhaecke J. Effect of late medication non-compliance on outcome after heart transplantation: a 5-year follow-up. *J Heart Lung Transplant*. (2004) 23:1245–51. doi: 10.1016/j.healun.2003.09.016

5. Currie CJ, Peyrot M, Morgan CL, Poole CD, Jenkins-Jones S, Rubin RR, et al. The impact of treatment non-compliance on mortality in people with type 1 diabetes. *J Diabetes Complicat*. (2013) 27:219–23. doi: 10.1016/j.jdiacomp.2012.10.006

6. Shahid R, Shoker M, Chu LM, Frehlick R, Ward H, Pahwa P. Impact of low health literacy on patients' health outcomes: a multicenter cohort study. *BMC Health Serv Res*. (2022) 22:1148. doi: 10.1186/s12913-022-08527-9

7. Pouliot A, Vaillancourt R, Stacey D, Suter P. Defining and identifying concepts of medication literacy: an international perspective. *Res Social Adm Pharm*. (2018) 14:797–804. doi: 10.1016/j.sapharm.2017.11.005

8. Marshall N, Butler M, Lambert V, Timon CM, Joyce D, Warters A. Health literacy interventions and health literacy-related outcomes for older adults: a systematic review. *BMC Health Serv Res*. (2025) 25:319. doi: 10.1186/s12913-025-12457-7

9. Raynor DK. Medication literacy is a 2-way street. *Mayo Clin Proc*. (2008) 83:520–2. doi: 10.4065/83.5.520

10. Parker CT. Literacy and misunderstanding prescription drug labels. *Ann Intern Med*. (2007) 147:280–1. doi: 10.7326/0003-4819-147-4-200708210-00016

11. Sørensen K, Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health*. (2012) 12:80. doi: 10.1186/1471-2458-12-80

12. Levic M, Stanojevic N, Krajnovic D. The instruments used to assess health literacy and pharmacotherapy literacy of diabetes mellitus type 2 patients: a scoping review. *Front Public Health*. (2021) 9:747807. doi: 10.3389/fpubh.2021.747807

13. Zheng F, Ding S, Lai L, Liu X, Duan Y, Shi S, et al. Relationship between medication literacy and medication adherence in inpatients with coronary heart disease in Changsha, China. *Front Pharmacol*. (2019) 10:1537. doi: 10.3389/fphar.2019.01537

14. Shi S, Shen Z, Duan Y, Ding S, Zhong Z. Association between medication literacy and medication adherence among patients with hypertension. *Front Pharmacol*. (2019) 10:822. doi: 10.3389/fphar.2019.00822

15. Wang W, Luan W, Zhang Z, Mei Y. Association between medication literacy and medication adherence and the mediating effect of self-efficacy in older people with multimorbidity. *BMC Geriatr*. (2023) 23:378. doi: 10.1186/s12877-023-04072-0

16. Qu J, Zhou T, Xue M, Sun H, Shen Y, Liu Y. Relationship between medication literacy and frailty in elderly inpatients with coronary heart disease: A cross-sectional study in China. *Front Pharmacol*. (2021) 12:691983. doi: 10.3389/fphar.2021.691983

17. Donaldson LJ, Kelley ET, Dhingra-Kumar N, Kieny MP, Sheikh A. Medication without harm: WHO'S third global patient safety challenge. *Lancet*. (2017) 389:1680–1. doi: 10.1016/s0140-6736(17)31047-4

18. Ma G, Luo A, Shen Z, Duan Y, Shi S, Zhong Z. The status of medication literacy and associated factors of hypertensive patients in China: a cross-sectional study. *Intern Emerg Med*. (2020) 15:409–19. doi: 10.1007/s11739-019-02187-0

19. Levic M, Bogavac-Stanojevic N, Ubavic S, Krajnovic D. Pharmacotherapy literacy level and predictors of low literacy among diabetes mellitus type 2 patients in Serbia. *BMC Public Health*. (2023) 23:1822. doi: 10.1186/s12889-023-16639-y

20. Ayaz-Alkaya S, Ozturk FO. Health literacy levels of women and related factors in Turkey. *J Nurs Res*. (2021) 29:e180. doi: 10.1097/jnr.0000000000000452

21. Neiva Pantuzza LL, Nascimento ED, Crepalde-Ribeiro K, Botelho SF, Parreiras Martins MA, de Souza C, et al. Medication literacy: A conceptual model. *Res Social Adm Pharm*. (2022) 18:2675–82. doi: 10.1016/j.sapharm.2021.06.003

22. Neiva Pantuzza LL, Reis AMM, Botelho SF, da Rocha ALP, Martins MAP, do Nascimento MMG, et al. Medication literacy test for older adults: psychometric analysis and standardization of the new instrument. *Int J Clin Pharm*. (2024) 46:1124–33. doi: 10.1007/s11096-024-01744-8

23. Gnägi R, Zúñiga F, Brunkert T, Meyer-Massetti C. Development of a medication literacy assessment instrument (MELIA) for older people receiving home care. *J Adv Nurs*. (2022) 78:4210–20. doi: 10.1111/jan.15429

24. Lu T, Yang Z, Chen P, Li J, Zheng C, Kong L, et al. Influencing factors of medication literacy among community-dwelling older adult patients with hypertension: a study based on social learning theory. *Front Pharmacol*. (2023) 14:1184701. doi: 10.3389/fphar.2023.1184701

25. Chen C, Song M. Visualizing a field of research: A methodology of systematic scientometric reviews. *PLoS One*. (2019) 14:e0223994. doi: 10.1371/journal.pone.0223994

26. Mingers J, Leydesdorff L. A review of theory and practice in scientometrics. *Eur J Oper Res*. (2015) 246:1–19. doi: 10.1016/j.ejor.2015.04.002

27. Chen C, Chen Y, Hou J, Liang Y. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J China Soc Sci Tech Inf*. (2009) 28:401–21. doi: 10.1002/asi.20317

28. Aria M, Cuccurullo C. *Bibliometrix*: an R-tool for comprehensive science mapping analysis. *J Informetr*. (2017) 11:959–75. doi: 10.1016/j.joi.2017.08.007

29. Singh VK, Singh P, Karmakar M, Leta J, Mayr P. The journal coverage of web of science, Scopus and dimensions: a comparative analysis. *Scientometrics*. (2021) 126:5113–42. doi: 10.1007/s11192-021-03948-5

30. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *FASEB J*. (2008) 22:338–42. doi: 10.1096/fj.07-9492LSF

31. AlRyalat SAS, Malkawi LW, Momani SM. Comparing bibliometric analysis using PubMed, Scopus, and web of science databases. *J Vis Exp*. (2019) 24:152. doi: 10.3791/58494

32. Chen CM, Ibekwe-SanJuan F, Hou JH. The structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis. *J Am Soc Inf Sci Technol*. (2010) 61:1386–409. doi: 10.1002/asi.21309

33. Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. (2010) 84:523–38. doi: 10.1007/s11192-009-0146-3

34. Chen CM. Predictive effects of structural variation on citation counts. *J Am Soc Inf Sci Technol*. (2012) 63:431–49. doi: 10.1002/asi.21694

35. Kleinberg J. Bursty and hierarchical structure in streams. *Data Min Knowl Discov*. (2003) 7:373–97. doi: 10.1023/a:1024940629314

36. Schillinger D, Grumbach K, Piette J, Wang F, Osmond D, Daher C, et al. Association of health literacy with diabetes outcomes. *JAMA*. (2002) 288:475–82. doi: 10.1001/jama.288.4.475

37. Davis TC, Wolf MS, Bass PF 3rd, Thompson JA, Tilson HH, Neuberger M, et al. Literacy and misunderstanding prescription drug labels. *Ann Intern Med*. (2006) 145:887–94. doi: 10.7326/0003-4819-145-12-200612190-00144

38. Davis TC, Wolf MS, Bass PF 3rd, Middlebrooks M, Kennen E, Baker DW, et al. Low literacy impairs comprehension of prescription drug warning labels. *J Gen Intern Med*. (2006) 21:847–51. doi: 10.1111/j.1525-1497.2006.00529.x

39. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*. (2011) 155:97–107. doi: 10.7326/0003-4819-155-2-201107190-00005

40. Zhang NJ, Terry A, McHorney CA. Impact of health literacy on medication adherence: a systematic review and meta-analysis. *Ann Pharmacother*. (2014) 48:741–51. doi: 10.1177/1060028014526562

41. Yin HS, Mendelsohn AL, Wolf MS, Parker RM, Fierman A, van Schaick L, et al. Parents' medication administration errors: role of dosing instruments and health literacy. *Arch Pediatr Adolesc Med*. (2010) 164:181–6. doi: 10.1001/archpediatrics.2009.269

42. Wolf MS, Davis TC, Curtis LM, Webb JA, Bailey SC, Shrank WH, et al. Effect of standardized, patient-centered label instructions to improve comprehension of prescription drug use. *Med Care*. (2011) 49:96–100. doi: 10.1097/MLR.0b013e3181f38174

43. Lee CH, Chang FC, Hsu SD, Chi HY, Huang LJ, Yeh MK. Inappropriate self-medication among adolescents and its association with lower medication literacy and substance use. *PLoS One*. (2017) 12:e0189199. doi: 10.1371/journal.pone.0189199

44. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. (2005) 353:487–97. doi: 10.1056/NEJMra050100

45. Kripalani S, Robertson R, Love-Ghaffari MH, Henderson LE, Praska J, Strawder A, et al. Development of an illustrated medication schedule as a low-literacy patient education tool. *Patient Educ Couns*. (2007) 66:368–77. doi: 10.1016/j.pec.2007.01.020

46. Wolf MS, Davis TC, Osborn CY, Skripkauskas S, Bennett CL, Makoul G. Literacy, self-efficacy, and HIV medication adherence. *Patient Educ Couns*. (2007) 65:253–60. doi: 10.1016/j.pec.2006.08.006

47. Persell SD, Osborn CY, Richard R, Skripkauskas S, Wolf MS. Limited health literacy is a barrier to medication reconciliation in ambulatory care. *J Gen Intern Med*. (2007) 22:1523–6. doi: 10.1007/s11606-007-0334-x

48. Quinlan P, Price KO, Magid SK, Lyman S, Mandl LA, Stone PW. The relationship among health literacy, health knowledge, and adherence to treatment in patients with rheumatoid arthritis. *HSS J*. (2013) 9:42–9. doi: 10.1007/s11420-012-9308-6

49. Bosworth HB, Granger BB, Mendys P, Brindis R, Burkholder R, Czajkowski SM, et al. Medication adherence: a call for action. *Am Heart J*. (2011) 162:412–24. doi: 10.1016/j.ahj.2011.06.007

50. Marzec LN, Maddox TM. Medication adherence in patients with diabetes and dyslipidemia: associated factors and strategies for improvement. *Curr Cardiol Rep*. (2013) 15:418. doi: 10.1007/s11886-013-0418-7

51. Tahkola A, Korhonen P, Kautiainen H, Niiranen T, Mäntyselkä P. Feasibility of a checklist in treating hypertension in primary care – base line results from a cluster-randomised controlled trial (check and support). *BMC Cardiovasc Disord*. (2018) 18:11. doi: 10.1186/s12872-018-0963-5

52. Kung PC, Yeh MC, Lai MK, Liu HE. Renal transplant recipients: the factors related to immunosuppressive medication adherence based on the health belief model. *J Nurs Res*. (2017) 25:392–7. doi: 10.1097/jnr.0000000000000181

53. Schüz B, Wurm S, Ziegelmann JP, Warner LM, Tesch-Römer C, Schwarzer R. Changes in functional health, changes in medication beliefs, and medication adherence. *Health Psychol*. (2011) 30:31–9. doi: 10.1037/a0021881

54. Richardson M, McCabe R, Priebe S. Are attitudes towards medication adherence associated with medication adherence behaviours among patients with psychosis? A systematic review and meta analysis. *Soc Psychiatry Psychiatr Epidemiol*. (2013) 48:649–57. doi: 10.1007/s00127-012-0570-1

55. Islam SMS, Maddison R, Karmarkar C, Rahman S. Can digital health help improve medication adherence in cardiovascular disease? *Expert Rev Med Devices*. (2024) 21:1071–5. doi: 10.1080/17434440.2024.2441954

56. Haga SB. Toward digital-based interventions for medication adherence and safety. *Expert Opin Drug Saf*. (2020) 19:735–46. doi: 10.1080/14740338.2020.1764935

57. Piña IL, Di Palo KE, Brown MT, Choudhry NK, Cvengros J, Whalen D, et al. Medication adherence: importance, issues and policy: A policy statement from the American Heart Association. *Prog Cardiovasc Dis*. (2021) 64:111–20. doi: 10.1016/j.pcad.2020.08.003

58. Tan JP, Cheng KKF, Siah RCJ. A systematic review and meta-analysis on the effectiveness of education on medication adherence for patients with hypertension, hyperlipidaemia and diabetes. *J Adv Nurs*. (2019) 75:2478–94. doi: 10.1111/jan.14025

59. Gentizon J, Hirt J, Jaques C, Lang PO, Mabire C. Instruments assessing medication literacy in adult recipients of care: A systematic review of measurement properties. *Int J Nurs Stud*. (2021) 113:103785. doi: 10.1016/j.ijnurstu.2020.103785

60. Sauceda JA, Loya AM, Sias JJ, Taylor T, Wiebe JS, Rivera JO. Medication literacy in Spanish and English: psychometric evaluation of a new assessment tool. *J Am Pharm Assoc (2003)*. (2012) 52:e231–40. doi: 10.1331/JAPhA.2012.11264

61. Emmerton LM, Mampallil L, Kairuz T, McKauge LM, Bush RA. Exploring health literacy competencies in community pharmacy. *Health Expect*. (2012) 15:12–22. doi: 10.1111/j.1369-7625.2010.00649.x

62. Jang SM, Jiang R, Grabe D, Pai AB. Assessment of literacy and numeracy skills related to non-steroidal anti-inflammatory drug labels. *SAGE Open Med*. (2019) 7:2050312119834119. doi: 10.1177/2050312119834119

63. Zhang N, Wang L, Ouyang YQ, Redding S. Survey on medication information literacy and influencing factors among pregnant Chinese women. *J Matern Fetal Neonatal Med*. (2021) 34:1619–26. doi: 10.1080/14767058.2019.1642869

64. Van den Oever F, Vasbinder E, Schrama Y, Koster E, Van den Bemt P, Van Gelder T. Limited pharmaceutical literacy in patients on hemodialysis in the Netherlands as assessed with the Ralph interview guide. *Nephrol Dial Transplant*. (2023) 38:I158-I. doi: 10.1093/ndt/gfad063b_3102

65. Zhong ZQ, Ma GY, Zheng F, Duan YL, Ding SQ, Luo AJ. Medication literacy in a cohort of Chinese patients discharged with essential hypertension. *Front Public Health*. (2020) 7:5. doi: 10.3389/fpubh.2019.00385

66. Jordan JE, Buchbinder R, Briggs AM, Elsworth GR, Busija L, Batterham R, et al. The health literacy management scale (HeLMS): a measure of an individual's capacity to seek, understand and use health information within the healthcare setting. *Patient Educ Couns*. (2013) 91:228–35. doi: 10.1016/j.pec.2013.01.013

67. Ge LX, Heng BH, Yap CW. Understanding reasons and determinants of medication non-adherence in community-dwelling adults: a cross-sectional study comparing young and older age groups. *BMC Health Serv Res*. (2023) 23:905. doi: 10.1186/s12913-023-09904-8

68. Semahegn A, Torpey K, Manu A, Assefa N, Tesfaye G, Ankomah A. Psychotropic medication non-adherence and its associated factors among patients with major psychiatric disorders: a systematic review and meta-analysis. *Syst Rev*. (2020) 9:17. doi: 10.1186/s13643-020-1274-3

69. Marks JR, Schectman JM, Groninger H, Plews-Ogan ML. The association of health literacy and socio-demographic factors with medication knowledge. *Patient Educ Couns*. (2010) 78:372–6. doi: 10.1016/j.pec.2009.06.017

70. Nutbeam D, McGill B, Premkumar P. Improving health literacy in community populations: a review of progress. *Health Promot Int*. (2018) 33:901–11. doi: 10.1093/heapro/dax015

71. Shen Z, Ding S, Shi S, Zhong Z. Association between social support and medication literacy in older adults with hypertension. *Front Public Health*. (2022) 10:987526. doi: 10.3389/fpubh.2022.987526

72. DiMatteo MR. Social support and patient adherence to medical treatment: a meta-analysis. *Health Psychol*. (2004) 23:207–18. doi: 10.1037/0278-6133.23.2.207

73. Brod M, Kongsø JH, Lessard S, Christensen TL. Psychological insulin resistance: patient beliefs and implications for diabetes management. *Qual Life Res*. (2009) 18:23–32. doi: 10.1007/s11136-008-9419-1

74. Brauer M, Roth GA, Aravkin AY, Zheng P, Abate KH, Abate YH, et al. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990-2021: a systematic analysis for the global burden of disease study 2021. *Lancet*. (2024) 403:2162–203. doi: 10.1016/s0140-6736(24)00933-4

75. Muluneh B, Upchurch M, Belayneh B, Mackler E, Bryant AL, Wood WA, et al. Design and implementation of a risk-adapted, longitudinal, theory-driven medication adherence intervention: A protocol for a multi-phasic, hybrid effectiveness-implementation trial. *Res Social Adm Pharm*. (2025) 21:444–52. doi: 10.1016/j.sapharm.2025.02.006

76. Lopes N, Rodrigues C, Pegado E. Medication literacy and its social contextuality. *Health (London)*. (2024) 28:858–76. doi: 10.1177/13634593231211520

77. Dang TH, Wickramasinghe N, Jayaraman PP, Burbury K, Alexander M, Whitechurch A, et al. Digital solution to support medication adherence and self-Management in Patients with Cancer (SAMSON): pilot randomized controlled trial. *JMIR Form Res*. (2025) 9:e65302. doi: 10.2196/65302

78. Kamal AK, Khalid W, Zulfiqar M, Muqeet A, Zaidi F, Gowani A, et al. The potential of m-health-based interventions to improve medication literacy and adherence in non-communicable diseases in Pakistan. *Int J Stroke*. (2019) 14:Np8–Np10. doi: 10.1177/1747493019827762

79. Qin N, Duan Y, Yao Z, Shi S, Liu H, Li X, et al. Psychometric properties and validation of the revised Chinese medication literacy scale for hypertensive patients (C-MLSHP-R). *Front Cardiovasc Med*. (2022) 9:976691. doi: 10.3389/fcvm.2022.976691

80. Saguban R, Shanmugam SR, Rondilla E, Buta J, Alatawi NAH, Maestrado R, et al. Self-efficacy, social support, and depression: mediators of medication adherence in Dialysis patients. *Healthcare (Basel)*. (2025) 13:425. doi: 10.3390/healthcare13040425

Physician professional motivation and online knowledge sharing for patient education: a perspective of motivation theory

Yun Huang[1†], Junping Guo[2†], Yan Wen[3†] and Qihui Fan[4]*

[1]School of Economics, Tianjin University of Finance and Economics, Tianjin, China, [2]Department of Pharmacy, Hubei Province Corps Hospital, The Chinese Armed Police Force (CAPF), Wuhan, China, [3]Department of General Medicine, Wuhan Third Hospital (Tongren Hospital of Wuhan University), Wuhan, China, [4]College of Management and Economics, Tianjin University, Tianjin, China

Aims: Drawing on motivation theory, this study aims to investigate the effect of professional motivation on online knowledge sharing for patient education with considering the contingencies of online experience and offline expertise.

Methods: Based on a panel dataset comprising 11,839 physicians with 24,389 physician-month observations selected from one of leading online health platforms in China, this study conducted the fixed hierarchical regression model to test the direct and moderating effects.

Results: The results show that professional motivation positively affects online knowledge-sharing quantity and quality. Meanwhile, online experience enhances the positive effects of professional motivation on the above two dimensions of online knowledge sharing. In addition, offline expertise hinders the benefits of professional motivation to online knowledge-sharing quantity.

Conclusion: This study makes contributions to the literatures of motivation theory, online knowledge sharing, online and offline contexts on online health platforms, and provides implications for physicians and platform managers.

KEYWORDS

professional motivation, online knowledge sharing, online experience, offline expertise, patient education

1 Introduction

The rapid development of digital technologies has accelerated the advancement of online healthcare services and transformed how medical knowledge is disseminated and accessed (1–3). Online health platforms (OHPs) play a pivotal role in people's health and disease management, as physicians increasingly engage in online knowledge-sharing activities to educate patients, which is beneficial to patient outcomes and ultimately improves societal healthcare outcomes (4–6). Physicians' knowledge sharing about OHPs is largely driven by their underlying motivations, such as their medical expertise and professional obligations (7). Therefore, motivation theory plays a critical role in understanding how different types of motivations (e.g., extrinsic and intrinsic motivation) affect physicians' online engagement (8, 9).

Existing research has extensively explored the role of intrinsic motivation such as interest or inherent satisfaction (10, 11) and extrinsic motivation such as specific expected outcomes or external rewards (12) in online knowledge sharing. Recently, recognizing that physicians with specialized and critical medical knowledge are often driven by a sense of obligation and professional identity, even when knowledge sharing requires significant time and effort (13),

scholars have shifted their attention toward professional motivation and found that they are motivated to share both free and paid health knowledge (7, 14, 15). However, this line of research has not systematically examined how professional motivation shapes the extent and quality of physicians' knowledge-sharing behaviors. In particular, little is known about how professional motivation influences both the quantity and quality of shared content—two critical dimensions for understanding the effectiveness of physicians' online engagement. Knowledge-sharing quantity refers to the volume of knowledge shared by physicians, while knowledge-sharing quality reflects the substantive value and utility of the shared content (16). Exploring the impact of professional motivation on knowledge-sharing quantity and quality is important in understanding the nuanced effects of professional motivation. Therefore, this study addresses this oversight by dissecting how professional motivation differentially influences these two dimensions of online knowledge sharing.

Additionally, contextual factors are increasingly recognized as critical moderators in the relationship between motivation and online knowledge sharing. According to the previous studies, online experience and offline expertise are important contingencies of physicians, which can shape their online knowledge-sharing behaviors (7, 17). Online experience refers to the accumulated knowledge and skills for using online platforms (6). Existing studies show that extensive online experience enhances physicians' technical proficiency and confidence in online interactions (6). Conversely, offline expertise refers to the experience developed after years of clinical work and professional assessment (17). Physicians with extensive offline expertise disseminate valuable information (18) while facing considerable workloads, which limit their time and energy (19). Online experience and offline expertise tend to shape the effects of professional motivation on online knowledge sharing. However, previous studies have not simultaneously considered how physicians' online experience and offline expertise affect the relationship between professional motivation and online knowledge sharing. Systematically considering two contingencies and guide physicians in knowledge sharing for patient education, thereby improving public health equity and welfare.

To address the above-mentioned gaps, this study aims to explore the effects of physicians' professional motivation on the quantity and the quality of online knowledge sharing for patient education while considering the contingencies of online experience and offline expertise. Specifically, we seek to answer the following research questions:

> (1) *How does physicians' professional motivation influence online knowledge-sharing quantity and quality for patient education?* (2) *What role do online experience and offline expertise play in these relationships?*

According to motivation theory, professional motivation is the driving force in physicians sharing health knowledge on OHPs for patient education. Physicians can obtain patients' recognition (15), enhance self-efficacy (7), pursue continuous self-growth (20), and challenge themselves (21), thereby increasing their motivation to share more knowledge online. Similarly, professional motivation, accompanied with a sense of obligation (15), confidence as authoritative guides (22), and creativity (23), also enhances physicians'

online knowledge-sharing quality. Additionally, physicians with substantial online experience demonstrate elevated perceptions of professional value due to their digital competencies and enhanced patient education capabilities (6). In this context, physicians with professional motivation are more likely to contribute a greater volume of knowledge with higher quality to support patient education. However, since distinct professional priorities are associated with varying levels of offline expertise (7, 19), the effects of professional motivation on online knowledge-sharing quantity and quality tend to be shaped by offline expertise.

The analysis in this study employed a panel dataset comprising 11,839 physicians with 24,389 physician-month observations selected from Haodf.com. The results indicate that most of the hypotheses are supported. Specifically, professional motivation is positively associated with both the quantity and quality of online knowledge sharing. Additionally, online experience further strengthens these positive effects, while offline expertise weakens the positive effect of professional motivation on online knowledge-sharing quantity. This study also makes several contributions. Firstly, this study contributes to the literature of motivation theory by focusing on professional motivation in the contexts of OHPs and identifying the formation mechanism of online knowledge sharing for patient education. Secondly, this study contributes to the literature on online knowledge sharing by revealing the positive effects of professional motivation on online knowledge-sharing quantity and quality. Thirdly, this study contributes to the literature on online and offline contexts in online knowledge sharing by uncovering the heterogeneous moderating effects of online experience and offline expertise about the impact of professional motivation on online knowledge-sharing quantity and quality.

2 Theory and hypotheses

2.1 Motivation theory

Individual needs and expectations play a crucial role in shaping behaviors (24). Drawing on this fundamental tenet, motivation defines the direction and underlying rationale of behavioral patterns, driving individuals to act in specific ways (6, 8). Motivation is commonly divided into two dimensions—extrinsic and intrinsic motivation—reflecting different attitudes and goal orientations (8). The former refers to engaging in an activity out of genuine interest or inherent satisfaction, while the latter involves undertaking an activity to achieve expected outcomes or external rewards (25).

As OHPs become increasingly prominent sources of medical knowledge (26), motivation theory has been widely applied to examine how different types of intrinsic and extrinsic motivations individually and collectively affect physicians' online engagement. Zhuo and Wang (27) found that intrinsic (e.g., competence and autonomy satisfaction) and extrinsic motivation (e.g., economic benefits) positively influence physicians' service behaviors. Wang et al. (9) demonstrated a crowding-out effect of informal payments on physicians' intrinsic motivation to participate in online consultation. Beyond this classic dichotomy, recent studies have highlighted the specific role of professional motivation—a key form of intrinsic motivation—in driving the dissemination of life-critical medical knowledge by trained professionals on OHPs (14, 15). Professional

motivation refers to a psychological process that influences professionals in fulfilling their goals and tasks. Zhang et al. (7) and Yang et al. (15) found that professional motivation fosters both free and paid knowledge-sharing behaviors on OHPs. However, although these studies addressed physicians' intention, its influence on the actual performance of online knowledge-sharing behaviors remain underexplored. Unlike general users, physicians, driven by irreplaceable medical expertise and strong professional ethics, demonstrate distinct patterns of engagement on OHPs (13). To fully assess their contributions, both the quantity and quality of online knowledge sharing should be considered (28, 29), as they jointly determine the effectiveness of patient education and the sustainable development of OHPs.

This study applied motivation theory to study how physicians' professional motivation influences both the quantity and quality of online knowledge sharing. In our framework, physicians driven by professional motivation tend to develop a stronger sense of responsibility for patient education, a desire for self-growth, and a willingness to embrace professional challenges (15), which encourages more frequent and higher-quality contributions on OHPs. Additionally, physicians' motivating behaviors are inevitably shaped by individual difference arising from their experiences in both online and offline environments (7). Online experience strengthens the relationship between motivation and engagement by lowering technical barriers and increasing confidence in online interactions (6, 30). Offline expertise, however, may exert a dual influence: Greater clinical responsibilities lead to limit the frequency of contributions (12), whereas advanced professional seniority with deeper knowledge and practical insights can enhance the quality of shared content (19). Therefore, based on motivation theory, this study aims to explore the effects of professional motivation on online knowledge-sharing quantity and quality by considering contingencies of online experience and offline expertise.

2.2 Professional motivation and online knowledge sharing

Professional motivation refers to a psychological process that influences professionals (e.g., physicians and teachers) in fulfilling their goals and tasks (15). As an intrinsic motivation, professional motivation affects physicians' online knowledge-sharing behaviors, encompassing two dimensions—online knowledge-sharing quantity and online knowledge-sharing quality (29). The former reflects the volume of knowledge shared by physicians on OHPs, while the latter pertains to the substantive value and utility of that content (16). According to motivation theory, professional motivation—characterized by heightened professional obligation, enhanced self-efficacy, and sustained persistence (7, 22)—significantly enhances physicians' knowledge-sharing quantity and quality on OHPs.

We expect physicians' professional motivation to be positively related to online knowledge-sharing quantity. First, professional motivation can induce physicians to use their expertise to obtain patients' recognition, for example, votes and readings (7). The resulting fulfillment and enjoyment motivate them to make unceasing contributions to OHPs. Second, professional motivation encourages physicians to pursue continuous self-growth (20). Engaging in knowledge sharing on OHPs enables them to help patients manage

health and deepen their own understanding of medical information, which, in turn, sparks their creative enthusiasm and increases the number of published articles. Third, guided by professional interests and internal goals, physicians tend to challenge themselves (21), remaining committed to frequently updating medical content and addressing patient inquiries (15), even at the cost of personal time. Based on these arguments, we propose the following hypothesis.

> *H1*: Professional motivation is positively related to physicians' online knowledge-sharing quantity.

Similarly, we expect physicians' professional motivation to be positively related to online knowledge-sharing quality. First, professional motivation creates a strong sense of obligation in physicians to disseminate medical knowledge more extensively and efficiently through OHPs (15). This commitment ensures that their contributions meet professional standards, enhancing online knowledge-sharing quality. Second, beyond responsibility-driven factors, professional motivation boosts physicians' confidence as authoritative guides in patient education (22), especially in the information-overloaded digital health landscape. By identifying misleading information and providing accurate clinic expertise (31), they further contribute to delivering online high-quality services. Third, professional motivation, as an intrinsically motivated orientation, is positively associated with creativity (23). Physicians with high creativity can develop innovative solutions to meet patients' diverse healthcare needs on OHPs (7), which improves the educational value and practical utility of their efforts. Based on these arguments, we propose the following hypothesis.

> *H2*: Professional motivation is positively related to physicians' online knowledge-sharing quality.

2.3 The moderating effect of online experience

Online experience refers to the professionals' accumulated knowledge and skills in using online platforms (6). Physicians with abundant online experience tend to perceive higher professional value, as they possess digital expertise and patient education skills (6). In this vein, the effects of professional motivation on online knowledge-sharing behaviors are contingent on online experience.

We expect online experience to strengthen the positive effect of professional motivation and physicians' online knowledge-sharing quantity. With accumulated online experience, physicians not only achieve greater self-actualization by providing targeted knowledge sharing in response to high-interest health concerns (20), but they also become more skilled at developing efficient content strategies that support ongoing skill development and personal growth (32). These dual benefits jointly amplify how professional motivation drives sustained content contributions (33). In addition, physicians with extensive online experience demonstrate strong platform familiarity and advanced self-regulation skills that help optimize their knowledge-sharing practices (34, 35). Accordingly, the internal drive stemming from professional motivation is more likely to result in consistent and high-volume knowledge-sharing outputs. Thus, we propose the following hypothesis.

H3a: Online experience strengthens the positive relationship between professional motivation and physicians' online knowledge-sharing quantity.

We also expect online experience to strengthen the positive effect of professional motivation and physicians' online knowledge-sharing quality. Physicians with greater online experience are inclined to build a favorable reputation and achieve robust online socialization through consistent knowledge sharing (36). This process fosters heightened feelings of professional responsibility and self-efficacy, which supports their commitment to providing effective patient education (36, 37). Moreover, proficiency in the features and affordances of OHPs provides experienced physicians with greater flexibility in content design and delivery (6, 33). Under such circumstances, they are better positioned to match their professional values with patients' expectations (20), further enhancing their intrinsic drive to contribute high-quality services. Thus, we propose the following hypothesis.

H3b: Online experience strengthens the positive relationship between professional motivation and physicians' online knowledge-sharing quality.

2.4 The moderating effect of offline expertise

Offline expertise refers to the experience in delivering healthcare services and medical proficiency gained within traditional hospital settings. This expertise is often embodied in physicians' clinical titles across four hierarchical levels—resident doctor, attending doctor, associate chief doctor, and chief doctor—signifying their professional capabilities and seniority (32). As professional priorities vary with levels of offline expertise (7, 19), offline expertise differentially moderates the relationship between physicians' professional motivation and their online knowledge-sharing behaviors on OHPs.

We expect offline expertise to weaken the positive effect of professional motivation on physicians' online knowledge-sharing quantity. Physicians with high levels of offline expertise (as indicated by senior clinical titles) devote greater attention to diagnosing complex medical cases to deliver specialized healthcare in offline channels (13, 38) while advancing their expertise through research and academic engagement (22). Consequently, routine health information sharing via OHPs with professional motivation is deprioritized, resulting in reduced levels of contributions. Besides, physicians in senior positions often face considerable workloads in offline hospitals, limiting the time and energy available for online engagement (19, 39), which leads to a decline in the frequency of their online knowledge sharing. Based on the above arguments, we propose the following hypothesis.

H4a: Offline expertise weakens the positive relationship between professional motivation and physicians' online knowledge-sharing quantity.

Conversely, we expect offline expertise to strengthen the positive effect of professional motivation on physicians' online knowledge-sharing quality. Motivated by a love of their work itself, accompanied with substantial professional expertise (7), physicians can cultivate a stronger sense of mission and greater confidence in their capabilities.

In this context, their intrinsic professional motivation can spur the dissemination of high-quality, valuable information and experiences on OHPs (18). Additionally, by drawing on extensive face-to-face clinical experience and unique insights, physicians are better able to anticipate patients' concerns and present in-depth medical knowledge on OHPs in a more readable manner (19). In other words, the effect of professional motivation on online knowledge-sharing quality is enhanced by offline expertise. Based on the above arguments, we propose the following hypothesis.

H4b: Offline expertise strengthens the positive relationship between professional motivation and physicians' online knowledge-sharing quality.

In summary, the research framework is presented in Figure 1.

3 Methods

3.1 Data collection

This study selected Haodf.com[1] as our data source to its objective, real-world data that effectively mitigates self-reporting biases (40, 41). As one of leading OHPs in China, Haodf.com offers several key advantages for our research. Firstly, it hosts an extensive network of physicians and patient interactions, ensuring a robust dataset. Secondly, physicians have a dedicated article section where they can independently publish accessible health-related articles without receiving any financial compensation from the platform. These knowledge-sharing activities are systematically tracked by the platform, which records several key indicators, including the number of votes received, the number of health-related articles published, and the number of article readings. In addition, physicians' homepages display relevant background information, such as their affiliated offline hospital, professional title, and online contribution experience, which provides valuable contextual data for this study.

To collect data, a Java-based web crawler was adopted to extract website statistics about physicians and patients over a six-month period from February to July 2017. The initial dataset included all physician profiles and their corresponding knowledge-sharing activity records available during this period. We excluded entries corresponding to physicians with incomplete profiles or without any recorded knowledge-sharing activity. This filtering process resulted in an unbalanced panel dataset comprising 11,839 physicians with 24,389 physician-month observations. This dataset enables longitudinal analysis of physicians' online activities while accounting for variations in participation over time.

3.2 Measures

3.2.1 Dependent variables

Online knowledge-sharing quantity refers to the volume of knowledge shared by physicians in online platforms (16), which was

---

1 www.haodf.com
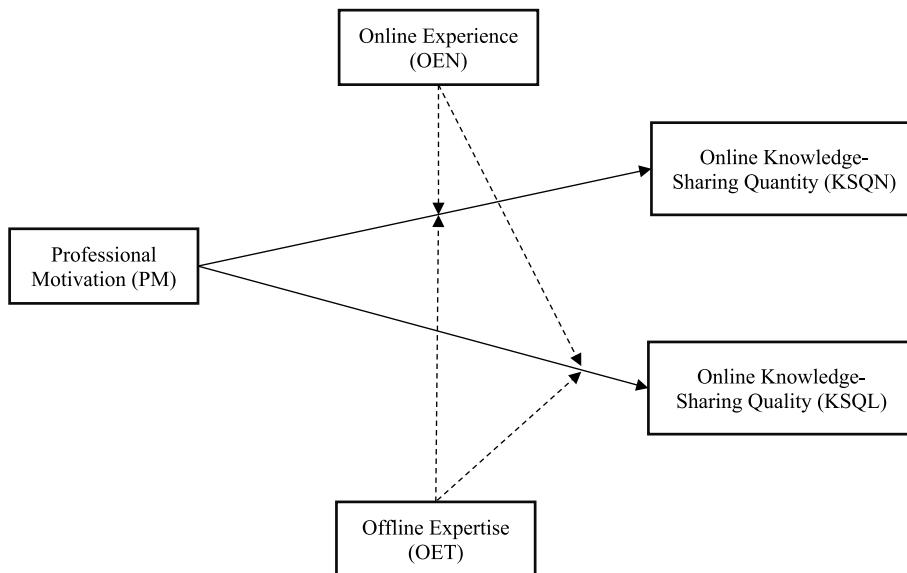
FIGURE 1
Research framework.

measured by the number of health-related articles published by physicians. Online knowledge-sharing quality reflects the substantive value and utility of the shared content in online platforms (16). It was measured by the number of article readings.

3.2.2 Independent variables and moderators

Professional motivation reflects a psychological process that influences professionals in fulfilling their goals and tasks (7, 15). The number of votes received by physicians was adopted to measure professional motivation. As physicians do not receive financial incentives for publishing online articles, voting reflects patients' recognition of their voluntary knowledge-sharing efforts (7). In this context, such behaviors are primarily driven by professional motivation, including a sense of obligation to contribute to patient education and a desire for self-development. Online experience reflects the professionals' accumulated knowledge and skills in using online platforms (6), which was measured by the opening time of physicians. Offline expertise refers to the experience in delivering healthcare services and medical proficiency gained within traditional hospital settings (17). It was measured by the offline titles of physicians. According to the hospital title hierarchy in China, we included three dummy variables for four levels, ordered from lowest to highest: resident physician, attending physician, associate chief physician, and chief physician (17).

3.2.3 Control variables

In studies on physician online knowledge sharing and patient education, previous scholars suggest that gifts and likes may be seen as factors influencing the online knowledge-sharing behaviors (17, 42); therefore, we choose them as control variables. Gifts were measured by the number of online gifts from patients. Likes were measured by the number of online loves from patients.

To reduce skewness, we used the logarithm of all variables except offline expertise. Table 1 presents an overview of all variables in this study.

3.3 Analytic strategy

Considering the inefficiency and estimated bias of the ordinary least squares regression model, this study conducted the fixed hierarchical regression model to test the direct and moderating effects (43, 44). To test our hypotheses, we introduced Equations 1, 2 to estimate the effects of professional motivation (PM) on online knowledge-sharing quantity (KSQN) and online knowledge-sharing quality (KSQL) with the contingencies of online experience (OEN) and offline expertise (OET).

$$KSQN_{it} = \alpha_0 + \alpha_1 Gift_{it} + \alpha_2 Like_{it} + \alpha_3 PM_{it} + \alpha_4 OEN_{it} \\ + \alpha_5 OET_{it} + \alpha_6 PM_{it} \times OEN_{it} + \alpha_7 PM_{it} \times OET_{it} + \mu_{it} \quad (1)$$

$$KSQL_{it} = \beta_0 + \beta_1 Gift_{it} + \beta_2 Like_{it} + \beta_3 PM_{it} + \beta_4 OEN_{it} \\ + \beta_5 OET_{it} + \beta_6 PM_{it} \times OEN_{it} + \beta_7 PM_{it} \times OET_{it} + \nu_{it} \quad (2)$$

where i indicates the number of observations, the α and β parameters are the coefficients that can be estimated in the hierarchical regression model, and the μ and ν parameters are the error terms in each equation.

4 Results

4.1 Regression analysis

Our hypotheses were tested using hierarchical regression, which is widely applied to test moderating effects (45). The results are presented in Table 2.

Model 1 shows that professional motivation is positively and significantly related to online knowledge-sharing quantity ($\beta = 0.084$, $p < 0.001$). Thus, H1 is supported. The coefficient of interaction term (PM × OEN) in Model 2 is positive and significant ($\beta = 0.067$, $p < 0.001$). Following the suggestions of Meyer et al. (46), we plotted

TABLE 1 The overview of all variables.

Variables	Measurements	Mean	SD	Min	Max
Online knowledge-sharing quantity	The number of health-related articles published by physicians	2.011	1.280	0.000	7.440
Online knowledge-sharing quality	The number of article readings	7.899	1.108	2.079	12.753
Professional motivation	The number of votes physicians received	2.634	1.239	0.000	6.911
Online experience	The opening time of physicians	7.270	0.798	3.332	8.010
Offline expertise	The offline titles of physicians	3.098	0.869	1.000	4.000
Gifts	The number of online gifts from patients	2.506	1.385	0.000	7.920
Likes	The number of online likes from patients	3.789	2.030	0.000	10.140

TABLE 2 Results of hierarchical regression.

Online knowledge-sharing quantity (KSQN)	Model 1	Model 2	Model 3	Model 4
Professional motivation (PM)	0.084***	-0.472***	0.081***	-0.497***
	(0.008)	(0.062)	(0.025)	(0.062)
Online experience (OEN)		0.150***		0.084***
		(0.021)		(0.023)
PM×OEN		0.067***		0.083***
		(0.008)		(0.009)
Offline expertise (OET)			0.198***	0.178***
			(0.020)	(0.022)
PM×OET			-0.017*	-0.039***
			(0.007)	(0.008)
Gifts	0.387***		0.370***	0.284***
	(0.012)		(0.012)	(0.013)
Likes	-0.126***		-0.101***	-0.034***
	(0.008)		(0.008)	(0.009)
Constant	1.298***		0.786***	0.294
	(0.019)		(0.061)	(0.152)
R^2	0.100	0.129	0.109	0.132
Online knowledge-sharing quality (KSQL)	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Professional motivation (PM)	0.151***	-0.168**	0.116***	-0.163**
	(0.007)	(0.055)	(0.022)	(0.055)
Online experience (OEN)		0.161***		0.171***
		(0.019)		(0.020)
PM×OEN		0.036***		0.036***
		(0.007)		(0.008)
Offline expertise (OET)			0.020	-0.034
			(0.018)	(0.019)
PM×OET			0.007	0.001
			(0.007)	(0.007)
Gifts	0.284***	0.200***	0.279***	0.200***
	(0.011)	(0.011)	(0.012)	(0.011)
Likes	-0.162***	-0.091***	-0.101***	-0.093***
	(0.007)	(0.007)	(0.011)	(0.007)
Constant	7.404***	6.317*	7.361***	6.331***
	(0.017)	(0.134)	(0.054)	(0.134)
R^2	0.075	0.099	0.075	0.099

$N = 11,839$; *$p < 0.050$, **$p < 0.010$, ***$p < 0.001$; standard errors are in parentheses.

the marginal effect of professional motivation on online knowledge-sharing quantity at different levels of online experience in Figure 2. As the value of online experience increases from low to high, the effect of professional motivation on knowledge-sharing quantity becomes stronger. Therefore, H3a is supported. Model 3 shows that interaction term (PM × OET) is negatively and significantly associated with online knowledge-sharing quantity ($\beta = -0.017$, $p < 0.050$). Figure 3 illustrates that the effect of professional motivation on online knowledge-sharing quantity becomes weaker with the value of offline expertise increasing from low to high. Therefore, H4a is supported.

Following the suggestions of Guo et al. (47), our study applied full models to further test our hypotheses testing in Model 4, and the results are consistent with Model 2 and Model 3.

Model 5 shows that professional motivation is positively and significantly related to online knowledge-sharing quality ($\beta = 0.151$, $p < 0.001$), thereby supporting H2. In Model 6, the interaction term (PM × OEN) is positively and significantly related to online knowledge-sharing quality ($\beta = 0.036$, p < 0.001). We also plotted the figure about the moderating effect of online experience. Figure 4 indicates that the positive relationship between professional
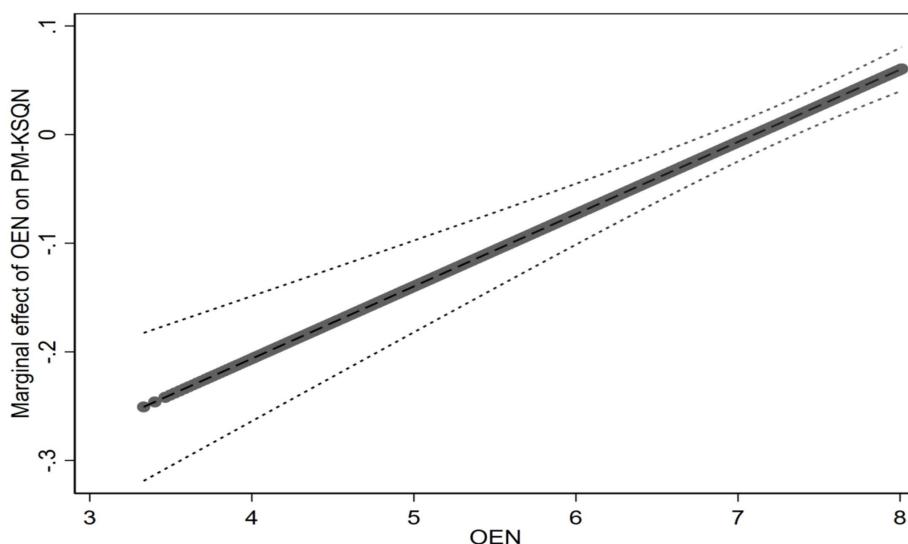

FIGURE 2
The moderating effect of online experience (OEN) on the relationship between professional motivation (PM) and online knowledge–sharing quantity (KSQN).
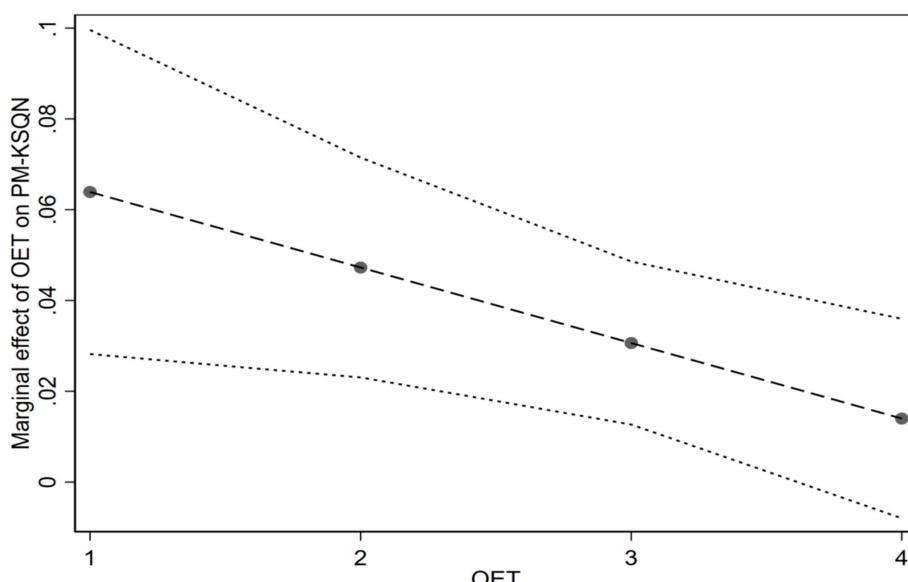

FIGURE 3
The moderating effect of offline expertise (OET) on the relationship between professional motivation (PM) and online knowledge-sharing quantity (KSQN).
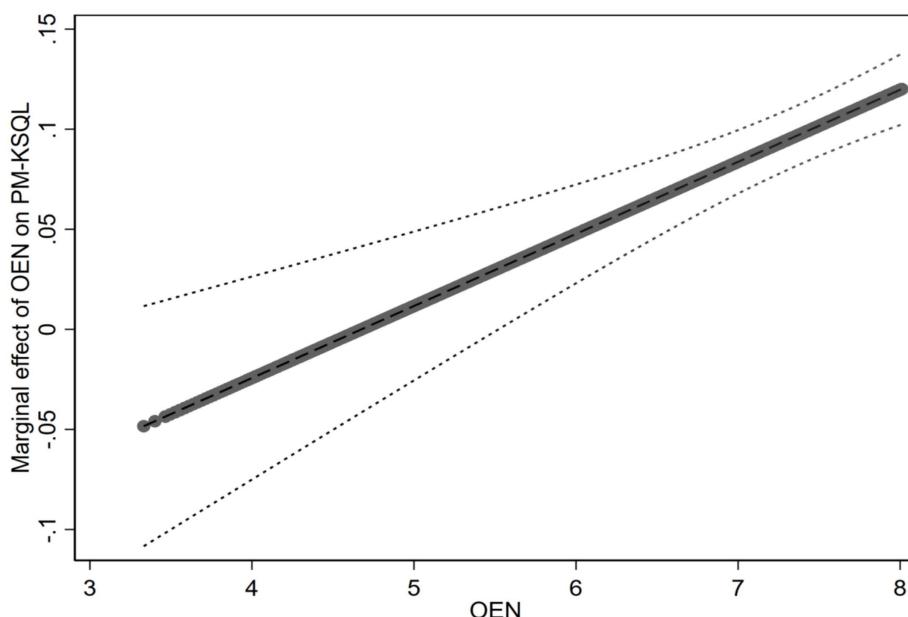
FIGURE 4
The moderating effect of online experience (OEN) on the relationship between professional motivation (PM) and online knowledge-sharing quality (KSQL).

motivation and online knowledge-sharing quality becomes stronger with an increase in online experience. Thus, H3b is supported. Model 7 shows that the coefficient of interaction term (PM × OET) on online knowledge-sharing quality is not significant ($\beta = 0.007$, $p > 0.050$). Thus, H4b is not supported. Full models were applied to test our hypotheses in Model 8, and the results are consistent with Model 6 and Model 7.

4.2 Supplementary analyses

To test the robustness of our regression results, according to whether online experience (OEN) and offline expertise (OET) are above or below the mean value, we divided the sample into subsamples (OEN_{high} vs. OEN_{low}, OET_{high} vs. OET_{low}) to conduct regression between the independent variable and dependent variable following the suggestion of previous studies (48, 49). For moderating effect of online experience, the effects of professional motivation on online knowledge-sharing quantity ($\beta_{high} = 0.228$, $p < 0.001$ vs. $\beta_{low} = 0.066$, $p < 0.001$) and online knowledge-sharing quality ($\beta_{high} = 0.158$, $p < 0.001$ vs. $\beta_{low} = 0.116$, $p < 0.001$) are stronger in high levels of online experience. For the moderating effect of offline expertise, the slope is flatter under higher levels of offline expertise in the relationship between professional motivation on online knowledge-sharing quantity ($\beta_{high} = -0.041$, $p < 0.010$ vs. $\beta_{low} = 0.107$, $p < 0.001$). All these results are consistent with the regression analysis, which further supports the interaction hypotheses.

To further test the robustness of our results, random effects regression models were conducted following previous studies (50, 51). The results are presented in Table 3. The effect of professional motivation on online knowledge-sharing quantity is positive and significant ($\beta = 0.082$, $p < 0.001$) in Model 9, which supports H1.

Model 13 indicates that professional motivation is positively and significantly related to online knowledge-sharing quality ($\beta = 0.151$, $p < 0.001$), thereby supporting H2. Model 10 and Model 14 verified the moderating roles of online experience about the effect professional motivation on online knowledge-sharing quantity ($\beta = 0.068$, $p < 0.001$) and online knowledge-sharing quality ($\beta = 0.036$, $p < 0.001$). Thus, H3a and H3b are supported. The interaction term (PM × OET) in Model 11 is significantly related to online knowledge-sharing quantity ($\beta = -0.016$, $p < 0.050$), while the interaction term (PM × OET) in Model 15 is insignificantly related to online knowledge-sharing quality ($\beta = 0.007$, $p > 0.050$). Therefore, H4a is supported but H4b is rejected. The full models in Model 12 and Model 16 align with the results from the respective separate models. In summary, the results are similar to fixed effects, and our results are robust.

5 Discussion

Drawing on motivation theory, this study examines the relationship between physicians' professional motivation for patient education and their online knowledge-sharing behaviors on OHPs. Using a six-moth panel dataset of 11,839 physicians, the findings confirm that professional motivation significantly enhances both the quantity and quality of online knowledge sharing, with online experience further strengthening these positive effects. Interestingly, while offline expertise weakens the positive effect of professional motivation on online knowledge-sharing quantity, this study finds no significant empirical support for its moderating role in online knowledge-sharing quality. One possible explanation lies in the hierarchical structure of Chinese hospitals, where physicians with extensive offline expertise often

TABLE 3 Results of regression models with random effects.

Online knowledge-sharing quantity (KSQN)	Model 9	Model 10	Model 11	Model 12
Professional motivation (PM)	0.082***	-0.486***	0.075***	-0.510***
	(0.008)	(0.062)	(0.025)	(0.062)
Online experience (OEN)		0.151***		0.084***
		(0.021)		(0.023)
PM×OEN		0.068***		0.084***
		(0.008)		(0.009)
Offline expertise (OET)			0.198***	0.178***
			(0.020)	(0.022)
PM×OET			-0.016*	-0.038***
			(0.007)	(0.008)
Gifts	0.394***	0.285***	0.377***	0.288***
	(0.012)	(0.013)	(0.012)	(0.013)
Likes	-0.136***	-0.044***	-0.110***	-0.041***
	(0.008)	(0.008)	(0.008)	(0.009)
Constant	1.320***	0.326*	0.806***	0.309*
	(0.019)	(0.153)	(0.061)	(0.153)
R^2	0.097	0.127	0.107	0.131
Online knowledge-sharing quality (KSQL)	Model 13	Model 14	Model 15	Model 16
Professional motivation (PM)	0.151***	-0.170**	0.115***	-0.165**
	(0.007)	(0.055)	(0.022)	(0.055)
Online experience (OEN)		0.161***		0.172***
		(0.019)		(0.020)
PM×OEN		0.036***		0.036***
		(0.007)		(0.008)
Offline expertise (OET)			0.021	-0.034
			(0.018)	(0.019)
PM×OET			0.007	0.002
			(0.007)	(0.007)
Gifts	0.286***	0.200***	0.280***	0.200***
	(0.011)	(0.011)	(0.011)	(0.011)
Likes	-0.164***	-0.092***	-0.157***	-0.094***
	(0.007)	(0.007)	(0.007)	(0.008)
Constant	7.406***	6.314*	7.363***	6.329***
	(0.017)	(0.134)	(0.054)	(0.134)
R^2	0.075	0.099	0.076	0.100

$N = 11,839$; *$p < 0.050$, **$p < 0.010$, ***$p < 0.001$; standard errors are in parentheses.

take on additional responsibilities, such as administrative duties and teaching tasks (19, 52), which divert their time and attention from refining shared knowledge. Consequently, even with high professional motivation, their online knowledge-sharing quality may not be shaped by offline expertise. In addition, using the number of article readings as a proxy for knowledge-sharing quality could introduce bias, as highly specialized content contributed by senior physicians, although often of high quality, may naturally attract fewer patients compared to common health topics (53). This limitation is likely to obscure the potential moderating effect of offline expertise. In the following sections, we will discuss the theoretical and practical implications, limitations, and future directions for research.

5.1 Theoretical implications

This study makes several theoretical contributions. First, it contributes to the literature on motivation theory by focusing on professional motivation in patient education and identifying the formation mechanism of online knowledge sharing. Previous studies have shown that physicians driven by professional motivation are more willing to engage in knowledge sharing on OHPs (7, 15). Our study confirms the significant intrinsic impact of professional motivation on physicians' online knowledge-sharing behaviors, supporting the existing literature on its role in patient education (13, 14). Moreover, our study explores how external factors, including both online experience and offline expertise, shape physicians'

internal motivational processes. Therefore, this study develops a comprehensive framework from the lens of motivation theory to better understand the formation mechanism of online knowledge sharing for patient education.

Second, this study enriches the literature of online knowledge sharing by revealing the positive effects of professional motivation on online knowledge-sharing quantity and quality. Although the extant research has widely explored their engagement with OHPs, it has primarily focused on the types of content shared (e.g., free vs. paid) or behavioral intentions (15, 27, 54), with relatively little attention given to the different dimensions of online knowledge sharing. In the context of OHPs, where patients rely on contributed information for making health-related decisions (55), understanding the scope and value of the knowledge being shared is particularly important (56). In this vein, our study quantifies online knowledge sharing in terms of quantity and quality, as well as examining how professional motivation influences each dimension. The results show that professional motivation significantly enhances both dimensions, although the strength of its effect varies across individuals (7). By incorporating the dual dimensions of quantity and quality into online knowledge sharing, this study provides a nuanced and outcome-oriented perspective on physician engagement in online knowledge sharing.

Third, this study underscores the importance of online and offline contexts in online knowledge sharing by uncovering the moderating effects of online experience and offline expertise. While OHPs serve as a complementary channel to offline healthcare, physicians often act as knowledge providers across both contexts (57). In this vein, their professional experiences in one domain may influence their engagement in the other, generating cross-contextual spillover effects (56). However, recognizing the limitations of isolated analyses of online and offline factors in prior research, we complement the literature by introducing online experience and offline expertise into a unified framework as co-existing contextual contingencies. Our findings show that online experience strengthens the effects of professional motivation on both the quantity and quality of online knowledge sharing, whereas offline expertise dampens its impact on quantity. In doing so, this study expands the boundary conditions of motivation theory by shedding light on the dynamic interplay between physicians' digital participation and their professional expertise.

5.2 Practical implications

This study also provides valuable implications for medical practitioners and platform managers. Given that professional motivation is positively related to both the quantity and quality of online knowledge sharing, physicians need to cultivate a strong professional identity. Specifically, they should recognize that online patient education—explaining medical concepts, addressing common misconceptions, and providing evidence-based health advice—is an integral part of their professional responsibilities rather than an additional burden (17). This shift can inspire greater engagement in online knowledge-sharing efforts. Also, physicians can boost their motivation by engaging with patient feedback online through responding to comments and participating in Q&A forums. Such interaction can help tailor content to address real concerns and reinforces their commitment to online knowledge sharing.

In addition, platforms are supposed to shoulder responsibilities for supporting physicians' adaptation to digital engagement (58),

particularly given the heterogeneity among physician groups. On the one hand, platform designers should implement targeted strategies to improve physicians' proficiency in online interactions in light of the positive moderating role of online experience. For example, by analyzing multi-dimensional patient feedback, these designers can leverage AI-driven algorithms to identify effective knowledge-sharing practices. By doing so, physicians can gain deeper insights into patient preferences, enabling them to publish high-quality articles efficiently with minimal time investment. Moreover, platform-based initiatives, such as mentorship programs that connect experienced online contributors with novices, and training modules designed to enhance digital communication skills for patient education, can further strengthen physicians' proficiency in online interactions and support sustained knowledge-sharing behaviors. On the other hand, platforms need to recognize and counterforce against the negative role of offline expertise on knowledge sharing. Teamwork serves as a viable solution to this challenge by, for example, allowing physicians to establish verified virtual team-based accounts (40). Within this framework, senior physicians with high clinic titles provide authoritative mentorship, sharing professional insights and practical experience, while the other junior physicians contribute by assisting with content creation and digital engagement. This division of labor creates a synergistic relationship, ultimately maintaining professional standards and online knowledge-sharing continuity.

5.3 Limitations and research directions

There are several limitations and open questions that are worthwhile to further research. First, the exclusive reliance on data from Haodf.com in China raises questions regarding the generalizability of our findings. Given the professional nature of physicians, the main finding that professional motivation positively influences online knowledge-sharing behaviors is likely to be broadly applicable across various OHPs. However, effects such as the moderating effect of offline expertise may be shaped by differences in healthcare system structures and professional hierarchies internationally. These potential variations call for future research to build upon this study by incorporating multiple OHPs across diverse digital ecosystems. Second, while our study mitigated endogeneity by employing several supplementary analyses, potential endogeneity concerns cannot be entirely eliminated. To strengthen causal identification, future studies could adopt methods such as difference-in-differences (DID), quasi-experimental designs, and synthetic control methods (59). Third, we primarily studied the contingent role from the perspective of physicians' attributes, particularly online experience and offline expertise. However, exploring alternative moderators may provide deeper insights into physicians' knowledge-sharing dynamics, for example, gamification settings (60) and income (15). They are other avenues for future studies. Finally, using article readings as a proxy for knowledge-sharing quality may underestimate the value of highly specialized content that naturally attracts a narrower patient audience, leading to potential measurement bias. Similarly, measuring online experience by physicians' opening time may fail to capture the intensity and nature of online engagement, such as

the frequency, recency, and diversity of online activities. Future research could incorporate multi-dimensional metrics or conduct experimental research designs to provide more nuanced assessments of these constructs.

Data availability statement

The datasets presented in this article are not readily available because the data is available for request. Requests to access the datasets should be directed to fanqihui_1019@tju.edu.cn.

Author contributions

YH: Supervision, Writing – original draft, Writing – review & editing. JG: Methodology, Resources, Writing – original draft, Writing – review & editing. YW: Data curation, Software, Writing – original draft, Writing – review & editing. QF: Conceptualization, Investigation, Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

1. Guo F, Zhang X, Lee P. Digital information for patient education. *Front Public Health*. (2023) 11:1211285. doi: 10.3389/fpubh.2023.1211285

2. Deng Z, Deng Z, Liu S, Evans R. Knowledge transfer between physicians from different geographical regions in China's online health communities. *Inf Technol Manag*. (2025) 26:83–100. doi: 10.1007/s10799-023-00400-3

3. Khan R, Khan S, Almohaimeed HM, Almars AI, Pari B. Utilization, challenges, and training needs of digital health technologies: perspectives from healthcare professionals. *Int J Med Inform*. (2025) 197:105833. doi: 10.1016/j.ijmedinf.2025.105833

4. Liu H, Zhang Y, Li Y, Albright K. Better interaction performance attracts more chronic patients? Evidence from an online health platform. *Inf Process Manag*. (2023) 60:103413. doi: 10.1016/j.ipm.2023.103413

5. Chen L, Rai A, Chen W, Guo X. Signaling effects under dynamic capacity in online matching platforms: evidence from online health consultation communities. *Inf Syst Res*. (2025) 36:202–24. doi: 10.1287/isre.2021.0150

6. Chen Q, Jin J, Yan X. Understanding physicians' motivations for community participation and content contribution in online health communities. *Online Inf Rev*. (2023) 47:604–29. doi: 10.1108/OIR-11-2021-0615

7. Zhang X, Guo F, Xu T, Li Y. What motivates physicians to share free health information on online health platforms? *Inf Process Manag*. (2020) 57:102166. doi: 10.1016/j.ipm.2019.102166

8. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp Educ Psychol*. (2000) 25:54–67. doi: 10.1006/ceps.1999.1020

9. Wang Q, Qiu L, Xu W. Informal payments and doctor engagement in an online health community: an empirical investigation using generalized synthetic control. *Inform Syst Res*. (2023) 35:706–26. doi: 10.1287/isre.2020.0475

10. Zhang X, Liu S, Chen X, Gong Y. Social capital, motivations, and knowledge sharing intention in health Q & A communities. *Manag Decis*. (2017) 55:1536–57. doi: 10.1108/MD-10-2016-0739

11. Oh S. The characteristics and motivations of health answerers for sharing information, knowledge, and experiences in online environments. *J Am Soc Inf Sci Technol*. (2012) 63:543–57. doi: 10.1002/asi.21676

12. Jing D, Jin Y, Liu J. The impact of monetary incentives on physician prosocial behavior in online medical consulting platforms: evidence from China. *J Med Internet Res*. (2019) 21:e14685. doi: 10.2196/14685

13. Imlawi J, Gregg D. Understanding the satisfaction and continuance intention of knowledge contribution by health professionals in online health communities. *Inform Health Soc Care*. (2019) 45:151–167. doi: 10.1080/17538157.2019.1625053

14. Li J, Li X, Zhang C. Understanding proactive knowledge sharing and knowledge withholding in physician-driven online health communities-a professional role identity perspective. *Inf Technol People*. (2024) 37:1983–2005. doi: 10.1108/ITP-06-2022-0443

15. Yang Y, Zhu X, Song R, Zhang X, Guo F. Not just for the money? An examination of the motives behind physicians' sharing of paid health information. *J Inf Sci*. (2023) 49:145–63. doi: 10.1177/0165551521991029

16. Chang HH, Chuang S-S. Social capital and individual motivations on knowledge sharing: participant involvement as a moderator. *Inf Manag*. (2011) 48:9–18. doi: 10.1016/j.im.2010.11.001

17. Guo F, Zhou A, Chang W, Sun X, Zou B. Is physician online information sharing always beneficial to patient education? An attention perspective. *Front Public Health*. (2022) 10:987766. doi: 10.3389/fpubh.2022.987766

18. Liu H, Perera SC, Wang J-J, Leonhardt JM. Physician engagement in online medical teams: a multilevel investigation. *J Bus Res*. (2023) 157:113588. doi: 10.1016/j.jbusres.2022.113588

19. Wang L, Yan L, Zhou T, Guo X, Heim GR. Understanding physicians' online-offline behavior dynamics: an empirical study. *Inf Syst Res*. (2020) 31:537–55. doi: 10.1287/isre.2019.0901

20. Sessions H, Pychlau S. Self-inconsistency or self-expansion from wearing multiple hats? The daily effects of enacting multiple professional identities on work meaningfulness. *J Appl Psychol*. (2024) 109:897–920. doi: 10.1037/apl0001176

21. Zhao CY, Zhu Q. Effects of extrinsic and intrinsic motivation on participation in crowdsourcing contest. *Online Info Rev*. (2014) 38:896–917. doi: 10.1108/OIR-08-2014-0188

22. Zhang X, Wilfried A, Saab N. Teachers' motivation to participate in continuous professional development: relationship with factors at the personal and school level. *J Educ Teach*. (2021) 47:714–31. doi: 10.1080/02607476.2021.1942804

23. Giurge LM, Woolley K. Working during non-standard work time undermines intrinsic motivation. *Organ Behav Hum Decis Process*. (2022) 170:104134. doi: 10.1016/j.obhdp.2022.104134

24. Safa NS, Von Solms R. An information security knowledge sharing model in organizations. *Comput Human Behav*. (2016) 57:442–51. doi: 10.1016/j.chb.2015.12.037

25. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. (2000) 55:68–78. doi: 10.1037//0003-066x.55.1.68

26. Zhang X, Guo X, Guo F, Lai K-H. Nonlinearities in personalization-privacy paradox in Mhealth adoption: the mediating role of perceived usefulness and attitude. *Technol Health Care*. (2014) 22:515–29. doi: 10.3233/thc-140811

27. Zhuo X, Wang W-T. Why are physicians willing to contribute knowledge? Evidence from online health communities. *Comput Hum Behav*. (2024) 152:108095. doi: 10.1016/j.chb.2023.108095

28. Suh A. The influence of self-discrepancy between the virtual and real selves in virtual communities. *Comput Hum Behav*. (2013) 29:246–56. doi: 10.1016/j.chb.2012.09.001

29. Suh A, Wagner C. How gamification of an enterprise collaboration system increases knowledge contribution: an affordance approach. *J Knowl Manag*. (2017) 21:416–31. doi: 10.1108/JKM-10-2016-0429

30. Zhang X, Guo X, Lai K-h, Guo F, Li C. Understanding gender differences in M-health adoption: a modified theory of reasoned action model. *Telemed J E-Health*. (2013) 20:39–46. doi: 10.1089/tmj.2013.0092

31. Atanasova S, Kamin T, Petrič G. Exploring the benefits and challenges of health professionals' participation in online health communities: emergence of (dis) empowerment processes and outcomes. *Int J Med Inform*. (2017) 98:13–21. doi: 10.1016/j.ijmedinf.2016.11.005

32. Tan H, Zhang X, Yang Y. Satisfaction or gratitude? Exploring the disparate effects of physicians' knowledge sharing on patients' service evaluation in online medical consultations. *Inf Syst J*. (2023) 33:1186–211. doi: 10.1111/isj.12440

33. Zhang X, Xiaocui H, Yuanyuan D, Fanbo M, Xitong G, Lin J. User acceptance of mobile health services from users' perspectives: the role of self-efficacy and response-efficacy in technology acceptance. *Inform Health Soc Care*. (2017) 42:194–206. doi: 10.1080/17538157.2016.1200053

34. Ye J, Jensen M. Effects of introducing an online community in a crowdsourcing contest platform. *Inf Syst J*. (2022) 32:1203–30. doi: 10.1111/isj.12397

35. Liu Q, Du Q, Hong Y, Fan W, Wu S. User idea implementation in open innovation communities: evidence from a new product development crowdsourcing community. *Inf Syst J*. (2020) 30:899–927. doi: 10.1111/isj.12286

36. Dodel M, Mesch G. Inequality in digital skills and the adoption of online safety behaviors. *Inf Commun Soc*. (2018) 21:712–28. doi: 10.1080/1369118X.2018.1428652

37. Guo S, G X, F Y, Vogel D. How doctors gain social and economic returns in online health-care communities: a professional capital perspective. *J Manag Inf Syst*. (2017) 34:487–519. doi: 10.1080/07421222.2017.1334480

38. Ramarajan L, Yen J. Defining who you are by whom you serve? Strategies for prosocial–professional identity integration with clients. *Adm Sci Q*. (2024) 69:515–67. doi: 10.1177/00018392241238407

39. Zhang X, Lai K-h, Guo X. Promoting China's Mhealth market: a policy perspective. *Health Policy Technol*. (2017) 6:383–8. doi: 10.1016/j.hlpt.2017.11.002

40. Zhang X, Pu J, Lu Y, Guo F. Team makes you better: evidence from online medical consultation platforms. *Inf Syst Res*. (2025). Available online at:). doi: 10.1287/isre.2022.0689

41. Zhang X, Guo X, Wu Y, Lai K-h, Vogel D. Exploring the inhibitors of online health service use intention: a status quo bias perspective. *Inf Manag*. (2017) 54:987–97. doi: 10.1016/j.im.2017.02.001

42. Liu Y, Ren C, Shi D, Li K, Zhang X. Evaluating the social value of online health information for third-party patients: is uncertainty always bad? *Inf Process Manag*. (2020) 57:102259. doi: 10.1016/j.ipm.2020.102259

43. Samila S, Sorenson O. Venture capital as a catalyst to commercialization. *Res Policy*. (2010) 39:1348–60. doi: 10.1016/j.respol.2010.08.006

44. Chen Z, Xie G. Esg disclosure and financial performance: moderating role of ESG investors. *Int Rev Financ Anal*. (2022) 83:102291. doi: 10.1016/j.irfa.2022.102291

45. Patel PC, Wolfe MT, Manikas AS. Logic is (somewhat) overrated: image-based versus concept-based rhetoric in crowdfunding narratives. *Entrep Theory Pract*. (2021) 45:600–31. doi: 10.1177/1042258720976310

46. Meyer KE, Van Witteloostuijn A, Beugelsdijk S. What's in a P? Reassessing best practices for conducting and reporting hypothesis-testing research. *J Int Bus Stud*. (2017) 48:535–51. doi: 10.1057/s41267-017-0078-8

47. Guo F, Zou B, Zhang X, Bo Q, Li K. Financial slack and firm performance of SMEs in China: moderating effects of government subsidies and market-supporting institutions. *Int J Prod Econ*. (2020) 223:107530. doi: 10.1016/j.ijpe.2019.107530

48. Song M, Droge C, Hanvanich S, Calantone R. Marketing and technology resource complementarity: an analysis of their interaction effect in two environmental contexts. *Strateg Manage J*. (2005) 26:259–76. doi: 10.1002/smj.450

49. Wang C, Qureshi I, Guo F, Zhang Q. Corporate social responsibility and disruptive innovation: the moderating effects of environmental turbulence. *J Bus Res*. (2022) 139:1435–50. doi: 10.1016/j.jbusres.2021.10.046

50. Chin T, Wang W, Yang M, Duan Y, Chen Y. The moderating effect of managerial discretion on blockchain technology and the firms' innovation quality: evidence from Chinese manufacturing firms. *Int J Prod Econ*. (2021) 240:108219. doi: 10.1016/j.ijpe.2021.108219

51. Luo P, Ma X, Zhang X, Liu J, He H. How to make money with credit information? Information processing on online accommodation-sharing platforms. *Tour Manag*. (2021) 87:104384. doi: 10.1016/j.tourman.2021.104384

52. Cheng X, Zhang S, Fan Q, Sun X, Zhang X. How to guarantee online service quality of physicians on the health platforms? *Ind Manag Data Syst*. (2025). doi: 10.1108/IMDS-10-2024-1010

53. Liu Y, Zhang X, Liu L, Lai K-h. Does voice matter? Investigating patient satisfaction on mobile health consultation. *Inf Process Manag*. (2023) 60:103362. doi: 10.1016/j.ipm.2023.103362

54. Ma X, Zhang P, Meng F, Lai K-h. How does physicians' educational knowledge-sharing influence patients' engagement? An empirical examination in online health communities. *Front Public Health*. (2022) 10:1036332. doi: 10.3389/fpubh.2022.1036332

55. Zhang F, Li X. How to find helpful health-related knowledge in the online healthcare community. *Inf Manag*. (2024) 61:104029. doi: 10.1016/j.im.2024.104029

56. Xia S, Zhang Z, Fu S, Chen X. Measuring knowledge contribution performance of physicians in online health communities: a Bp neural network approach. *J Inf Sci*. (2024) 50:1382–99. doi: 10.1177/01655515221121946

57. Chen S, Guo X, Wu T, Ju X. Exploring the online doctor-patient interaction on patient satisfaction based on text mining and empirical analysis. *Inf Process Manag*. (2020) 57:102253. doi: 10.1016/j.ipm.2020.102253

58. Wang L, Wu T, Guo X, Zhang X, Li Y, Wang W. Exploring mhealth monitoring service acceptance from a service characteristics perspective. *Electron Commerce Res Appl*. (2018) 30:159–68. doi: 10.1016/j.elerap.2018.06.003

59. Hill A, Johnson S, Greco L, O'Boyle E, Walter S. Endogeneity: a review and agenda for the methodology-practice divide affecting micro and macro research. *J Manage*. (2020) 47:014920632096053. doi: 10.1177/0149206320960533

60. Liu J, Zhang X, Meng F, Lai K-h. Deploying gamification to engage physicians in an online health community: an operational paradox. *Int J Prod Econ*. (2020) 228:107847. doi: 10.1016/j.ijpe.2020.107847

Tools/instruments for assessing YouTube videos on surgical procedures for patient/consumer health education: a systematic review

Manasa Pavuloori[1]*, Amy Lin[1] and Misa Mi[2]

[1]Oakland University William Beaumont School of Medicine, Rochester, MI, United States, [2]Department of Foundational Medical Studies, Oakland University William Beaumont School of Medicine, Rochester, MI, United States

Background: YouTube is becoming an increasingly popular platform for health education; however, its reliability for surgical patient education remains largely unexplored. Given the global prevalence of preoperative anxiety, it becomes essential to ensure accurate information online.

Objectives: The objective is to assess tools/instruments used to evaluate YouTube videos on surgical procedures created to educate patients or health consumers.

Methods: In June 2023, a comprehensive literature search was conducted on PubMed, PsycINFO, CINAHL, and Scopus. Primary studies with empirical data that evaluate English YouTube videos to educate patients about surgical procedures in all specialties were included. Two reviewers independently completed title/abstract and full text screening, and data extraction in duplicate. The data extracted includes the number of videos evaluated, assessment tools, outcomes of significance, specific objectives, and features examined.

Results: A total of 41 studies were included in the review. The most commonly used evaluation tools were DISCERN (21 studies), the Global Quality Scale (11 studies), and the JAMA benchmark criteria (11 studies). Notably, 23 studies used a unique assessment instrument, and several studies employed more than one tool concurrently. Of the total studies included, 88% of the articles determined that patients were not adequately educated by YouTube videos per the ratings of the assessment tools, and 19 out of 41 articles mentioned that videos from professional sources were most useful.

Conclusions: This systematic review suggests that the educational qualities in YouTube videos are substandard. Patients should be cautious when relying solely on YouTube videos for medical guidance. Surgeons and medical institutions are encouraged to direct patients to high-quality patient education sources and create accessible medical content. As there is variability in the quality assessment tools used for evaluation, a standardized approach to creating and assessing online medical videos would improve patient education.

Introduction

In 2024, the number of active users on YouTube exceeded 2.56 billion (1). More than 500 h of content are uploaded to YouTube every minute (2). Approximately 25% of adults in the United States stated that they rely on YouTube as a regular source for obtaining news (3).

In recent years, YouTube has emerged as a popular platform for patient and health consumer education. In 2020, 40.8% of U.S. adults used YouTube to watch health–related videos (4). Recent literature published in 2022 has shown that YouTube is not a reliable source for medical and health-related information and there has only been one systematic review investigating the reliability of YouTube as a source of knowledge for surgical patients (5, 6). With the increasing availability of surgical videos on YouTube, it is crucial to assess the tools or instruments used to evaluate the quality and educational value of such content.

Surgical education plays a vital role in empowering patients to make informed decisions about their healthcare and enhance their understanding of complex medical interventions. YouTube offers an easily accessible and visually engaging platform to deliver such educational content. As preoperative anxiety remains a critical issue, occurring in ∼48% of surgical patients globally (7), it is vital to ensure that accessible information online on surgical procedures is accurate and regulated to prevent unnecessary confusion.

The purpose of this systematic review is to assess the tools or instruments employed for evaluating YouTube videos focused on surgical procedures with the intent of educating patients or health consumers. The findings of this review will have implications for healthcare providers, educators, and content creators involved in patient education. Understanding the strengths and weaknesses of existing evaluation tools will facilitate the development of standardized guidelines and best practices for assessing the quality and educational impact of YouTube videos on surgical procedures. Ultimately, this systematic review aims to contribute to the improvement of patient education materials available on YouTube, ensuring that patients and health consumers have access to reliable, accurate, and informative content that enhances their surgical knowledge and decision-making abilities.

Methods

Literature search

A comprehensive literature search (MM) was conducted using PubMed, PsycINFO, CINAHL and Scopus from each database's inception to June 6, 2023. A combination of index terms and keywords were used to represent key concepts of "patient education," "YouTube video," "psychometrics," "quality assessment," and "surgical procedure" (See Appendix A for a sample search strategy for PubMed). A hand search of the reference lists of all identified studies were examined for additional studies.

Abbreviations: GQS, Global Quality Scale; HON, Health on the Net; JAMA, Journal of the American Medical Association; PEMAT, Patient Education Materials Assessment Tool and the Usefulness Scoring System.

Eligibility criteria

Articles were selected based on specific inclusion criteria. The review included original, full-text primary studies published in the English language that provided empirical data evaluating English YouTube videos created for patient and health consumer education. Videos encompassed information regarding surgical procedures in all surgical specialties. Reviews, duplicate articles, comments, editorials, letters, and abstracts lacking full content articles were excluded. Studies analyzing videos in other languages, from differing social media sites, and targeting health professional education were also excluded.

Data selection

All search results were imported into Covidence for screening and data extraction. Covidence is a web-based software for managing and streamlining systematic reviews. Two reviewers (MP and AL) first screened titles and abstracts against the selection criteria, followed by full text screening done in duplicate and independently. Any discrepancies in screening by the two reviewers were discussed and resolved to reach consensus. The third author (MM) assessed any variances and determined their inclusion.

Data extraction

A standardized data collection form was created on the Covidence platform, and the authors (MP and AL) completed data extraction in duplicate and independently. All discrepancies were discussed and resolved with the third author (MM). The parameters consisted of study aims, surgeries evaluated, type of quality tools used, number of videos analyzed, primary source of videos, types of video characteristics studied, educational quality based on author's judgment, study limitations, future recommendations, and video sources deemed the most useful. The video sources were divided into four categories: commercials, patients, professional entities (e.g., created by physicians, hospitals), educational institutions (association, organization, society, and others). A rating scale was developed to assess the educational quality of the videos: "poor," "moderate" or "good" based on the articles' direct analyses, and the reasoning was noted.

Results

Using an initial dataset of 125 studies, articles were excluded based on specific inclusion criteria (Figure 1). Ultimately, 41 studies remained in the review for data extraction and analysis.

Video information data

Using an initial dataset of 125 studies, articles were excluded based on specific inclusion criteria (Figure 1). Ultimately, 41 studies remained in the review for data extraction and analysis.

FIGURE 1
PRISMA flow diagram of article selection process.
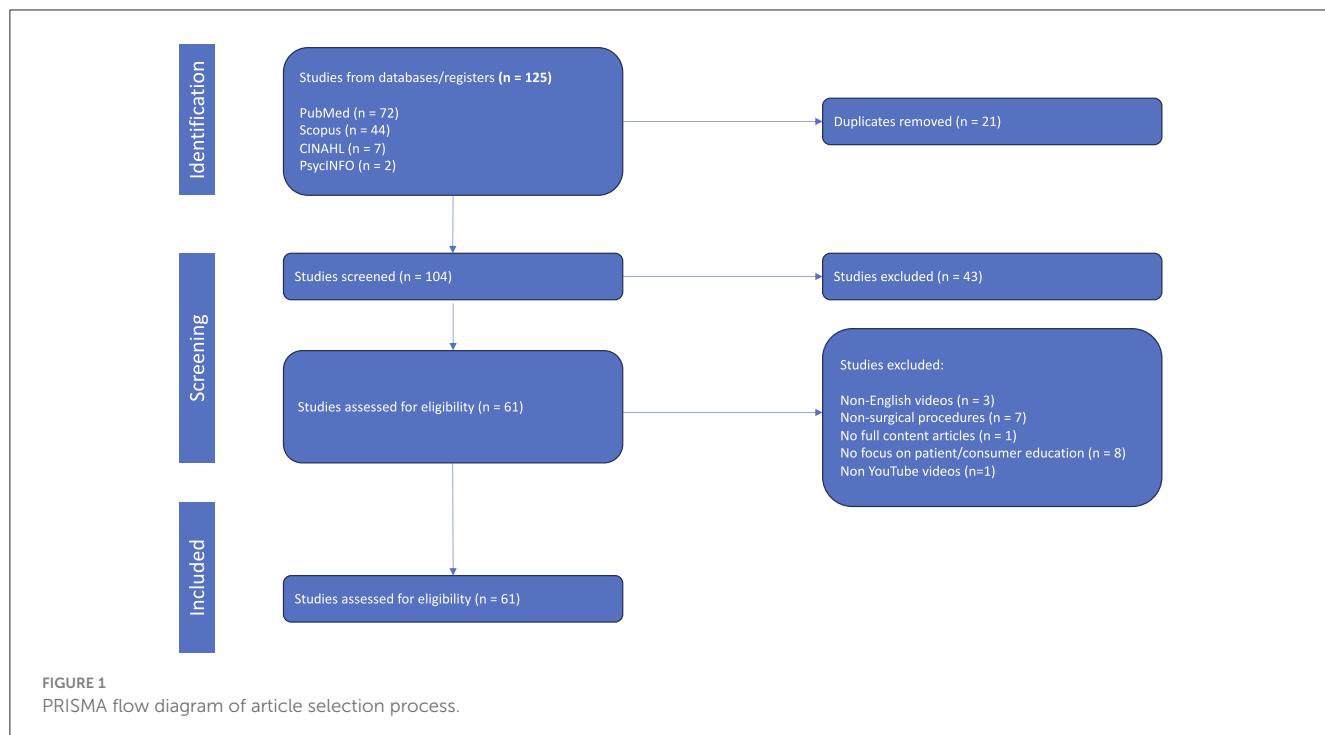
These articles assessed an average of 98.8 videos per study, ranging from 16 to 523 videos evaluated per study. The parameters that quality assessment tools analyzed included YouTube video views, video duration, likes and dislikes, time on YouTube, comments, and other features (Figure 2). Studies were published from 2013 to 2023, with 2021 as the median publication year with a notable increase after 2020. A trend was noted for an increasing amount of research articles evaluating the quality of YouTube videos, reflecting growing reliance on social media and digital platforms for patient education. Videos encompassed information regarding surgical procedures in all surgical specialties, such as general surgery, oral and maxillofacial surgery, cardiac surgery, orthopedic surgery, dental and endodontic surgery, obstetric surgery, gynecology, urology, ophthalmology, neurosurgery, plastic and reconstructive surgery, neonatal surgery, and colorectal surgery.

Video source characteristics

All of the YouTube videos included professional sources and creators, such as physicians, hospitals, educational institutions, and societies. The majority (68%) featured patients and their testimonials. Commercial content (59%) was present in over half of the videos, while other diverse sources were also frequently utilized (73%). Of the 23 studies reporting sources that provided the most useful data for patient and health consumer education, 20 recognized videos from professional sources, two noted patient sources, and one described commercials as the most helpful source. These findings suggest a strong association between source professionalism and perceived educational value.

Use of quality assessment tools

The DISCERN reliability instrument was the predominant tool for video evaluation (8–16, 18–29). Other prevalent tools were the Journal of the American Medical Association (JAMA) Benchmark Criteria (8–18), Global Quality Scale (GQS) Criteria (8–10, 14, 17–19, 26, 27, 30, 31), Health on the Net (HON) Code of Conduct (12, 13), Patient Education Materials Assessment Tool (PEMAT) and the Usefulness Scoring System (20, 29). A notable portion of the articles used standardized reliability instruments developed by the authors, from previous studies, physicians and medical organizations (10, 15–17, 27, 30–47). The YouTube Video Assessment Criteria and Ensuring Quality Information for Patients Score were utilized less frequently (Table 1). While standardized tools provided a consistent framework, variation in implementation and scoring limited comparability. Notably, all studies that noted adequate patient education or moderate quality based on article author analysis used multiple assessment tools (e.g., DISCERN + JAMA + GQS).

Educational value

Educational quality, as rated by the respective quality assessment instruments, was predominantly low. Based on the articles' analyses of the educational quality of the YouTube videos, 33 studies were rated as poor quality, seven as moderate quality, and one as good quality. Five of the seven studies that determined moderate quality utilized GQS, four used DISCERN, three used JAMA, and four used other types of assessment tools, suggesting some consistency among these tools in identifying informative videos (8–10, 26, 30, 46, 48). The article that reported the videos
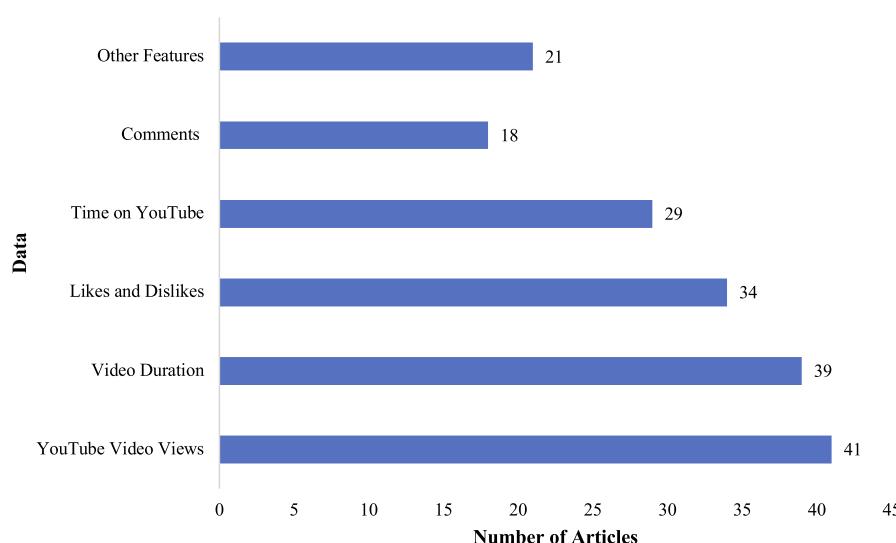
FIGURE 2
Data assessed by quality assessment tools with "Other Features," including title of video, universal resource locator, number of channel subscribers, video power index, country of origin, percentage positivity (proportion of likes to total likes plus dislikes), presence of subtitles, viewer interaction index, video title, target audience, presence of animation, video/audio quality, and daily viewing rate.

TABLE 1 Quality assessment tools ranked by frequency of their use across all articles (*n* = 41).

Type of assessment tool	Number of videos
DISCERN	21
JAMA	11
GQS	11
HON	2
PEMAT	2
Likert 5-point	1
Other	24

were of good quality used JAMA, GQS, and its own quality assessment tool as well (17).

Three studies concluded that surgical YouTube videos adequately educated patients, and 36 yielded contrasting data. Many articles (44%) mentioned that YouTube videos were missing vital information regarding surgical procedures, such as treatment alternatives and potential risks. Others (27%) noted that the scores associated with the reliability instruments were low and 20% of the articles saw an increasing prevalence of source bias with the videos. Several studies (15%) detailed that there were notable issues with videography, like difficulties with music, overall quality, and narration. The three other studies noted high scores among the reliability instruments (9), reputable sources creating and distributing content (10), and easily understandable videos with adequate procedure descriptions, which are not systematically captured by existing assessment instruments (8). Interestingly, among the five videos related to ophthalmology, two were reported to have adequately educated patients.

Common deficiencies in video content

Several recurring deficiencies in YouTube surgical video content were identified across the included studies. Most notably, eighteen studies (44%) reported that videos often omitted critical information such as treatment alternatives, potential risks, or post-operative expectations, limiting their utility for comprehensive patient education. Additionally, eleven studies (27%) documented consistently low scores across validated reliability instruments, including DISCERN and JAMA, reflecting concerns about content accuracy and trustworthiness. Eight studies (20%) highlighted the presence of source bias or overt promotional messaging, which may compromise the objectivity of the information presented. Furthermore, six studies (15%) described production-related limitations, such as poor narration quality, distracting background music, or inadequate video resolution, which could detract from viewer comprehension.

Discussion

Patient and health consumer education is rapidly evolving, with digital platforms and social media resources becoming increasingly prominent. This shift requires a reevaluation of how surgical patients receive information about their conditions and procedures. Validated evaluation tools are essential for assessing the quality and accuracy of educational videos to identify videos that provide clear information and align with current medical standards.

Video sources and content

Since all 41 articles evaluated YouTube videos that included professional sources, the content is likely to be credible and created

with expertise. However, despite this professional endorsement, the majority of the videos were deemed to provide poor educational quality. The prevalence of patient testimonials and commercial content contributes to further complications as these types of videos often prioritize personal experiences and promotional content over comprehensive educational information.

Quality assessment tools and quality analysis

To assess video quality, the authors predominantly utilized the DISCERN reliability instrument, along with other established tools such as the JAMA Benchmark Criteria and the GQS Criteria.

The DISCERN tool is a validated questionnaire designed to assess the quality of written consumer health information on treatment choices, consisting of 15 items rated on a 5-point scale and culminating in an overall quality score. Though comprehensive, DISCERN can be time-consuming and requires training for consistent use (49). In the context of video content, DISCERN has been adapted by some researchers to evaluate scripted narration or on-screen information, but its written-format origins may limit applicability to visual, interactive, or audiovisual cues that influence viewer perception and comprehension.

The JAMA Benchmark Criteria provide a more objective evaluation of online health information, assessing four elements: authorship, attribution, disclosure, and currency. While useful for gauging source credibility, the binary scoring system does not assess content accuracy, completeness, or audiovisual clarity—factors highly relevant in video-based media. As such, the JAMA criteria are often used as a supplemental tool rather than a standalone measure when evaluating videos (50).

The GQS criteria uses a 5-point Likert scale to assess the overall quality, flow, and usefulness of online content, especially videos, ranging from poor (1) to excellent (5). Though fast and intuitive, GQS is subjective and lacks detailed evaluative criteria, limiting its diagnostic utility (51). The HONcode certification, developed by the Health On the Net Foundation, was another credibility-focused tool that evaluated websites based on eight ethical principles, including authority, complementarity, privacy, attribution, justifiability, transparency, financial disclosure, and advertising policy. While HONcode was useful for identifying trustworthy health websites, it did not assess content depth, accuracy, or readability. As of December 15, 2022, the HONcode certification service has been discontinued, limiting its utility for future website evaluations (52, 53).

Eighty percent of articles rated the video educational quality as poor, suggesting a need to improve video quality. While these are established standardized tools, they are not designed to assess medical videos. These gaps suggest that the current tools are not entirely sufficient for ensuring high-quality educational content in dynamic online environments like YouTube (7). Therefore, a standardized tool should be created to assess video quality to ensure consistency across video evaluations.

In comparison to DISCERN, JAMA, and GQS, specialized frameworks such as the Instructional Videos in Otorhinolaryngology by YO-IFOS(IVORY) and LAParoscopic surgery Video Educational GuidelineS (LAP-VEGaS) guidelines have been developed to evaluate surgical videos intended for healthcare professional training (54, 55). These tools are more rigorous and procedure-specific in that they incorporate detailed criteria related to surgical technique, anatomical accuracy, intraoperative decision-making, step-by-step procedural clarity, video speed, camera angles, presentation clarity, and audio-visual delivery. Although these tools are designed for surgical training, they could be adapted to enhance the evaluation of surgical videos intended for patient education.

While the majority of the wide variety of assessment tools utilized by the articles indicated poor overall video quality, they also highlighted other problematic issues, including the omission of vital information, such as treatment alternatives, potential risks, low reliability scores, and increasing source bias. Several studies also pointed out technical issues, such as poor videography, suboptimal audio quality, and ineffective narration, which further detract from the educational value of the videos. Moreover, some studies identified misinformation and outdated content as critical problems, emphasizing the need for continuous updating and verification of online medical content. These deficiencies highlight another gap in the current use of YouTube as an educational tool for patients, and suggest that many videos fail to provide comprehensive information, which is essential for informed patient decision-making.

Contrasting findings

Interestingly, three studies (8–10) rated the YouTube videos as adequate educational tools, citing high scores on quality assessment instruments, reputable sources, and clear, understandable content. This discrepancy indicates that while the general trend points toward inadequate educational quality, there are exceptions where videos meet high standards. These positive examples can serve as benchmarks for creating better educational content in the future.

Limitations

This review is subject to several limitations. First, the included studies were assessed from the perspective of patients and healthcare consumers. While this approach is relevant to understanding public accessibility and perceived educational value, it may not fully reflect clinical accuracy or high educational quality. Additionally, it is important to acknowledge that YouTube functions primarily as an entertainment and social media platform rather than a formal educational resource. Consequently, many of the videos uploaded may not be intended for, or suitable as, educational content, limiting its suitability for patient and health consumer education. Another limitation is the exclusion of studies published in languages rather than English, which could introduce a potential selection bias, leading to incomplete or inaccurate conclusions, as studies published in other languages may contain crucial information that is not available in English-language sources.

Future recommendations

This review highlights the need for improved standards in the creation of surgical educational videos on YouTube. A new standardized tool should be developed that incorporates the strengths of widely used current tools, while addressing the unique challenges of assessing medical videos. Key criteria should account for dynamic audiovisual elements (e.g., clarity of narration, visual accuracy of demonstrations, use of animations or overlays), content accuracy, source credibility, and viewer engagement strategies. It should also consider accessibility features such as closed captions, language simplicity, and cultural sensitivity.

Advancements in artificial intelligence (AI), particularly natural language processing and deep learning, present promising opportunities for moderating health-related video content. For example, real-time misinformation detection using machine learning has proven effective during the COVID-19 pandemic (56). To implement these innovations, we propose a multidisciplinary task force, composed of clinicians, AI researchers, digital media experts, public health officials, and patient advocates, to develop validated scoring systems and collaborate directly with platforms, such as YouTube. Integration strategies may include voluntary quality tagging by verified content creators, peer-review-based content badges, and platform-endorsed health information panels. These features may help elevate trustworthy content while guiding users toward evidence-based information in an increasingly decentralized and saturated media landscape.

By directing patients to high-quality educational resources, surgeons can significantly enhance patient understanding and preparedness for surgical procedures. Given the overwhelming amount of online medical information, surgeons must guide patients toward reputable websites, vetted educational videos, and institutionally approved resources. They should also be aware of the quality assessment tools available to evaluate the quality of video content, ensuring that the materials they endorse are of the highest standard.

Conclusions

Though YouTube has indubitably transformed patient and health consumer education, the reliability and educational quality of its patient education videos remain a concern, particularly with surgical procedures. This systematic review finds that, despite their perceived credibility, quality assessment tools have determined that many videos from professional sources offer limited educational value. With improved patient education materials, the medical community can improve health consumer education, ultimately enhancing patient understanding, reducing anxiety, and potentially improving clinical outcomes in surgical settings.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2025.1575801/full#supplementary-material

References

1. L Ceci, Youtube - Statistics and Facts. Statista. February 16, 2024. Available online at: https://www.statista.com/topics/2019/youtube/#editorsPicks (accessed September 19, 2024).

2. Oxford economics. The State of the Creator Economy. Assessing the Economic, Cultural, and Societal Impact of YouTube in the US in 2022. Available online at: https://kstatic.googleusercontent.com/files/721e9ce19a4fe59b40135ad7b06a61d5 89468bf7df4b2d654c1be32aca225939b6dac1e16d1de1c0d0fe5bb496e82f5920afa7e ceb7dc867e4f71751bc233261 (accessed June 10, 2023).

3. Pew Research Center. Social Media and News Fact Sheet. Pew Research Center; September 20, (2022). Available online at: https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/ (accessed June 10, 2023).

4. Lee J, Turner K, Xie Z, Kadhim B, Hong YR. Association between health information-seeking behavior on YouTube and physical activity among U.S. adults. *AJPM Focus.* (2022) 2:100035. doi: 10.1016/j.focus.2022.100035

5. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med. Educ.* (2022) doi: 10.21203/rs.3.rs-1047141/v1

6. Javidan A, Nelms MW, Li A, Lee Y, Zhou F, Kayssi A, et al. Evaluating YouTube as a source of education for patients undergoing surgery: a systematic review. *Ann Surgery.* (2023). 278:e712-e718. doi: 10.1097/SLA.0000000000005892

7. Guler MA, Aydin EO. Development and validation of a tool for evaluating YouTube-based medical videos. *Ir J Med Sci.* (2022) 191:1985–1990. doi: 10.1007/s11845-021-02864-0

8. Heisinger S, Huber D, Matzner MP. TLIF online videos for patient education-evaluation of comprehensiveness, quality, and reliability. *Int J Environ Res Public Health.* (2023) 20:4626. doi: 10.3390/ijerph20054626

9. Karataş ME, Karataş G. Evaluating the reliability and quality of the upper eyelid blepharoplasty videos on YouTube. *Aesthetic Plast Surg.* (2022) 46:754–759. doi: 10.1007/s00266-021-02504-z

10. Kalayci M, Cetinkaya E, Suren E, Yigit K, Erol MK. Are YouTube videos useful in informing patients about keratoplasty? *Semin Ophthalmol.* (2021) 36:469–474. doi: 10.1080/08820538.2021.1890145

11. Yasin S, Altunisik E. Quality analysis of YouTube videos on mechanical thrombectomy for acute ischemic stroke. *J Stroke Cerebrovasc Dis.* (2023) 32:106914. doi: 10.1016/j.jstrokecerebrovasdis.2022.106914

12. Noll D, Green L, Asokan G. Is YouTube™ a good source of information for patients to understand laparoscopic fundoplication? *Ann R Coll Surg Engl.* (2023) 105:365–371. doi: 10.1308/rcsann.2022.0039

13. Gupta AK, Kovoor JG, Ovenden CD, Cullen HC. Paradigm shift: Beyond the COVID-19 era, is YouTube the future of education for CABG patients? *J Card Surg.* (2022) 37:2292–2296. doi: 10.1111/jocs.16617

14. Sayin O, Altinkaynak H, Adam M, Dirican E, Agca O. Reliability of YouTube videos in vitreoretinal surgery. *Ophthalmic Surg Lasers Imaging Retina.* (2021) 52:478–483. doi: 10.3928/23258160-20210817-01

15. Ng MK, Emara AK, Molloy RM, Krebs VE, Mont M, Piuzzi NS. YouTube as a source of patient information for total knee/hip arthroplasty: quantitative analysis of video reliability, quality, and content. *J Am Acad Orthop Surg.* (2021) 29:e1034-e1044. doi: 10.5435/JAAOS-D-20-00910

16. Cassidy JT, Fitzgerald E, Cassidy ES. YouTube provides poor information regarding anterior cruciate ligament injury and reconstruction. *Knee Surg Sports Traumatol Arthrosc.* (2018) 26:840–845. doi: 10.1007/s00167-017-4 514-x

17. Jadhav G, Mittal P, Bajaj NC, Banerjee S. Are YouTube videos on regenerative endodontic procedure reliable source for patient edification? *Saudi Endodontic J.* (2021) 12:43–49. doi: 10.4103/sej.sej_111_21

18. Ozturkmen C, Berhuni M. YouTube as a source of patient information for pterygium surgery. *Ther Adv Ophthalmol.* (2023) 15:25158414231174143. doi: 10.1177/25158414231174143

19. Zaliznyak M, Masterson JM, Duel B. YouTube as a source for information on newborn male circumcision: is YouTube a reliable patient resource? *J Pediatr Urol.* (2022) 18:678.e1-678.e7. doi: 10.1016/j.jpurol.2022. 07.011

20. Lang JJ, Giffen Z, Hong S. Assessing vasectomy-related information on YouTube: sn snalysis of the quality, understandability, and actionability of information. *Am J Mens Health.* (2022) 6:15579883221094716. doi: 10.1177/15579883221094716

21. Er N, Çanakçi FG. Temporomandibular joint arthrocentesis videos on YouTube: are they a good source of information? *J Stomatol Oral Maxillofac Surg.* (2022) 123:e310-e315. doi: 10.1016/j.jormas.2022.03.011

22. Thomas C, Westwood J, Butt GF. Qualitative assessment of YouTube videos as a source of patient information for cochlear implant surgery. *J Laryngol Otol.* (2021) 135:671–674. doi: 10.1017/S0022215121001390

23. Ward M, Abraham ME, Craft-Hacherl C. Neuromodulation, deep brain stimulation, and spinal cord stimulation on YouTube: a content-quality analysis of search terms. *World Neurosurg.* (2021) 151:e156-e162. doi: 10.1016/j.wneu.2021.03.151

24. Tripathi S, ReFaey K, Stein R. The reliability of Deep Brain Stimulation YouTube videos. *J Clin Neurosci.* (2020) 74:202–204. doi: 10.1016/j.jocn.2020.02.015

25. Ward M, Ward B, Abraham M. The educational quality of neurosurgical resources on YouTube. *World Neurosurg.* (2019) 130:e660-e665. doi: 10.1016/j.wneu.2019.06.184

26. Oremule B, Patel A, Orekoya O, Advani R, Bondin D. Quality and reliability of YouTube videos as a source of patient information on rhinoplasty. *JAMA Otolaryngol Head Neck Surg.* (2019) 145:282–283. doi: 10.1001/jamaoto.2018.3723

27. Bayazit S, Ege B, Koparal M. Is the YouTube™ a useful resource of information about orthognathic surgery? A cross-sectional study. *J Stomatol Oral Maxillofac Surg.* (2022) 123:e981-e987. doi: 10.1016/j.jormas.2022.09.001

28. Safa A, De Biase G, Gassie K, Garcia D, Abode-lyamah K, Chen S. Reliability of YouTube videos on robotic spine surgery for patient education. *J. Clin. Neurosci.* (2023) 109:6–10. doi: 10.1016/j.jocn.2022.12.014

29. Roberts B, Kobritz M, Nofi C. Social media, misinformation, and online patient education in emergency general surgical procedures. *J Surg Res.* (2023) 287:16–23. doi: 10.1016/j.jss.2023.01.009

30. Coban G, Buyuk SK. YouTube as a source of information for craniofacial distraction osteogenesis. *J Craniofac Surg.* (2021) 32:2005–2007. doi: 10.1097/SCS.0000000000007478

31. Sader N, Kulkarni AV, Eagles ME, Ahmed S, Koschnitzky JE, Riva-Cambrin J. The quality of YouTube videos on endoscopic third ventriculostomy and endoscopic third ventriculostomy with choroid plexus cauterization procedures available to families of patients with pediatric hydrocephalus. *J Neurosurg Pediatr.* (2020) 25:607–614. doi: 10.3171/2019.12.PEDS19523

32. Lee KN, Son GH, Park SH, Kim Y, Park ST. YouTube as a source of information and education on hysterectomy. *J Korean Med Sci.* (2020) 35:e196. doi: 10.3346/jkms.2020.35.e196

33. Özdal Zincir Ö, Bozkurt AP, Gaş S. Potential patient education of YouTube videos related to wisdom tooth surgical removal. *J Craniofac Surg.* (2019) 30:e481-e484. doi: 10.1097/SCS.0000000000005573

34. Jain N, Abboudi H, Kalic A, Gill F, Al-Hasani H. YouTube as a source of patient information for transrectal ultrasound-guided biopsy of the prostate. *Clin Radiol.* (2019) 74:79.e11–79.e14. doi: 10.1016/j.crad.2018.09.004

35. Camm CF, Russell E, Ji Xu A, Rajappan K. Does YouTube provide high-quality resources for patient education on atrial fibrillation ablation? *Int J Cardiol.* (2018) 272:189–193. doi: 10.1016/j.ijcard.2018.08.066

36. Bae SS, Baxter S. YouTube videos in the English language as a patient education resource for cataract surgery. *Int Ophthalmol.* (2018) 38:1941–1945. doi: 10.1007/s10792-017-0681-5

37. Tan ML, Kok K, Ganesh V, Thomas SS. Patient information on breast reconstruction in the era of the world wide web. A snapshot analysis of information available on youtube.com. *Breast.* (2014) 23:33–37. doi: 10.1016/j.breast.2013. 10.003

38. Strychowsky JE, Nayan S, Farrokhyar F, MacLean J. YouTube: a good source of information on pediatric tonsillectomy? *Int J Pediatr Otorhinolaryngol.* (2013) 77:972–975. doi: 10.1016/j.ijporl.2013.03.023

39. Sorensen JA, Pusz MD, Brietzke SE. YouTube as an information source for pediatric adenotonsillectomy and ear tube surgery. *Int J Pediatr Otorhinolaryngol.* (2014) 78:65–70. doi: 10.1016/j.ijporl.2013.10.045

40. Larouche M, Geoffrion R, Lazare D. Mid-urethral slings on YouTube: quality information on the internet? *Int Urogynecol J.* (2016) 27:903–908. doi: 10.1007/s00192-015-2908-1

41. Nason K, Donnelly A, Duncan HF. YouTube as a patient-information source for root canal treatment. *Int Endod J.* (2016) 49:1194–1200. doi: 10.1111/iej.12575

42. Hegarty E, Campbell C, Grammatopoulos E, DiBiase AT, Sherriff M, Cobourne MT. YouTube™ as an information resource for orthognathic surgery. *J Orthod.* (2017) 44:90–96. doi: 10.1080/14653125.2017.1319010

43. Gray MC, Gemmiti A, Ata A, et al. Can you trust what you watch? An assessment of the quality of information in aesthetic surgery videos on YouTube. *Plast Reconstr Surg.* (2020) 146:329e-336e. doi: 10.1097/PRS.0000000000007213

44. Castillo J, Wassef C, Wassef A, Stormes K, Berry AE. YouTube as a source of patient information for prenatal repair of myelomeningocele. *Am J Perinatol.* (2021) 38:140–144. doi: 10.1055/s-0039-1694786

45. Cho HY, Park SH. Evaluation of the quality and influence of YouTube as a source of information on robotic myomectomy. *J Pers Med.* (2022) 12:1779. doi: 10.3390/jpm12111779

46. Azer SA, AlKhawajah NM, Alshamlan YA. Critical evaluation of YouTube videos on colostomy and ileostomy: can these videos be used as learning resources? *Patient Educ Couns.* (2022) 105:383–389. doi: 10.1016/j.pec.2021.05.023

47. Lee KN, Joo YJ, Choi SY, et al. Content analysis and quality evaluation of cesarean delivery-related videos on YouTube: cross-sectional study. *J Med Internet Res.* (2021) 23:e24994. doi: 10.2196/24994

48. Brooks FM, Lawrence H, Jones A, McCarthy MJ. YouTube™ as a source of patient information for lumbar discectomy. *Ann R Coll Surg Engl.* (2014) 96:144–146. doi: 10.1308/003588414X13814021676396

49. Charnock D, Shepperd S, Needham G, Gann R. DISCERN an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health.* (1999) 53:105–11. doi: 10.1136/jech.53.2.105

50. Mac OA, Thayre A, Tan S, Dodd RH. Web-based health information following the renewal of the cervical screening program in Australia: evaluation of readability, understandability, and credibility. *J Med Internet Res.* (2020) 22:e16701. doi: 10.2196/16701

51. Gudapati JD, Franco AJ, Tamang S, et al. A study of global quality scale and reliability scores for chest pain: an instagram-post analysis. *Cureus.* (2023) 15:e45629. doi: 10.7759/cureus.45629

52. Boyer C, Selby M, Scherrer JR, Appel RD. The Health on the net code of conduct for medical and health web sites. *Comput Biol Med.* (1998) 28:603–10. doi: 10.1016/S0010-4825(98)00037-7

53. Tan DJY, Ko TK, Fan KS. The readability and quality of web-based patient information on nasopharyngeal carcinoma: quantitative content analysis. *JMIR Form Res.* (2023) 7:e47762. doi: 10.2196/47762

54. Simon F, Peer S, Michel J, Bruce IA, Cherkes M, Denoyelle F, et al. IVORY Guidelines (Instructional Videos in Otorhinolaryngology by YO-IFOS): a consensus on surgical videos in ear, nose, and throat. *Laryngoscope.* (2021) 131:E732–7. doi: 10.1002/lary.29020

55. Celentano V, Smart N, Cahill RA, Spinelli A, Giglio MC, et al. Development and validation of a recommended checklist for assessment of surgical videos quality: the LAParoscopic surgery Video Educational GuidelineS (LAP-VEGaS) video assessment tool. *Surg Endosc.* (2021) 35:1362–9. doi: 10.1007/s00464-020-07 517-4

56. Al-Rakhami MS, Al-Amri AM. Detecting fake news and misinformation on COVID-19 using NLP and machine learning. *Int J Inf Manage Data Insights.* (2020) 1:100007. doi: 10.2196/38756

Assessing video-based health education in African contexts: a systematic review

Pinto Francisco Impito[1]* and José Azevedo[2,3]

[1]Department of Arts and Humanities, Licungo University, Quelimane, Mozambique, [2]Department of Communication and Information Sciences, Laboratory for Integrative and Translational Research in Population Health (ITR), EPIUnit-Institute of Public Health, University of Porto, Porto, Portugal, [3]Faculty of Arts and Humanities, University of Porto, Porto, Portugal

Introduction: Health education is vital for empowering patients with knowledge about their health conditions, treatment options, and self-care, enabling them to make informed decisions. Video-based Health Education (VbHE) has proven to be a powerful tool for enhancing health literacy, often outperforming traditional methods in various circumstances. Previous research has focused on specific diseases in low-resource settings, but a broader understanding of video applications and implementation challenges remains necessary. This systematic review assessed the use of video for health education purposes in the least developed African countries from 2020 to 2024, exploring insights into the types of health content delivered through video, their effectiveness, and implementation obstacles.

Methods: We searched the ScienceDirect, PubMed, Scopus, Google Scholar, PLOS, and AJOL databases, limiting our review to the UN's 33 least-developed African countries. Studies were evaluated based on country, participants, sample sizes, intervention methods, duration, video delivery, educational emphasis, objectives, outcomes, and findings. Adhering to PRISMA guidelines ensured a structured methodology.

Results: From the 218 records identified, 15 studies met our inclusion criteria. These studies highlighted video interventions with a focus on cancer ($n = 5$, 33.3%), HIV ($n = 3$, 20%), maternity care ($n = 5$, 33.3%), and other areas ($n = 2$, 13.3%, including COVID-19 and spinal anesthesia procedures). Most studies ($n = 13$, 86.6%) showed videos positively influenced knowledge retention and health outcomes, while the remaining two ($n = 2$, 13.3%) noted no significant difference from traditional methods. Video dissemination methods featured mobile devices, TV screens, and online platforms, with intervention durations ranging from 30 min to 36 months. Of the included studies, video has proven effective in enhancing health education in the least-developed African nations.

Discussion: Challenges including connectivity issues, cultural adaptation, and digital literacy remain significant. This review highlights the critical role of mobile video health communication in empowering underserved populations and informs future strategies to optimize health outcomes in resource-constrained African Contexts.

1 Introduction

The use of video for information and education has experienced substantial growth (Chatterjee et al., 2021). Among various digital formats, video has proven to be an effective medium for delivering health education due to its ability to enhance knowledge retention and accessibility. However, there are significant challenges; health professionals and ordinary web users often share health-related content within the same online spaces, raising concerns about the reliability of the information (Karlsen et al., 2013), and deviant behavior (Zhou et al., 2022). This is particularly concerning given the increasing trend of individuals posting self-produced health videos on social media, shifting responsibility onto the viewers if adverse outcomes happen. Despite these concerns, health professionals continue to rely on user-generated videos instead of producing original material, which impacts the quality and relevance of the education provided (Farrell et al., 2014; Snelson, 2018).

In recent years, research from the least-developed African countries has shown a shift toward health workers producing their video content for educational purposes (Ferla et al., 2023; Scott et al., 2022). This change signals a growing awareness of the significance of original video content in boosting patient education and involvement (Chatterjee et al., 2021; Adam SAM et al., 2019; WHO, 2024). This awareness became even more pronounced during the COVID-19 pandemic, further emphasizing the importance of digital health tools. During this period, many African nations have successfully utilized mobile devices to disseminate educational videos, illustrating the potential for enhanced health outcomes through digital innovation (Kinkade et al., 2022). However, while these tools have helped share information during the pandemic, they have also hosted rumors that led to misinformation, with traditional media being the only means capable of countering it by promoting critical thinking and limiting the spread of health-related rumors (Guo et al., 2023). Providing clear, credible information can foster more thoughtful engagement and increase user acceptance (Luo et al., 2021).

To encourage the use of digital health tools, the World Health Organization (WHO) has urged 47 countries in Africa to develop and implement digital health strategy plans (Alegana et al., 2023). While many of these nations acknowledge the benefits of such technologies and have started incorporating them into their health systems, a notable number ($n = 14$) still do not have comprehensive plans (Alegana et al., 2023). This initiative reflects a growing recognition across the continent of the value of digital technologies for various applications and emphasizes the need for their effective integration into health care. As part of the 2030 Agenda for the African region, WHO's recommendations on digital health interventions support the pursuit of Universal Health Coverage (UHC) and disease prevention, reinforcing the digital health strategy for 2020–2025 (WHO, 2023). In the context of video utilization, the WHO advocates for mHealth based on initiatives such as 'virtual care,' 'remote monitoring,' 'digital therapeutics' (WHO, 2021), 'video-based training,' 'video recordings,' and leveraging YouTube for health videos (WHO, 2024) to improve health outcomes.

Recent reports highlight the constraints of video-based health education in the least developed African nations (Kabukye et al., 2021; Mamuye et al., 2022; Till et al., 2023), where some, if not many, local institutions struggle to maintain their promotion. Barriers such as limited access to affordable mobile data, low smartphone usage, and insufficient digital literacy contribute to a digital divide, hindering the reach and efficacy of video interventions. The WHO indicates that, although some countries have made significant strides in adopting digital health tools, many still require 'institutional support' for developing and strengthening their digital health strategies (WHO, 2021). Beyond infrastructure, emphasizing the usefulness of these tools and fostering attitudes through trust and engaging content can significantly enhance adoption (Zhang et al., 2014).

This systematic review aimed to assess the use of video-based health education in the least developed African countries, focusing on studies published from 2020 to 2024 and examining the impact, obstacles, and opportunities for improvement in under-resourced settings. Since African low-resource settings are often facing systemic barriers, such as underfunded health systems and a lack of trained professionals in using digital media tools for health education, traditional health education methods frequently fall short or are unsustainable. In this context, our study highlights these countries as a strategic and relevant area for evaluating the use of video-based health education as a cost-effective and scalable solution. The goal is not only to assess its applicability and impact but also to increase awareness of its potential and promote further research into integrating it into health training programs in similar environments.

2 Methods

2.1 Search strategy and database overview

A comprehensive literature review was conducted using the ScienceDirect, PubMed, Scopus, Google Scholar, PLOS, and AJOL databases. The search was limited to studies conducted in the 33 least-developed African countries as defined by the UN, and adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). The keywords and search terms included combinations of: "Health videos," "Educational videos," "Video-based health education," "Digital health," "Health education," and "Low-resource settings." Boolean operators (AND, OR) were utilized to refine the search results, with filters applied to select studies published between 2020 and 2024, in English, and meeting the specified inclusion criteria.

During full-text review and data extraction, it became apparent that the included studies could be meaningfully categorized based on the primary population addressed by the interventions, namely, Patients and/or citizens (e.g., general public, disease-specific populations, adolescents, pregnant women, rural communities), and Health workers (e.g., doctors, nurses, community health workers, program managers, drug distributors). The latter group highlights the importance of multi-stakeholder approaches grounded in supportive policies to guide the development and promotion of digital health content (Zhang et al., 2017). In these situations, stakeholders are typically dedicated professionals who volunteer their time without pay, reflecting the realities of low-resource environments (Yang et al., 2021). Although the division between Patients and/or Citizens and Health workers was not an initial inclusion criterion, it was identified inductively due to apparent differences in intervention design, educational content, and reported outcomes. Consequently, after screening and extraction, studies were grouped according to the target population (patients/citizens vs. health workers). Subsequent synthesis

and reporting were structured along these two categories to provide a more nuanced understanding of how video-based health education approaches differ according to the intended audience.

2.2 Ligibility and exclusion criteria

The review encompassed studies focused on enhancing, evaluating, designing, advocating for, testing, or investigating the use of video as a medium for health education. The eligibility criteria centered on studies published between 2020 and 2024 that examined Video-based Health Education in low-income or least-developed African nations. Studies lacking educational video content, those conducted outside the least-developed African countries, and those published outside the designated timeframe were excluded.

2.3 Data extraction and synthesis

Data extraction was performed using a predefined data extraction form to ensure uniformity across studies. Key information extracted included study characteristics (such as country, participants, sample size, intervention method, research duration, and video delivery approach) along with a performance summary (including educational contents, objectives, outcomes, and findings). To ensure the accuracy and integrity of all extracted data, two independent reviewers (PFI and JA) initially tested the extraction form on a random sample of included studies. They refined variable definitions until achieving an agreement of over 95%. Subsequently, each reviewer independently extracted data for all studies. Discrepancies were identified through systematic cross-checking in a shared spreadsheet and then resolved through discussion. Both reviewers adjudicated any remaining disagreements. Audit trails documented all changes, and the authors recalculated records to verify consistency. The gathered data were qualitatively synthesized, highlighting the effectiveness of Video-based Health Education interventions. A narrative synthesis was employed to summarize and interpret the results, with a focus on identifying observed patterns, differences, and common themes across the studies. The findings were then organized in tables, and key outcomes were discussed in relation to the research objectives (Table 1).

Following initial study selection, included studies were subsequently grouped by the population targeted by the intervention (patients/citizens vs. health workers) based on data extracted from the full texts.

TABLE 1 The path followed on the method.

Step	Description
Screening & Inclusion	All studies of Video-based Health Education were included, regardless of target population.
Data Extraction	During the review, information was extracted about the population focus (patients/citizens or health workers).
Data Synthesis	Studies were grouped *post hoc* based on the intervention population; results were reported by group.

3 Results

3.1 Study selection

The database search yielded 218 papers, of which 15 met our selection criteria after excluding duplicates ($n = 3$). We excluded 197 articles for the following reasons: (1) no use of educational video ($n = 57$), (2) research was not conducted in the least-developed African countries ($n = 63$), and (3) studies fell outside the 2020–2024-year range ($n = 77$). Participants' demographics included health workers who received video training ($n = 4$), patients, and/or citizens ($n = 11$), all of whom were assessed for their responses to video-based health education (Figure 1).

3.2 Study characteristics

Only 24 countries were identified as implementing video-based health education. These include Tanzania ($n = 4$), Ghana ($n = 3$), Malawi ($n = 2$), Ethiopia ($n = 2$), Nigeria ($n = 2$), Kenya ($n = 1$), Botswana ($n = 1$), Uganda ($n = 1$), Rwanda ($n = 1$), Burundi ($n = 1$), Niger ($n = 1$), the Democratic Republic of the Congo ($n = 1$), São Tomé and Príncipe ($n = 1$), Benin ($n = 1$), Burkina Faso ($n = 1$), Gambia ($n = 1$), Guinea ($n = 1$), Guinea-Bissau ($n = 1$), Liberia ($n = 1$), Mali ($n = 1$), Mauritania ($n = 1$), Senegal ($n = 1$), Sierra Leone ($n = 1$), and Togo ($n = 1$). In this review, Tanzania and Ghana emerged as the countries with the most studies, followed by Malawi, Ethiopia, and Nigeria, as illustrated in Figure 2.

3.2.1 Sample size

The study's sample size varied from 9 to 7,648 participants. Four studies reported an average age range of 22–42 (Holst et al., 2021; Masiano et al., 2021; Godana Boynito et al., 2023; Obasola, 2021); three had an average age of 18 (Martei et al., 2024; Drokow et al., 2021; Kanyeki et al., 2022), while the remaining eight did not provide age data (Ferla et al., 2023; Scott et al., 2022; Tilly et al., 2022; Cooper et al., 2021; Ampofo et al., 2020; Mengistu et al., 2021; Holst et al., 2022; Ansari et al., 2024). The search exclusively targeted the least-developed African countries between 2020 and 2024. According to the UN classification, 33 countries are categorized as least developed (UN, 2024).

3.2.2 Dissemination, method, and duration

The studies utilized various media to present videos, including TV monitors or projectors ($n = 6$), mobile devices ($n = 7$) such as tablets, mobile phones, and portable projectors, online media ($n = 1$) through platforms such as YouTube, Web, and WhatsApp, and other formats ($n = 1$), such as TV, DVD, or compact disks. These video-based health education initiatives occurred in health facilities, schools, and occasionally in residential settings through door-to-door visits. The duration of the education programs varied from 30 to 40 min up to 36 months.

Three of the 15 studies reviewed employed true randomized control designs, assessing both control and intervention groups (Kanyeki et al., 2022; Holst et al., 2022; Ansari et al., 2024). The other 12 were single-group studies, often conducted in focus groups or with individuals, and utilized pre- and post-intervention assessments (Ferla et al., 2023; Scott et al., 2022; Holst et al., 2021; Masiano et al., 2021; Godana Boynito et al., 2023; Obasola, 2021; Martei et al., 2024;
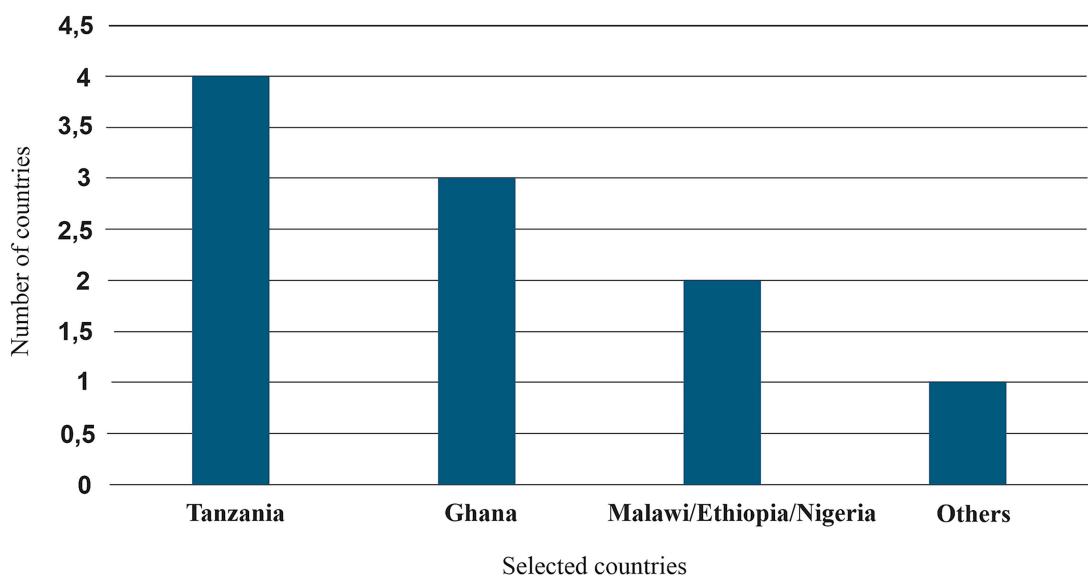
FIGURE 1
Number of studies per identified country.

Drokow et al., 2021; Tilly et al., 2022; Cooper et al., 2021; Ampofo et al., 2020; Mengistu et al., 2021) (Table 2).

3.3 Performance summary

3.3.1 Educational contents and outcomes

In general, the studies focused on video education initiatives covering topics such as cancer ($n = 5$, 33.3%; Martei et al., 2024; Drokow et al., 2021; Tilly et al., 2022; Cooper et al., 2021; Ampofo et al., 2020), HIV ($n = 3$, 20%; Holst et al., 2021; Masiano et al., 2021; Holst et al., 2022), maternity care ($n = 5$, 33.3%; Ferla et al., 2023; Godana Boynito et al., 2023; Obasola, 2021; Mengistu et al., 2021; Ansari et al., 2024), and other areas ($n = 2$, 13.3%, including COVID-19 and spinal anesthesia procedures; Scott et al., 2022; Kanyeki et al., 2022). In terms of preliminary outcomes, most studies ($n = 13$, 86.6%) found that the use of video proved effective, and in the remaining ($n = 2$, 13.3%), no significant differences were noted between video usage and no video.

The studies focusing on patients and/or citizens ($n = 11$) covered a range of topics. Of these, five ($n = 5$) focused on cancer-related subjects, using videos to discuss common myths, diagnostic procedures, bodily diseases, treatments, reactions to aftereffects, and quality of life (Martei et al., 2024). They aimed to raise awareness and knowledge about the screening and risks associated with human papillomavirus infection, the treatment and experiences of breast cancer survivors, and to promote vaccination, awareness, and counseling for the Pap smear test. Three studies ($n = 3$) addressed HIV-related issues (Holst et al., 2021; Masiano et al., 2021; Holst et al., 2022), through videos that highlighted the importance of antiretroviral therapy and addressed significant concerns regarding HIV/AIDS, tuberculosis, Taenia Solium (neuro) cysticercosis, taeniasis, and anthrax. Finally, the last three studies ($n = 3$) covered maternal and child health (Godana Boynito et al., 2023; Ansari et al., 2024), as well as spinal anesthesia procedures, including a video that reviews the pre-operative anxiety state (Kanyeki et al., 2022).

Similarly, studies focusing on health workers ($n = 4$) featured content related to maternity care topics that included testimonials from mothers (Mengistu et al., 2021), COVID-19 education, showcasing a video that guided handwashing, mask-wearing, and maintaining social distancing (Scott et al., 2022), motherhood, neonatal, and maternal child healthcare (Obasola, 2021), and finally, counselors offering guidance on nurturing care for caregivers (Ferla et al., 2023).

3.3.2 Finding summary

Eleven of the fifteen selected studies focused on patients and/or citizens, including students and women. Nine of these studies reported increased (1) knowledge (Holst et al., 2021; Drokow et al., 2021; Tilly et al., 2022; Cooper et al., 2021), (2) awareness and information retention (Masiano et al., 2021; Godana Boynito et al., 2023; Ampofo et al., 2020), and overall, (3) the video content was found to be acceptable, usable, and recommendable (Holst et al., 2021; Martei et al., 2024). Two studies reported "no statistical difference" in outcomes with or without video use, yet indicated increased knowledge (Ansari et al., 2024) and no discernible benefits (Kanyeki et al., 2022) (Table 3). Meanwhile, the remaining four studies, which involved health workers comprising health professionals, healthcare providers, program managers, and community health workers, showed improvements and clarity in understanding health content presented in video format (Ferla et al., 2023; Scott et al., 2022; Obasola, 2021; Mengistu et al., 2021). These studies recognized the potential of video for delivering effective health content, and overall, video-based health education interventions were reported to have a positive impact on the target audience (Table 4).

4 Discussion

This review systematically compiles evidence on video-based health education initiatives in the 24 least-developed African countries from 2020 to 2024. Among 15 eligible studies, most (13 out of 15, 86.6%) reported positive effects on knowledge, awareness, or self-efficacy among patients,

FIGURE 2
PRISMA flow diagram used in the study description.
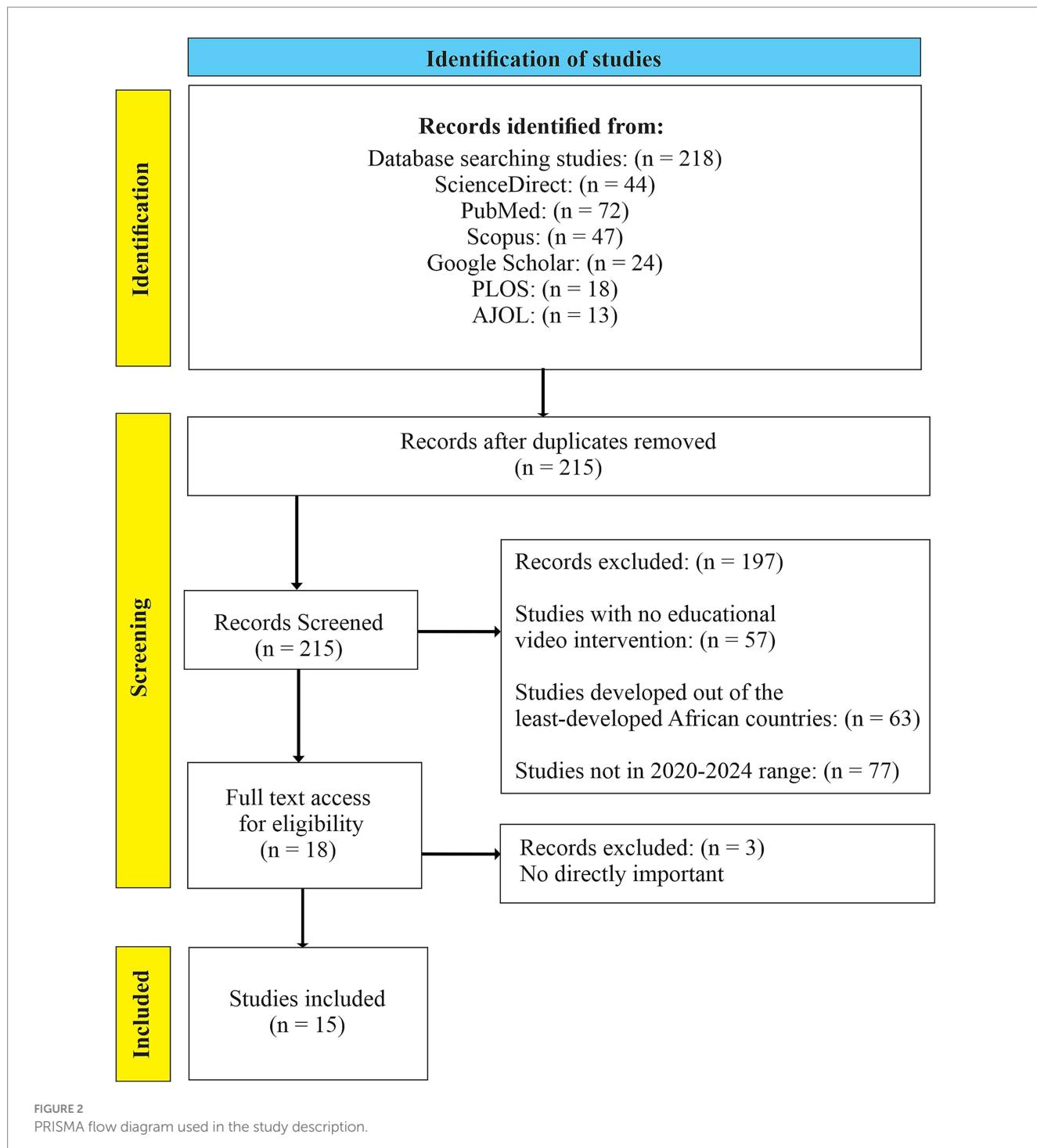
citizens, and health workers. These results support global findings that multimodal media, such as video, can enhance understanding and memory (Ferguson, 2012; Robert et al., 2023; Yang et al., 2021), particularly in settings with low literacy levels (Ferguson, 2012; Liu et al., 2025).

4.1 Effectiveness and reach

Video-based health education is an impactful method for engaging patients effectively across various healthcare intervention sectors, including diagnosis, treatment, and self-care (Ploderer et al., 2022). Research conducted in some of the least-developed African nations highlighted this potential. Ansari et al. (2024) and Masiano et al. (2021) reported that video is a valuable tool for promoting health education among pregnant women regarding maternal and childbirth health (Ansari et al., 2024), and improving maternal retention in HIV care and the uptake of antiretroviral therapy (Masiano et al., 2021), respectively. Furthermore, Tilly et al. (2022) demonstrate that video is a cost-effective tool. Their study reveals that video-based health education saves patients time and money compared to traveling by

TABLE 2 Study characteristics.

Author/Year	Study participants	Sample size	Intervention method	Research duration	Video delivery
Tilly et al. (2022)	Patients and/or Citizens	100	Pre and post	6 months	TV monitor or projector
Masiano et al. (2021)		132	Pre and post	1 month	TV monitor or projector
Cooper et al. (2021)		760	Pre and post	2–5 days	TV monitor or projector
Ampofo et al. (2020)		50	Pre and post	1 session	TV monitor or projector
Kanyeki et al. (2022)		38	Cg. and Ig.	8 months	Tablet and compact disc
Holst et al. (2021)		12	Pre and post	24 months	Mobile phone
Holst et al. (2022)		302	Cg. and Ig.	12 months	Tablet
Ansari et al. (2024)		7,684	Cg. and Ig.	36 months	Android handsets
Martei et al. (2024)		106	Pre and post	8 months	Tablet
Drokow et al. (2021)		645	Pre and post	12 months	TV monitor or projector
Godana Boynito et al. (2023)		Focus group of 10–12	Pre and post	12 months	Mobile projector
Mengistu et al. (2021)	Health workers	106	Pre and post	27 months	TV monitor or projector
Scott et al. (2022)		52	Pre and post	1 month	WhatsApp, YouTube Web and PCs
Ferla et al. (2023)		113	Pre and post	10 months	Mobile devices
Obasola (2021)		9	Pre and post	30–40 min. Per session	TV/DVD

Cg, Control group; Ig, Intervention group.

bus or other means to attend awareness sessions at health centers. Patients can watch videos from home as often as they like using smartphones or other internet-enabled devices. The rapid sharing and widespread availability of health-related videos online have established them as a leading medium for health education and promotion (Ghozali, 2023).

Understanding health conditions helps diminish the barriers between doctors and patients, transforming doubt into informed dialogue. This improved understanding leads to higher patient satisfaction, whether it comes from real-time online support delivered by a physician (Tan et al., 2023), or from a pre-recorded health education video that engages both visual and auditory senses (Anasi, 2004). Further to the above, research from Tilly et al. (2022), Cooper et al. (2021), Holst et al. (2021), and Drokow et al. (2021) demonstrated how effective videos were in increasing knowledge among their audiences, primarily focusing on cancer, human papillomavirus infection, and HIV/AIDS. This increase in knowledge encourages patients to engage with the content anytime, anywhere, thereby enhancing comprehension retention (Naderyanfar et al., 2019; Gusdorf et al., 2023) and positively influencing memory, understanding, and emotions. These cognitive and emotional attitudes are primarily driven by how the information is tailored to the audience and how effectively it is delivered (Wang et al., 2018).

Numerous systematic reviews evaluating the effects of video-based health education in the least-developed African countries have

shown steady growth in its application throughout the continent (Till et al., 2023). This approach has been shown to enhance knowledge among individuals with low literacy levels markedly (Mbanda et al., 2021); it acts as an effective decision-making tool (Mwase et al., 2022) and improves patients' comprehension of health-related information (Boynito et al., 2024). Yet, several challenges hinder the implementation of video-based health education in these regions, which local governments need to tackle. These challenges include inadequate funding, limited technological literacy, a scarcity of technological resources (Kabukye et al., 2021), a lack of a health information exchange policy (Mamuye et al., 2022), and an overreliance on text-based communication (Till et al., 2023). Nevertheless, at least four studies from the least-developed African countries indicate that health institutions actively promote health education through videos: (1) Mengistu et al. conducted a survey with multidisciplinary health professionals, finding video-based education to be a powerful method for enhancing care (Mengistu et al., 2021), (2) Scott et al., through their study of program managers and drug distributors, discovered that educational videos with health content effectively reinforce messages and are easy to understand (Scott et al., 2022), (3) Ferla et al. explored video-based health education with community health workers, who reported improved counseling performance and enhanced mentorship (Ferla et al., 2023), and finally, (4) Obasola investigated healthcare providers' perspectives, finding that implementing information and communication technology with

TABLE 3 Performance summary (patients and/or citizens).

Author/Year	Educational contents	Objectives	Outcomes	Findings
Tilly et al. (2022)	Cancer	Assess change in cancer knowledge and care experience.	A low-cost tool that educates and empowers patients.	Significant improvement in knowledge.
Masiano et al. (2021)	HIV	Improve maternal retention in HIV care and ART uptake.	Acceptable, feasible, and supported decision-making tool.	Reduced anxiety and fear, and an increase in hope.
Cooper et al. (2021)	HPV/Cervical Cancer	Measure the impact of the HPV/cervical cancer educational video.	Knowledge acquisition improved.	Significant improvement in knowledge.
Ampofo et al. (2020)	Cervical Cancer	Evaluate the video's impact on awareness of screening.	Raised awareness, understanding, and belief.	Significant increase in awareness.
Kanyeki et al. (2022)	Spinal anesthesia procedure	Test video information on pre-procedure anxiety.	No benefit from the video.	No statistical difference
Holst et al. (2021)	HIV/AIDS, TB, Cysticercosis, anthrax	Demonstrate a multi/interdisciplinary design process.	An animated video developed key health messages.	Recommended for multidisciplinary design.
Holst et al. (2022)	HIV/AIDS, TB, Cysticercosis, anthrax	Assess knowledge uptake and retention in the rural community.	Animated video significantly improved health knowledge.	Significant improvement in knowledge.
Ansari et al. (2024)	Maternal and Child Health	Improve maternal and child health through "edutainment" videos.	Knowledge and attitudes improved, but not significantly.	No statistical change in attitudes.
Martei et al. (2024)	Breast Cancer	Develop and assess a survivor-narrative video for treatment support.	Highly acceptable and usable.	Highly acceptable and usable.
Drokow et al. (2021)	HPV vaccination /Cervical Carcinoma	Promote HPV vaccination and Pap tests via behavior-change videos.	Influenced perception, self-efficacy, and understanding.	Significant improvement in knowledge.
Godana Boynito et al. (2023)	Maternal and Infant Community Care	Test the acceptability and feasibility of community-based video.	Acceptable and feasible for promoting behavior change.	Acceptable and feasible intervention tool.

HIV, Human Immunodeficiency Virus; ART, Antiretroviral Therapy; TB, Tuberculosis; HPV, Human Papillomavirus.

TABLE 4 Performance summary (health workers).

Author/Year	Educational emphasis	Objectives	Outcomes	Findings
Mengistu et al. (2021)	Maternity care	Describe the development and implementation of improving respectful maternity care.	Testimonial-based video education is an effective way to improve care.	Significant improvement in knowledge.
Scott et al. (2022)	COVID-19	Describe the development and evaluation of video job aids to support the safe delivery of SMC during COVID-19 and improve seasonal malaria chemoprevention quality.	Users easily understand video-based health education and are reinforced by key messages by allowing for repeated viewing.	Significant improvement in the quality of message delivery.
Ferla et al. (2023)	Nurturing care/Early Child Development counselling	Evaluate the effect of ECD video with mentorship and supervision to improve CHW nurturing care/ECD counselling.	Video-based health education, enhanced supervision, mentorship, and improved CHW performance in providing counselling	Significant improvement in education.
Obasola (2021)	Maternal Child Health	Explore healthcare providers' experiences with using ICT tools (e.g., videos) for MCH and optimize their application.	The growing demand for ICT, particularly educational videos in health services, has led to demonstrable improvements in education outcomes.	Significantly relevant.

SMC, Seasonal malaria chemoprevention; ECD, Early Child Development; CHW, Community Health Worker; ICT, Information and Communication Technology.

video usage positively impacted health education interventions and received strong recommendations for its effectiveness (Obasola, 2021).

Compatibility issues with legacy devices, such as DVD players, outdated smartphones, and analog TVs, as well as the lack of standardized educational platforms in low-resource settings, continue to limit the adoption of AR/VR tools (Mondal and Mondal, 2025). Moreover, the rapid obsolescence of such devices, coupled with high maintenance costs, contributes to financial strain. Within this context, specialized outreach programs can capitalize on the widespread availability of smartphones, observed as early as 2015, by distributing smartphone-based VR kits like Google Cardboard. This strategy enables the delivery of immersive training experiences without requiring significant infrastructure investments (Mondal and Mondal, 2025). The Pragati study conducted in rural India demonstrates this approach by using Google Cardboard-mounted phones to train community health workers in maternal and child health (Sorathia et al., 2017). Adopting 360° guided audiovisual modules that simulated an Assamese home environment (Sorathia et al., 2017), participants experienced a training method distinct from those employed by Ferla and Obasola in training health workers across various African contexts (Ferla et al., 2023; Obasola, 2021).

Beyond the Google Cardboard approach, Neuwirth and Ros proposed the use of 180° video viewed through VR head-mounted displays (HMDs) or 3D-capable computer monitors (Neuwirth and Ros, 2021). These were deemed affordable and practical solutions, particularly suitable for training healthcare professionals in Africa. Compared to educational videos delivered via mobile phones, televisions, projectors, or DVDs, this approach is more effective in fostering engagement. Unlike the least-developed countries in Africa, more developed nations, including the United States, Europe, and parts of Asia, have successfully utilized VR-embedded video as a medium for health education (Costa et al., 2024; Jia et al., 2021). In the United States, for example, this broad application has led to the establishment of online public databases containing videos on various health education topics (Media GH, 2024; Plus, 2024), as well as VR websites dedicated to health education, such as medvr.education (USA), exr.education (England), SIMx (USA), among others.

4.2 Core findings

The assessment of selected studies in this systematic review was divided into two data subsets: "study design" and "performance summary" (Tables 2–4). These datasets were crucial for illustrating the role of video in health education in the least developed African countries from 2020 to 2024. The intervention methodologies employed in these studies to reach their audience were essential for evaluating the impact of videos designed for health education. Randomized controlled trial studies ($n = 3$) consisted of two groups (e.g., control group with no video and intervention group with video) to compare the effects of video-based health education versus traditional methods. Single-group studies ($n = 12$) assessed the effectiveness of video by examining a single group through pre- and post-intervention measures.

All included studies reported positive outcomes, although two ($n = 2$) found no significant difference between video-based methods and conventional practices. Research involving patients and citizens ($n = 9$) revealed increased knowledge, awareness, and retention of information related to cancer, HIV, human papillomavirus, and

maternal and child health care. Furthermore, these studies indicated that video-based health education was deemed acceptable, feasible, and usable. Various sharing tools were employed to evaluate video effectiveness, including TV monitors, projectors, mobile phones, tablets, compact disks, PCs, and digital platforms such as WhatsApp, YouTube, and websites. Additionally, four studies ($n = 4$) involving health workers highlighted video as an effective educational tool due to its comprehensive nature. They reported understanding health content better when presented in video format, acknowledging its potential for effective health education delivery.

4.3 Future directions

Future research should focus on longitudinal studies to determine whether observed knowledge gains result in lasting behavioral changes, such as regular cancer screenings or consistent adherence to ART. Comparative studies assessing the cost-effectiveness of video versus traditional methods, as well as the level of engagement generated by live-action versus animated videos, are also recommended. Technological innovations, such as AI-driven adaptive videos and gamified features, could further enhance engagement and effectiveness. Equity-oriented research questions remain urgent, including how female participants respond and how to deliver health video education to nomadic and conflict-affected populations.

Ultimately, this review serves as a crucial reminder for the least-developed African nations about the efficacy of digital solutions in overcoming educational obstacles and enhancing health education strategies. Importantly, although limitations and challenges, this review concludes that significant attention is being given to the delivery of Video-based Health Education in most target countries ($n = 24$, 72.7%).

5 Conclusion

Video-based Health Education has demonstrated substantial potential in enhancing health literacy, knowledge retention, and patient engagement across various health issues in the least-developed African nations. This systematic review highlights the successful use of videos for educational health content, particularly in areas such as cancer, HIV, COVID-19, Human papillomavirus infection, spinal anesthesia procedures, and maternal health. The findings reveal that Video-based Health Education can help bridge gaps left by traditional healthcare delivery, offering a more accessible and engaging resource for patients, health workers, and the general public to enhance their understanding of health matters.

However, several challenges still need to be addressed, such as limited access to digital infrastructure, low levels of technological literacy, and inconsistent internet connectivity. These issues create a digital divide that hinders the broad adoption of video education approaches. Furthermore, the absence of institutional backing and inadequate incorporation of digital health strategies in many African nations restricts the full realization of its benefits. To address these challenges, further research is needed to assess the scalability, sustainability, and long-term effects of the Video-based Health Education initiative. For instance, comparing offline and online delivery methods in both rural and urban areas can help assess the health outcomes, such as knowledge acquisition and behavior change,

alongside operational metrics like cost per user, training duration, and technology uptime. Incorporating formal cost-effectiveness and cost–benefit analysis into these studies will provide police makers with essential ROI (Return on investment) data to justify infrastructure investment. Longitudinal cohort studies that combine quantitative surveys and detailed qualitative interviews can also shed light on how digital literacy efforts for patients and healthcare workers influence engagement with video-based health education, as well as identify relevant facilitators or barriers in different contexts.

In conclusion, when carefully adapted and implemented, Video-based Health Education can substantially bridge knowledge and healthcare service gaps in marginalized African populations, empower patients, and provide scalable solutions for the development of health workers. Nevertheless, its effectiveness remains context-dependent, and continued evaluation is needed to maximize its transformative potential while addressing persistent challenges in resource-constrained settings.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

PI: Conceptualization, Data curation, Writing – original draft, Writing – review & editing. JA: Data curation, Formal analysis, Methodology, Supervision, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Correction note

A correction has been made to this article. Details can be found at: 10.3389/fcomm.2025.1699250.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Adam SAM, M., Prober, C., and Bärnighausen, T. (2019). Human-centered design of video-based health education: an iterative, collaborative, community-based approach. *Adv. Digit. Health Open Sci.* 21. doi: 10.2196/12128

Alegana, A., Victor LJF, A., Makubalo, L. E., Kalu, A. A., and Impouma, B. (2023). Digital health in the African region should be integral to the health system's strengthening. *Mayo Clin. Proc. Digit. Health* 1, 425–434. doi: 10.1016/j.mcpdig.2023.06.003

Ampofo, A. G., SAGOpoku, N. S., Owusu, S. O., and Ibitoye, M. B. (2020). A pilot study of a video-based educational intervention and knowledge of cervical cancer among senior high school students in Ghana: a before-after study. *J. Cancer Policy* 24:6. doi: 10.1016/j.jcpo.2020.100220

Anasi, S. N. I. (2004). Audiovisual media capabilities and the dissemination of the health information. *Libr. Focus* 22. doi: 10.29087/2011.3.4.08

Ansari, U., Omer, K., Aziz, A., Gidado, Y., Mudi, H., Jamaare, I. S., et al. (2024). Added value of video edutainment on android handsets in home visits to improve maternal and child health in Bauchi state, Nigeria: secondary analysis from a cluster randomised controlled trial. *Digit. Health* 10:20552076241228408. doi: 10.1177/20552076241228408

Boynito, W. G., Pauwels, N. S., Otayto, K., De Henauw, S., and Abbeddou, S. (2024). Effects of community-based educational video interventions on nutrition, health, and use of health services in low- and middle-income countries: systematic review and meta-analysis. *Nutr. Rev.* 83, 201–216. doi: 10.1093/nutrit/nuae004

Chatterjee, A. G. S., Meinert, E., Milne-Ives, M., Halkes, M., and Wyatt-Haines, E. (2021). The use of video for patient information and education: a scoping review of the

variability and effectiveness of interventions. *Patient Educ. Couns.* 104, 2189–2199. doi: 10.1016/j.pec.2021.02.009

Cooper, E. C., Maher, J. A., Naaseh, A., Crawford, E. W., Chinn, J. O., Runge, A. S., et al. (2021). Implementation of human papillomavirus video education for women participating in mass cervical cancer screening in Tanzania. *Am. J. Obstet. Gynecol.* 224, 105.e1–105.e9. doi: 10.1016/j.ajog.2020.07.018

Costa, JBD, Leyva, RGE, S, PHODF, and Souto, EJP, editors. (2024). "Enhancing virtual reality experiences through embedded 3D models in video content." in *2024 IEEE International Conference on Consumer Electronics (ICCE); 6–8 Jan. 2024.*

Drokow, E. K., Effah, C. Y., Agboyibor, C., Sasu, E., Amponsem-Boateng, C., Akpabla, G. S., et al. (2021). The impact of video-based educational interventions on cervical cancer, pap smear and HPV vaccines. *Front. Public Health* 9:3. doi: 10.3389/fpubh.2021.681319

Farrell, E. H., Phillips, K., Morgan, B., Savage, K., Lewis, V., Kelly, M., et al. (2014). Systematic review and meta-analysis of audio-visual information aids for informed consent for invasive healthcare procedures in clinical practice. *Patient Educ. Couns.* 94, 20–32. doi: 10.1016/j.pec.2013.08.019

Ferguson, L. A. (2012). Implementing a video education program to improve health literacy. *J. Nurse Pract.* 8, e17–e22. doi: 10.1016/j.nurpra.2012.07.025

Ferla, J. P., Gill, M. M., Komba, T., Abubakar, A., Remes, P., Jahanpour, O., et al. (2023). Improvement of community health worker counseling skills through early childhood development (ECD) videos, supervision and mentorship: a mixed methods pre-post

evaluation from Tanzania. *PLoS Glob. Public Health* 3:e0001152. doi: 10.1371/journal.pgph.0001152

Ghozali, M. T. (2023). Is integrating video into tech-based patient education effective for improving medication adherence? A review. *Paladyn J. Behav. Robot.* 14:109. doi: 10.1515/pjbr-2022-0109

Godana Boynito, W., Tessema, G. Y., Temesgen, K., De Henauw, S., and Abbeddou, S. (2023). Acceptability and feasibility of video-based health education for maternal and infant health in Dirashe District, South Ethiopia: a qualitative study. *PLoS Glob. Health* 3:e0000821. doi: 10.1371/journal.pgph.0000821

Guo, F., Zhou, A., Zhang, X., Xu, X., and Liu, X. (2023). Fighting rumors to fight COVID-19: investigating rumor belief and sharing on social media during the pandemic. *Comput. Hum. Behav.* 139:107521. doi: 10.1016/j.chb.2022.107521

Gusdorf, R. E., Shah, K. P., Triana, A. J., McCoy, A. B., Pabla, B., Scoville, E., et al. (2023). A patient education intervention improved rates of successful video visits during rapid implementation of telehealth. *J. Telemed. Telecare* 29, 607–612. doi: 10.1177/1357633X211008786

Holst, C., Isabwe, G. M. N., Sukums, F., Ngowi, H., Kajuna, F., Radovanović, D., et al. (2021). Development of digital health messages for rural populations in Tanzania: multi- and interdisciplinary approach. *JMIR Mhealth Uhealth* 9:e25558. doi: 10.2196/25558

Holst, C., Stelzle, D., Diep, L. M., Sukums, F., Ngowi, B., Noll, J., et al. (2022). Improving health knowledge through provision of free digital health education to rural communities in Iringa, Tanzania: nonrandomized intervention study. *J. Med. Internet Res.* 24:e37666. doi: 10.2196/37666

Jia, F., Xu, S., and Huo, J. (2021). Embedded VR video image control system in the ideological and political teaching system based on artificial intelligence. *Mob. Inf. Syst.* 2021, 1–12. doi: 10.1155/2021/7839821

Kabukye, J. K., Kakungulu, E., Keizer, N., and Cornet, R. (2021). Digital health in oncology in Africa: a scoping review and cross-sectional survey. *Int. J. Med. Inform.* 158:104659. doi: 10.1016/j.ijmedinf.2021.104659

Kanyeki, T. V. M., Bal, R., and Odaba, D. (2022). Effect of video-based information on preoperative state trait anxiety inventory scores in adult patients presenting for elective caesarean section: a randomized controlled trial. *Afr. Health Sci.* 22, 117–124. doi: 10.4314/ahs.v22i3.14

Karlsen, R, Morell, JEB, Luque, LF, and Salcedo, VT editors. (2013). "A domain-based approach for retrieving trustworthy health videos from YouTube." in *14th World Congress on Medical and Health Informatics (MEDINFO); 2013 Aug 20–23; Copenhagen, DENMARK*.

Kinkade, C., Russpatrick, S., Potter, R., Saebo, J., Sloan, M., Odongo, G., et al. (2022). Extending and strengthening routine DHIS2 surveillance systems for COVID-19 responses in Sierra Leone, Sri Lanka, and Uganda. *Emerg. Infect. Dis.* 28, S42–S48. doi: 10.3201/eid2813.220711

Liu, X., Susarla, A., and Padman, R. (2025). Promoting health literacy with human-in-the-loop video understandability classification of YouTube videos: development and evaluation study. *J. Med. Internet Res.* 27:e56080. doi: 10.2196/56080

Luo, P., Ma, X., Zhang, X., Liu, J., and He, H. (2021). How to make money with credit information? Information processing on online accommodation-sharing platforms. *Tour. Manag.* 87:104384. doi: 10.1016/j.tourman.2021.104384

Mamuye, A. L., Yilma, T. M., Abdulwahab, A., Broomhead, S., Zondo, P., Kyeng, M., et al. (2022). Health information exchange policy and standards for digital health systems in africa: a systematic review. *PLoS Digit. Health* 1:e0000118. doi: 10.1371/journal.pdig.0000118

Martei, Y. M., Mokokwe, L., Ngwako, N., Kebuang, K., Setlhako, D. I., Gabaatlhole, G., et al. (2024). Development, acceptability and usability of culturally appropriate survivor narrative videos for breast cancer treatment in Botswana: a pilot study. *BMJ Open* 14:e073867. doi: 10.1136/bmjopen-2023-073867

Masiano, S., Machine, E., Mphande, M., Markham, C., Tembo, T., Chitani, M., et al. (2021). Video-based intervention for improving maternal retention and adherence to HIV treatment: patient perspectives and experiences. *Int. J. Environ. Res. Public Health* 18:1737. doi: 10.3390/ijerph18041737

Mbanda, N., Dada, S., Bastable, K., Ingalill, G. B., and Ralf, W. S. (2021). A scoping review of the use of visual aids in health education materials for persons with low-literacy levels. *Patient Educ. Couns.* 104, 998–1017. doi: 10.1016/j.pec.2020.11.034

Media GH. (2024). Our Videos. Available online at: https://globalhealthmedia.org/about-us/ (Accessed October 14, 2024).

Mengistu, B., Alemu, H., Kassa, M., Zelalem, M., Abate, M., Bitewulign, B., et al. (2021). An innovative intervention to improve respectful maternity care in three districts in Ethiopia. *BMC Pregnancy Childbirth* 21:541. doi: 10.1186/s12884-021-03934-y

Mondal, H., and Mondal, S. (2025). Adopting augmented reality and virtual reality in medical education in resource-limited settings: constraints and the way forward. *Adv. Physiol. Educ.* 49, 503–507. doi: 10.1152/advan.00027.2025

Mwase, C., Nkhoma, K., and Allsop, M. J. (2022). The role of digital health in palliative care for people living with HIV in sub-Saharan Africa: a systematic review. *Digit. Health* 8:205520762211337. doi: 10.1177/20552076221133707

Naderyanfar, F., Shahrakimoghadam, E., Heidari, M. A., and Soleimani, M. (2019). Evaluation of the effect of video-based education on self-care of patients with type II diabetes. *J. Diabetes Nurs.* 7, 672–682.

Neuwirth, L. S., and Ros, M. (2021). Comparisons between first person point-of-view 180° video virtual reality head-mounted display and 3D video computer display in teaching undergraduate neuroscience students stereotaxic surgeries. *Front. Virtual Real.* 2:706653. doi: 10.3389/frvir.2021.706653

Obasola, O. (2021). Experiences of health care providers using information and communications technology for maternal and child health care in selected health facilities in Nigeria. *Libr. Philos. Pract.* 2021, 1–22.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71

Ploderer, B., Rezaei Aghdam, A., and Burns, K. (2022). Patient-generated health photos and videos across health and well-being contexts: scoping review. *J. Med. Internet Res.* 24:e28867. doi: 10.2196/28867

Plus, M. (2024). Trusted health information for you. Available online at: https://medlineplus.gov/anatomyvideos.html (Accessed August 7, 2024).

Robert, H, Fernando, S, Kishanthini, Tan, E, Jin, Y, Moe, S, et al. (2023). Effectiveness of video-based learning on students' memory: a randomized controlled trial.

Scott, S., Camara, B. S., Hill, M., Lama, E. K., Barry, L., Ogouyemi-Hounto, A., et al. (2022). The use of video job-aids to improve the quality of seasonal malaria chemoprevention delivery. *PLoS Digit. Health* 1:e0000165. doi: 10.1371/journal.pdig.0000165

Snelson, C. (2018). Video production in content-area pedagogy: a scoping study of the research literature. *Learn. Media Technol.* 43, 294–306. doi: 10.1080/17439884.2018.1504788

Sorathia, K., Sharma, K., Bhowmick, S., and Kamidi, P. (2017). "Pragati - a mobile based virtual reality (VR) platform to train and educate community health workers" in Human-computer interaction – INTERACT 2017 (Cham: Springer International Publishing).

Tan, H., Zhang, X., and Yang, Y. (2023). Satisfaction or gratitude? Exploring the disparate effects of physicians' knowledge sharing on patients' service evaluation in online medical consultations. *Inf. Syst. J.* 33, 1186–1211. doi: 10.1111/isj.12440

Till, S., Mkhize, M., Farao, J., Shandu, L. D., Muthelo, L., Coleman, T. L., et al. (2023). Digital health Technologies for Maternal and Child Health in Africa and other low- and middle-income countries: cross-disciplinary scoping review with stakeholder consultation. *J. Med. Internet Res.* 25:e42161. doi: 10.2196/42161

Tilly, A. E., Ellis, G. K., Chen, J. S., Manda, A., Salima, A., Mtangwanika, A., et al. (2022). Implementation and evaluation of educational videos to improve Cancer knowledge and patient empowerment. *JCO Glob. Oncol.* 8:e2100315. doi: 10.1200/GO.21.00315

UN. (2024). UN list of developed countries. Available online at: https://unctad.org/topic/least-developed-countries/list (Accessed October 8, 2024).

Wang, L., Wu, T., Guo, X., Zhang, X., Li, Y., and Wang, W. (2018). Exploring mHealth monitoring service acceptance from a service characteristics perspective. *Electron. Commer. Res. Appl.* 30, 159–168. doi: 10.1016/j.elerap.2018.06.003

WHO. (2021). Global strategy on digital health. Available online at: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf (Accessed July 25, 2024).

WHO. (2023). Vision, strategy and special initiatives, 2023–2030. Available online at: https://www.afro.who.int/sites/default/files/2023-09/Ending%20disease%20in%20Africa_ENDISA_ENG_0.pdf (Accessed July 31, 2024).

WHO. (2024). WHO compendium of innovative health technologies for low-resource settings Available online at: https://iris.who.int/bitstream/handle/10665/378235/9789240095212-eng.pdf?sequence=1 (Accessed July 25, 2024).

WHO. (2024). Tactics to apply to make your communications understandable: Make it visual. Available online at: https://www.who.int/about/communications/understandable/visuals.

Yang, Q., Wu, Z., Xie, Y., Xiao, X., Wu, J., Sang, T., et al. (2021). The impact of health education videos on general public's mental health and behavior during COVID-19. *Glob. Health Res. Policy* 6:37. doi: 10.1186/s41256-021-00211-5

Yang, Y., Zhu, X., Song, R., Zhang, X., and Guo, F. (2021). Not just for the money? An examination of the motives behind physicians' sharing of paid health information. *J. Inf. Sci.* 49:016555152199102. doi: 10.1177/0165551521991029

Zhang, X., Guo, X., Guo, F., and Lai, K. H. (2014). Nonlinearities in personalization-privacy paradox in mHealth adoption: the mediating role of perceived usefulness and attitude. *Technol. Health Care* 22, 515–529. doi: 10.3233/THC-140811

Zhang, X., Lai, K.-h., and Guo, X. (2017). Promoting China's mHealth market: a policy perspective. *Health Policy Technol.* 6, 383–388. doi: 10.1016/j.hlpt.2017.11.002

Zhou, C., Li, K., and Zhang, X. (2022). Why do i take deviant disclosure behavior on internet platforms? An explanation based on the neutralization theory. *Inf. Process. Manag.* 59:102785. doi: 10.1016/j.ipm.2021.102785

Correction: Assessing video-based health education in African contexts: a systematic review

Pinto Francisco Impito[1]* and José Azevedo[2,3]

[1]Department of Arts and Humanities, Licungo University, Quelimane, Mozambique, [2]Department of
Communication and Information Sciences, Laboratory for Integrative and Translational Research in
Population Health (ITR), EPIUnit-Institute of Public Health, University of Porto, Porto, Portugal, [3]Faculty
of Arts and Humanities, University of Porto, Porto, Portugal

KEYWORDS

VbHE, digital health, health education, educational video, African context

A Correction on

Assessing video-based health education in African contexts: a
systematic review

by Impito, P. F., and Azevedo, J. (2025). *Front. Commun.* 10:1545306.
doi: 10.3389/fcomm.2025.1545306

There was a mistake in the published captions of **Figures 1** and **2**. The caption of **Figure 1** belongs to **Figure 2** and the caption of **Figure 2** belongs to **Figure 1**.

They were previously published as:

Figure 1. PRISMA flow diagram used in the study description.

Figure 2. Number of studies per identified country.

The corrected captions of **Figures 1** and **2** appear below.

Figure 1. Number of studies per identified country.

Figure 2. PRISMA flow diagram used in the study description.

The original version of this article has been updated.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Check for updates

Comparative performance of large language models for patient-initiated ophthalmology consultations

Mingxue Huang[1†], Xiaoyan Wang[1†], Shiqi Zhou[2], Xinyu Cui[2], Zilin Zhang[2], Yanwu Xu[2], Weihua Yang[3*] and Wei Chi[3*]

[1]School of Nursing, Southwest Medical University, Luzhou, China, [2]School of Future Technology, South China University of Technology, Guangzhou, China, [3]Shenzhen Eye Hospital, Shenzhen Eye Medical Center, Southern Medical University, Shenzhen, China

Background: Large language models (LLMs) are increasingly accessed by lay users for medical advice. This study aims to conduct a comprehensive evaluation of the responses generated by five large language models.

Methods: We identified 31 ophthalmology-related questions most frequently raised by patients during routine consultations and subsequently elicited responses from five large language models: ChatGPT-4o, DeepSeek-V3, Doubao, Wenxin Yiyan 4.0 Turbo, and Qwen. A five-point likert scale was employed to assess each model across five domains: accuracy, logical consistency, coherence, safety, and content accessibility. Additionally, textual characteristics, including character, word, and sentence counts, were quantitatively analyzed.

Results: ChatGPT-4o and DeepSeek-V3 achieved the highest overall performance, with statistically superior accuracy and logical consistency ($p < 0.05$). Existing safety evaluations indicate that both Doubao and Wenxin Yiyan 4.0 Turbo exhibit significant security deficiencies. Conversely, Qwen generated significantly longer outputs, as evidenced by greater character, word, and sentence counts.

Conclusion: ChatGPT-4o and DeepSeek-V3 demonstrated the highest overall performance and are best suited for laypersons seeking ophthalmic information. Doubao and Qwen, with their richer clinical terminology, better serve users with medical training, whereas Wenxin Yiyan 4.0 Turbo most effectively supports patients' pre-procedural understanding of diagnostic procedures. Prospective randomized controlled trials are required to determine whether integrating the top-performing model into pre-consultation triage improves patient comprehension.

KEYWORDS

large language model, healthcare, consultation, ophthalmology, patient education

1 Introduction

Advances in deep learning have enabled large language models (LLMs) to achieve substantial breakthroughs in natural language processing, demonstrating broad utility across text generation, semantic comprehension, translation, and inferential reasoning ([1], [2]). Recently, generative artificial intelligence has exhibited considerable promise within the healthcare sector, particularly in standardized examination simulation and clinical

documentation, thereby invigorating contemporary medical practice (3–5). Advanced LLMs, exemplified by ChatGPT and DeepSeek, are now systematically deployed across diverse medical specialties and have demonstrated early efficacy in disease recognition, diagnostic support, and evidence-based clinical decision-making (6–8). LLMs have demonstrated high diagnostic accuracy and decision-making efficacy in subspecialties such as neuro-ophthalmology, glaucoma, and thyroid eye disease, underscoring their substantial application potential (9–11). Furthermore, large language models such as Qwen, Doubao, and Wenxin Yiyan exhibit substantial translational promise across clinical and research workflows (12, 13). These systems enhance healthcare service efficiency, mitigate clinician workload, and foster patient health literacy and equitable access to care (14).

Nevertheless, the deployment of LLMs in clinical settings faces several challenges, among which 'model hallucination' is particularly pronounced (15, 16). Such models may produce outputs that are structurally coherent yet factually erroneous, a limitation arising from outdated training corpora and restricted access to contemporary medical guidelines, ultimately compromising the comprehensiveness and authority of their knowledge bases (17, 18). Although initiatives such as DeepSeek seek to mitigate the black-box problem through enhanced transparency and interpretability, the medical community retains circumspection regarding their reliability (19, 20). The growing utilization of LLMs for unsupervised health self-diagnosis may expose lay users to inaccurate or unsafe information, thereby amplifying potential harms (21, 22). Besides, ophthalmology necessitates exceptionally high diagnostic precision, as even marginal deviations can adversely affect patient prognosis (23). Therefore, comprehensive performance evaluations within ophthalmological contexts are urgently required prior to their widespread clinical adoption (24). Existing studies focus primarily on different versions of ChatGPT, leaving a scarcity of comparative analyses across models (25).

This study systematically evaluates five LLMs (ChatGPT-4o, DeepSeek-V3, Qwen, Doubao, and Wenxin Yiyan 4.0 Turbo) and focuses on their responses to ophthalmology-related questions from patients. Model outputs will be comprehensively assessed across five domains: accuracy, logical consistency, coherence, safety, and content accessibility. Additionally, quantitative text metrics (character, word, and sentence counts) will be extracted from Chinese-language outputs to elucidate their practical utility for patient education and clinical decision support.

2 Method

2.1 Ethical statement

This cross-sectional evaluation compared responses generated by five LLMs to 31 frequently encountered consultation questions in ophthalmology. The questions were derived from routine clinical inquiries collected by healthcare providers during patient encounters. Crucially, the study involved no patient-level data or personally identifiable information, thereby fully preserving individual anonymity and privacy.

2.2 Model selection

We purposefully selected five state-of-the-art LLMs: ChatGPT-4o, DeepSeek-V3, Qwen, Doubao, and Wenxin Yiyan 4.0 Turbo. Selection criteria encompassed recent benchmark performance, public accessibility, the developers' institutional credibility, and demonstrated suitability for medical question-answering.

2.2.1 ChatGPT-4o

ChatGPT-4o[1] is OpenAI's newest transformer-based large-scale language model. It leverages deep-learning techniques to deliver advanced generative and comprehension capabilities, and its multimodal architecture ensures robust performance across heterogeneous input modalities, encompassing text and images.

2.2.2 DeepSeek-V3

DeepSeek-V3[2] is engineered for high-performance information retrieval and open-domain question answering, integrating deep-learning and reinforcement-learning techniques to optimize retrieval efficiency and accuracy.

2.2.3 Qwen

Qwen[3] is a conversational LLM optimized for interactive question-answering, emphasizing user engagement and real-time feedback.

2.2.4 Doubao

Doubao[4] is specifically optimised for Chinese-language tasks, employing multi-layer attention mechanisms to capture nuanced semantics and cultural contexts.

2.2.5 Wenxin Yiyan 4.0 Turbo

Wenxin Yiyan 4.0 Turbo[5] is tailored for Chinese natural-language processing, exhibiting strong generative and semantic-understanding capabilities.

2.3 Study design

We conducted a cross-sectional benchmarking study evaluating how the five selected LLMs respond to 31 frequently asked consultation questions covering retinal diseases, macular degeneration, glaucoma, dry eye and associated procedures. Questions were classified as definitional, causal, comparative, or procedural and reflect typical patient queries.

On 6 March 2025, two investigators jointly recorded the answer generated by each model in a single submission. Each question was submitted separately through the online platforms corresponding to the five models. No system prompts were provided, and responses were generated *de novo* from the query. Following response generation, the chat histories were manually reset to prevent carryover of context.

All outputs were independently verified by two researchers and transcribed into a Microsoft Excel spreadsheet. Character, word, and

---

1 https://openai.com/index/hello-gpt-4o/

2 https://chat.deepseek.com/

3 https://qwen.org/chat

4 https://www.doubao.com/chat/

5 https://yiyan.baidu.com/

TABLE 1 Performance scores of five large language models across accuracy, logic, coherence, safety, and content accessibility.

Metric	Chat GPT-4o	DeepSeek-V3	Qwen	Doubao	Wenxin Yiyan 4.0 turbo	H-value	p-value
Accuracy	5.0 (4.0,5.0)	5.0 (4.0,5.0)	4.0 (4.0,5.0)	4.0 (3.0,4.0)	4.0 (4.0,4.0)	50.90	<0.05
Logic	5.0 (4.0,5.0)	5.0 (5.0,5.0)	4.0 (4.0,5.0)	4.0 (4.0,4.0)	4.0 (4.0,5.0)	29.82	<0.05
Coherence	4.0 (4.0,5.0)	4.0 (3.0,4.0)	4.0 (4.0,5.0)	4.0 (4.0,4.0)	4.0 (4.0,4.0)	11.69	<0.05
Safety	4.0 (4.0,4.0)	3.0 (3.0,4.0)	4.0 (3.0,4.0)	3.0 (2.0,3.0)	3.0 (3.0,4.0)	52.30	<0.05
Content accessibility	4.0 (3.0,5.0)	4.0 (3.0,5.0)	3.0 (3.0,4.0)	3.0 (2.0,4.0)	4.0 (3.0,5.0)	12.54	<0.05

sentence counts were automatically derived using the online text analytics tool Xiezuocat.[6]

Two board-certified vitreoretinal attending physicians with equivalent seniority (each with ≥5 years of subspecialty experience) independently rated each response across five domains: accuracy, logical consistency, coherence, safety, and content accessibility, using a five-point Likert scale (1 = poor, 5 = excellent). Detailed scoring criteria and the full question list are provided in the Supplementary material. All interactions were conducted within a controlled online environment following standardized operating procedures to maximize reproducibility.

2.4 Data analysis

All analyses were conducted in SPSS software (version 27.0). Inter-rater consistency of total scores was assessed with the intraclass correlation coefficient (ICC). Normality was evaluated using the Shapiro–Wilk test. Normally distributed continuous variables were expressed as mean ± SD; non-normally distributed variables as median (IQR). Homogeneity of variances was evaluated using Levene's test. Parametric comparisons among the five models employed one-way analysis of variance (ANOVA); non-parametric analyses utilized the Kruskal–Wallis H test. Where significant differences were detected, Bonferroni-corrected *post-hoc* pairwise comparisons were performed. $p < 0.05$ was deemed statistically significant.

3 Results

3.1 Comparative performance of five LLMs

The ICC between the two raters was 0.87. Table 1 summarizes the median scores of the five LLMs across five domains: accuracy, logical consistency, coherence, safety, and content accessibility. Accuracy: ChatGPT-4o and DeepSeek-V3 attained the maximum median score of 5.0, significantly surpassing the remaining models ($H = 50.90$, $p < 0.05$). Logical consistency: Likewise, ChatGPT-4o and DeepSeek-V3 achieved a median of 5.0, significantly exceeding the others ($H = 29.82$, $p < 0.05$). Coherence: Scores differed modestly; nevertheless, ChatGPT-4o and DeepSeek-V3 exhibited marginally higher stability ($H = 11.69$, $P<0.05$). Safety: ChatGPT-4o scored

highest (4.0), whereas Doubao and Wenxin Yiyan 4.0 Turbo recorded the lowest (3.0), with significant between-group differences ($H = 52.30$, $p < 0.05$). Content accessibility: ChatGPT, DeepSeek-V3 and Wenxin Yiyan 4.0 Turbo performed best (4.0), while Qwen and Doubao scored lower (3.0); these differences were statistically significant ($H = 12.54$, $p < 0.05$). Detailed differences are provided in Table 2 and Figure 1.

3.2 Output length characteristics

Table 3 and Figure 2 present descriptive statistics for character, word, and sentence counts. Qwen produced the longest responses (1,380.58), significantly exceeding ChatGPT-4o (826.48) and DeepSeek-V3 (636.90) ($p < 0.05$). Wenxin Yiyan 4.0 Turbo generated 916.45 words, approximating ChatGPT-4o. Similarly, Qwen yielded the highest token count (639.16), substantially surpassing DeepSeek-V3 (314.45) and ChatGPT-4o (417.55) ($p < 0.05$). Doubao and Wenxin Yiyan 4.0 Turbo produced fewer tokens (428.03 and 465.00, respectively). Qwen also generated the greatest number of sentences (53.06), significantly exceeding DeepSeek-V3 (33.16) and ChatGPT-4o (36.00) ($p < 0.05$). Conversely, Doubao and Wenxin Yiyan 4.0 Turbo produced the fewest sentences (20.97 and 32.90, respectively). Collectively, Qwen generated significantly more characters, words, and sentences than all other models ($p < 0.05$). Comprehensive pairwise comparisons are presented in Table 3 and Figure 2.

4 Discussion

As LLMs are increasingly adopted in ophthalmology, where diagnostic precision is paramount, their accuracy, safety, and clarity directly affect clinical decision support and patient education (26). Patients now commonly seek online health information and may obtain LLM-based advice without clinician oversight; therefore, these systems must meet rigorous quality standards before healthcare implementation.

Our findings demonstrate statistically significant inter-model heterogeneity, with ChatGPT-4o and DeepSeek-V3 achieving superior overall performance. Consistent with earlier reports (23, 27), ChatGPT-4o exhibits near-expert proficiency in ophthalmological question-answering, while DeepSeek-V3 matches ChatGPT-4o in accuracy; both significantly outperform the remaining three models. This superiority may be attributable to: (1) the increased complexity of open-ended questions relative to prior multiple-choice formats; (2)

6 https://xiezuocat.com/

TABLE 2 Pairwise comparisons between models.

Comparison	Accuracy	Logic	Coherence	Safety	Content accessibility	Characters	Words	Sentences
ChatGPT-4o vs. DeepSeek-V3	1.000	1.000	**0.010***	**0.002***	1.000	0.091	**0.037***	1.000
ChatGPT-4.o vs. Qwen	0.063	1.000	1.000	1.000	1.000	**0.002***	**0.006***	0.312
ChatGPT-4o vs. Doubao	**0.000***	**0.000***	0.866	**0.000***	0.072	1.000	1.000	**0.000***
ChatGPT-4o vs. Wenxin Yiyan 4.0 turbo	**0.000***	0.431	0.927	**0.002***	1.000	1.000	1.000	1.000
DeepSeek-V3 vs. Qwen	**0.032***	0.234	0.195	0.057	0.848	**0.000***	**0.000***	0.077
DeepSeek-V3 vs. Doubao	**0.000***	**0.000***	1.000	0.065	**0.021***	0.083	**0.017***	**0.002***
DeepSeek-V3 vs. Wenxin Yiyan 4.0 turbo	**0.000***	0.078	1.000	1.000	1.000	**0.001***	**0.000***	1.000
Qwen vs. Doubao	0.060	0.077	1.000	**0.000***	1.000	**0.003***	**0.014***	**0.000***
Qwen vs. Wenxin Yiyan 4.0 turbo	1.000	1.000	1.000	0.052	1.000	0.180	0.369	0.058
Doubao vs. Wenxin Yiyan 4.0 turbo	1.000	0.231	1.000	0.071	0.121	1.000	1.000	**0.004***

* denotes statistical significance at the 0.05 level.

delayed updates in competing models; and (3) advanced algorithmic architectures and curated training corpora employed by ChatGPT-4o and DeepSeek-V3.

Coherence scores were comparable across models, yet ChatGPT-4o and Qwen exhibited marginally superior stability (28), suggesting that architectural heterogeneity influences medical reasoning construction. We additionally assessed the inclusion of disclaimers intended to mitigate medical and legal risk. ChatGPT-4o and Qwen frequently appended disclaimers (e.g., "seek prompt medical attention" or "consult a qualified clinician"), indicating stronger safety-control mechanisms than their counterparts.

When addressing different query types, all models provided comprehensive descriptions of disease-related content, particularly for definitional questions. For diagnostic tasks, DeepSeek-V3 and Wenxin Yiyan 4.0 Turbo supplied extensive clinical context and complication analyses, whereas ChatGPT-4o remained concise yet superficial (14). Previous studies have not reported that Doubao and Qwen display broader stylistic variation than other models (29), whereas ChatGPT-4o, although clear and concise, shows limited stylistic flexibility.

Upon addressing the query "How does diabetes induce retinal damage?," ChatGPT-4o first defined diabetic retinopathy and summarized its pathophysiology, then listed preventive measures (glycemic control, annual retinal screening, optimization of lipids and blood pressure, smoking cessation, limited alcohol intake, and supplementation with lutein, vitamins C and E, and ω-3 fatty acids). DeepSeek-V3 more deeply into the underlying molecular mechanisms

while simultaneously elaborating on disease progression and clinical manifestations. Qwen and Doubao concentrate on a hierarchical analysis of pathological mechanisms, whereas clinical management recommendations are comparatively sparse. Wenxin Yiyan 4.0 Turbo first described the disease, then detailed relevant examinations such as optical coherence tomography. The examples of this study indicate that ChatGPT-4o and DeepSeek-V3 are better suited for the general public seeking disease information, whereas Qwen, Wenxin Yiyan 4.0 Turbo, and Doubao employ more complex medical terminology that benefits clinical trainees but may hinder comprehension among non-specialists. Such complexity may impede information acquisition, emotional support, and interpersonal rapport among patients (30, 31).

Converging evidence from our multi-dimensional assessment described above suggests that the observed balance of accuracy, conciseness, and safety renders these models operationally viable for eye-care pathways.

Previous studies have demonstrated that ChatGPT demonstrates diagnostic accuracy comparable to, or even exceeding, that achieved by ophthalmology residents in distinguishing primary from secondary glaucoma (10). This study further demonstrated that ChatGPT-4o rapidly identified patients requiring immediate referral versus routine follow-up, consistent with earlier studies (32). Within hierarchical diagnostic and treatment settings, chatbots demonstrate a superior capacity to identify acute and severe conditions (33), substantially enhancing patient satisfaction and the overall care experience (34). In circumstances where a patient cannot attend a hospital or clinic in person, or requires expeditious triage to ascertain the urgency of
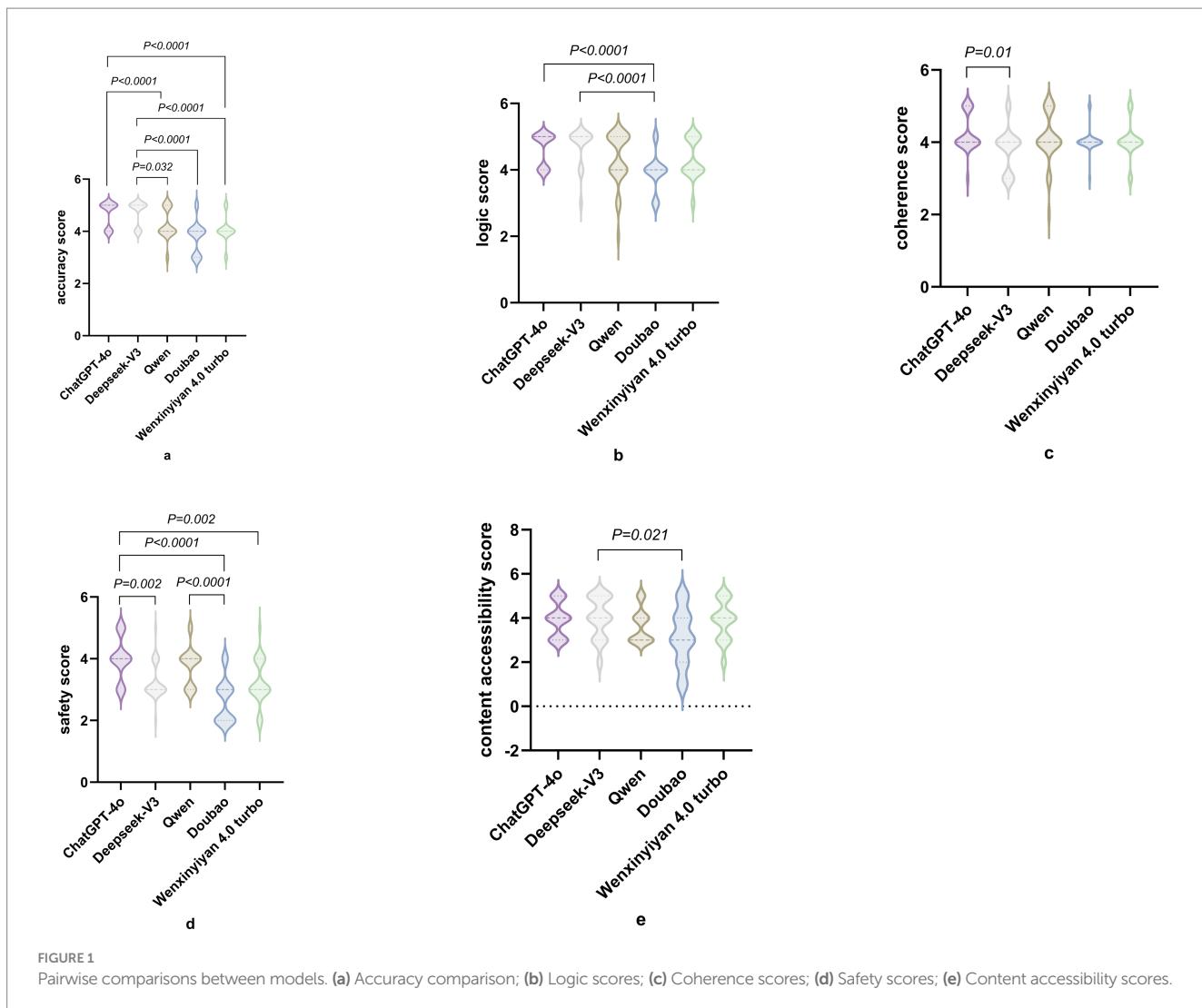
FIGURE 1
Pairwise comparisons between models. **(a)** Accuracy comparison; **(b)** Logic scores; **(c)** Coherence scores; **(d)** Safety scores; **(e)** Content accessibility scores.

TABLE 3 Response lengths of five large language models to 31 ophthalmology related queries.

Metric	Chat GPT-4o	DeepSeek-V3	Qwen	Doubao	Wenxin Yiyan 4.0 turbo	H-value	p-value
Characters	826.48 ± 240.62	636.90 ± 213.66	1380.58 ± 584.93	833.29 ± 285.08	916.45 ± 237.45	41.94	<0.05
Words	417.55 ± 115.32	314.45 ± 96.95	639.16 ± 247.72	428.03 ± 140.17	465.00 ± 114.65	42.29	<0.05
Sentences	36.00 ± 13.63	33.16 ± 12.689	53.06 ± 25.77	20.97 ± 7.71	32.90± 11.80	41.52	<0.05

professional medical attention, LLMs can be leveraged to provide case-specific recommendations (34).

The application of LLMs in ophthalmology is rapidly expanding across medical education, clinical support, research, and patient education (35). However, persistent challenges (inconsistent performance, algorithmic bias, hallucinations, data-privacy risks, and ethical dilemmas) remain (36). Patients with ophthalmic concerns should continue to consult certified eye-care professionals, ensuring adequate human oversight in clinical decision-making (26, 37). Future initiatives must prioritize iterative model refinement and interdisciplinary ethical governance to ensure responsible clinical deployment (24, 25). Empirical evidence confirms that well-crafted prompts enhance

both output accuracy and contextual relevance (38–40), although prompt variation exerts limited influence on accuracy, it substantially modifies textual readability (41, 42). Consequently, readability remains pivotal for effective patient communication even when accuracy gains are marginal.

LLMs trained with domain-specific ophthalmological expertise outperform those trained on general corpora (43). Future validation pipelines for ophthalmology-focused LLMs should span multi-center, multi-tier institutions and establish an iterative cycle of fine-tuning, validation, and governance. Interdisciplinary experts in ophthalmology, law, and ethics will craft an adaptive governance framework, while curated multi-center datasets drive continuous model refinement. The integration of this model into medical
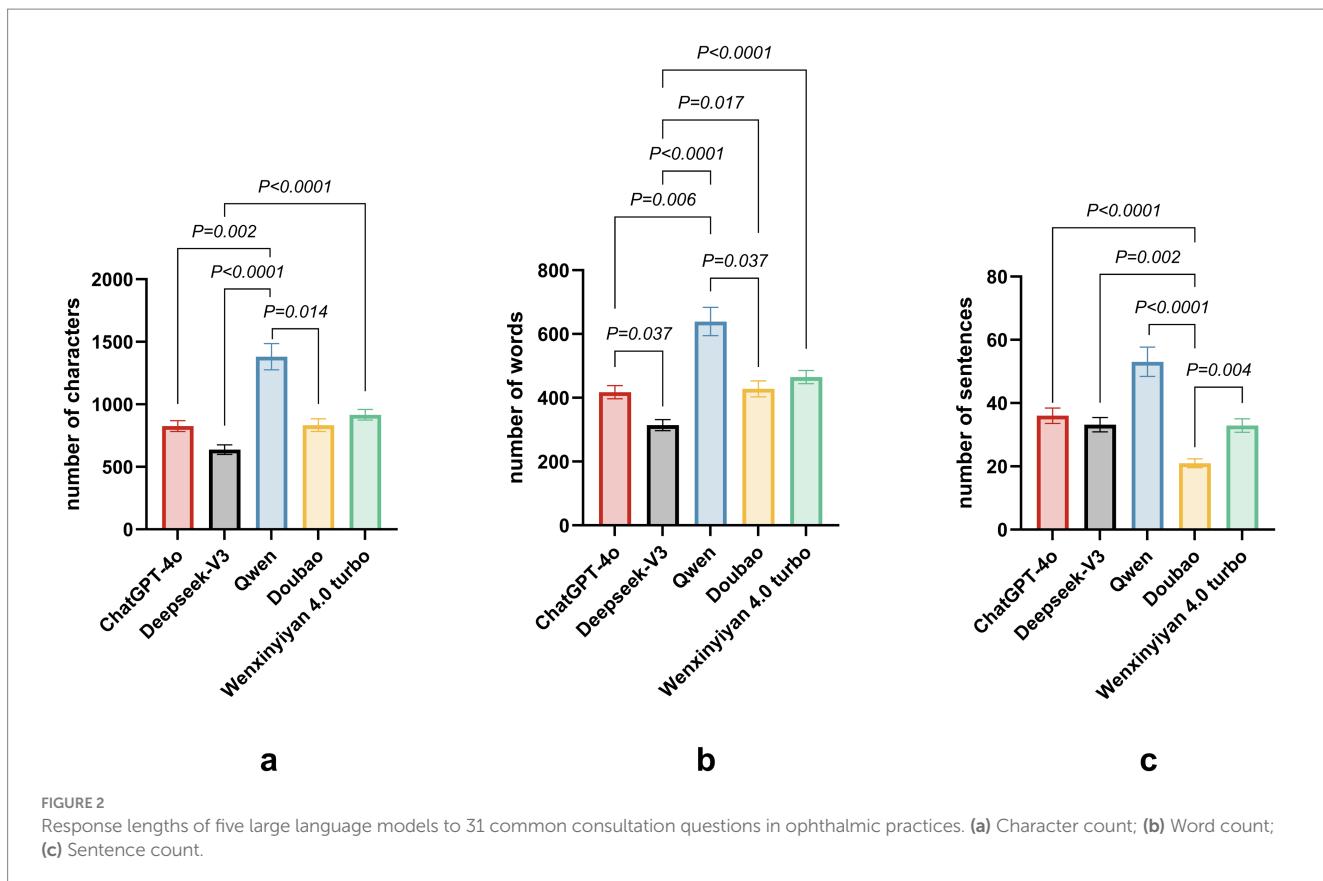
FIGURE 2
Response lengths of five large language models to 31 common consultation questions in ophthalmic practices. **(a)** Character count; **(b)** Word count; **(c)** Sentence count.

education platforms can be used to generate immersive virtual patient cases that significantly bridge the gap between theory and clinical practice (44), while alleviating the healthcare burden in resource-limited regions (45, 46). We therefore recommend that the platform adopt a two-pronged strategy: first, encourage physicians to participate as cohesive teams to leverage peer-learning and collaborative mechanisms for enhancing overall service quality; second, embed robust privacy-preserving safeguards within personalized services so that patients can fully benefit from precision medicine without concerns about data security.

Our study has several limitations. First, each query was presented only once without priming or real-world outcome validation, potentially underestimating model capabilities. Second, analyses were restricted to Chinese-language responses, limiting generalizability. Third, we focused on the most common ophthalmic conditions, which may not fully capture the breadth of LLM functions. Future work should incorporate diverse, real-time datasets and develop validated tools for assessing linguistic complexity in Chinese LLMs to improve reliability, and should expand evaluation to additional models to clarify domain-specific strengths and limitations.

5 Conclusion

This study systematically evaluated five mainstream LLMs on ophthalmology question-answering tasks, revealing inter-model differences in accuracy, logical consistency, coherence, safety, and

content accessibility. ChatGPT-4o and DeepSeek-V3 consistently outperformed the others, particularly in accuracy and logical consistency. Qwen produced the longest and most lexically rich outputs. Qwen, Wenxin Yiyan 4.0 Turbo, and Doubao employed complex medical terminology that may hinder comprehension among non-specialists. Continued technological advances and mitigation of current limitations will substantially enhance the clinical utility of LLMs.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

MH: Data curation, Validation, Writing – original draft. XW: Data curation, Formal analysis, Writing – original draft. SZ: Validation, Software, Writing – original draft. XC: Software, Visualization, Writing – original draft. ZZ: Software, Writing – original draft. YX: Methodology, Software, Writing – original draft. WY: Conceptualization, Methodology, Project administration, Supervision, Writing – review & editing. WC: Conceptualization, Methodology, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by National Natural Science Foundation of China (82571272) and Sanming Project of Medicine in Shenzhen (No. SZSM202411007).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2025.1673045/full#supplementary-material

References

1. Jin Q, Leaman R, Lu Z. PubMed and beyond: biomedical literature search in the age of artificial intelligence. *EBioMedicine*. (2024) 100:104988. doi: 10.1016/j.ebiom.2024.104988

2. Health TLD. Large language models: a new chapter in digital health. *Lancet Digit Health*. (2024) 6:e1. doi: 10.1016/S2589-7500(23)00254-6

3. Stribling D, Xia Y, Amer MK, Graim KS, Mulligan CJ, Renne R. The model student: GPT-4 performance on graduate biomedical science exams. *Sci Rep*. (2024) 14:5670. doi: 10.1038/s41598-024-55568-7

4. Peng Y, Rousseau JF, Shortliffe EH, Weng C. AI-generated text may have a role in evidence-based medicine. *Nat Med*. (2023) 29:1593–4. doi: 10.1038/s41591-023-02366-9

5. Ali SR, Dobbs TD, Hutchings HA, Whitaker IS. Using ChatGPT to write patient clinic letters. *Lancet Digit Health*. (2023) 5:e179–81. doi: 10.1016/S2589-7500(23)00048-1

6. Perlis RH, Fihn SD. Evaluating the application of large language models in clinical research contexts. *JAMA Netw Open*. (2023) 6:e2335924. doi: 10.1001/jamanetworkopen.2023.35924

7. Jin X, Zhang J, Yang J, Yang S, Xue D, Zhang Z. Every cloud has a silver lining: DeepSeek's light through acute respiratory distress syndrome shadows. *J Thorac Dis*. (2025) 17:1109–13. doi: 10.21037/jtd-2025-381

8. Wei J, Wang X, Huang M, Xu Y, Yang W. Evaluating the performance of ChatGPT on board-style examination questions in ophthalmology: a meta-analysis. *J Med Syst*. (2025) 49:94. doi: 10.1007/s10916-025-02227-7

9. Yeganeh M, Mohammad D, Priscilla A L, Joseph W F. ChatGPT assisting diagnosis of neuro-ophthalmology diseases based on case reports. (2024). Available online at: https://journals.lww.com/jneuro-ophthalmology/fulltext/9900/chatgpt_assisting_diagnosis_of_neuro_ophthalmology.723.aspx (Accessed April 10, 2025)

10. Delsoz M, Raja H, Madadi Y, Tang AA, Wirostko BM, Kahook MY, et al. The use of ChatGPT to assist in diagnosing Glaucoma based on clinical case reports. *Ophthalmol Ther*. (2023) 12:3121–32. doi: 10.1007/s40123-023-00805-x

11. Milad D, Antaki F, Milad J, Farah A, Khairy T, Mikhail D, et al. Assessing the medical reasoning skills of GPT-4 in complex ophthalmology cases. *Br J Ophthalmol*. (2024) 108:1398–405. doi: 10.1136/bjo-2023-325053

12. Liu Y, Yuan Y, Yan K, Li Y, Sacca V, Hodges S, et al. Evaluating the role of large language models in traditional Chinese medicine diagnosis and treatment recommendations. *NPJ Digit Med*. (2025) 8:466. doi: 10.1038/s41746-025-01845-2

13. Kang D, Wu H, Yuan L, Shen W, Feng J, Zhan J, et al. Evaluating the efficacy of large language models in guiding treatment decisions for pediatric refractive error. *Ophthalmol Ther*. (2025) 14:705–16. doi: 10.1007/s40123-025-01105-2

14. Lim ZW, Pushpanathan K, Yew SME, Lai Y, Sun CH, Lam JSH, et al. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google bard. *EBioMedicine*. (2023) 95:104770. doi: 10.1016/j.ebiom.2023.104770

15. Farquhar S, Kossen J, Kuhn L, Gal Y. Detecting hallucinations in large language models using semantic entropy. *Nature*. (2024) 630:625–30. doi: 10.1038/s41586-024-07421-0

16. Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, et al. ChatGPT and other large language models are double-edged swords. *Radiology*. (2023) 307:e230163. doi: 10.1148/radiol.230163

17. Alberts IL, Mercolli L, Pyka T, Prenosil G, Shi K, Rominger A, et al. Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be? *Eur J Nucl Med Mol Imaging*. (2023) 50:1549–52. doi: 10.1007/s00259-023-06172-w

18. Tian S, Jin Q, Yeganova L, Lai P-T, Zhu Q, Chen X, et al. Opportunities and challenges for ChatGPT and large language models in biomedicine and health. *Brief Bioinform*. (2024) 25:bbad493. doi: 10.1093/bib/bbad493

19. Normile D. (2025) Chinese firm's faster, cheaper AI language model makes a splash. Available online at: https://www.science.org/content/article/chinese-firm-s-faster-cheaper-ai-language-model-makes-splash (Accessed March 27, 2025)

20. Peng Y, Malin BA, Rousseau JF, Wang Y, Xu Z, Xu X, et al. From GPT to DeepSeek: significant gaps remain in realizing AI in healthcare. *J Biomed Inform*. (2025) 163:104791. doi: 10.1016/j.jbi.2025.104791

21. Temsah A, Alhasan K, Altamimi I, Jamal A, Al-Eyadhy A, Malki KH, et al. DeepSeek in healthcare: revealing opportunities and steering challenges of a new open-source artificial intelligence frontier. *Cureus*. (2025) 17:e79221. doi: 10.7759/cureus.79221

22. Kuehn BM. More than one-third of US individuals use the internet to self-diagnose. *JAMA*. (2013) 309:756–7. doi: 10.1001/jama.2013.629

23. Giannuzzi F, Carlà MM, Hu L, Cestrone V, Caputo CG, Sammarco MG, et al. Artificial intelligence with ChatGPT 4: a large language model in support of ocular oncology cases. *Int Ophthalmol*. (2025) 45:59–17. doi: 10.1007/s10792-024-03399-w

24. Moreno AC, Bitterman DS. Toward clinical-grade evaluation of large language models. *Int J Radiat Oncol Biol Phys*. (2024) 118:916–20. doi: 10.1016/j.ijrobp.2023.11.012

25. Agnihotri AP, Nagel ID, Artiaga JCM, Guevarra MCB, Sosuan GMN, Kalaw FGP. Large language models in ophthalmology: a review of publications from top ophthalmology journals. *Ophthalmol Sci*. (2024) 5:100681. doi: 10.1016/j.xops.2024.100681

26. Cappellani F, Card KR, Shields CL, Pulido JS, Haller JA. Reliability and accuracy of artificial intelligence ChatGPT in providing information on ophthalmic diseases and management to patients. *Eye*. (2024) 38:1368–73. doi: 10.1038/s41433-023-02906-0

27. Mishra V, Lurie Y, Mark S. Accuracy of LLMs in medical education: evidence from a concordance test with medical teacher. *BMC Med Educ*. (2025) 25:443–8. doi: 10.1186/s12909-025-07009-w

28. Wu X, Cai G, Guo B, Ma L, Shao S, Yu J, et al. A multi-dimensional performance evaluation of large language models in dental implantology: comparison of ChatGPT, DeepSeek, Grok, Gemini and Qwen across diverse clinical scenarios. *BMC Oral Health*. (2025) 25:1272. doi: 10.1186/s12903-025-06619-6

29. Pan Y, Tian S, Guo J, Cai H, Wan J, Fang C. Clinical feasibility of AI doctors: evaluating the replacement potential of large language models in outpatient settings for central nervous system tumors. *Int J Med Inform*. (2025) 203:106013. doi: 10.1016/j.ijmedinf.2025.106013

30. Matin RN, Linos E, Rajan N. Leveraging large language models in dermatology. *Br J Dermatol*. (2023) 189:253–4. doi: 10.1093/bjd/ljad230

31. ElHawary H, Gorgy A, Janis JE. Large language models in academic plastic surgery: the way forward. *Plast Reconstr Surg Glob Open*. (2023) 11:e4949. doi: 10.1097/GOX.0000000000004949

32. Waisberg E, Ong J, Zaman N, Kamran SA, Sarker P, Tavakkoli A, et al. GPT-4 for triaging ophthalmic symptoms. *Eye*. (2023) 37:3874–5. doi: 10.1038/s41433-023-02595-9

33. Zandi R, Fahey JD, Drakopoulos M, Bryan JM, Dong S, Bryar PJ, et al. Exploring diagnostic precision and triage proficiency: a comparative study of GPT-4 and bard in addressing common ophthalmic complaints. *Bioengineering*. (2024) 11:120. doi: 10.3390/bioengineering11020120

34. Betzler BK, Chen H, Cheng CY, Lee CS, Ning G, Song SJ, et al. Large language model and its impact in ophthalmology. *Lancet Digit Health*. (2023) 5:e917–24. doi: 10.1016/S2589-7500(23)00201-7

35. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. *JAMA*. (2023) 330:866–9. doi: 10.1001/jama.2023.14217

36. Cai W. Feasibility and Prospect of privacy-preserving large language models in radiology. *Radiology*. (2023) 309:e232335. doi: 10.1148/radiol.232335

37. Yi C, Niu G, Zhang Y, Rao J, Liu G, Yang W, et al. Advances in artificial intelligence in thyroid-associated ophthalmopathy. *Front Endocrinol*. (2024) 15:1356055. doi: 10.3389/fendo.2024.1356055

38. Lee D, Palmer E. Prompt engineering in higher education: a systematic review to help inform curricula. *Int J Educ Technol High Educ*. (2025) 22:1–22. doi: 10.1186/s41239-025-00503-7

39. Patel D, Raut G, Zimlichman E, Cheetirala SN, Nadkarni GN, Glicksberg BS, et al. Evaluating prompt engineering on GPT-3.5's performance in USMLE-style medical calculations and clinical scenarios generated by GPT-4. *Sci Rep*. (2024) 14:17341–10. doi: 10.1038/s41598-024-66933-x

40. Ranade N, Saravia M, Johri A. Using rhetorical strategies to design prompts: a human-in-the-loop approach to make AI useful. *AI Soc*. (2025) 40:711–32. doi: 10.1007/s00146-024-01905-3

41. Akkan H, Seyyar GK. Improving readability in AI-generated medical information on fragility fractures: the role of prompt wording on ChatGPT's responses. *Osteoporos Int*. (2025) 36:403–10. doi: 10.1007/s00198-024-07358-0

42. Campbell DJ, Estephan LE, Sina EM, Mastrolonardo EV, Alapati R, Amin DR, et al. Evaluating ChatGPT responses on thyroid nodules for patient education. *Thyroid*. (2024) 34:371–7. doi: 10.1089/thy.2023.0491

43. Luo MJ, Pang J, Bi S, Lai Y, Zhao J, Shang Y, et al. Development and evaluation of a retrieval-augmented large language model framework for ophthalmology. *JAMA Ophthalmol*. (2024) 142:798–805. doi: 10.1001/jamaophthalmol.2024.2513

44. Sadeq MA, Ghorab RMF, Ashry MH, Abozaid AM, Banihani HA, Salem M, et al. AI chatbots show promise but limitations on UK medical exam questions: a comparative performance study. *Sci Rep*. (2024) 14:18859–11. doi: 10.1038/s41598-024-68996-2

45. Li Z, Wang Z, Xiu L, Zhang P, Wang W, Wang Y, et al. Large language model-based multimodal system for detecting and grading ocular surface diseases from smartphone images. *Front Cell Dev Biol*. (2025) 13:1600202. doi: 10.3389/fcell.2025.1600202

46. Chen S, Lai K h, Guo X, Zhang X. The influence of digital health technology on the allocation of regional medical resources in China. *Health Policy Technol*. (2025) 14:101013. doi: 10.1016/j.hlpt.2025.101013

*CORRESPONDENCE
Lihua Wang
✉ 2721322182@qq.com
Jun He
✉ 649977437@qq.com

Content accuracy and reliability of pulmonary nodule information on social media platforms: a cross-platform study of YouTube, Bilibili, and TikTok

Yunanji Zhou[1], Xiang Zeng[2], Ting Yuan[2], Qian Wang[1], Siqi Wu[3], Lei Du[2], Lihua Wang[3]* and Jun He[3]*

[1]Qi Huang Chinese Medicine Academy, Jiangxi University of Chinese Medicine, Nanchang, China, [2]School of Clinical Medicine, Jiangxi University of Chinese Medicine, Nanchang, China, [3]Pulmonary Disease Department, Affiliated Hospital of Jiangxi University of Traditional Chinese Medicine, Nanchang, China

Background: Pulmonary nodules (PNs) are often overlooked, potentially leading to health risks. Social media platforms are increasingly used for health information dissemination. This study evaluates the quality and engagement of PN-related videos on YouTube, Bilibili, and TikTok.

Methods: On March 1, 2025, we searched each platform using "pulmonary nodule" or its Chinese equivalent. After screening, 271 videos were analyzed. Video characteristics were documented, and quality was assessed using PEMAT, VIQI, GQS, and mDISCERN tools. Inter-rater reliability was high ($\kappa = 0.81$).

Results: The final sample included 98 (YouTube), 74 (Bilibili), and 99 (TikTok) videos. TikTok videos were the shortest (median 114 s) yet had the highest engagement. Nonprofit organizations dominated YouTube uploads; physicians were most common on Bilibili and TikTok. Treatment was the most covered topic. YouTube scored highest in comprehensibility and actionability (PEMAT-T/A), while Bilibili and TikTok scored higher in production quality (VIQI, GQS). Video quality did not differ significantly between professional and non-professional uploaders. Most quality metrics showed weak correlation with audience engagement.

Conclusion: Long-form platforms (YouTube and Bilibili) offer higher-quality PN information but lower engagement, whereas short-form platforms (TikTok) show high interaction but lower informational depth. Social media can play a supportive role in public PN education. We provide recommendations for creators, platforms, and viewers to improve the quality and reliability of medical content.

KEYWORDS

pulmonary nodule, public health, social media, online video, information quality

1 Introduction

A pulmonary nodule (PN) is a localized, rounded, high-density shadow observed on chest imaging (e.g., computed tomography [CT] or X-ray), typically measuring less than 3 cm in diameter with well-defined or hazy borders (1). Based on density, PNs are classified into three categories: solid nodules, part-solid nodules, and pure ground-glass nodules (2). The majority

of benign solid nodules are associated with tuberculosis (TB) or inflammatory conditions (3). Malignant solid nodules can be precursors to lung cancer. Malignancy should be suspected in solid nodules exhibiting features such as spiculation, lobulation, or irregular margins (4). Part-solid nodules, also termed mixed ground-glass nodules, contain both solid and ground-glass components and exhibit heterogeneous density on CT (2). These nodules have a relatively high malignancy probability, particularly when the solid component increases, which often indicates progression (1). This often signifies early-stage lung cancer, necessitating prompt management and close monitoring (5). After follow-up, most part-solid nodules may resolve spontaneously or remain stable. However, persistent or growing nodules with increasing density carry a high malignancy risk (6).

Clinically, identifying individuals with lung nodules is challenging due to the frequent absence of overt symptoms. The detection rate of asymptomatic lung nodules has increased dramatically in China, largely due to the widespread adoption of low-dose CT screening, which offers substantial potential for the early diagnosis of lung cancer (7). Early and accurate diagnosis of lung nodule type is crucial for improving patient prognosis and reducing lung cancer mortality, as lung cancer remains a leading cause of cancer-related death worldwide (8). Consequently, improving public understanding of the various types of lung nodules and the distinctions between benign and malignant forms is critical.

Social media platforms have become increasingly popular for sharing and exchanging knowledge. YouTube is the dominant video platform in international markets (9), whereas TikTok and Bilibili are the most popular in China (10, 11). Video-based learning about current events and knowledge has gained popularity, and videos play a significant role in health communication and intervention. However, video quality varies considerably due to diverse content sources and differing levels of platform content regulation. A recent report indicated that the top 100 COVID-19 videos on TikTok garnered over 1.19 billion views, yet the most frequently discussed topics were not the most relevant to public health information (12). This suggests that videos under pertinent keywords may disseminate undesirable emotions and often lack scientific or expert medical knowledge. Moreover, the motives behind physicians' sharing of health information, which can range from altruism to financial incentives, significantly impact the quality and reliability of content (13). Understanding these dynamics is crucial for framing effective platform policies and incentive structures. This study aims to assess and analyze PN-related videos on YouTube, Bilibili, and TikTok with the goals of raising public awareness and providing actionable recommendations to platforms for enhancing the reliability of medical content.

Clinically, it is challenging to identify people with lung nodules because they typically do not exhibit any overt symptoms. The detection rate of asymptomatic lung nodules has dramatically grown in recent years due to China's extensive use of low-dose CT screening, which offers a substantial potential for the early diagnosis of lung cancer (7). Early and accurate diagnosis of the kind of lung nodules is crucial for improving patient prognosis and lowering lung cancer mortality, as lung cancer is one of the world's leading causes of cancer-related fatalities (8). As a result, it is critical that the general public understands the many

kinds of lung nodules and the differences between benign and malignant ones.

2 Method

2.1 Search strategy

Before collecting video feature information, we searched for "pulmonary nodule," "lung nodule," "solitary pulmonary nodule," "ground glass nodule," and "coin lesion" on three platforms and found that all search results were similar to those obtained using only "pulmonary nodule." Furthermore, using "pulmonary nodule" yielded the highest number of search results. On March 1, 2025, we searched on YouTube using the English keyword "Pulmonary Nodule," and on Bilibili and TikTok using the term "肺结节" ("Pulmonary Nodule" in Chinese).

To minimize the bias of the platforms' algorithmic recommendations, we created new accounts and deleted browser caches and search histories. The included videos were required to be posted by February 15, 2025. The top 100 videos for each platform were retrieved without the use of filtering criteria, and the results are shown in "default order" across all platforms. Due to the regional applicability of the three video platforms, we found that almost all non-English/Chinese videos were reposted (not original) from existing publishers. Videos featuring commercials, videos in languages other than English and Chinese, videos regarding infants and young children, duplicates, and unnecessary videos were all disqualified (Supplementary file 1). One individual gathered and downloaded all of the videos; two investigators classified the categories of videos and the uploaders. If discrepancies arose, the authors collectively reviewed the different videos to determine whether to retain them or not.

2.2 Video content

A number of video attributes, including length, duration, views, thumbs up, comments, collections, shares, coin-operated, upload time, upload source, author, and video attributes, were methodically recorded by two investigators on the same day (Supplementary file 1). However, for (1) TikTok views and (2) YouTube collections and shares, the following data were not accessible. The uploader's ID, number of followers, authentication status, and kind are among the many details gathered. Using preset criteria, professionals and credentials were located (Supplementary file 1). Videos in rough form, translations, or direct copies are not considered original. The many forms of video filming include solo narration, Q&A, PPT or class, animation/action, and medical scenarios.

2.3 Video review and classification

Two researchers looked over the listed videos independently between March 3 and 5, 2025, removing a large number of videos that were identical or unrelated (Supplementary file 1). The topics

---

Abbreviations: GQS, Global Quality Score; mDISCERN, Modified DISCERN; PEMAT, Patient Education Materials Assessment Tool; PN, Pulmonary Nodule; TCM, Traditional Chinese Medicine; TB, Tuberculosis; VIQI, Video Information and Quality Index.

covered in the videos were etiology/prevention, anatomy, pathology, epidemiology, symptoms, exams, diagnosis, and prognosis. The number of subjects each video addressed was counted since several movies were relevant to the broad spectrum of issues. Videos that did not cover these topics were judged superfluous and were taken off.

2.4 Video quality assessment

Video characteristics were systematically recorded. Two independent reviewers classified uploader type and video content. Between March 6 and 9, 2025, two respiratory disease specialists performed a double-blind evaluation of video quality using four validated tools: PEMAT, VIQI, GQS, and mDISCERN. The inter-rater reliability was excellent (Cohen's κ = 0.81). Disagreements were resolved through discussion or by a third senior expert. All three respiratory disease specialists have over 10 years of clinical experience and extensive involvement in health education (e.g., publishing articles, recording educational videos, and providing educational services in schools and communities).

The research team used Cohen's kappa coefficient (κ) for reliability analysis to measure the degree of inter-assessor agreement. Its determination criterion was based on international norms: $\kappa > 0.8$ indicates excellent agreement, 0.6–0.8 indicates good agreement, 0.4–0.6 indicates moderate agreement, and ≤ 0.4 indicates insufficient agreement (14).

The Modified DISCERN Scale (mDISCERN) was used to evaluate the information's credibility (15), the Global Quality Scoring System (GQS) concentrates on the overall production standards (16), the Video Information Quality Index (VIQI) examines the content's quality from a multidimensional perspective (16), and the Patient Education Materials Assessment Tool (PEMAT) focuses on the comprehensibility and actionability of health education materials (17) (Supplementary file 2). It is important to note that these evaluation instruments have demonstrated their reliability in a number of worldwide studies and have a wide range of uses, particularly in the area of social media health information assessment (15–18).

2.5 Statistical analysis

The normality of continuous variables was assessed using the Shapiro–Wilk test. Due to the non-normal distribution of the data, descriptive statistics are presented as median (min-max) and interquartile range [IQR] (P25, P75). Group comparisons for non-normally distributed continuous variables were performed using the Mann–Whitney U test. Categorical variables were compared using Chi-square tests, with continuity correction or Fisher's exact test applied as appropriate. Correlations between quantitative variables were assessed using Spearman's rank correlation coefficient. A two-sided p-value of < 0.05 was considered statistically significant. For Spearman's correlation, a positive association was indicated by $r > 0$, and a negative association by $r < 0$. The strength of the link was classified. All statistical analyses were performed using R software (version 4.4.2; R Foundation for Statistical Computing).

## 3 Results

3.1 Video characteristics

After removing duplicates and irrelevant content, the final dataset comprised 98 YouTube videos, 74 Bilibili videos, and 99 TikTok videos (Figure 1). All videos across these platforms were either in Chinese or English or provided bilingual subtitles. The Shapiro–Wilk test revealed that all continuous variables exhibited non-normal distributions. Table 1 and Figure 2 provide specific details about the videos from each platform. Notably, TikTok videos (114 [19–1,321] seconds) were significantly shorter than those on YouTube (394.5 [26–3,743] seconds) and Bilibili (273.5 [54–3,983] seconds). TikTok also demonstrated more frequent updates based on release dates. TikTok had the most likes and comments out of the three sites. YouTube displayed the least amount of interaction from users, whereas Bilibili had the most extensive interactive components.

3.2 Uploader profiles and certification status

This research video included 85 YouTube uploaders, 51 Bilibili uploaders, and 68 TikTok uploaders. The categories of uploaders on different platforms vary greatly (Table 2; Figure 3). Doctors account for 91.2% on TikTok, 76.5% on Bilibili, and only 16.5% on YouTube. Non-profit organizations accounted for 50.6% of uploaders on YouTube, but were absent on Bilibili and TikTok. Bilibili uploaders posted videos more regularly than the less active YouTubers. The certification rates were 38.8% on YouTube, 82.4% on Bilibili, and 98.5% on TikTok. We examined certified uploaders to determine the video's authority. Only a small percentage of YouTube-certified doctors were included in this study, despite having the highest subscribers. Furthermore, 15 acknowledged Traditional Chinese Medicine (TCM) physicians posted to TikTok, and 21 recognized TCM physicians uploaded to Bilibili. Among the videos, one titled 'Heart Art' received a high number of likes.

3.3 Content categories and presentation styles

Table 3 and Figure 4 describe the video categories. TikTok (100%) and YouTube (100%) contained more raw materials than Bilibili (97.3%). The videos on Bilibili cover a wider range of topics than those on YouTube and TikTok, indicating that there are differences in the diversity of video topics across different platforms. Interestingly, after a thorough review, we found that the number of topics is correlated with longer video lengths. The most viewed videos on all three platforms were those with therapeutic content. The difference is that TikTok and Bilibili were more interested in etiology/prevention and pathology, while the most popular topic on YouTube was examination/diagnosis. Treatment options for PNs were covered in a number of videos, all of which emphasized the significance of keeping a positive psychological attitude after being detected. According to several video presentations, the majority of PNs are benign, and the most commonly advised therapy strategy is routine follow-up exams. Certain medical experts added clinical case demonstrations using chest CT scans to

FIGURE 1
Search strategy for videos on pulmonary nodule.
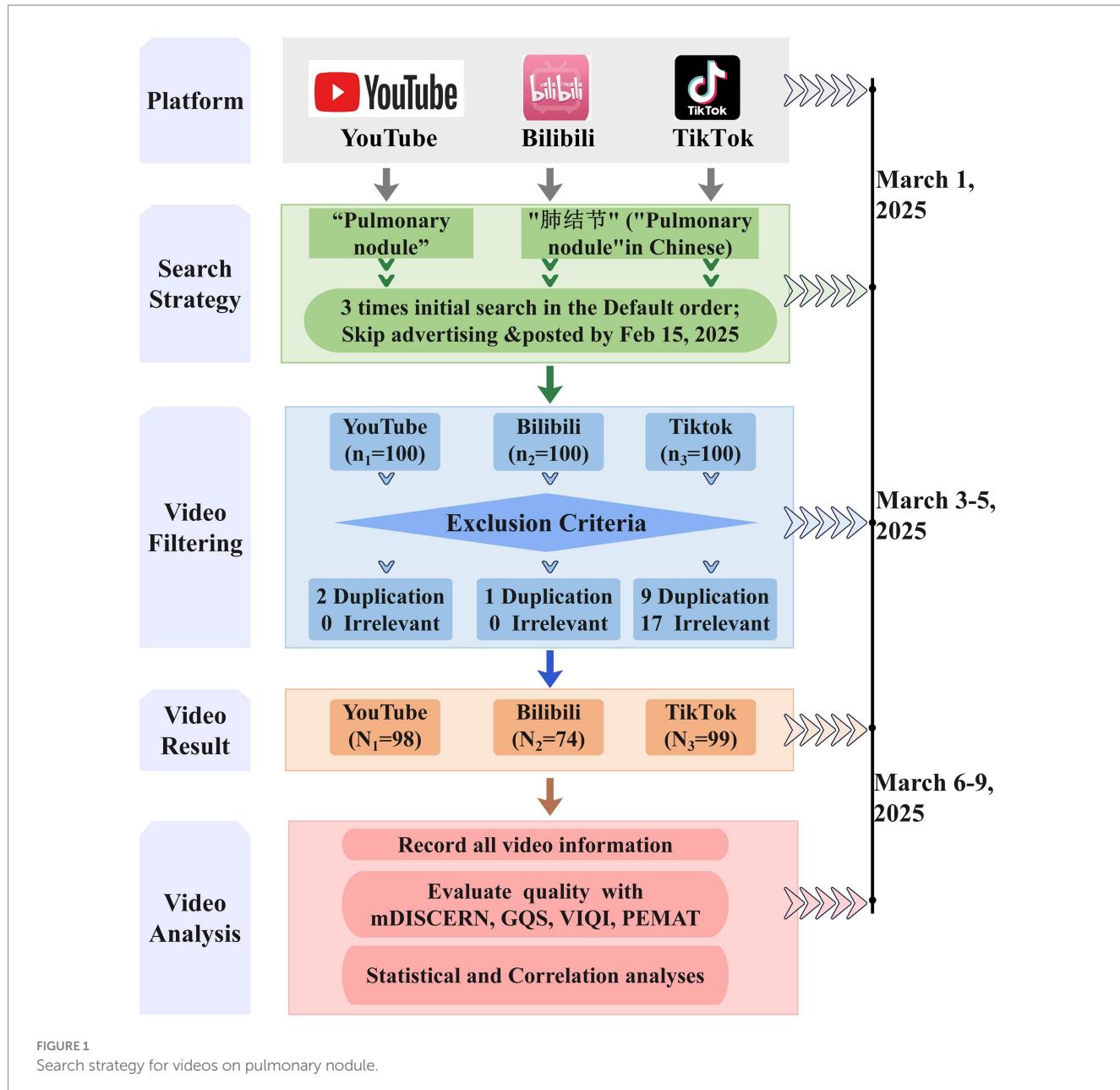
their videos. Understanding anatomy and etiology has helped the public recognize the need for prevention. Platform-to-platform variations in content presentation patterns were substantial (Table 3; Figure 5). Solo narration predominated on TikTok and Bilibili, while PPT or class presentations were most prevalent on YouTube. In addition, TikTok is skilled at using medical scenarios.

3.4 Quality assessment of videos

Inter-rater reliability was assessed using Cohen's kappa, which indicated substantial agreement ($\kappa = 0.81$). The results of the video quality evaluation are presented in Table 4; Figure 6. YouTube demonstrated significantly higher PEMAT-T (comprehensibility) and PEMAT-A (operability) scores compared with Bilibili and TikTok ($p < 0.001$). In contrast, PEMAT-U (usability) ratings followed an

inverse pattern. TikTok and Bilibili showed substantially higher VIQI and GQS ratings than YouTube ($p < 0.001$). No significant differences were observed in mDISCERN total scores (information reliability) across the three platforms. As shown in Table 5, no significant differences were found in PEMAT, VIQI, GQS, or mDISCERN scores between videos from professional and non-professional uploaders ($p > 0.05$).

3.5 Correlation between video quality and audience engagement

Video quality and audience engagement did not significantly correlate (Table 6; Figure 7). Correlation studies showed that Bilibili's Views and VIQI had the highest index (0.701, $p < 0.001$), suggesting a strong connection with audience support actions.

TABLE 1 Characteristics of video about laryngeal carcinoma on YouTube/ Bilibili/ TikTokk.

Characteristic	YouTube (N₁ = 98)			Bilibili (N₂ = 74)			TikTok (N₃ = 99)			p-value		
	M	Min–Max	P25-P75	M	Min–Max	P25-P75	M	Min–Max	P25-P75	$P_{(Y-B)}$	$P_{(B-T)}$	$P_{(Y-T)}$
Video length(s)	394.5	26–3,743	117.75–1,177	273.5	54–3,983	134.5–466.5	114	19–1,321	70–205.5	0.115	**<0.001**	**<0.001**
Duration(day)	1457.5	169–5,621	755–2608.25	392.5	1–1,573	288.25–771	141	1–1,058	51–315.5	**<0.001**	**<0.001**	**<0.001**
Views	7,082	46–695,780	1,631–25153.25	13,500	193–357,000	2452.25–40,500	-	-	-	0.122	-	-
Thumbs up	63	0–5,803	7–252.5	174.5	5–15,000	55–787.75	11,000	29–450,000	2254.5–28,000	**<0.001**	**<0.001**	**<0.001**
Commentsª	2	0-674	0-25.5	22	0–1,073	5–104	633	1–17,000	144.5–1,642	**<0.001**	**<0.001**	**<0.001**
Collections	-	-	-	154	3–5,921	45–668.5	2,735	7–214,000	543.5–12,000	-	**<0.001**	-
Shares	-	-	-	79	3–3,361	18.5–456	2,285	2–188,000	396–7,037	-	**<0.001**	-
Coin-operated	-	-	-	26.5	0–1,319	7–125.75	-	-	-			

All the p-values were obtained from Mann–Whitney U test.

$P_{(Y-B)}$: YouTube versus Bilibili; $P_{(B-T)}$: Bilibili versus TikTok; $P_{(Y-T)}$: YouTube versus TikTok. Bold text means the p-value < 0.05.

ªExcluded: 16 videos on YouTube turned off the function of comments.
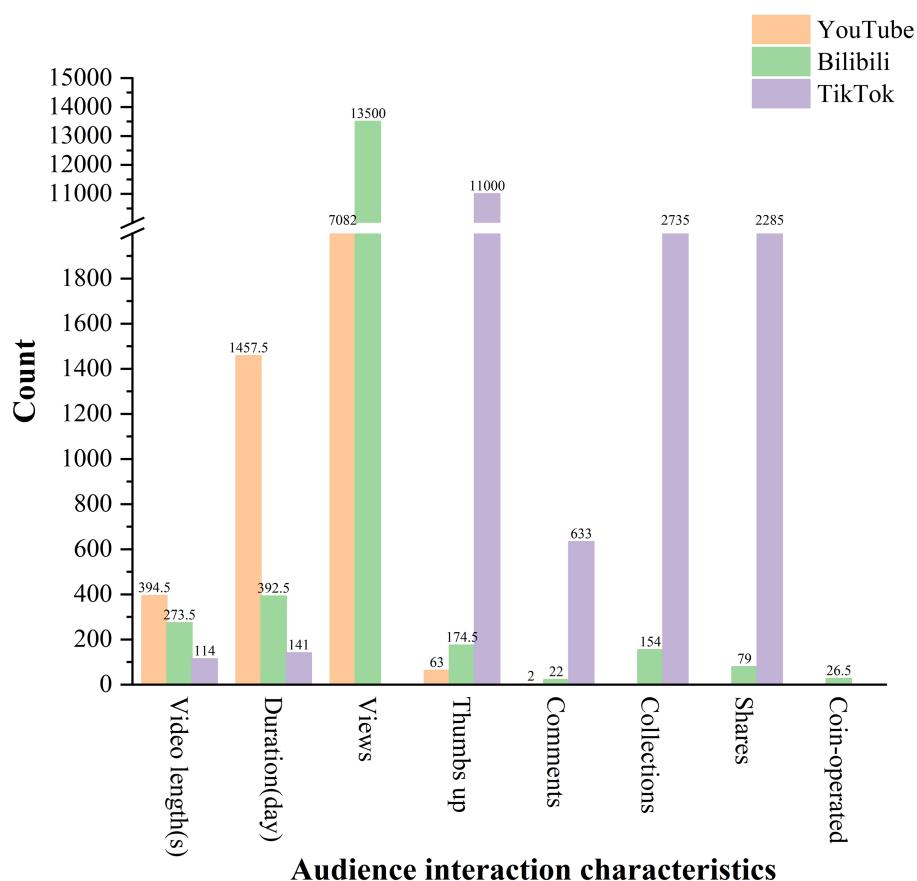
FIGURE 2
Characteristics of video viewer interaction on different platforms (mean).

TABLE 2 Characteristics of video uploaders about laryngeal carcinoma on YouTube/ Bilibili/ TikTok.

Platform	YouTube	Bilibili	TikTok	P(Y-B)	P(B-T)	P(Y-T)
Number of uploaders	85	51	68	-	-	-
Followers, median [P25, P75]	10,000 [1880, 52,600]	29,000 [3,605, 75,750]	269,000 [68,750, 1,048,750]	0.079[a]	**<0.001[a]**	**<0.001[a]**
Number of videos per person, Mean±SD, Median[P25, P75]	1.15 ± 0.42, 1 [1, 1]	1.48 ± 0.93, 1 [1, 2]	1.46 ± 1.20, 1 [1, 1]	0.021[a]	0.328[a]	0.215[a]
Type of uploaders, *n* (%)						
Doctor	14 (16.5%)	39 (76.5%)	62 (91.2%)	**<0.001[b]**	**0.027[b]**	**<0.001[b]**
Other medical worker/ student	19 (22.4%)	5 (9.8%)	0	0.063[b]	**0.030[c]**	**<0.001[b]**
Non-profit organization	43 (50.6%)	0	0	**<0.001[b]**	-	**<0.001[b]**
Company with profit	4 (4.7%)	0	0	0.294[c]	-	0.193[c]
Official media	3 (3.5%)	3 (5.9%)	5 (7.4%)	0.829[c]	1.000[c]	0.490[c]
Self-media	2 (2.3%)	4 (7.8%)	1 (1.4%)	0.281[c]	0.210[c]	1.000[c]
Doctor of TCM, *n* (%)	0	21 (41.2%)	15 (22.1%)	**<0.001[b]**	**0.025[b]**	**<0.001[b]**
Authentication, *n* (%)	33 (38.8%)	42 (82.4%)	67 (98.5%)	**<0.001[b]**	**0.005[c]**	**<0.001[b]**
Verify the type of uploader, *n* (%)						
Doctor	2 (6.1%)	34 (81.0%)	62 (92.5%)	**<0.001[b]**	0.069[b]	**<0.001[b]**
Other medical worker/ student	3 (9.1%)	4 (9.5%)	0	1.000[c]	**0.040[c]**	0.034[d]
Non-profit organization	24 (72.7%)	0	0	**<0.001[b]**	-	**<0.001[b]**
Company with profit	1 (3.0%)	0	0	0.440[d]	-	0.330[d]
Official media	3 (9.1%)	3 (7.1%)	5 (7.5%)	1.000[c]	1.000[c]	1.000[c]
Self-media	0	1 (2.4%)	0	1.000[d]	0.385[d]	-

P(Y-B): YouTube versus Bilibili; P(B-T): Bilibili versus TikTok; P(Y-T): YouTube versus TikTok. Bold text means the *p*-value < 0.05.
[a]Mann-Whitney U test.
[b]Chi-squared test.
[c]Continuity correction.
[d]Fisher's exact test.

Approximately, most of the quality assessment results revealed a mild to moderate positive connection with audience engagement. In contrast to YouTube, Bilibili and TikTok showed varied degrees of negative association between some quality and engagement metrics. PEMAT scores showed a negative link with all audience interaction measures in Bilibili, with the exception of a positive correlation with Views. In Bilibili, on the other hand, VIQI had a negative correlation with Coin-operated and Comments. Surprisingly though, GQS was negatively connected with Comments and Collections on TikTok, while Views and mDISCERN displayed the most negative link in Bilibili ($r = -0.483$, $p < 0.001$).

4 Discussion

To identify statistically significant interplatform differences and elucidate variations in audience demographics, we analyzed both long-form (e.g., YouTube, Bilibili) and short-form (e.g., TikTok) video platforms. We employed a comprehensive propensity score matching analysis to delineate the distinctive features of YouTube, Bilibili, and TikTok. Using four distinct scoring methodologies, we derived the following detailed conclusions. This analysis yielded four principal findings: First, significant disparities in content distribution and quality were observed across platforms, with YouTube generally hosting higher-quality productions. Second, uploader identity and platform certification mechanisms significantly influence perceived content reliability. Third, video formats and thematic preferences exhibit substantial regional and cultural variations. Finally, we found no strong correlation between audience interaction metrics and objective video quality measures. Our findings underscore that algorithmic curation, professional certification, cultural context, and user engagement collectively shape video quality and dissemination efficacy. Specifically, platform architecture dictates content format, algorithmic filtering governs content dissemination, professional involvement enhances information quality, while user engagement often diverges from clinical value. A comprehensive summary of these key findings is shown in Table 7.

Social media, informational videos, and online platforms have become increasingly prominent sources of information. The prevalence of video platforms continues to grow, driven by a public preference for visually presented information over text-based or auditory formats. Notwithstanding the limitations in public health communication, video platforms can augment public health education by assisting medical personnel in improving prescreening efficacy.
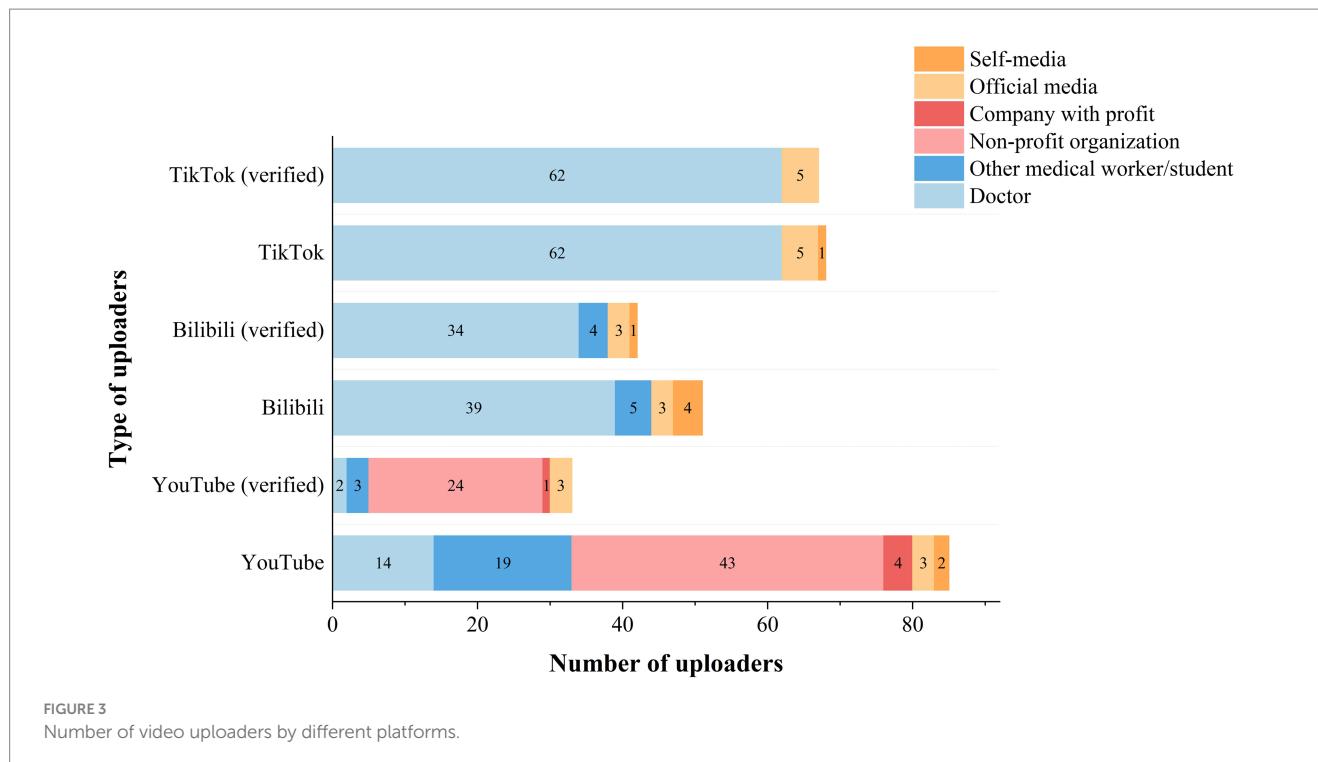
FIGURE 3
Number of video uploaders by different platforms.

TABLE 3 Categorization of videos about laryngeal carcinoma on YouTube/ Bilibili/ TikTok.

| Platform	YouTube (N$_1$ = 98)	Bilibili (N$_2$ = 74)	TikTok (N$_3$ = 99)	P$_{(Y-B)}$	P$_{(B-T)}$	P$_{(Y-T)}$
Originality, n (%)	98 (100%)	72 (97.3%)	99 (100%)	0.184[d]	-	0.182[d]
Number of topics per video, Median [P25, P75]	1.5 (1–2)	3 (2–4)	2 (2–3)	**<0.001[a]**	0.239[a]	**<0.001[a]**
Type of topics, n (%)						
Anatomy	8 (8.4%)	22 (29.7%)	21 (21.1%)	**<0.001[b]**	0.164[b]	**0.010[b]**
Etiology/Prevention	17 (17.3%)	41 (55.4%)	52 (52.5%)	**<0.001[b]**	0.567[b]	**<0.001[b]**
Pathology	7 (7.1%)	37 (50%)	55 (55.6%)	**<0.001[b]**	0.589[b]	**<0.001[b]**
Epidemiology	9 (9.2%)	8 (10.8%)	11 (11.1%)	0.723[b]	1.000[b]	0.654[b]
Symptoms	7 (7.1%)	20 (27.0%)	16 (16.1%)	**<0.001[b]**	0.066[b]	**0.049[b]**
Examinations/Diagnosis	81 (82.7%)	33 (44.6%)	30 (30.3%)	**<0.001[b]**	**0.038[b]**	**<0.001[b]**
Treatment	35 (35.7%)	51 (68.9%)	66 (66.7%)	**<0.001[b]**	0.563[b]	**<0.001[b]**
Prognosis	14 (14.3%)	10 (13.5%)	23 (23.2%)	0.885[b]	0.126[b]	0.108[b]
TCM, n (%)	0	30 (40.5%)	17 (17.1%)	**<0.001[b]**	**<0.001[b]**	**<0.001[b]**
Style of video shooting, n (%)						
Solo narration	29 (29.6%)	49 (66.2%)	74 (74.7%)	**<0.001[b]**	0.336[b]	**<0.001[b]**
Q & A	12 (12.3%)	4 (5.4%)	5 (5.1%)	0.140[b]	1.000[c]	0.012[b]
PPT or Class	40 (40.8%)	6 (8.1%)	2 (2.0%)	**<0.001[b]**	0.118[c]	**<0.001[b]**
Animation / Action	3 (3.1%)	1 (1.4%)	1 (1.0%)	0.842[c]	1.000[d]	0.403[c]
Medical scenarios	7 (7.1%)	10 (13.5%)	15 (15.2%)	0.147[b]	0.818[b]	0.295[b]
TV programs/ documentaries	7 (7.1%)	2 (2.7%)	1 (1.0%)	0.363[c]	0.780[c]	**0.022[c]**
Other	0	2 (2.7%)	1 (1.0%)	0.178[d]	0.780[c]	1.000[d]

P$_{(Y-B)}$: YouTube versus Bilibili; P$_{(B-T)}$: Bilibili versus TikTok; P$_{(Y-T)}$: YouTube versus TikTok. Bold text means the p-value < 0.05.
[a]Mann-Whitney U test.
[b]Chi-squared test.
[c]Continuity correction.
[d]Fisher's exact test.
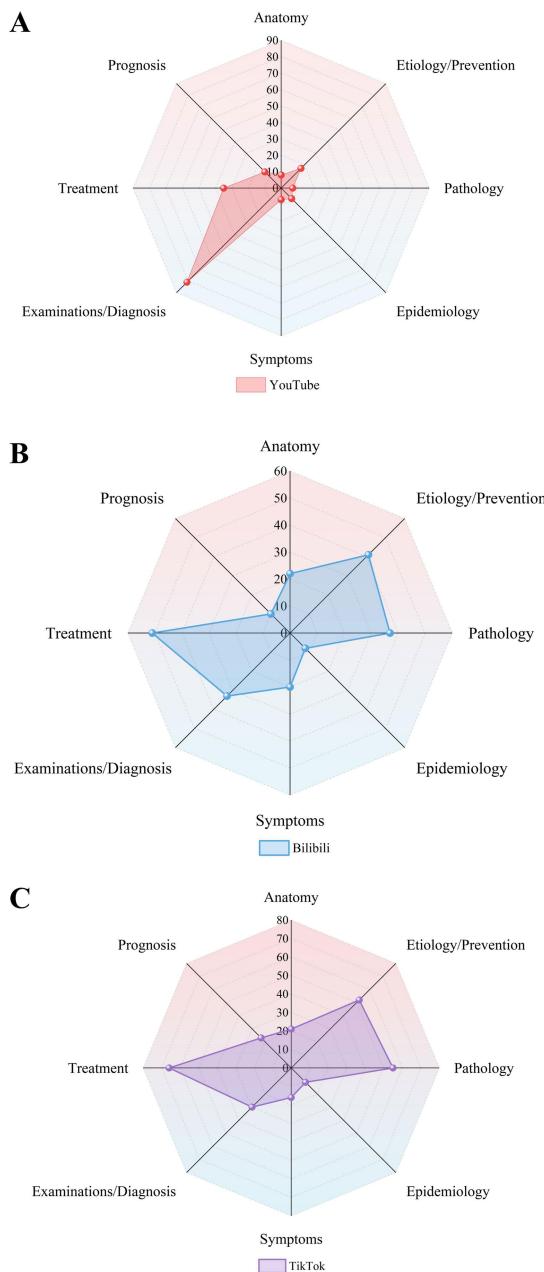
FIGURE 4
Types of topics for video content on different platforms. (**A** data source YouTube; **B** data source Bilibili; **C** data source TikTok).

Recent studies have investigated video content related to asthma (19), pulmonary rehabilitation (20), inspiratory muscle training (21), COPD (22, 23), allergies (24), and general respiratory health on platforms like YouTube and TikTok. However, these studies have consistently identified significant content quality issues, including inconsistent quality across platforms, the absence of standardized evaluation methodologies, and suboptimal presentation of medical information. Although these video tools show promise for health education initiatives, concerns persist among medical professionals regarding content authenticity and potential misinformation. To ensure reliability and educational value, greater involvement of

certified healthcare professionals in the creation and review of medical video content is imperative (25, 26).

Beyond the challenges associated with radiological detection, PNs pose a significant public health threat due to the neglect experienced by a substantial proportion of patients (27). Publications on PN-related video content remain scarce, with the exception of Han et al.'s (28) study. However, that study was limited to an analysis of the top 30 videos on a single Chinese social media platform and did not perform any correlation analysis on video content or quality metrics. Our study is situated within the context of the rapid proliferation of video platforms in contemporary society (29). We implemented a multi-platform (covering domestic and international sites), multi-perspective, and multi-dimensional quality assessment framework. Furthermore, correlation analyses were conducted to examine relationships between key variables.

4.1 Significant interplatform disparities in content distribution and quality

YouTube consistently demonstrated superior video quality, particularly in structure and clarity as measured by PEMAT, although its content predominantly focused on examination and diagnosis. In contrast, Bilibili and TikTok offered broader coverage of etiology, prevention, and treatment, but with greater quality variability and occasional inaccuracies. The short-form nature of TikTok content often resulted in superficial coverage of topics, and some videos exaggerated health anxieties to attract views, thereby compromising educational utility.

Video platforms utilize big data analytics to track user behavior patterns, adjusting their algorithms to influence video length and content relevance. Content relevance is fundamentally determined by the specific screening mechanisms employed by each platform. YouTube, which pioneered video sharing in 2005, processes billions of uploads and views daily, hosting diverse content ranging from disease information to disaster reports and personal narratives (9). However, its default "relevance" search function has significant shortcomings: medical content lacks standardized classification and quality assessment systems, search algorithms are suboptimally calibrated, and video review processes lack rigorous oversight (30). TikTok (10) and Bilibili (11) employ complex algorithms that integrate user profiles, search relevance, and interaction patterns to recommend content. However, effectively managing personalized multi-dimensional learning models and trending topics remains challenging. For instance, our analysis revealed that some videos devoid of practical health information gained disproportionate popularity in search results. These videos frequently depict patients experiencing heightened anxiety following chest CT scans that reveal lung nodules, often portraying accompanying emotional distress and depressive symptoms. This suggests that some creators prioritize viewership over all else, disregarding the potential utility of their content for health promotion. Currently popular topics (e.g., celebrities, variety shows, commerce) may be overshadowed by such content. This phenomenon is commonly reported in studies across various disease contexts (10, 31).

From a public health perspective, this phenomenon can influence public perceptions and health behaviors and may exacerbate health misconceptions or anxiety. For instance, although most lung nodules are benign, one-sided or exaggerated content can induce unnecessary panic. Conversely, insufficient dissemination of high-quality
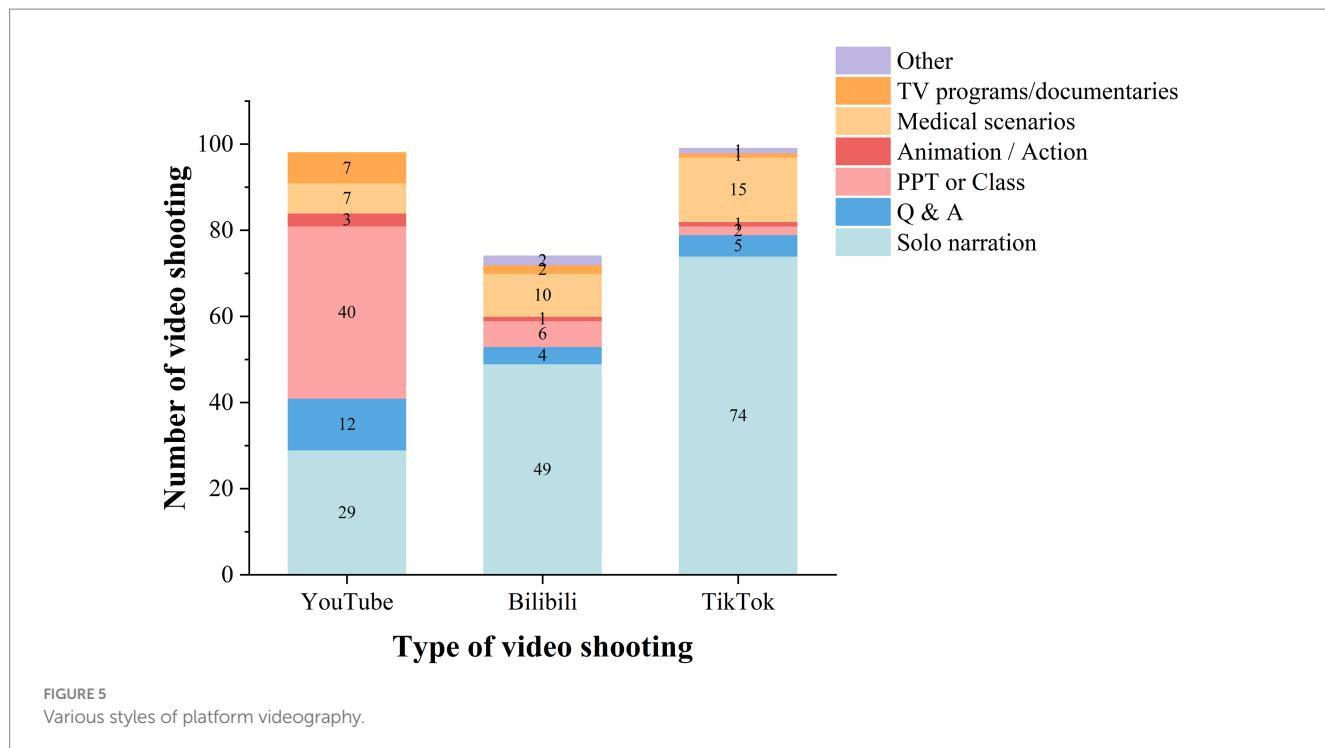
FIGURE 5
Various styles of platform videography.

TABLE 4 Quality assessment of videos about pulmonary nodule on YouTube/ Bilibili/ TikTok.

Platform characteristic	YouTube (N_1 = 98)			Bilibili (N_2 = 74)			TikTok (N_3 = 99)			*p*-value		
	M	Min–Max	P25-P75	M	Min–Max	P25-P75	M	Min–Max	P25-P75	P(Y-B)	P(B-T)	P(Y-T)
PEMAT-T	9.5	4–15	8–12	9	4–16	8–11	9	6–12	8–10	0.787[a]	0.071[a]	**0.033[a]**
PEMAT-U	3	2-4	3–3	4	3–4	4–4	4	4–4	4–4	**<0.001[a]**	**0.020[a]**	**<0.001[a]**
PEMAT-A	6	2-12	5–9	5	1–12	4.45–7	5	2–8	4–6	**0.008[a]**	**0.040[a]**	**<0.001[a]**
VIQI-sum	13	6–18	12–15	15	10–19	14–17	15	10–20	14–16	**<0.001[a]**	0.588[a]	**<0.001[a]**
VIQI-1	2	1–4	2–3	3	1–5	2–3	4	1–5	3–4	**0.001[a]**	**<0.001[a]**	**<0.001[a]**
VIQI-2	4	1–5	3.25–4	5	3–5	4–5	4	2–5	4–5	**<0.001[a]**	0.062[a]	**<0.001[a]**
VIQI-3	3	1–5	2–4	3	1–5	2–4	2	2–5	2–3	0.352[a]	**<0.001[a]**	**0.003[a]**
VIQI-4	4	2–5	3–4	5	2–5	5–5	5	3–5	5–5	**<0.001[a]**	0.905[a]	**<0.001[a]**
GQS	3	1–5	3–4	5	2–5	4–5	5	3–5	4–5	**<0.001[a]**	0.740[a]	**<0.001[a]**
mDISCERN-sum	3	1–5	3–3	3	3–5	3–3	3	3–4	3–3	0.357[a]	**0.042[a]**	0.647[a]
mDISCERN-1	96 (96.7%)			74 (100%)			99 (100%)			0.507[d]	-	0.246[d]
mDISCERN-2	97 (97.8%)			74 (100%)			99 (100%)			1.000[d]	-	0.497[d]
mDISCERN-3	87 (87.9%)			74 (100%)			97 (98.0%)			**0.008[c]**	0.508[d]	**0.009[b]**
mDISCERN-4	20 (20.2%)			15 (20.3%)			7 (7.1%)			0.982[b]	**0.010[b]**	**0.006[b]**
mDISCERN-5	15 (15.1%)			1 (1.4%)			4 (4.0%)			**0.002[b]**	0.558[c]	**0.007[b]**

P(Y-B): YouTube versus Bilibili; P(B-T): Bilibili versus TikTok; P(Y-T): YouTube versus TikTok. Bold text means the *p*-value < 0.05.
[a]Mann-Whitney U test.
[b]Chi-squared test.
[c]Continuity correction.
[d]Fisher's exact test.

information may delay early screening among high-risk groups. According to current clinical guidelines, the probability of malignancy is less than 1% for nodules smaller than 6 mm in diameter, and 1–2% for those measuring 6–8 mm (3). For nodules measuring 6–8 mm, follow-up chest CT scans every 6–12 months are recommended (3).

The specific follow-up interval should be individualized based on a comprehensive assessment, including the patient's risk factors, imaging features suggestive of malignancy, the clinician's judgment, and patient preferences. For solid nodules larger than 8 mm, management strategies depend on three key factors: (1) the predicted probability of malignancy,

FIGURE 6

Quality assessment of videos about pulmonary nodule on YouTube/ Bilibili/ TikTok. (**A** for VIQI-sum score; **B** for mDISCERN-sum score; **C** for GQS score; **D** for PEMAT-sum score; red for YouTube; blue for Bilibili; purple for TikTok).
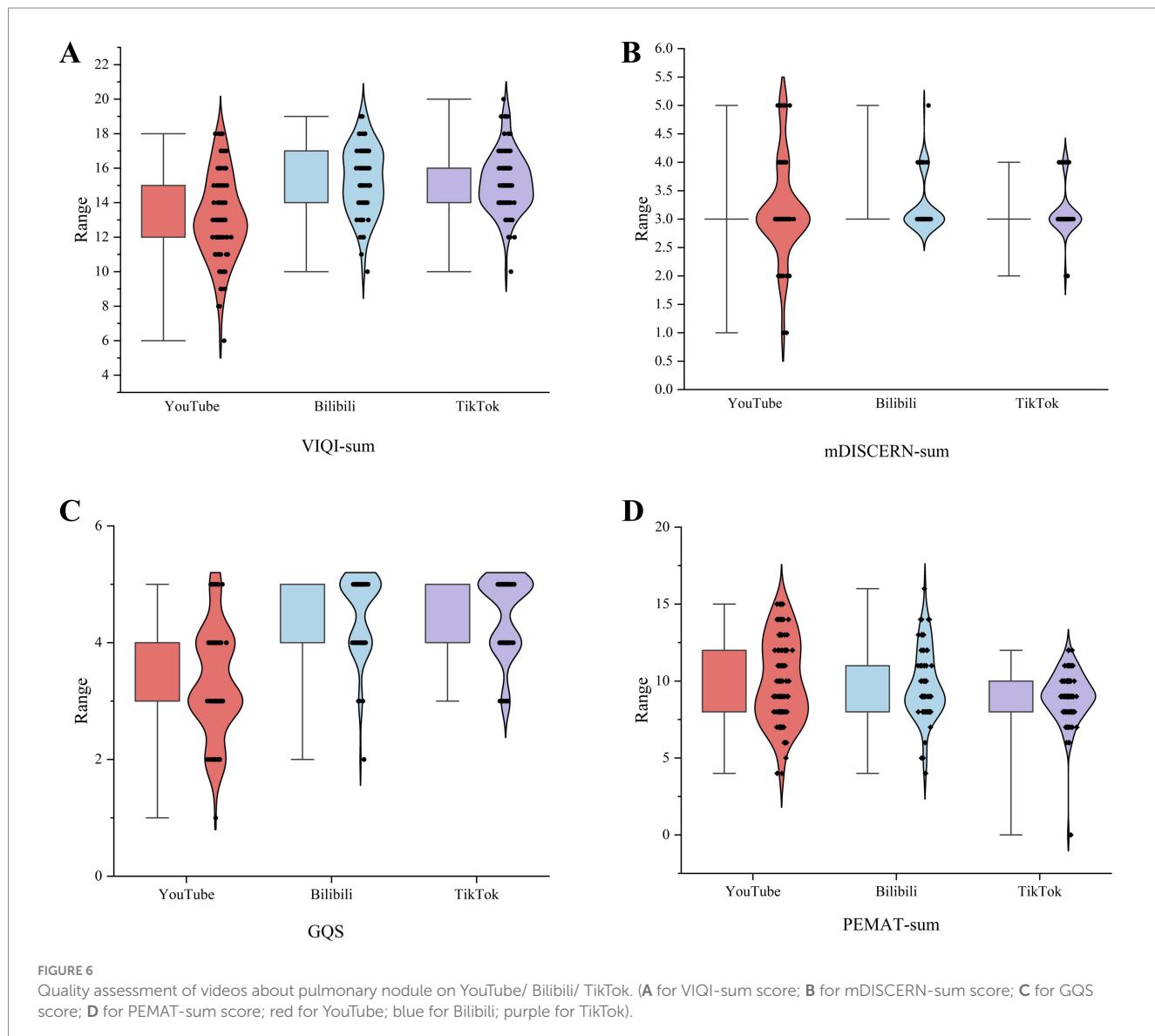
(2) the presence of comorbidities (e.g., coronary heart disease or COPD), and (3) patient preferences. Current management options include surgical resection, non-surgical biopsy (e.g., bronchoscopy or transthoracic needle biopsy), positron emission tomography–computed tomography (PET-CT), or active surveillance with serial CT scans to monitor for growth (1, 4). For part-solid nodules, the size of the solid component directly informs management strategy, with larger solid components indicating a higher malignancy risk. Persistent pure ground-glass nodules exceeding 10 mm in diameter have a 10–50% probability of malignant transformation (32). Notably, pure ground-glass nodules typically exhibit slow growth even if malignant. The sensitivity of current diagnostic techniques (e.g., transthoracic biopsy or bronchoscopy) for lung cancer ranges from 70 to 90% (24, 32).

4.2 Uploader identity and authentication mechanisms influence content reliability

Although certification methods vary across platforms, independently verifying uploader credentials can help prevent the dissemination of low-quality medical information. Multiple studies indicate that TikTok enforces strict qualification reviews, permitting only chief physicians, associate chief physicians, and registered residents from tertiary Grade A hospitals to use the professional title "doctor" (17, 33). The platform's algorithm prioritizes videos with high like counts, which can narrow the gap between popularity and perceived quality (34). Consequently, even with stringent certification standards, the correlation between a video's popularity and its actual quality is not guaranteed. In contrast, Bilibili employs more lenient review standards, allowing grassroots doctors, medical students, and even general users to upload content. This approach diversifies perspectives but also introduces risks to information reliability (25). The public increasingly relies on social media to access health information, encompassing topics from disease surveillance and health education to behavioral change, professional development, and doctor-patient communication (35). Thus, we recommend that all healthcare professionals actively pursue platform certification. Certification can enhance public trust through greater transparency regarding medical expertise. Additionally, certified users often receive greater platform support and visibility. Recent studies on online

TABLE 5 Quality comparison between the videos uploaded by professionals and non-professionals.

Scores	Professionals (N₁ = 263)			Non-professionals (N₂ = 8)			p-value
	M	Min−Max	P25-P75	M	Min−Max	P25-P75	
PEMAT-T	9	4–16	8–11	9	5–13	6.5–10	0.317[a]
PEMAT-A	4	2–4	3–4	4	3–4	3.75–4	0.290 [a]
PEMAT-U	5	1–12	4.5–7	5.5	1–9	2.75–6	0.618[a]
VIQI-1	3	1–5	2–4	3	1–4	1.75–3.25	0.716[a]
VIQI-2	4	1–5	4–5	4	2–5	3.75–5	0.539[a]
VIQI-3	3	1–5	2–4	3.5	1–5	2–4.25	0.453[a]
VIQI-4	5	3–5	4–5	4	2–5	3.5–5	0.155[a]
VIQI-sum	15	8–20	13–16	14	6–19	11.75–17	0.762[a]
GQS	4	1–5	3–5	3.5	2–5	2.75–5	0.310 [a]
mDISCERN-sum	3	1–5	3–3	3	2–4	2.75–3	0.170[a]
mDISCERN-1	260 (98.9%)			8 (100.0%)			1.000[d]
mDISCERN-2	260 (98.9%)			8 (100.0%)			1.000[d]
mDISCERN-3	248 (94.3%)			6 (75.0%)			0.083[d]
mDISCERN-4	41 (15.6%)			1 (12.5%)			1.000[d]
mDISCERN-5	20 (7.6%)			0			1.000[d]

Bold text means the p-value < 0.05.
[a]Mann-Whitney U test.
[d]Fisher's exact test.

medical consultation platforms further emphasize the importance of proper incentive structures and team-based collaborations among physicians, which not only improve service quality but also enhance patient satisfaction and trust (36, 37). These findings suggest that platforms could benefit from designing mechanisms that promote collaborative content creation and knowledge sharing among certified professionals. TCM has recently garnered significant attention in disease diagnosis and treatment (38). However, the proportion of TCM practitioners specializing in PNs remains low. Our analysis reveals limited TCM-focused content on PNs within TikTok and Bilibili. Despite this limited presence, some content creators are exploring TCM's role in managing PNs and innovatively incorporating it, sometimes alongside other diseases, into their diagnostic and therapeutic discussions (10). Therefore, engaging more professional TCM practitioners is crucial to enhancing the quality and representativeness of TCM-related content on these platforms.

Notably, no significant difference was observed in overall video quality between professional and non-professional creators (all $p > 0.05$; see Table 5). This finding challenges the conventional reliance on uploader credentials as a sole proxy for content reliability, demonstrating that professional qualifications do not automatically translate into effective public communication. Similarly, mDISCERN scores showed no significant inter-platform differences, indicating that major platforms share common challenges in conveying medical information reliability. Thus, platforms should implement risk warnings for potentially misleading content (e.g., flagging "controversial viewpoints") and bolster their accountability. Concurrently, medical schools have incorporated health communication courses to foster interdisciplinary "medicine + communication" skills (39). They also

encourage physicians to collaborate with media professionals to translate specialized knowledge into engaging formats, as exemplified by initiatives like the "Medical Science Popularization Alliance" at United Hospital. In conclusion, this study highlights three key recommendations: (1) the public should cultivate critical thinking and diversify information sources; (2) platforms must optimize review mechanisms and balance traffic incentives with social responsibility; and (3) healthcare practitioners should adapt to the digital media landscape and enhance their communication skills.

4.3 Regional and cultural variations in video format and thematic preferences

Significant differences exist in both the breadth and focus of topics between YouTube and the Chinese platforms (TikTok, Bilibili). YouTube content focuses predominantly on examination and diagnosis, whereas etiology, prevention, and treatment are the most prevalent topics on TikTok and Bilibili. This preference aligns with a "prevention-first" health philosophy prevalent in Chinese culture (40), where understanding disease origins is valued for preventing progression. Content on Chinese platforms (TikTok, Bilibili) frequently employs monologs and medical scene imagery, emphasizing simplicity and emotional resonance. In contrast, YouTube often utilizes classroom-style or PowerPoint presentations, prioritizing structural clarity and objectivity. Furthermore, medical scene imagery was identified as a particularly effective format for disseminating information on TikTok. Given the diversity of creators and topics, there is a pressing need for platforms and uploaders to

TABLE 6 Spearman correlation between video quality and audience interaction on YouTube/ Bilibili/ TikTok.

r, p-value	YouTube (N₁ = 98)				Bilibili (N₂ = 74)				TikTok (N₃ = 99)			
	PEMAT	VIQI	GQS	mDISCERN	PEMAT	VIQI	GQS	mDISCERN	PEMAT	VIQI	GQS	mDISCERN
Views	0.256, **0.011**	0.605, **<0.001**	0.267, **0.008**	0.191, 0.059	0.421, **<0.001**	0.701, **<0.001**	0.056, 0.579	−0.483, **<0.001**	-	-	-	-
Thumbs-up	0.318, **0.001**	0.607, **<0.001**	0.341, **<0.001**	0.257, **0.011**	−0.144, 0.156	0.238, **0.017**	0.233, **0.020**	0.175, 0.083	0.152, 0.133	0.392, **<0.001**	0.054, 0.596	0.064, 0.532
Comments[a]	0.273, **0.006**	0.560, **<0.001**	0.378, **<0.001**	0.265, **0.016**	−0.424, **<0.001**	−0.172, 0.088	0.108, 0.285	0.435, **<0.001**	0.072, 0.478	0.299, **0.003**	−0.221, **0.028**	0.117, 0.248
Collections	-	-	-	-	−0.065, 0.525	0.201, **0.046**	0.315, **0.002**	0.227, **0.024**	0.205, **0.041**	0.405, **<0.001**	−0.017, 0.868	0.064, 0.530
Shares	-	-	-	-	−0.185, 0.066	0.106, 0.295	0.312, **0.001**	0.284, **0.004**	0.199, **0.049**	0.452, **<0.001**	0.015, 0.884	0.052, 0.610
Coin-operated	-	-	-	-	−0.317, **0.001**	−0.067, 0.508	0.286, **0.004**	0.500, **<0.001**	-	-	-	-

Bold text means the p-value < 0.05.

|r| ≤ 0.2 no relationship; 0.2 < |r| ≤ 0.4 weak relationship; 0.4 < |r| ≤ 0.6 moderate relationship; 0.6 < |r| ≤ 0.8 strong relationship; |r| > 0.8 very strong relationship.

[a]Excluded: 16 videos on YouTube turned off the function of comments.
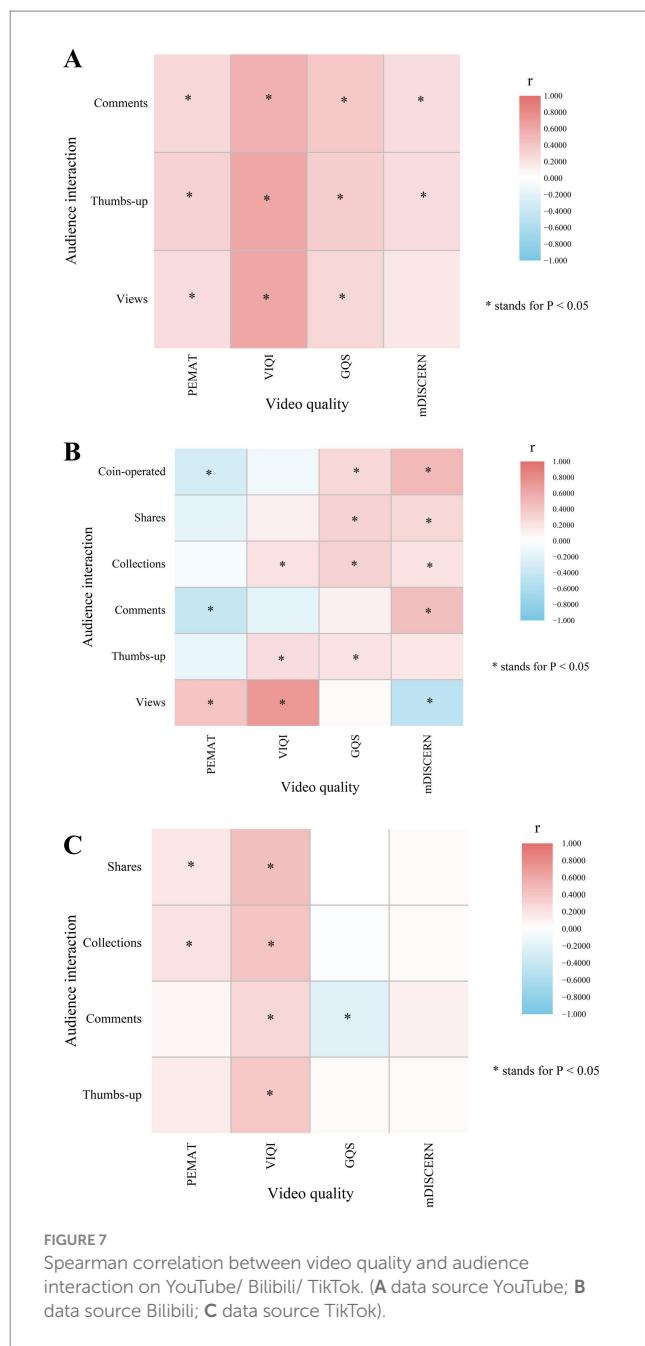
FIGURE 7
Spearman correlation between video quality and audience interaction on YouTube/ Bilibili/ TikTok. (**A** data source YouTube; **B** data source Bilibili; **C** data source TikTok).

integrate multimodal formats to enhance the accessibility and comprehensibility of medical information.

4.4 Weak correlation between audience engagement and video quality

With the exception of Bilibili, where view count significantly correlated with VIQI ($r = 0.701$, $p < 0.001$), overall user engagement metrics (e.g., likes, comments, and shares) demonstrated only weak-to-moderate correlations with objective quality measures. This finding is consistent with earlier studies on health science video consumption (41), which suggest that most viewers cannot accurately assess the professional quality of health content, leading to the weak correlations

TABLE 7 Summary of key findings on video quality and engagement across platforms.

Key finding	Data support	Interpretation
No significant difference in video quality between professional and non-professional uploaders	Table 5: All $p > 0.05$ for PEMAT, VIQI, GQS, mDISCERN	Professional credentials do not guarantee higher-quality public health communication.
No major inter-platform differences in information reliability (mDISCERN)	Table 4: mDISCERN-sum scores similar across platforms ($p > 0.05$)	All platforms face similar challenges in conveying reliable medical information.
Weak-to-moderate correlation between engagement and quality metrics	Table 6: Most r-values between 0.2–0.4; highest correlation: Bilibili Views vs. VIQI ($r = 0.701$, $p < 0.001$)	Audience engagement is poorly predictive of video quality; high views ≠ high quality.
YouTube leads in comprehensibility and actionability (PEMAT)	Table 4: PEMAT-T and PEMAT-A significantly higher on YouTube ($p < 0.001$)	YouTube videos are more structured and easier to understand and act upon.
Bilibili and TikTok lead in production quality (VIQI, GQS)	Table 4: VIQI and GQS significantly higher on Bilibili and TikTok ($p < 0.001$)	Short-form platforms excel in visual and production quality, but may lack depth.
TikTok has highest engagement despite shortest videos	Table 1: Median length = 114 s; highest likes, comments, shares	Algorithmic promotion and format favor high interaction, not necessarily quality.

typically observed between engagement and quality (17). This phenomenon can be attributed to several platform-specific factors: (1) cultural and behavioral user group differences, where a preference for entertainment-oriented content creates a "high quality, low interaction" paradox; (2) symbolic interaction behaviors that contrast with the "low quality, high interaction" pattern; and (3) public preference for fragmented viewing, compounded by algorithms that prioritize "high-stimulus" content (e.g., suspenseful thumbnails, rapid editing). The logical rigor inherent to professional medical information is often incompatible with the ultra-short video format. Furthermore, the reflective engagement required for understanding professional medical content fundamentally conflicts with platform algorithms optimized for "community weighting" and "completion rate." This discrepancy creates a vicious cycle: passive content consumption (rather than active real-time interaction) leads algorithms to classify these users as "low-activity," consequently reducing the content's exposure and recommendations. Moreover, content production strategies are heavily influenced by underlying platform business models. YouTube's ad revenue-sharing model, closely linked to viewing duration and user retention, incentivizes creators to produce higher-quality content to ensure stable income. In contrast, the diversified monetization strategies of TikTok and Bilibili often prioritize traffic-driven content, which can compromise quality. This analysis suggests that the positive correlation between quality and user stickiness on YouTube stems from a synergy of "deep content ecosystem + quality-aligned algorithms + user value recognition."

These findings align with broader research into the acceptance of mHealth services, which highlights the mediating role of perceived usefulness and user attitude in technology adoption (42), as well as the influence of service characteristics on continuous engagement (43). Conversely, the challenges observed on TikTok and Bilibili reflect systemic contradictions between "fragmented consumption + emotion-driven interaction + traffic-optimized algorithms."

These findings reveal fundamental structural tensions in health communication on commercial video platforms, encompassing conflicts between professional rigor and algorithmic engagement, between cultural preferences and universal standards, and between creator credibility and communicative effectiveness. If unaddressed, these disparities risk exacerbating public misinformation and health anxiety, particularly in clinically nuanced areas like PNs.

Therefore, coordinated multi-stakeholder efforts are required to enhance the quality and accessibility of medical video content. Improving the reliability of online health information is not merely an issue of content moderation; it is a foundational step toward leveraging digital technology to optimize the allocation of medical resources and mitigate regional health disparities (44). Platforms should implement more robust review systems, introduce quality-rating schemes or risk warnings, and adjust algorithms to prioritize scientifically accurate content. Healthcare professionals require training in public communication and should collaborate with media specialists to enhance audience engagement without compromising scientific accuracy. Finally, users should be empowered to develop critical media literacy skills, such as cross-referencing information sources and identifying potential biases. From a policy perspective, promoting a sustainable and reliable mHealth market requires coordinated efforts that address both technological infrastructure and user trust (45). Additionally, the disparate effects of physicians' knowledge-sharing behaviors—whether driven by satisfaction or gratitude—can significantly influence patients' evaluation of online medical services (46, 47), underscoring the need for platforms to foster genuine and informative interactions.

Theoretically, this study contributes to health communication scholarship by demonstrating how the interplay of platform algorithms, cultural contexts, and creator credentials shapes medical content quality and dissemination. Methodologically, it provides a validated, multi-dimensional framework for evaluating health-related videos. Practically, it offers actionable insights for platforms, content creators, and educators seeking to improve the effectiveness and reliability of digital health communication.

4.5 Limitations

Variations in evaluator expertise and assessment capabilities may introduce subjective bias, despite the use of a respiratory specialist team employing four multi-functional assessment tools. The cross-sectional design captures video attributes related to "Pulmonary Nodule" only at a specific point in time. Consequently, dynamic changes in platform algorithms and content update rates may limit the temporal validity of the findings.

The search strategy relied solely on the single keyword "Pulmonary Nodule," excluding auxiliary information (e.g., user comments, video descriptions). This approach may have omitted relevant content and reduced search comprehensiveness. Finally, analyzing only the top 100 results per platform may inadequately represent the overall content

ecosystem. This approach introduces selection bias, as ranking algorithms prioritize high-engagement (though not necessarily high-quality) content, thereby limiting the generalizability of the conclusions.

These limitations highlight the need for future research to employ multi-platform longitudinal monitoring, natural language processing, and user behavior analytics to develop more comprehensive health information quality assessment models.

5 Conclusion

This study offers the first thorough cross-sectional evaluation of lung nodule-related videos across three distinct video platforms. It gives trustworthy information for the general public to comprehend the present situation of online videos connected to PNs on Internet platforms. These insights may help platforms, content creators, and the public at large. While short-form video platforms like TikTok have a high index of audience engagement but poor informativeness, long-form video platforms like YouTube and Bilibili are informative but have low traffic. YouTube videos have the best average quality out of the three platforms, but they can still be made better. This study suggests translating and sharing high-quality information between various platforms for interoperability and learning. A growing number of healthcare professionals are being requested to obtain certification and become proficient in video capturing techniques to enhance the quality of PN-related movies and provide high-quality medical information in an intelligible and clear manner. The platform has developed a complex method for evaluating the content's quality, improving the recommendation system's weighting elements, and building a more reliable environment for the spread of health information.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors. The data analyzed in this study (such as video content, user comments, and public metadata) were sourced from the following publicly available online video platforms: https://www.youtube.com; https://www.bilibili.com; https://www.douyin.com.

Ethics statement

Ethical approval was not required for the study involving human data in accordance with the local legislation and institutional requirements. Written informed consent was not required, for either participation in the study or for the publication of potentially/indirectly identifying information, in accordance with the local legislation and institutional requirements. The social media data was accessed and analyzed in accordance with the platforms' terms of use and all relevant institutional/national regulations.

Author contributions

YZ: Methodology, Software, Visualization, Writing – original draft. XZ: Data curation, Writing – original draft. TY: Data curation, Writing – review & editing. QW: Methodology, Writing – review & editing. SW:

Data curation, Writing – review & editing. LD: Software, Writing – review & editing. LW: Funding acquisition, Supervision, Writing – review & editing. JH: Supervision, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2025.1613526/full#supplementary-material

References

1. Mac Mahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the fleischner society 2017. *Radiology*. (2017) 284:228–43. doi: 10.1148/radiol.2017161659

2. Naidich DP, Bankier AA, MacMahon H, Schaefer-Prokop CM, Pistolesi M, Goo JM, et al. Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the fleischner society. *Radiology*. (2013) 266:304–17. doi: 10.1148/radiol.12120628

3. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner society: glossary of terms for thoracic imaging. *Radiology*. (2008) 246:697–722. doi: 10.1148/radiol.2462070712

4. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med*. (2013) 369:910–9. doi: 10.1056/NEJMoa1214726

5. Mets OM, de Jong PA, Chung K, Lammers J-WJ, van Ginneken B, Schaefer-Prokop CM. Fleischner recommendations for the management of subsolid pulmonary nodules: high awareness but limited conformance - a survey study. *Eur Radiol*. (2016) 26:3840–9. doi: 10.1007/s00330-016-4249-y

6. Dyer SC, Bartholmai BJ, Koo CW. Implications of the updated lung CT screening reporting and data system (lung-RADS version 1.1) for lung cancer screening. *J Thorac Dis*. (2020) 12:6966–77. doi: 10.21037/jtd-2019-cptn-02

7. van Riel SJ, Ciompi F, Winkler Wille MM, Dirksen A, Lam S, Scholten ET, et al. Malignancy risk estimation of pulmonary nodules in screening CTs: comparison between a computer model and human observers. *PLoS One*. (2017) 12:e0185032. doi: 10.1371/journal.pone.0185032

8. Li R, Chen J, Li Z, Li P, Qiu B, Tan F. Reassessment of the prognostic implication of ground-glass opacity: an investigation into hypermetabolic clinical stage IA lung adenocarcinoma. *Thorac Cancer*. (2025) 16:e15530. doi: 10.1111/1759-7714.15530L

9. Yuksel B, Cakmak K. Healthcare information on you tube: pregnancy and COVID-19. *Int J Gynaecol Obstet*. (2020) 150:189–93. doi: 10.1002/ijgo.13246

10. Ostrovsky AM, Chen JR. Tik Tok and its role in COVID-19 information propagation. *J Adolesc Health*. (2020) 67:730. doi: 10.1016/j.jadohealth.2020.07.039

11. Niu Z, Hao Y, Yang F, Jiang Q, Jiang Y, Zhang S, et al. Quality of pancreatic neuroendocrine tumor videos available on Tik Tok and bilibili: content analysis. *JMIR Form Res*. (2024) 8:e60033. doi: 10.2196/60033

12. Basch CH, Hillyer GC, Jaime C. COVID-19 on TikTok: harnessing an emerging social media platform to convey important public health messages. *Int J Adolesc Med Health*. (2022) 34:367–9. doi: 10.1515/ijamh-2020-0111

13. Yang Y, Zhu X, Song R, Zhang X, Guo F. Not just for the money? An examination of the motives behind physicians' sharing of paid health information. *J Inf Sci*. (2021) 49:145–63. doi: 10.1177/0165551521991029

14. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med*. (2012) 22:276–82. doi: 10.11613/BM.2012.031

15. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health*. (1999) 53:105–11. doi: 10.1136/jech.53.2.105

16. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the world wide web. *Am J Gastroenterol*. (2007) 102:2070–7. doi: 10.1111/j.1572-0241.2007.01325.x

17. Liu Z, Chen Y, Lin Y, Ai M, Lian D, Zhang Y, et al. YouTube/ bilibili/ TikTok videos as sources of medical information on laryngeal carcinoma: cross-sectional content analysis study. *BMC Public Health*. (2024) 24:1594. doi: 10.1186/s12889-024-19077-6

18. Shoemaker SJ, Wolf MS, Brach C. Development of the patient education materials assessment tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ Couns*. (2014) 96:395–403. doi: 10.1016/j.pec.2014.05.027

19. Diers CS, Remvig C, Meteran H, Thomsen SF, Sigsgaard T, Høj S, et al. The usefulness of YouTube videos as a source of information in asthma. *J Asthma*. (2023) 60:737–43. doi: 10.1080/02770903.2022.2093218

20. Choudhary OA, Dhawan J, Sohrabipour S, Semenchuk J, Silva TD, Masthan MI, et al. Assessment of online YouTube videos as a source of information and instruction for pulmonary rehabilitation. *Respir Med*. (2025) 238:107968. doi: 10.1016/j.rmed.2025.107968

21. Sohrabipour S, Ibrahim A, Dhawan J, Choudhary OA, Luu B, Shore J, et al. Evaluation of online videos and websites on inspiratory muscle training for individuals with chronic lung disease. *Respir Med*. (2025) 238:107967. doi: 10.1016/j.rmed.2025.107967

22. Yoo M, Bang MH, Jang CW. Evaluation of YouTube videos as a source of information on pulmonary rehabilitation for COPD. *Respir Care*. (2022) 67:534–42. doi: 10.4187/respcare.09651

23. Stellefson M, Chaney B, Ochipa K, Chaney D, Haider Z, Hanik B, et al. YouTube as a source of chronic obstructive pulmonary disease patient education: a social media content analysis. *Chron Respir Dis*. (2014) 11:61–71. doi: 10.1177/1479972314525058

24. Høj S, Thomsen SF, Høj L, Sigsgaard T, Meteran H. Investigating the scientific accuracy of YouTube videos on food allergy. *Ann Allergy Asthma Immunol*. (2024) 133:342–4. doi: 10.1016/j.anai.2024.05.017

25. Guler MA, Aydın EO. Development and validation of a tool for evaluating YouTube-based medical videos. *Ir J Med Sci*. (2022) 191:1985–90. doi: 10.1007/s11845-021-02864-0

26. Wang X, Sun Y, Qu X, Lei T, Song Y, Deng J. Is YouTube™ a useful source of information on home parenteral nutrition? *BMC Public Health*. (2025) 25:684. doi: 10.1186/s12889-025-21929-8

27. Feragalli B, Guido F, Larici AR, Storto ML, Bonomo L. Pulmonary nodule. *Radiol Med*. (2005) 110:294–314. doi: 10.1007/s00117-005-1276-0

28. Han J, Shi Y, Ma H. Assessment of videos related to lung nodules in China. *Front Surg*. (2022) 9:1019212. doi: 10.3389/fsurg.2022.1019212

29. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ*. (2022) 22:382. doi: 10.1186/s12909-022-03446-z

30. Helming AG, Adler DS, Keltner C, Igelman AD, Woodworth GE. The content quality of YouTube videos for professional medical education: a systematic review. *Acad Med*. (2021) 96:1484–93. doi: 10.1097/ACM.0000000000004121

31. Liang Y, Xia J, Huo W, Liu B, Wang Z, Ding Y, et al. Video quality assessment and analysis of gastroesophageal reflux disease on TikTok and bilibili: cross-sectional study. *J Multidiscip Healthc*. (2024) 17:5927–39. doi: 10.2147/JMDH.S485781

32. Mazzone PJ, Lam L. Evaluating the patient with a pulmonary nodule: a review. *JAMA*. (2022) 327:264–73. doi: 10.1001/jama.2021.24287

33. Sun F, Zheng S, Wu J. Quality of information in gallstone disease videos on TikTok: cross-sectional study. *J Med Internet Res*. (2023) 25:e39162. doi: 10.2196/39162

34. Li J, Shi W. Accessing the impact of TikTok's algorithm on regional inequality in health information. *Health Commun*. (2024) 40:1636–44. doi: 10.1080/10410236.2024.2414882

35. Kanchan S, Gaidhane A. Social media role and its impact on public health: a narrative review. *Cureus*. (2023) 15:e33737. doi: 10.7759/cureus.33737

36. Zhang X, Xie K, Gu B, Guo X. Engaging physicians with introductory incentives: references to online and offline income. *MIS Q*. (2022) 49:6484. doi: 10.2139/ssrn.5106484

37. Zhang X, Pu J, Lu Y, Guo F. Team makes you better: evidence from online medical consultation platforms. *Inform Syst Res*. (2025) 26:689. doi: 10.1287/isre.2022.0689

38. Huang K, Zhang P, Zhang Z, Youn JY, Wang C, Zhang H, et al. Traditional Chinese medicine (TCM) in the treatment of COVID-19 and other viral infections: efficacies and mechanisms. *Pharmacol Ther*. (2021) 225:107843. doi: 10.1016/j.pharmthera.2021.107843

39. Bullard MJ, Fox SM, Heffner AC, Bullard CL, Wares CM. Unifying resident education: 12 interdisciplinary critical care simulation scenarios. *MedEdPORTAL*. (2020) 16:11009. doi: 10.15766/mep_2374-8265.11009

40. Wu X-Q, Zhang W-N, Hao M-Z, Liu X-P, Xiao J, Wang T-F, et al. How Chinese herbal medicine prevents epidemics: from ancient pestilences to COVID-19 pandemic. *Am J Chin Med*. (2021) 49:1017–44. doi: 10.1142/S0192415X2150049X

41. Zeng F, Zhang W, Wang M, Zhang H, Zhu X, Hu H. Douyin and bilibili as sources of information on lung cancer in China through assessment and analysis of the content and quality. *Sci Rep*. (2024) 14:20604. doi: 10.1038/s41598-024-70640-y

42. Zhang X, Guo X, Guo F, Lai K-H. Nonlinearities in personalization-privacy paradox in mHealth adoption: the mediating role of perceived usefulness and attitude. *Technol Health Care*. (2014) 22:515–29. doi: 10.3233/THC-140811

43. Wang L, Wu T, Guo X, Zhang X, Li Y, Wang W. Exploring mHealth monitoring service acceptance from a service characteristics perspective. *Electron Commer Res Appl*. (2018) 30:159–68. doi: 10.1016/j.elerap.2018.06.003

44. Chen S, Lai K, Guo X, Zhang X. The influence of digital health technology on the allocation of regional medical resources in China. *Health Policy Technol*. (2025) 14:101013. doi: 10.1016/j.hlpt.2025.101013

45. Zhang X, Lai K, Guo X. Promoting China's mHealth market: a policy perspective. *Health Policy Technol*. (2017) 6:383–8. doi: 10.1016/j.hlpt.2017.11.002

46. Tan H, Zhang X, Yang Y. Satisfaction or gratitude? Exploring the disparate effects of physicians' knowledge sharing on patients' service evaluation in online medical consultations. *Inf Syst J*. (2023) 33:1186–211. doi: 10.1111/isj.12440

47. Zhang W, Jiang M, Zhao W, Li S, Li F, Feng F, et al. Evaluation of the effectiveness of using flipped classroom in puncture skills teaching. *BMC Med Educ*. (2024) 24:176. doi: 10.1186/s12909-024-05132-8

Check for updates

Reveal the dynamics of mobile health services continuance intention: effects of expectation, confirmation, and chronic disease

Xiumei Ma[1], Yanxia Li[2]* and Ao Suo[3]

[1]School of Public Administration, Sichuan University, Chengdu, China, [2]School of Economics and Management, Yanshan University, Qinhuangdao, China, [3]Clinical Medical School, Gansu Medical College, Pingliang, China

Introduction: The sustainable development of mobile health (mHealth) services relies on the continuous use by users. Most studies consider users' intention to continue using mHealth services as a static measure rather than one that changes dynamically over time, often neglecting the impact of individual differences, such as the presence of chronic disease, on usage patterns.

Methods: Drawing on the expectation-confirmation model, this study investigates the dynamic nature of the intention to continue using mHealth services, with a particular focus on the role of chronic disease. Conducting a longitudinal study with three rounds of online survey, we analyzed data collected from 236 completed respondents using a latent growth model.

Results: The results indicate that users' intention to continue using mHealth services is not static and tends to decrease over time. Expectation accelerates the descent, while confirmation mitigates it. Furthermore, expectation shows a stronger impact on users without chronic disease compared to those with chronic conditions.

Discussion: This study advances the understanding of the continuous use of mHealth services by incorporating time-based dynamics and the influence of chronic disease. This study also extends the traditional expectation-confirmation model into a dynamic framework.

KEYWORDS

mHealth services, continuous use, latent growth model, expectation-confirmation model, chronic disease

1 Introduction

Mobile health (mHealth) services have flourished in recent decades due to advancements in mobile communications and network technologies (1, 2). Compared to earlier electronic health services accessed via desktop computers (3–5), mHealth offers a more convenient way for patients to obtain remote health consultation, health monitoring, and customized information services (6). As an innovative approach to delivering health services, mHealth has been empirically proven to be effective in promoting health management, particularly by enabling patients with chronic disease to actively engage in self-care (7). Due to these advantages and conveniences, mHealth services have garnered increasing attention and widespread popularity in recent years.

Although a growing number of studies have made initial efforts to investigate continuance issues within mHealth, most have viewed this phenomenon as static rather than dynamic (8).

This is evident as many existing studies on mHealth continuance rely on cross-sectional surveys conducted at a single point in time (9, 10). However, users' beliefs and use intentions are dynamic and vary over time (11). Initially, users may have a high intention to continue using mHealth services due to high expectations. However, the intention is likely to decrease or even disappear if the service performance does not meet their expectations. Although a few studies have emphasized the dynamic nature of mHealth continuance intention (12, 13), they have not empirically explored the change trajectory. Thus, incorporating a time dimension into empirical studies that examine changes in users' intention to continue using will address this research gap, providing insights into users' engagement, and promote the sustainable development of mHealth. Therefore, we propose the first research question:

RQ1: How does users' intention to continue using mHealth services change over time?

The expectation-confirmation model (ECM) is widely used to understand individuals' continuous use behavior and intention (14, 15). The original ECM posits that users' expectation and confirmation influence satisfaction, which further leads to behavior or intention (16). Building on the ECM, Bhattacherjee and Premkumar (17) verified temporal pattern of technology usage and investigated how expectation and confirmation at different stages influence continuous use intention through longitudinal testing. However, they only emphasized that users' beliefs and attitudes vary over time, failing to reveal the trajectory of change, especially how expectation and confirmation affect the change trajectory. Thus, while an expectation-confirmation paradigm is suggested to explain the continuous use of technologies, no study to date has examined how expectation and confirmation affect the change trajectory of continuous use intention. To fill this gap, the second research question is proposed:

RQ2: How do expectation and confirmation influence the change in users' intention to continue using mHealth services?

Moreover, individual differences may also influence the continuous use of mHealth services. Research has indicated that individual traits such as gender (18), professional seniority (19), income (20), and health condition (21) interact with technology to shape user behavior. Specifically, individuals' vulnerability to chronic disease has a contingent effect on mHealth use (22). This maybe because a fundamental component of mHealth services is assisting in chronic disease management. Thus, the demand for and use of mHealth services may differ between individuals with and without chronic disease. We are particularly interested in whether chronic disease involvement affects the change trajectory of mHealth continuous use over time. Specifically, we aim to understand how chronic disease interacts with users' expectation and confirmation, as its role remains unclear. Therefore, the third research question of this study is:

RQ3: What role does chronic disease play in the change of users' intention to continue using mHealth services?

To answer these three questions, we conducted a longitudinal test through three successive surveys. Utilizing the latent growth model, this study explores how use intention evolves over time, how expectation and confirmation influence the change trajectory, and the role of chronic disease in this process. This study sheds light on current literature and practice. First, this study views continuous use intention as dynamic and clarifies how it changes over time in the mHealth context. Given that most previous studies have treated individual

intention to use as static (23, 24), this study expands the literature by diagnosing time-based dynamics of use intention. Second, this study provides a new understanding of changes in continuous use intention from the lens of expectation-confirmation paradigm. In particular, this work is among the first to combine the ECM with the trajectory of continuous use intention, presenting a clear picture of the impact of expectation and confirmation on the evolution of mHealth continuous use intention. Third, through examining the role of chronic disease, this study expands the literature on the ECM by exploring its applicability in contexts involving both presence and absence of chronic disease. Fourth, this study offers methodological guidance for future longitudinal studies. Finally, results in this study may inspire further research into the impact of chronic disease in the mHealth context and provide practitioners with actionable strategies for retaining mHealth users, both with and without chronic disease.

2 Theoretical background and hypothesis development

2.1 The dynamics of mHealth services continuous use intention

Continuous use is defined as the behavioral patterns reflecting sustained use of a technology by individuals over long time after adoption (25), which has attracted extensive attention in information systems and social science field (26). Continuous use of mHealth can not only ensure the economic return of an mHealth business but also improve the users' health management. Most research about the continuous use of mHealth services have investigated determinant factors influencing users' continuance intention from viewpoints of system features and user perceptions, seperately or concurrently (27–29). For example, drawing on the technology acceptance model and investment model, Cho et al. (30) investigated the role of perceived usefulness, perceived ease of use, and investment size in predicting mHealth continuance intention (67). Showed that individuals' continuous intention was linked to service assurance, hedonic benefits, efficiency, reliability, and content quality. Such studies in the mHealth services context have generally viewed continuous use as unchanging or static, failing to understand how continuous use changes over time and explore the trajectory or the rate of the change.

However, continuous use is far from static: it is a dynamic process in which users may stop and readopt the technology, and studies as early as Bhattacherjee (31) have emphasized its dynamic aspects and nature. More recent studies have shown the importance of investigating the dynamics of continuous use in the mHealth context. For example, conducting a longitudinal study, Meyer et al. (32) discovered that users often break and abandon their mHealth devices during long-term use. Shen et al. (13) indicated that since users may behave intermittently, continuous use is not a fixed decision but a dynamic process. Thus, we argue that mHealth services continuous use intention decreases over time. This would align with one previous longitudinal study on technology use, in which use intention in a later stage was found to be lower than in an earlier stage (32). We therefore propose that:

> *H1*: Individuals' continuous use intention is dynamic and decreases over time.

2.2 Expectation confirmation model

Bhattacherjee (31) proposed the expectation confirmation model (ECM), a postadoption model addressing that individuals' continuous use of information technologies or services. The ECM mainly focuses on the post-adoption variables and posits that confirmation of the initial expectation influences perceived performance and then influence users' satisfaction and intention to continue using. In recent years, more scholars paid notice to ECM and applied this model in mHealth services continuance research (30, 33, 34).

However, despite the valuable theoretical insights the ECM has provided regarding continuous use, several research blanks remain. First, the ECM has widely been drawn upon to explain users' continuous use behavior and intention related to a technology or service, but it has never been revoiced to address the dynamic development of continuous behavior and intention (35, 36). Given that some more recent studies have been stressing the importance of understanding how continuous use behavior and intention change during the whole use process (13), the ECM needs to be extended through considering the dynamic change of continuous use. Second, most current studies drawing on the ECM only focus on post-adoption variables (e.g., perceived performance, confirmation, satisfaction, and continuous use), while they have lost sight of pre-adoption variables (e.g., expectation) (37–39). Recent research has alerted researchers to investigate the whole adoption process, including the ways in which pre-adoption expectations shape post-adoption perceptions/behaviors (40, 41). Thus, in order to understand the whole use process, especially the initial use intention and changes in intention in the long-term use stage (e.g., the rate of change), it is critical to involve expectation and confirmation in the model.

Expectation refers to a user's anticipated or predicted belief about the attributes or characteristics of technologies or services (42). When users believe technologies or services will perform well in the future, they are more likely to use them. The unified theory of acceptance and use of technology has also asserted that the initial usage intention is affected by the user's expectations about the technology's performance (43). That is, high-level expectation will increase user's continuous intention to use mHealth services. In addition to influencing the initial use intention, expectation may have a significant effect on subsequent use, such as changes in continuous intention. Expectation, as users' pre-use belief, serves as an anchor or a baseline for evaluating the technology or service. A higher baseline expectation is harder to satisfy, and thus users are more likely to incur a psychological loss during the subsequent use process. According to the theory of risk aversion, individuals will take measures to avoid risk, especially when risk aversion increases with time (44, 45). Specific to this context, high-level expectation is likely to bring a risk of psychological loss, and in response to this loss, users may reduce their continuous intention. Thus, we infer that the higher the initial expectations, the faster the continuous intention will decrease. Therefore, we propose that:

> *H2*: Expectation is positively related to the initial continuous use intention.
>
> *H3*: Expectation is positively related to the rate of change in continuous use intention.

Confirmation refers to users' perceptions of the congruity between their expectations of a technology's performance and its real performance (31). In this regard, confirmation is an outcome of comparing a user's pre-adoption expectation and post-adoption perceived performance. A high-level confirmation indicates that technologies' or services' performance satisfies the user's expectations, which should positively influence the user's intention to use them, at least initially (31). Studies have demonstrated that confirmation significantly improves users' attitude towards technologies and services as well as their intentions to continue using technologies and services (37, 40). In this study, we expect use intention to decrease over time during the use process (H1), perhaps substantially due to fatigue or tiredness. In this regard, confirmation may play a role in the change of use intention. Individuals who get high confirmation, and therefore high satisfaction, may experience a drop in use intention, though not as much or as quickly as other users. Individuals who get low confirmation (and low satisfaction) should see a faster decrease in their use intention. Thus, we put forward that:

> *H4*: Confirmation is negatively related to the rate of change in continuous use intention.

2.3 The role of chronic disease

Chronic disease is often incurable and requires long-term management, which may include addressing diet and exercise behaviors, monitoring and recording real-time health data, and asking for medication advice from physicians (46). Chronic disease degrades individuals' overall health condition and quality of life, so it negatively affects them both physiologically and psychologically (47). Physiological effects include poor health and drug dependence, and psychological effects include gloom and panic caused by poor health condition and psychological dependence on physician support. Due to these negative impacts, there are significant differences in health service needs between individuals with chronic disease and those without chronic disease (48).

Although mHealth can offer health management services and information to both chronic patients and individuals with no chronic disease, these two groups may have different demands and expectations regarding the function and performance of mHealth. The expectation and confirmation paradigm has been widely examined in mHealth services studies (41); however, the role of chronic disease has rarely been investigated. According to the expectation confirmation theory, expectation negatively correlates to confirmation since a higher expectation is harder to meet (49). Compared to individuals with no chronic disease, chronic patients have more and urgent needs as well as higher requirements for mHealth services to solve their problems (50). In this situation, their expectation is harder to satisfy and confirm than others'. Thus, expectation may lead to lower confirmation for chronic patients.

Moreover, the differing expectations of individuals with and without chronic disease may have differing impacts on their initial mHealth services use. For example, chronic patients tend to expect mHealth to perform well in solving their health problems, while individuals with no chronic disease are likely to use mHealth to learn more about their health conditions. Chronic patients using mHealth services may therefore place more emphasis on their performance expectations, while in contrast, users with no chronic disease may pay more attention to their interaction with the technology (22). In other words, chronic patients have higher use intention when they have high expectation while users with no chronic disease have higher use

intention when they experience valuable interaction. That is, the impact of expectation before using mHealth is stronger for chronic patients. Thus, it is reasonable to argue that compared to individuals with no chronic disease, mHealth services expectation may play a more important role in motivating initial continuous use intention for chronic patients. Hence, we put forward that:

H5: The relationship between expectation and confirmation is stronger for chronic patients.

H6: The relationship between expectation and initial continuous use intention is stronger for chronic patients.

According to Maslow's hierarchy of needs, protecting ourselves from disease and staying healthy is the basic and pressing need driving individuals to take action (51). For a chronic patient, using mHealth services is an efficient way to fulfill this need by managing their disease, and it has significant potential to improve their health. Thus, chronic patients tend to have more pressing mHealth needs. According to demand theory, more demand can cause a higher level of preferences and more quantity of services that individuals would like to choose (52). That is to say, compared to individuals with no chronic disease, chronic patients are more likely to rely on mHealth both physiologically and psychologically due to their demands. For individuals who have less demands (e.g., individuals with no chronic disease), they show less stickiness to the service (53). For this reason, users with no chronic disease tend to change their usage intention due to their mutable evaluations of mHealth services. In this study context, expectation and confirmation correlate to users' perception and evaluation of mHealth services, which are both likely to show stronger impacts on the change of use intention for individuals with no chronic disease. Thus, we put forward that:

H7: The relationship between expectation and rate of change in continuous use intention is stronger for individuals with no chronic disease.

H8: The relationship between confirmation and rate of change in continuous use intention is stronger for individuals with no chronic disease.

To investigate the dynamics of mHealth services use intention from the ECM perspective and explain the role of chronic disease, this study employs a latent growth model. We build the research model based on ECM and empirically test the model with longitudinal survey data. The research model and hypothesis relationships are described in Figure 1.

3 Methods

3.1 Research setting

Most previous studies have examined expectation through asking respondents to recall their past expectations; however, memories of expectation are not accurate. Therefore, we collected data using multi-wave surveys which measured respondents' perceptions in real time. To achieve this objective, we conducted this study in cooperation with a major hospital in Beijing, China. At the time, the hospital was launching an mHealth application that offered services such as health information education, online consultation, health index record and monitoring, and advice notification, similar to other mHealth services on the market. Both chronic patients and individuals with no chronic disease can use these services for self-health management. Individuals who came to the hospital's health center for a physical examination were selected as target participants. These individuals are appropriate for this study since they did not come to the hospital for any specific treatment. Thus, their chronic disease status is random, and they are potential users of mHealth services.

3.2 Data collection

Participants who did not previously use the hospital's mHealth application were recruited. We collected data at three time points
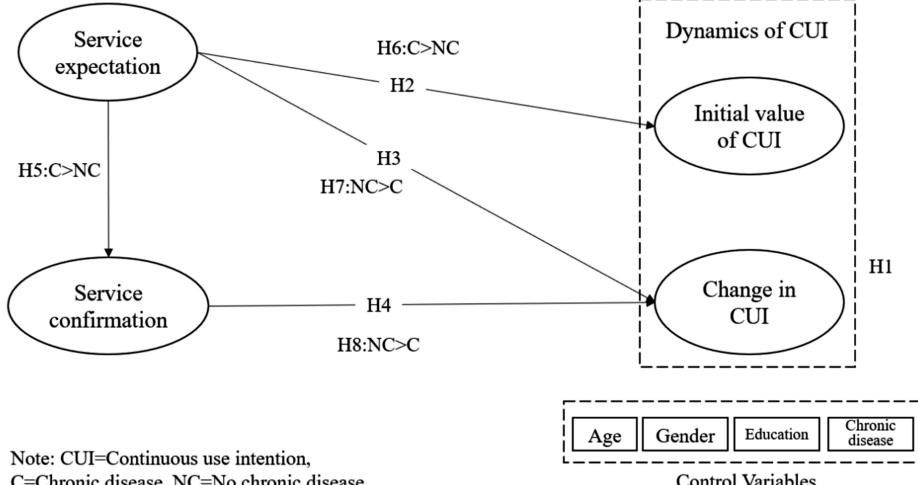

FIGURE 1
Research model.

(T1-T3) to explore the initial value and dynamics of users' beliefs and their use intention. At the beginning, we introduced our research and the mHealth services to potential participants and invited them to join the study. Patients who agreed were immediately invited to download the application to their mobile phone and complete a survey (at time T1). The survey contained two parts: one part investigated their expectation and continuous use intention regarding the mHealth services, and the other collected their basic demographic information. Two months later, another survey link was sent to examine their confirmation and continuous use intention (at time T2). The last survey investigating users' continuous use intention was disseminated after another 2 months (at time T3). Respondents participated in the research voluntarily and were able to quit the study anytime. All research activities had received institutional review board approval and followed approved institutional review board guidelines. Considering the high attrition rate in multi-time-point studies, we recruited 500 participants, of whom 462 respondents completed the first-round survey. For the second-round survey, 350 of the 462 participants responded to the questionnaire. Finally, 236 respondents completed all three surveys. To test the nonresponse bias, we compared the initial respondents and final valid respondents in terms of gender (t = −1.43, $p > 0.1$), age (t = 0.72, $p > 0.1$), education (t = −1.28, $p > 0.1$), and chronic disease (t = 0.31, $p > 0.1$). No significant difference was found across these characteristics, indicating nonresponse bias is not a concern. Of these 236 valid respondents, 53.4% were male and 44.1% had a chronic disease. The demographic information is shown in Table 1.

3.3 Measurements

All measurements were adapted from previous studies and modified to the mHealth context as shown in Table 2. We employed seven-point Likert scales ranging from "1 = strongly disagree" to "7 = strongly agree" to measure multiple items of latent constructs.

TABLE 1 Demographic statistics.

Variables	Category	Frequency	Percentage (%)
Gender	Male	126	53.4
	Female	110	46.6
Age	30 or younger	36	15.3
	31–40	84	35.6
	41–50	60	25.4
	51–60	38	16.1
	61 or older	18	7.6
Education	High school or lower	62	26.3
	College	44	18.6
	Undergraduate	78	33.1
	Postgraduate or higher	52	22.0
Chronic disease	Yes	104	44.1
	No	132	55.9

Specifically, continuous use intention was measured with three items adapted from Bhattacherjee (31). Expectation was measured with three items adapted from Lin et al. (49) and confirmation was measured with items adapted from Bhattacherjee (31). Chronic disease was measured as a dummy variable based on respondents' yes/no answers to the question "Are you suffering from one or more chronic diseases?" Variables such as age, gender, and education were measured and collected due to control purposes. Because the respondents are Chinese, we translated all items into Chinese with back-translation.

4 Results

This study aims to explore the effects of users' expectation and confirmation of mHealth services on the change in their continuous use intention. The latent growth model approach is widely applied to examine how independent variables influence the change trajectory of dependent variables over time based on three or more longitudinal data sets (54). Thus, a latent growth model is employed for data analysis in this study. There are two latent factors involved in the latent growth model, namely intercept and slope. The intercept is a constant indicating the initial mean level of constructs measured at time T1, while the slope indicates the extent to which the mean value of the construct changes over time. In the latent growth model, the intercept and the slope were estimated as latent variables (55).

AMOS 24.0 is one of the various programs used to analyze the latent growth model and thus is employed to conduct data analysis in this study. Analyzing the latent growth involves two steps (56). The first step is to estimate the unconditional model which measures the dependent variable at three times. If the unconditional model is acceptable, the second step is conducted to estimate the conditional model, which evaluates how independent variables affect the slope of the dependent variable. Moreover, the bootstrapping approach was employed to validate moderating effects in the research model.

4.1 Unconditional latent growth model

There are two types of change trajectory in the unconditional latent growth model: one is the linear change growth (the slop loadings are labeled as 1,2,3…) and the other is non-specified cumulative growth (the first slope loading was fixed to 0 and the last was fixed to 1). To adopt the most suitable model in this study, we compared the fitness of these two models. The results showed that the non-specified cumulative growth model fit for the three stages of continuous use intention was ($\chi2 = 14.43$, df = 2, $p = 0.001$, TLI = 0.505, NFI = 0.645, CFI = 0.670, and RMSEA = 0.163). In contrast, the linear model fit was ($\chi2 = 2.339$, df = 3, $p = 0.311$, TLI = 0.987, NFI = 0.943, CFI = 0.991, and RMSEA = 0.027). Comparing the fit of these two models, we found that the linear growth model presented a higher-quality goodness-of-fit index and was superior to the non-specified cumulative growth model. Thus, we choose the linear change model as the latent growth model for analyzing continuous use intention in this study.

After adopting this model, we estimated the intercept (initial value) and the slope (rate of change) of continuous use intention as follows. The intercept of continuous use intention was 6.042 and statistically significant. The average of the slope of use intention was

TABLE 2 Construct measures.

Construct	Item	Questions	Related research
Continuous use intention	CUI1	I intend to continue using the mHealth services rather than discontinue its use.	(31)
	CUI2	My intentions are to continue using the mHealth services than use any alternative means (traditional medical service).	
	CUI3	If I could, I would like to discontinue my use of the mHealth services (reverse coded).	
Service expectation	SE1	I expected the service quality of the mHealth services to be good.	(49)
	SE2	I expected the service quality of the mHealth services to be stable.	
	SE3	I expected the service quality of the mHealth services to be effective.	
Service confirmation	SC1	My experience with using the mHealth services was better than what I expected.	(31)
	SC2	The service level provided by the mHealth services was better than what I expected.	
	SC3	Overall, most of my expectations from using the mHealth services were confirmed.	

−0.273, indicating that use intention significantly decreases with a negative slope. This indicates that individuals' intention to continue use mHealth services decreases significantly over time, supporting H1. Additionally, both the variance of intercept and of slope are significant, showing that individual differences exist in initial value and rate of change of use intention (see Table 3).

4.2 Reliability and validity analysis

A confirmative factor analysis (CFA) was conducted to assess reliability and validity of construct measurement using collected data. The reliability can be estimated by evaluating composite reliability and Cronbach's alpha. Table 4 showed that all composite reliabilities and Cronbach's alpha values were higher than the suggested threshold value of 0.7 (57). Thus, all constructs showed a good reliability. To test convergent validity, we applied two criteria, factor loadings and average variance extracted (AVE), as suggested by Fornell and Larcker (57). Table 4 showed that all factor loadings were significant and greater than 0.7 and all AVEs were over 0.5, suggesting a good convergent validity of all constructs. To evaluate discriminant validity, we compared the square root of AVE for a given construct and the correlation coefficients between this construct and other constructs. As shown in Table 5, for each construct, the square root of AVE was greater than the correlation coefficients, suggesting the discriminant validity is acceptable.

4.3 Conditional latent growth model

To test our hypotheses, we constructed a conditional latent growth model including all independent and dependent variables such as expectation, confirmation, continuance intention intercept and continuance intention slope. The unstandardized loadings of continuance intention on all intercept factors were fixed at 1.0. The unstandardized loadings on the slope factors by T1, T2, and T3 were fixed at 0, 1, and 2 to set a linear trend for the change in continuance intention. The relationship between variables is shown in Figure 2.

First, the model fit was tested. The criteria for a good model fit includes chi-square/degrees of freedom ($\chi2/df$) less than 5, NFI, TLI, and CFI greater than 0.9, and RMSEA less than 0.08 (58). The results

showed a satisfying model fit for the indices: $\chi2 = 61.56$, $\chi2/df = 1.466$, NFI = 0.968, TLI = 0.980, CFI = 0.989, and RMSEA = 0.023. All fit indices satisfied the criteria.

Next, we examined how expectation and confirmation influence the intercept (initial value) and the slope (rate of change). We found that expectation positively predicted the intercept of continuance intention ($\beta = 0.517$, $t = 6.702$, $p < 0.001$) and positively related to the slope as well ($\beta = 0.208$, $t = 3.642$, $p < 0.001$), supporting H2 and H3. In terms of the effect of confirmation on the slope, the result showed a significant negative effect ($\beta = -0.100$, $t = 4.785$, $p < 0.001$), lending support to H4. We also found that expectation significantly decreased confirmation ($\beta = -0.304$, $t = 3.420$, $p = 0.001$). In terms of control variables, chronic disease was found to positively affect the intercept ($\beta = 0.186$, $t = 2.169$, $p < 0.05$) and negatively affect the slope ($\beta = -0.187$, $t = -2.740$, $p < 0.01$). Figure 2 shows the conditional latent growth model results.

To assess the moderating role of chronic disease, we manipulated a multi-group invariance test through nested model comparisons to check whether the path coefficients for the chronic disease group and those for the no chronic disease group statistically differed from each other. Parameters of the unconstrained model (M0) were freely estimated, and equality constraints were imposed, particularly for the parameters of the nested models (M1-M4). In this test, the chi-square difference between these two models should indicate whether the equality constraint holds across different groups (59). Table 6 showed the differences between the chronic disease group and no disease chronic group in the path coefficients of expectation-confirmation ($p < 0.001$), expectation-intercept ($p < 0.05$), expectation-slope ($p < 0.01$), and confirmation-slope ($p = 0.098$). Based on the results, H5-H7 were supported but H8 was not supported.

5 Discussion

5.1 Key findings

We examined the dynamics of users' long-term use intentions regarding mHealth services. From the point view of ECM, we investigated the impacts of expectation and confirmation on the change in users' intentions, as well as the role of chronic disease. A

TABLE 3 Mean and variance analysis.

Mean and variance	Unstandardized estimate	S.E.
Intercept mean	6.042***	0.052
Slope mean	−0.273***	0.033
Intercept variance	−0.205***	0.065
Slope variance	0.143***	0.034

***$p < 0.001$.

latent growth model was used to analyze longitudinal survey data, and most of the hypotheses were supported. Here are a few vital findings.

First, this study offers several insights derived from its first research objective. Through a longitudinal analysis, the research demonstrated how users' intention to continue using mHealth services is dynamic and decreases over time with a negative slope. This result supports the notion that users' beliefs and intentions change across time (11). While users may have high expectations of mHealth services and a high initial intention to use them, the intention is likely to decline in subsequent use. Thus, these findings seem to buttress the mechanisms observed in cognitive dissonance theory (60), which indicates that individuals' behavior and cognition will adjust as time progresses to keep behavior and cognition in balance.

Second, in light of its second research objective, this study clarified the impacts of expectation and confirmation on the change in users' intention to continue using mHealth services. Consistent with prior research (61), this study validates that expectation is positively related to initial continuous use intention. Additionally, the results indicate that expectation shows positive impacts on the rate of change in continuous use intention. That is to say, when users have high expectations for mHealth services, their intention to continue using decreases more quickly. Contrary to expectation, however, this study finds that confirmation is negatively related to intention change rate, indicating that for users who have high confirmation, their intention declines more slowly. These results are consistent with the mechanism of the ECM. According to a previous study, confirmation leads to higher satisfaction and relatively higher use intention (31) and thus will slow the decline in continuous use intention. However, high expectation will cause low confirmation (41), ultimately leading to a faster decline in continuous use intention.

Third, considering the third research objective, the results suggested that chronic disease plays a vital role in mHealth services use. Specifically, compared to users with no chronic disease, chronic patients have higher initial continuous intention and their continuous use intention declines more slowly. Moreover, the results indicated that expectation showed a stronger impact on initial use intention but showed a weaker impact on the rate of change in intention for users with chronic disease than for users without chronic disease. In regard to the impacts of confirmation on the rate of change in intention, no significant difference exists between the two user groups. This finding may indicate that confirmation is critical for users both with and without chronic disease. Furthermore, this study found that there is a difference between the two groups with regard to the impact of expectation on confirmation: expectation has significant negative impact on confirmation for users with chronic disease, but it has no significant impact for users with no chronic disease.

5.2 Theoretical implications

First, this study advances the understanding of continuous use, particularly in the mHealth services context, by accounting for time-based dynamics. Although plenty of extant studies have contributed to an understanding of the post-adoption and continuous use of mHealth (62), they have primarily examined the continuous use intention at one time point under the implicit assumption that this intention is static. However, since users' experience with mHealth is dynamic, their attitude regarding the service may not be static, and continuous use intention may change during a long-term use process (34). For example, a user may initially have a high continuous intention but may then change their decision or even stop using the mHealth services after a period of time. Examining the change in continuous use intention helps us more deeply understand the sustainability of mHealth use from a whole process perspective. This study opens new theoretical views for research on the conceptualization of continuous mHealth services use. In other words, when researchers investigate the continuous use of mHealth, they are suggested to consider both users' initial continuous intentions and the change in their subsequent intentions.

Second, this study extends the ECM to a new dynamic model through examining the effects of expectation and confirmation on continuous use dynamics. As previously mentioned, several research gaps exist in previous studies that draw upon the ECM to explain continuous use of technologies, including a failure to address the dynamic nature of continuous use intention and a joint consideration of the impacts of pre- and post-beliefs. In this regard, the current conceptualization of the ECM cannot sufficiently capture the concept of development and change in users' intentions. Thus, we build a research model to investigate impacts of pre-use expectation and post-use confirmation on initial intention and changes of intention in multiple time windows. This research not only expands the ECM by involving changes of continuous use intention but reveals the different impacts of expectation and confirmation. Thus, this study advances in study of Bhattacherjee and Premkumar (17) and improves our understanding of ECM from a dynamic perspective. Moreover, this study also responds to the appeal of Lee et al. (11), who called for studies on the ECM to consider the dynamic nature of processes within information technology usage. The findings advocate that future research on ECM would be well served by paying more attention to both users' pre- and post-beliefs and to the dynamics of their continuous use intentions.

Third, this study investigates the role of chronic disease in continuous use related to mHealth services, which has rarely been examined. Although prior studies have explored the moderating effect of perceived vulnerability to chronic disease on users' beliefs and preferences regarding mHealth channels (22), none have empirically tested the role of chronic disease in long-term mHealth use. It is therefore necessary to diagnose the differences between users with chronic disease and users with no chronic disease in terms of the dynamics of their mHealth use intentions. The results of this study demonstrate that chronic disease leads to higher use intention and plays nuanced moderating roles in relationships between expectation, confirmation, and changes in use intention. These findings suggest that differences in users' health status (e.g., chronic disease) should be studied alongside user's beliefs on mHealth services. Furthermore, this research enriches the ECM in the

Variables and measurement items		Standardized factor loading	Cronbach's alpha	Composite reliability	Average variance extracted
SE	SE1	0.943	0.904	0.940	0.840
	SE2	0.874			
	SE3	0.930			
SC	SC1	0.973	0.967	0.978	0.937
	SC2	0.960			
	SC3	0.971			
CUI [T1]	CUI1[T1]	0.919	0.923	0.942	0.844
	CUI2[T1]	0.929			
	CUI3[T1]	0.908			
CUI [T2]	CUI1[T2]	0.956	0.952	0.968	0.909
	CUI2[T2]	0.950			
	CUI3[T2]	0.954			
CUI [T3]	CUI1[T3]	0.825	0.870	0.911	0.773
	CUI2[T3]	0.917			
	CUI3[T3]	0.893			

SE, service expectation; SC, service confirmation; CU, continuous use intention.

TABLE 5 Discriminant validity.

Variables and measurement items	SE	SC	CUI [T1]	CUI [T2]	CUI [T3]
SE	**0.919**				
SC	0.114	**0.968**			
CUI [T1]	0.396	0.040	**0.919**		
CUI [T2]	0.134	0.519	−0.005	**0.953**	
CUI [T3]	0.122	0.333	0.130	0.388	**0.879**

SE, service expectation; SC, service confirmation; CU, continuous use intention.
Bold values refer to the square roots of AVE.

mHealth context through introducing chronic disease as a contextual factor. The results support the viewpoint of Hong et al. (63), who stated that integrating contextual factors into research models is able to cause new explanations and elaborations of theoretical understanding.

Fourth, using a latent growth model, this study provides a methodological guidance for future research about the dynamics of user intentions and behaviors. Since it is well suited for measuring the dynamics of constructs, the method employed here can be used to examine longitudinal adoption and use of technologies and services. Unlike prior longitudinal studies, which investigated constructs in two time windows and analyzed them as cross-sectional without considering the time-based dynamics (64, 65), this study provides guidance for research investigating dynamics of constructs with multiple time windows and additional factors that may influence the dynamics. Furthermore, the statistical methodology inspires scholars to establish more dynamic models examining users' beliefs and technology use, thus making a theoretical contribution to the information systems and management discipline.

5.3 Practical implications

This research shows insightful significance for practice regarding ways to increase continuous use and improve the sustainability of mHealth. First, the results indicate that mHealth service providers should recognize that mHealth use intention is not static and decreases over time. Users may show great passion for using mHealth services but lose their enthusiasm for it gradually. Given that the high dropout rate is one of the most serious threats to the sustainability of mHealth (66), mHealth service providers should invest more effort in retaining existing users in addition to attracting new ones. For example, mHealth service providers can regularly investigate users' attitudes about the service and continually monitor their use behavior. Then, when the provider detects that a user's continuous use intention is dropping and that the user may become inactive, they can take immediate steps to retain them, such as surveying their expectation and optimizing the service to improve their confirmation.

Second, mHealth managers should note that although high expectation leads to high initial continuous intention, it accelerates the decrease of intention in the subsequent use process. Moreover, expectation is negatively related to confirmation, while confirmation is beneficial to continuous use. Thus, managers may want to strike a balance between building up users' expectations and forming confirmation. Especially in the advertising phase, setting very high expectations of mHealth services is not the best option. We recommend that mHealth managers shape users' expectations appropriately at the initial stage and then subsequently pay closer attention to improving users' confirmation. For instance, mHealth developers can update mHealth applications, provide some new functions, and optimize the interface to improve user experience and the users' confirmation of the expected performance.
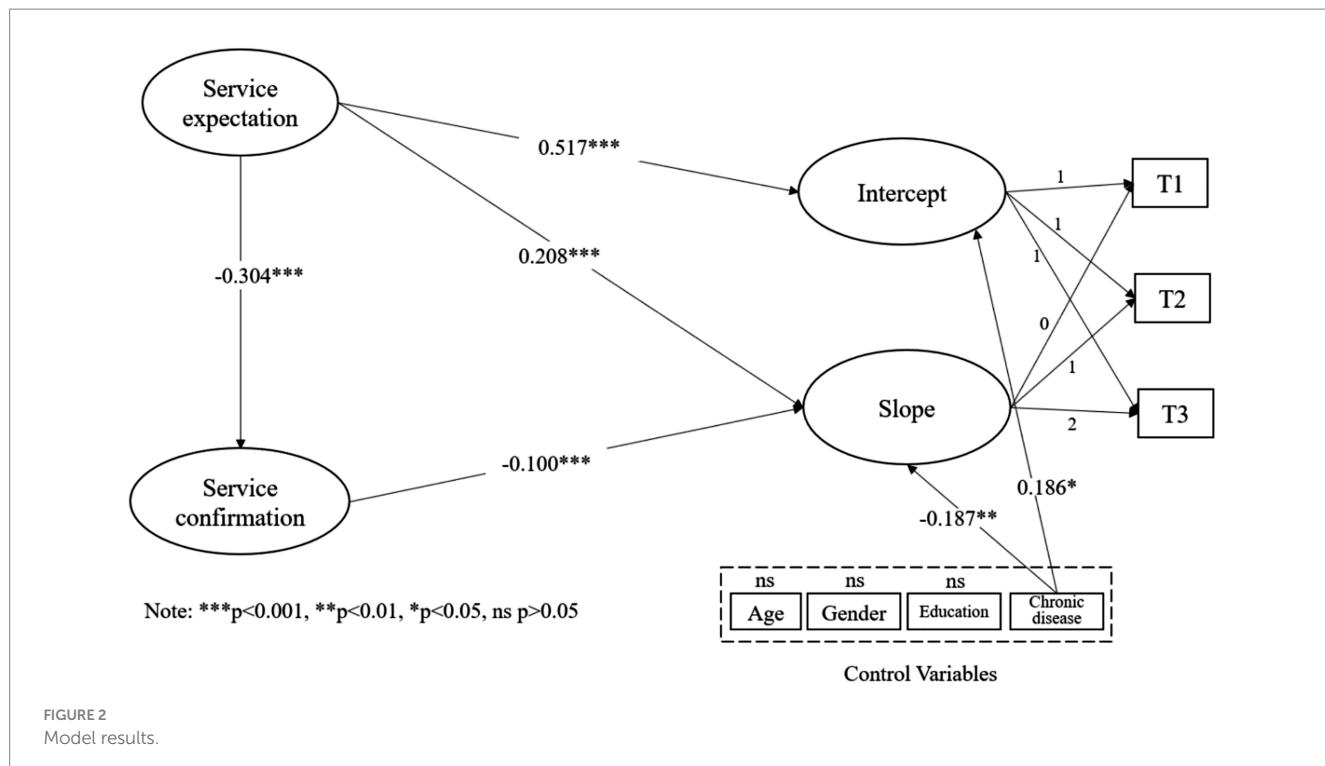
FIGURE 2
Model results.

TABLE 6 Nested model comparisons.

Model	Path	Coefficients of chronic disease/no chronic disease groups	$\Delta\chi^2$	Δdf	p-Value	Hypothesis support
M0 (χ^2 = 127.9, df = 72)						
M1	SE → SC	−0.447 / -0.291	12.7	1	0.000***	H5 Supported
M2	SE → Intercept	0.526 / 0.384	6.46	1	0.011*	H6 Supported
M3	SE → Slope	0.167 / 0.323	10.81	1	0.001**	H7 Supported
M4	SC → Slope	−0.102 / -0.107	2.74	1	0.098 ns	H8 Not supported

SE = service expectation, SC = service confirmation, ***p < 0.001, **p < 0.01, *p < 0.05, ns p > 0.05.

Third, mHealth service providers and designers should consider individual differences in the process of providing services. Results in this study indicate that individuals with chronic disease and those with no chronic disease engage differently with mHealth services in the long term. Specifically, users with no chronic disease are prone to cut back on the service, especially when their initial expectations are high. Given that the mHealth industry is shifting from health management to health enhancement, this study suggests mHealth service providers and designers to provide personalized services for different users. For example, specific artifact designs and retention strategies can be aligned with individual differences.

5.4 Limitations and future research

We acknowledge that there are some limitations to be addressed in future research. First, the data were collected from users of a single mHealth platform in China. Since people living in different cultural contexts may have different experiences of and beliefs about mHealth services, the generalizability of this study's findings for other populations remains to be explored in the future. Second, since the study mainly focused on the impacts of pre-adoption belief (e.g., expectation) and post-adoption belief (e.g., confirmation), only the direct effects of these two factors and the moderating role of chronic disease were examined. Future research could incorporate more elements; for instance, it could investigate other contextual factors in the research model. Third, this study was conducted in three time-windows and confirmation was examined in the second time window. We admit that there is a possibility that confirmation will change across time and we failed to identify the change. Future research could try to increase the number of data collection time-windows (at least four times) to measure the change of confirmation during this longer process. Finally, we examined users' intention to continue using mHealth services rather than actual use behaviors, which could be the target of future investigation.

6 Conclusion

This study investigates the dynamics of continuous use of mHealth services. Building a latent growth model with a longitudinal analysis, this study identifies the change trajectory of users' continuous use intention. Moreover, drawing upon the ECM, this study examines the impacts of expectation and confirmation on changes in continuous use. The results indicate that expectation is positively related to both initial use and the long-term changes in use, while confirmation is negatively related to changes in use. The analysis also shows that there are individual differences in the dynamics of continuous use as well as in the relationships between expectation, confirmation, and changes in use. The findings theoretically contribute to understanding the dynamic aspect of users' continuous use of mHealth services and expanding the ECM through introducing chronic disease as a moderator. This study also provides guidance for practitioners, suggesting that they should pay more attention to retaining current mHealth services users and developing strategies to facilitate their continuous use.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Sichuan University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

XM: Data curation, Project administration, Methodology, Conceptualization, Investigation, Funding acquisition, Writing – original draft. YL: Investigation, Writing – review & editing, Formal analysis, Visualization. AS: Resources, Investigation, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Chen S, Lai K-H, Guo X, Zhang X. The influence of digital health technology on the allocation of regional medical resources in China. *Health Policy Technol.* (2025) 14:101013. doi: 10.1016/j.hlpt.2025.101013

2. Zhang X, Lai K-H, Guo X. Promoting China's mHealth market: a policy perspective. *Health Policy Technol.* (2017) 6:383–8. doi: 10.1016/j.hlpt.2017.11.002

3. Zhang X, Pu J, Lu Y, Guo F. Team makes you better: evidence from online medical consultation platforms. *Inf Syst Res.* (2025). doi: 10.1287/isre.2022.0689

4. Zhang X., Wu Y., Valacich J., Jenkins J. (2025). Transforming Patient-Physician Interaction Through Asynchronous Online Health Interaction: A Relational Communication Perspective.

5. Zhang Y, Sun F, Zhang X, Ma X. Are multiple information sources better? The effect of multiple physicians in online medical teams on patient satisfaction. *Inf Process Manag.* (2025) 62:103889. doi: 10.1016/j.ipm.2024.103889

6. Wang L, Wu T, Guo X, Zhang X, Li Y, Wang W. Exploring mHealth monitoring service acceptance from a service characteristics perspective. *Electron Commer Res Appl.* (2018) 30:159–68. doi: 10.1016/j.elerap.2018.06.003

7. Zhang X, Guo X, Guo F, Lai K-H. Nonlinearities in personalization-privacy paradox in mHealth adoption: the mediating role of perceived usefulness and attitude. *Technol Health Care.* (2014) 22:515–29. doi: 10.3233/THC-140811

8. Zhang X, Yin C, Guo X, Lai K-H. The role of previous affective responses in the continuance of mHealth monitoring services: a longitudinal investigation. *IEEE Trans Eng Manag.* (2023) 71:6982–94. doi: 10.1109/TEM.2023.3275924

9. Guo X, Chen S, Zhang X, Ju X, Wang X. Exploring patients' intentions for continuous usage of mHealth services: elaboration-likelihood perspective study. *JMIR Mhealth Uhealth.* (2020) 8:e17258. doi: 10.2196/17258

10. Zhang X, Guo X, Wu Y, Lai K-H, Vogel D. Exploring the inhibitors of online health service use intention: a status quo bias perspective. *Inf Manag.* (2017) 54:987–97. doi: 10.1016/j.im.2017.02.001

11. Lee K, Yan A, Joshi K. Understanding the dynamics of users' belief in software application adoption. *Int J Inf Manag.* (2011) 31:160–70. doi: 10.1016/j.ijinfomgt.2010.07.009

12. Babar KA. A longitudinal perspective of mHealth services: the impact of trust and perceived usefulness. *Int J Bus Inf Syst.* (2024) 47:149–72. doi: 10.1504/IJBIS.2024.142288

13. Shen X.-L., Yang Y., Sun Y., Li Y.-J. (2020). "Intermittent Continuance of Smart Health Devices: A Zone-of-Tolerance Perspective." in Proceedings of the 53rd Hawaii International Conference on System Sciences.

14. Alshammari T, Messom C, Cheung Y. M-government continuance intentions: an instrument development and validation. *Inf Technol Dev.* (2022) 28:189–209. doi: 10.1080/02681102.2021.1928589

15. Venkatesh V, Goyal S. Expectation disconfirmation and technology adoption: polynomial modeling and response surface analysis. *MIS Q.* (2010) 34:281–303. doi: 10.2307/20721428

16. Oliver RL. A cognitive model of the antecedents and consequences of satisfaction decisions. *J Mark Res.* (1980) 17:460–9. doi: 10.1177/002224378001700405

17. Bhattacherjee A, Premkumar G. Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test. *MIS Q.* (2004) 28:229–54. doi: 10.2307/25148634

18. Zhang X, Guo X, Lai K-H, Guo F, Li C. Understanding gender differences in m-health adoption: a modified theory of reasoned action model. *Telemed e-Health.* (2014) 20:39–46. doi: 10.1089/tmj.2013.0092

19. Li Y, Ma X, Song J, Yang Y, Ju X. Exploring the effects of online rating and the activeness of physicians on the number of patients in an online health community. *Telemed e-Health.* (2019) 25:1090–8. doi: 10.1089/tmj.2018.0192

20. Zhang X, Xie K, Gu B, Guo X. Engaging physicians with introductory incentives: references to online and offline income. *MIS Q.* (2025) 49:917–52. doi: 10.25300/MISQ/2024/18172

21. Yang H, Li D. Understanding the dark side of gamification health management: a stress perspective. *Inf Process Manag.* (2021) 58:102649. doi: 10.1016/j.ipm.2021.102649

22. Chen L, Baird A, Rai A. Mobile health (mHealth) channel preference: an integrated perspective of approach-avoidance beliefs and regulatory focus. *J Assoc Inf Syst.* (2019) 20:1743–1773.

23. Ma X, Sun Y, Guo X, Lai K-H, Luo P. Understanding first aid knowledge adoption on social media with an extended information adoption model. *Internet Res.* (2025) 35:567–93. doi: 10.1108/INTR-08-2023-0651

24. Zhang X, Guo X, Lai K-H, Yin C, Meng F. From offline healthcare to online health services: the role of offline healthcare satisfaction and habits. *J Electron Commer Res.* (2017) 18:138–54.

25. Yoon C, Rolland E. Understanding continuance use in social networking services. *J Comput Inf Syst.* (2015) 55:1–8. doi: 10.1080/08874417.2015.11645751

26. Zhang X, Yan X, Cao X, Sun Y, Chen H, She J. The role of perceived e-health literacy in users' continuance intention to use mobile healthcare applications: an exploratory empirical study in China. *Inf Technol Dev.* (2018) 24:198–223. doi: 10.1080/02681102.2017.1283286

27. Huang G, Ren Y. Linking technological functions of fitness mobile apps with continuance usage among Chinese users: moderating role of exercise self-efficacy. *Comput Hum Behav.* (2020) 103:151–60. doi: 10.1016/j.chb.2019.09.013

28. Meng F, Guo X, Peng Z, Ye Q, Lai K-H. Trust and elderly users' continuance intention regarding mobile health services: the contingent role of health and technology anxieties. *Inf Technol People.* (2021) 35:259–80. doi: 10.1108/ITP-11-2019-0602

29. Yan M, Filieri R, Raguseo E, Gorton M. Mobile apps for healthy living: factors influencing continuance intention for health apps. *Technol Forecast Soc Change.* (2021) 166:120644. doi: 10.1016/j.techfore.2021.120644

30. Cho J. The impact of post-adoption beliefs on the continued use of health apps. *Int J Med Inform.* (2016) 87:75–83. doi: 10.1016/j.ijmedinf.2015.12.016

31. Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. *MIS Q.* (2001) 25:351–70. doi: 10.2307/3250921

32. Meyer J., Schnauber J., Heuten W., Wienbergen H., Hambrecht R., Appelrath H.-J., et al. (2016). "Exploring longitudinal use of activity trackers." in 2016 Ieee International Conference on Healthcare Informatics (Ichi). IEEE, pp. 198–206.

33. Hsiao J-L, Chen R-F. Understanding determinants of health care professionals' perspectives on mobile health continuance and performance. *JMIR Med Inform.* (2019) 7:e12350. doi: 10.2196/12350

34. Shen X-L, Li Y-J, Sun Y. Wearable health information systems intermittent discontinuance: a revised expectation-disconfirmation model. *Ind Manag Data Syst.* (2018) 118:506–23. doi: 10.1108/IMDS-05-2017-0222

35. Bawack RE, Ahmad MO. Understanding business analytics continuance in agile information system development projects: an expectation-confirmation perspective. *Inf Technol People.* (2021) 34:1551–69. doi: 10.1108/ITP-10-2020-0681

36. Tam C, Santos D, Oliveira T. Exploring the influential factors of continuance intention to use mobile apps: extending the expectation confirmation model. *Inf Syst Front.* (2020) 22:243–57. doi: 10.1007/s10796-018-9864-5

37. Huang Y-M. Examining students' continued use of desktop services: perspectives from expectation-confirmation and social influence. *Comput Hum Behav.* (2019) 96:23–31. doi: 10.1016/j.chb.2019.02.010

38. Jumaan IA, Hashim NH, Al-Ghazali BM. The role of cognitive absorption in predicting mobile internet users' continuance intention: an extension of the expectation-confirmation model. *Technol Soc.* (2020) 63:101355. doi: 10.1016/j.techsoc.2020.101355

39. Park E. User acceptance of smart wearable devices: an expectation-confirmation model approach. *Telemat Inform.* (2020) 47:101318. doi: 10.1016/j.tele.2019.101318

40. Gupta A, Yousaf A, Mishra A. How pre-adoption expectancies shape post-adoption continuance intentions: an extended expectation-confirmation model. *Int J Inf Manag.* (2020) 52:102094. doi: 10.1016/j.ijinfomgt.2020.102094

41. Serrano CI, Shah V, Abràmoff MD. Use of expectation disconfirmation theory to test patient satisfaction with asynchronous telemedicine for diabetic retinopathy detection. *Int J Telemed Appl.* (2018) 2018:14.

42. Thong JY, Hong S-J, Tam KY. The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance. *Int J Hum Comp Stud.* (2006) 64:799–810. doi: 10.1016/j.ijhcs.2006.05.001

43. Chopdar PK, Korfiatis N, Sivakumar V, Lytras MD. Mobile shopping apps adoption and perceived risks: a cross-country perspective utilizing the unified theory of acceptance and use of technology. *Comput Hum Behav.* (2018) 86:109–28. doi: 10.1016/j.chb.2018.04.017

44. Guiso L, Sapienza P, Zingales L. Time varying risk aversion. *J Financ Econ.* (2018) 128:403–21. doi: 10.1016/j.jfineco.2018.02.007

45. Ma X, Sun Y, Guo X, Lai K-H, Wang X. Learning first aid knowledge not only for the self but also for others: toward a collective protection motivation theory. *Health Commun.* (2025) 40:1907–20. doi: 10.1080/10410236.2024.2428880

46. Hu H, Liang H, Wang H. Longitudinal study of the earliest pilot of tiered healthcare system reforms in China: will the new type of chronic disease management be effective? *Soc Sci Med.* (2021) 285:114284. doi: 10.1016/j.socscimed.2021.114284

47. Megari K. Quality of life in chronic disease patients. *Health Psychol Res.* (2013) 1:932. doi: 10.4081/hpr.2013.932

48. Ho SY, Guo X, Vogel D. Opportunities and challenges in healthcare information systems research: caring for patients with chronic conditions. *Commun Assoc Inf Syst.* (2019) 44:852–73. doi: 10.17705/1CAIS.04439

49. Lin C-P, Tsai YH, Chiu C-K. Modeling customer loyalty from an integrative perspective of self-determination theory and expectation–confirmation theory. *J Bus Psychol.* (2009) 24:315–26. doi: 10.1007/s10869-009-9110-8

50. Watkins JOTA, Goudge J, Gómez-Olivé FX, Griffiths F. Mobile phone use among patients and health workers to enhance primary healthcare: a qualitative study in rural South Africa. *Soc Sci Med.* (2018) 198:139–47. doi: 10.1016/j.socscimed.2018.01.011

51. Mcleod S. (2007). Maslow's hierarchy of needs. Simply psychology, 1.

52. Hicks JR. A revision of demand theory. Oxford: Oup Catalogue (1986).

53. Lien C-H, Cao Y, Zhou X. Service quality, satisfaction, stickiness, and usage intentions: an exploratory evaluation in the context of WeChat services. *Comput Hum Behav.* (2017) 68:403–10. doi: 10.1016/j.chb.2016.11.061

54. Duncan TE, Duncan SC, Strycker LA. An introduction to latent variable growth curve modeling: Concepts, issues, and application. Mahwah, NJ: Routledge (2013).

55. Kwon J, Lee H. Why travel prolongs happiness: longitudinal analysis using a latent growth model. *Tour Manag.* (2020) 76:103944. doi: 10.1016/j.tourman.2019.06.019

56. Kline RB. Software review: software programs for structural equation modeling: Amos, Eqs, and Lisrel. *J Psychoeduc Assess.* (1998) 16:343–64. doi: 10.1177/073428299801600407

57. Fornell C, Larcker DF. Structural equation models with unobservable variables and measurement error: algebra and statistics. Los Angeles, Ca: Sage (1981).

58. Gerbing DW, Anderson JC. Monte Carlo evaluations of goodness of fit indices for structural equation models. *Sociol Methods Res.* (1992) 21:132–60. doi: 10.1177/0049124192021002002

59. Byrne BM. Testing for multigroup invariance using Amos graphics: a road less traveled. *Struct Equ Model.* (2004) 11:272–300. doi: 10.1207/s15328007sem1102_8

60. Harmon-Jones E, Harmon-Jones C. Cognitive dissonance theory after 50 years of development. *Z Sozialpsychol.* (2007) 38:7–16. doi: 10.1024/0044-3514.38.1.7

61. Joo YJ, Park S, Shin EK. Students' expectation, satisfaction, and continuance intention to use digital textbooks. *Comput Hum Behav.* (2017) 69:83–90. doi: 10.1016/j.chb.2016.12.025

62. Chiu W, Cho H, Chi CG. Consumers' continuance intention to use fitness and health apps: an integration of the expectation–confirmation model and investment model. *Inf Technol People.* (2020) 34:978–98. doi: 10.1108/ITP-09-2019-0463

63. Hong W, Chan FK, Thong JY, Chasalow LC, Dhillon G. A framework and guidelines for context-specific theorizing in information systems research. *Inf Syst Res.* (2014) 25:111–36. doi: 10.1287/isre.2013.0501

64. Chen X, Li Y, Davison RM, Liu Y. The impact of imitation on Chinese social commerce buyers' purchase behavior: the moderating role of uncertainty. *Int J Inf Manag.* (2021) 56:102262. doi: 10.1016/j.ijinfomgt.2020.102262

65. Ding Y. Modelling continued use of information systems from a forward-looking perspective: antecedents and consequences of hope and anticipated regret. *Inf Manag.* (2018) 55:461–71. doi: 10.1016/j.im.2017.11.001

66. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. *J Med Internet Res.* (2020) 22:e20283. doi: 10.2196/20283

66. Alzahrani AI, AL-Samarraie H, Eldenfria A, Dodoo JE, Alalwan N. Users' intention to continue using mHealth services: A DEMATEL approach during the COVID-19 pandemic. *Technology in society.* (2022) 68:101862.

Evaluating the quality of educational TikTok videos on diabetic retinopathy: a cross-sectional study

Ligang Jiang[1], Wencan Wu[2,3,4], Xin Jiang[5]* and Fangzheng Jiang[1]*

[1]Department of Ophthalmology, Quzhou Affiliated Hospital of Wenzhou Medical University, Quzhou People's Hospital, Quzhou, Zhejiang, China, [2]The Eye Hospital, School of Ophthalmology & Optometry, Wenzhou Medical University, Wenzhou, China, [3]Oujiang Laboratory (Zhejiang Lab for Regenerative Medicine Vision and Brain Health), Wenzhou, Zhejiang, China, [4]Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou, China, [5]Quzhou College of Technology, Quzhou, Zhejiang, China

Background: Diabetic retinopathy (DR) is a leading cause of vision loss among working-age adults, and enhancing public health literacy through effective education is crucial for its prevention and management. With the rise of mobile internet and short video platforms such as TikTok, new opportunities have emerged for disseminating medical knowledge. However, concerns remain regarding the accuracy and quality of this content.

Methods: A cross-sectional search was conducted on September 18, 2025. A total of 200 Mandarin-language TikTok videos directly relevant to DR were included after screening. Baseline characteristics, uploader type, and user engagement metrics were extracted. Video quality was assessed independently by two trained reviewers using the DISCERN tool and the Patient Education Materials Assessment Tool for Audiovisual Materials (PEMAT-A/V). Content coverage was evaluated against the American Academy of Ophthalmology (AAO) Preferred Practice Pattern®. Inter-rater reliability was measured by intraclass correlation coefficients (ICCs). Group comparisons and correlation analyses were performed.

Results: Significant differences were observed in quality scores across uploader categories (one-way ANOVA, $p < 0.001$). Non-profit organizations achieved the highest DISCERN scores (59.4 ± 8.2) and PEMAT-A/V understandability (88.5%), while for-profit accounts had the lowest DISCERN scores (23.0 ± 6.5; understandability 61.5%). Videos from non-profit sources also demonstrated balanced coverage across six core DR themes (14–20% per theme). Inter-rater reliability was excellent for all tools (ICC range 0.825–0.933). Engagement metrics were strongly correlated with DISCERN scores (likes $r = 0.76$, comments $r = 0.64$, favorites $r = 0.73$, shares $r = 0.71$; all $p < 0.05$), whereas video duration showed no significant correlation with quality ($p > 0.05$).

Conclusion: The quality of DR-related educational short videos on TikTok varies widely, with the source of the video emerging as the key determinant. High-quality content from non-profit organizations and medical professional users not only demonstrates greater reliability but also fosters comprehensive health education. Strengthening professional participation, platform regulation, and evidence-informed communication strategies is essential to maximize the potential of short videos in DR health education and ultimately improve patient outcomes.

1 Introduction

Diabetic Retinopathy (DR), a principal microvascular complication of diabetes, represents the foremost cause of irreversible blindness in working-age adults globally (1, 2). As the global prevalence of diabetes continues to climb, the incidence and morbidity of DR have correspondingly increased, creating a substantial public health and socioeconomic burden (3–5). Nevertheless, the progression of DR can be significantly delayed or prevented through early screening, prompt intervention, and tight control of both glycemic and blood pressure levels (6–8). Therefore, it is imperative that both patients and the general public possess adequate awareness and health literacy concerning DR. Consequently, the provision of high-quality, comprehensible, and readily accessible public health education is of paramount importance for the prevention and control of DR.

With the recent and rapid advancement of mobile internet technologies, short videos, such as those on YouTube and TikTok have emerged as the most rapidly expanding medium for information dissemination (9–13). short videos hold prominent advantages over traditional formats such as text with images or long-form videos, characterized by their rapid dissemination, extensive reach, bite-sized content, and engaging visual presentation. This medium has substantially reduced the threshold for obtaining health information (14, 15), establishing itself as a crucial new avenue for the public to gain medical knowledge on topics including DR (16–18). Moreover, physicians and healthcare organizations have progressively leveraged these platforms for science popularization to enhance public health consciousness (19).

Despite the great potential of short videos in education (20), the quality, accuracy, and reliability of their content face severe challenges (10, 21). Due to the lack of strict content review mechanisms, these platforms are inundated with a large volume of short videos produced by non-professionals or for commercial marketing purposes (22). This content may contain misleading information, inaccurate diagnostic and treatment advice, or even erroneous health concepts (23–25), which, for DR patients and their families who need to follow long-term and precise management plans, could lead to anxiety, delayed medical consultation, or the adoption of improper treatments, ultimately endangering their vision.

However, to date, there have been no systematic and objective quality assessment studies on health education content about the specific and important disease of DR on mainstream short video platforms. This study aims to use a cross-sectional design to systematically retrieve and evaluate the information quality and reliability of education short videos about DR on current mainstream short video platforms.

2 Materials and methods

2.1 Search strategy and video selection

On September 18, 2025, to minimize interference from personalized recommendation algorithms, we cleared the browser history and cookies before the search. We entered the keyword "diabetic retinopathy" (糖尿病视网膜病变) into TikTok (the mainland China version). In accordance with previous research practices and to ensure a representative sample, the top 210 videos sorted by the platform's default ranking were selected for screening. The resulting videos were then assessed for eligibility against predefined inclusion criteria by two researchers, with those directly relevant to DR being included for final analysis (Figure 1).

- Inclusion criteria: Videos were included if they described any of the following aspects of DR: definition, classification, symptoms, risk factors, diagnostic, treatment, management, or prognosis.
- Exclusion criteria: Videos were excluded if they were duplicates, unrelated to DR education, not in Mandarin Chinese, silent, or not targeted at a patient audience.
- Screening process: Two researchers independently assessed each video for eligibility. Any disagreements were resolved through discussion or, if necessary, by a third researcher to reach a consensus.

2.2 Baseline data extraction

Two trained researchers independently and systematically extracted baseline characteristics from each included video and recorded the data in a pre-designed, structured spreadsheet. Extracted variables included the unique video identifier (URL), upload date, uploader source, video duration, account registration country and region, and video title. User engagement metrics were also recorded, including the number of likes, comments, favorites, and shares. All information was manually extracted from publicly available TikTok webpages. Only aggregated, non-identifiable data (e.g., counts of likes, comments, favorites, and shares) were recorded; no personal identifiers or user-level data were collected or stored. The study procedures were conducted in compliance with TikTok's Terms of Service and Privacy Policy, and the analysis was limited to publicly accessible, de-identified data.

Based on the uploader's account name and verification status, videos were categorized into four groups: Medical professional users (ophthalmologists or other healthcare professionals). Non-medical users (general public and patients). For-profit organizations (e.g., corporations, commercial media). Non-profit organizations (e.g., healthcare institutions, news agencies). This classification was based on the account's verification information and profile description.

2.3 Video assessment

To ensure the scientific validity of the evaluation, this study used the *Diabetic Retinopathy Preferred Practice Pattern*® published by the American Academy of Ophthalmology (AAO) as the standard for content assessment (26). The videos were analyzed based on two dimensions: information coverage and quality. Two researchers with backgrounds in DR research first jointly reviewed the top 10 ranked
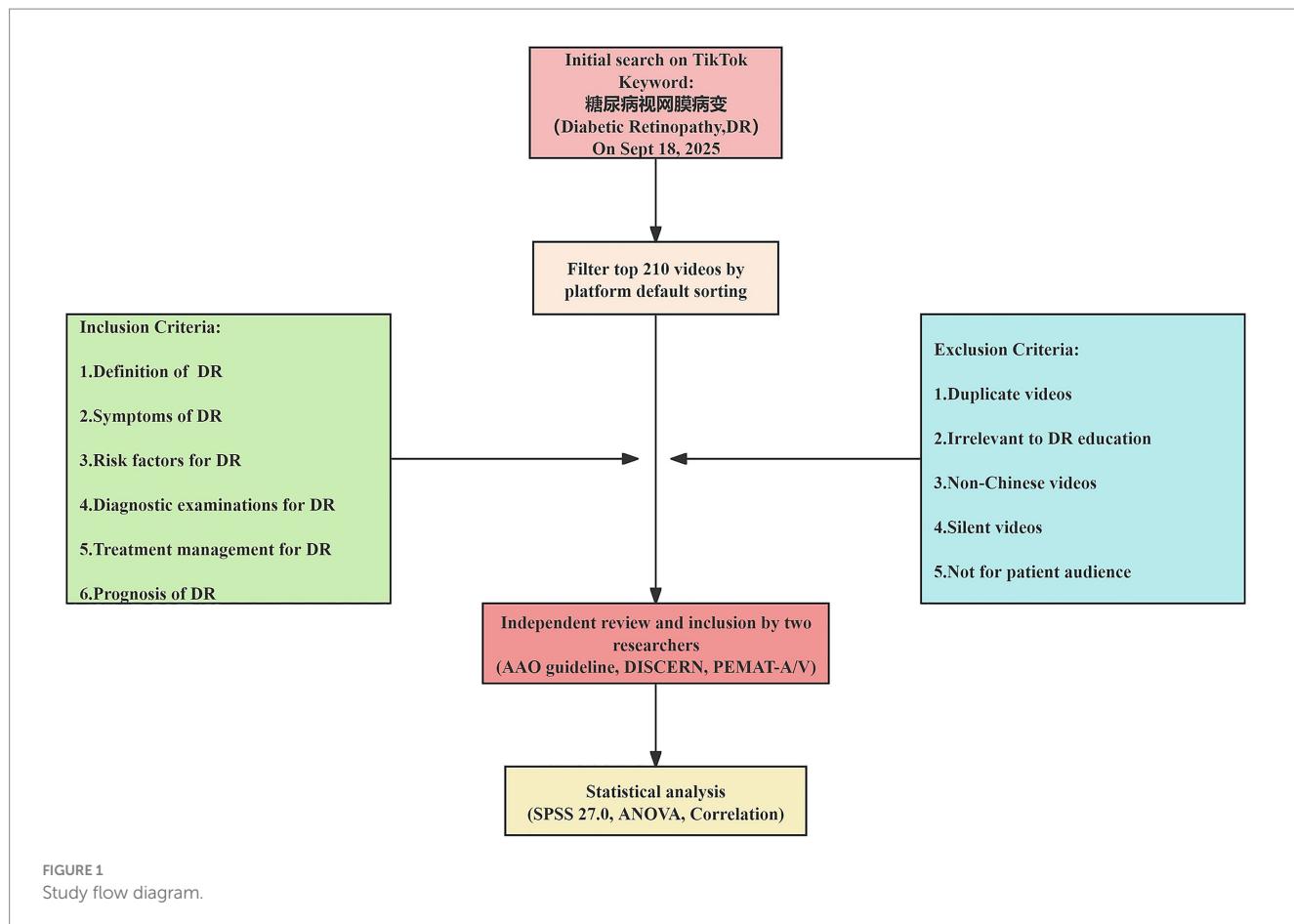
FIGURE 1
Study flow diagram.

videos to familiarize themselves with the framework and standardize their understanding, after which they independently conducted the formal assessment for all included videos. Content completeness was assessed using a framework adapted from Goobie et al. (27) for evaluating educational videos, which covered six core aspects: DR definition, DR classification, DR symptoms, DR risk factors, DR diagnostic methods, DR treatment and management.

Video quality was assessed using two validated tools: The DISCERN tool (16 items, divided into reliability, treatment information, and overall quality) (28, 29); The Patient Education Materials Assessment Tool for Audiovisual Materials (PEMAT-A/V) (17 items, divided into understandability and actionability), which was scored as "agree," "disagree," or "not applicable" (11, 30). Two assessors independently scored all videos. Before scoring, they received standardized training and referred to the latest AAO guidelines to ensure consistency. The average of the two assessors' scores was used for the final statistical analysis, and the intraclass correlation coefficient (ICC) was calculated.

2.4 Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, Version 27.0 (IBM Corp., Armonk, NY, USA). The Shapiro–Wilk test was used to evaluate the normality of continuous variables. Normally distributed continuous variables were expressed as mean ± standard deviation (SD), whereas non-normally distributed

variables were presented as median and interquartile range (IQR). Categorical data are presented as counts and percentages (n, %). Inter-rater reliability between the two reviewers was evaluated using the ICC. Differences in quality scores across uploader categories were first analyzed using ANOVA. When the assumption of homogeneity of variances was violated, Welch's ANOVA or the Kruskal–Wallis test was applied, as appropriate. Whenever an omnibus test was statistically significant, pairwise *post hoc* comparisons between uploader groups were conducted using Bonferroni-adjusted significance levels to control for type I error (for parametric tests) or Dunn's post hoc test with Bonferroni correction (for non-parametric data). Pearson's correlation coefficient was used to assess the relationship between video characteristics and quality scores for normally distributed variables, whereas Spearman's rank correlation coefficient was used for non-normally distributed variables. A two-tailed p value < 0.05 after adjustment for multiple comparisons was considered statistically significant for all analyses.

3 Results

3.1 Analysis of basic video parameters

A total of 210 TikTok videos were initially identified using the keyword "diabetic retinopathy." After excluding 10 videos that did not meet the inclusion criteria (4 duplicates, 2 videos unrelated to DR education, 3 non-Mandarin videos, and 1 silent video without

narration), 200 DR-related educational short videos were included in the final analysis. All included videos were verified by the platform. The majority of these videos were uploaded by medical users (174/200, 87.0%), followed by non-profit organizations (10/200, 5.0%). Non-medical users and for-profit organizations each accounted for the fewest videos (8/200, 4.0% each). Across all videos, the median numbers of likes, comments, favorites, and shares were 274.0, 15.0, 87.0, and 52.0, respectively, and the median duration was 70.5 s. Overall, there were significant differences in user engagement based on the uploader type, with videos from non-profit organizations and medical users showing higher engagement metrics (see Table 1 for details).

3.2 Video quality analysis

The inter-rater agreement was high for video reliability (ICC = 0.886), treatment choice (ICC = 0.933), DISCERN tool scores (ICC = 0.871), overall quality score (ICC = 0.879), understandability (ICC = 0.839), and actionability (ICC = 0.825) (all $p < 0.001$). In the subgroup analysis, which assessed video quality by uploader source, the results showed significant differences among user types across all quality assessment dimensions ($p < 0.001$) (Table 2). For video reliability, DISCERN tool scores, and overall quality score, videos from non-profit organizations scored the highest (32.20, 59.40, and 4.00, respectively), followed by those from medical users (25.00, 48.00, and 3.00, respectively). Videos from for-profit organizations had the lowest scores on these metrics (12.00, 23.00, and 1.00, respectively), See Table 2 and Figures 2A–D for details. A similar trend was observed in the assessment using the PEMAT-A/V tool. Videos from non-profit organizations had the highest understandability scores (88.46%), closely followed by those from medical users (84.62%). In terms of actionability, videos from medical users and non-profit organizations tied for the highest score (both 75.00%), whereas videos from for-profit organizations scored significantly lower (33.33%) (Table 2 and Figures 2E,F).

As shown in Figure 3, clear differences were observed in the proportional distribution of scores across uploader groups. Regarding Reliability, the Non-profit and Medical groups accounted for 32.5 and 25.6% of the total, respectively, whereas the Non-medical and For-profit groups contributed only 17.8 and 12.1%. For Treatment choice, the Non-profit and Medical groups contributed 30.0 and 26.1%, compared with only 17.1 and 14.8% for the Non-medical and For-profit groups. For Understandability, the Non-profit and Medical groups accounted for 26.0 and 24.1%, whereas the Non-medical and For-profit groups accounted for 20.6 and 17.2%. For Actionability, the Non-profit and Medical groups contributed 28.5 and 25.7%, while the Non-medical and For-profit groups contributed only 21.5 and 12.4%. Across all four dimensions, Non-profit and Medical uploaders consistently accounted for a larger proportion of the quality scores, whereas Non-medical and For-profit uploaders lagged behind.

3.3 Video content analysis

There were significant differences in the content theme distribution strategies among the different video sources. Specifically, the content strategies of medical professional users, non-medical users, and for-profit organizations were similar, with all three groups tending to prioritize practical topics. In content from medical professional users, "Management" (28.4%) and "Symptom" (24.8%) were the main components, whereas content from non-medical and for-profit sources was more focused, with the "Management" theme accounting for over 35% for both groups. In contrast, content from non-profit sources had the most balanced distribution, with coverage rates for the six themes all falling between 14 and 20%, indicating a more comprehensive educational strategy. In summary, videos from medical professional users, non-medical users, and for-profit organizations prioritized practical topics like "Management" and "Symptom," while content from non-profit sources was more comprehensive and balanced across all basic knowledge themes (Figure 4).

3.4 Correlation analysis

Correlation analysis revealed significant positive correlations among the engagement metrics for DR-related videos on TikTok. Specifically, the number of likes was strongly and positively correlated with the number of comments ($r = 0.83$, $p < 0.05$), favorites ($r = 0.94$, $p < 0.05$), and shares ($r = 0.89$, $p < 0.05$). video duration showed a weak positive correlation with all engagement

TABLE 1 Basic parameters of videos.

Video parameters	Individual user		Organizational user		Overall
	Medical ($n = 174$)	Non-medical ($n = 8$)	For-profit ($n = 8$)	Non-profit ($n = 10$)	
Likes	304.50 (106.25, 855.00)	18.50 (14.25, 31.50)	17.00 ± 6.48	18,000.00 (10,680.25, 59,250.00)	274.00 (87.25, 906.50)
Comments	16.00 (8, 50.25.00)	3.5 (1.25, 12.00)	1.38 ± 0.92	633.00 (345.75, 3,882.50)	15.00 (7.00, 54.75)
Favorites	100.00 (36.00, 258.25)	5.75 ± 4.30	2.75 ± 2.32	5,292.50 (3,135.00, 17,000.00)	87.00 (26.25, 300.25)
Shares	57.50 (16.75, 213.00)	3.50 ± 2.83	4.75 ± 3.85	6,215.50 (1,726.00, 29,000.00)	52.00 (12.25, 232.75)
Times	71.00 (47.00, 123.00)	80.25 ± 35.78	48.25 ± 15.92	121.50 (51.50, 218.00)	70.50 (47.00, 122.50)

TABLE 2 Scores by different video sources.

Variables	Individual users		Organization users		p Value
	Medical (n = 174)	Non-medical (n = 8)	For-profit (n = 8)	Non-profit (n = 10)	
Videos reliability (n = 200)	25.00 (24.00, 27.00)	17.63 ± 1.77	12.00 ± 0.71	32.20 ± 2.04	<0.001
Treatment choice (n = 200)	20.00 (19.00, 21.00)	13.00 (13.00, 13.75)	11.40 ± 0.89	23.20 ± 2.20	<0.001
DISCERN tool scores (n = 200)	48.00 (47.00, 50.00)	32.38 ± 2.13	23.00 (21.50, 25.00)	59.40 ± 1.90	<0.001
Overall quality score (n = 200)	3.00 (2.00, 4.00)	1.50 (1.00, 2.00)	1.00 (1.00, 2.00)	4.00 ± 0.67	<0.001
PEMAT-A/V understandability (%)	84.62 (76.92, 92.31)	70.19 ± 12.63	58.65 ± 16.41	88.46 ± 6.54	<0.001
PEMAT-A/V actionability (%)	75.00 (66.67, 75.00)	61.46 ± 15.39	33.33 (27.08, 45.83)	75.00 (75.00, 100.00)	<0.001

FIGURE 2
Intergroup comparison among four different video source users. (A) Reliability score. (B) Treatment choice score. (C) Overall quality score.
(D) DISCERN tool score. (E) Understandability score. (F) Actionability score.
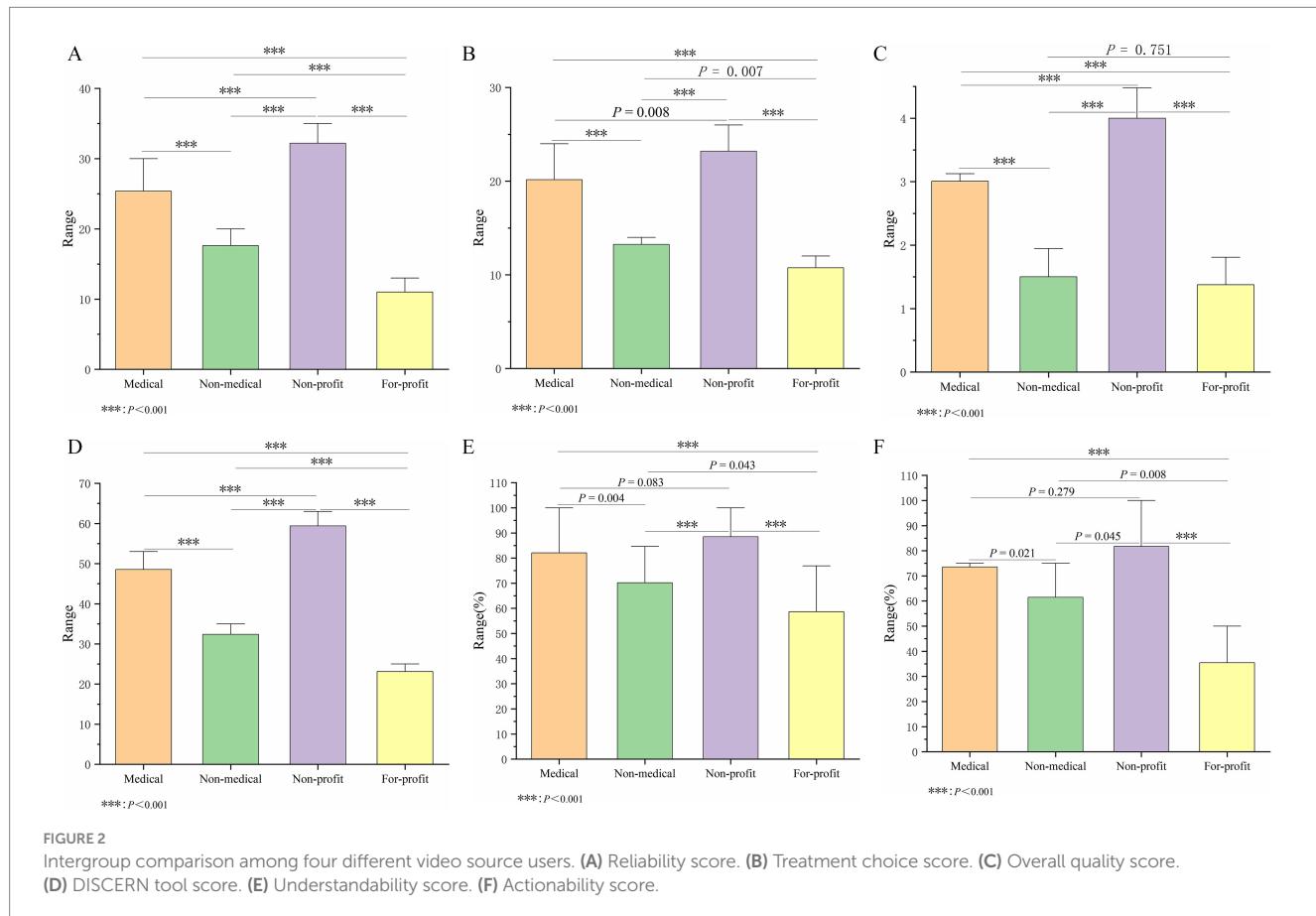
metrics ($r = 0.32$–0.39, $p < 0.05$). Among the medical information quality metrics, video reliability was moderately and positively correlated with the number of likes ($r = 0.60$, $p < 0.05$), comments ($r = 0.45$, $p < 0.05$), favorites ($r = 0.57$, $p < 0.05$), and shares ($r = 0.52$, $p < 0.05$), but it showed no significant correlation with video duration. The overall quality score, understandability, and actionability were all weakly and positively correlated with video engagement metrics ($p < 0.05$). The DISCERN tool score was strongly and positively correlated with the number of likes ($r = 0.76$, $p < 0.05$), comments ($r = 0.64$, $p < 0.05$), favorites ($r = 0.73$, $p < 0.05$), and shares ($r = 0.71$, $p < 0.05$), but had no significant correlation with video duration (Figure 5).

4 Discussion

To the best of our knowledge, this is the first study to systematically assess the quality and reliability of educational short video content concerning DR on the TikTok platform. The results indicate that although DR-related educational short videos are active in terms of user engagement and dissemination, their overall quality varies significantly. Notably, videos published by for-profit accounts generally had low information accuracy and educational value. In contrast, videos from non-profit organizations and medical professional users scored higher in terms of reliability, understandability, and actionability, demonstrating greater potential for health education.

Kong et al. (31) noted in their study of diabetes-related videos on TikTok that videos uploaded by non-profit organizations performed best on the DISCERN score, while those from commercial organizations scored the lowest. Analogously, the analysis by Wu et al. (32) of videos concerning hypertension and diabetes treatment on WeChat and TikTok revealed that the majority of videos were of low quality and did not satisfy established standards, indicating a pressing need to enhance the scientific rigor of medical information, even on platforms with vast numbers of users. The existing literature widely

agrees that social media and short video platforms have great potential for disseminating medical and health information (33), but the quality of their content is inconsistent. In particular, videos produced by non-medical users or commercial accounts often contain one-sided or even erroneous information that may mislead patients (34). A similar trend was observed in a study on educational short videos about dry eye disease (35): although the videos had high understandability, their actionability was insufficient, and the coverage of basic knowledge was incomplete.

The findings of the present study align with the conclusions of previous research. We observed that videos concerning DR that were uploaded by non-profit organizations were superior to those from other sources in several aspects; they not only received the highest scores on both the DISCERN and PEMAT-A/V tools but also provided more thorough content coverage. In contrast to content from medical or commercial sources that frequently concentrated on pragmatic subjects like "management" and "symptoms," non-profit organizations showed a greater tendency to systematically address the six crucial themes of DR, including its definition, classification, symptoms, risk factors, diagnosis, and management. Such a holistic and well-rounded educational approach is capable of balancing scientific rigor with effective dissemination, which aids in building a complete conceptual understanding of the disease for patients and consequently holds greater value in enhancing health literacy.

This study further revealed a moderate-to-strong positive correlation between user engagement indicators (e.g., likes, comments, favorites, shares) and information quality scores, which indicates that higher-quality content is more readily endorsed and shared by users (36). This result implies that the audience has some capacity for discernment; however, the extent of a video's dissemination is likely still heavily influenced by its visibility and the platform's algorithms (37). It is noteworthy that video duration were not significantly correlated with


FIGURE 3
Comparison of the quality of the videos from four sources.

FIGURE 4
Videos from different sources cover various aspects of DR.
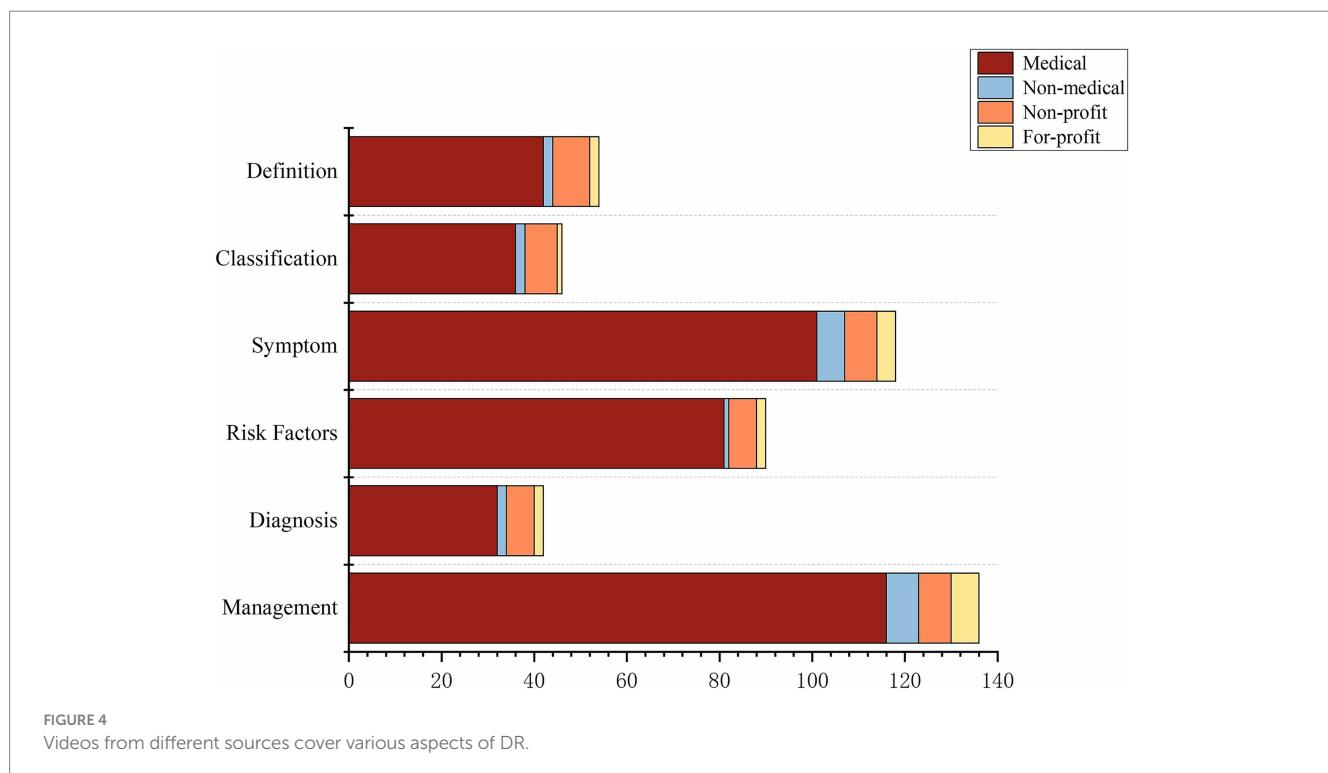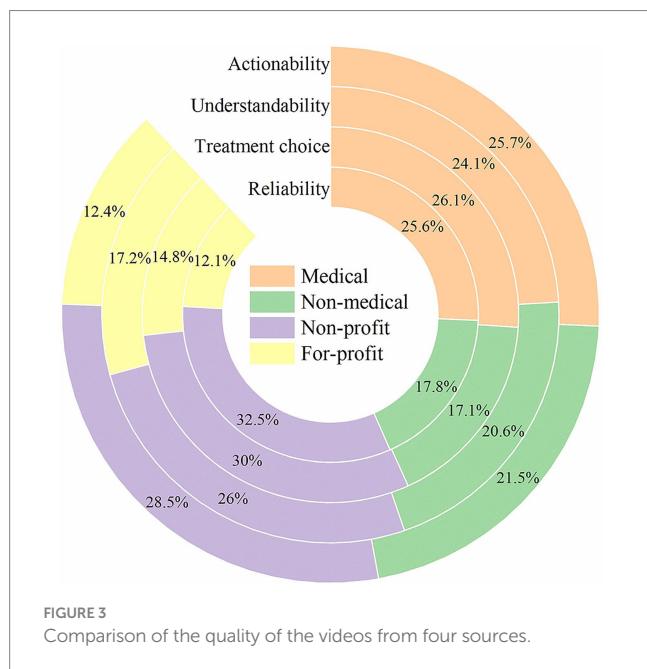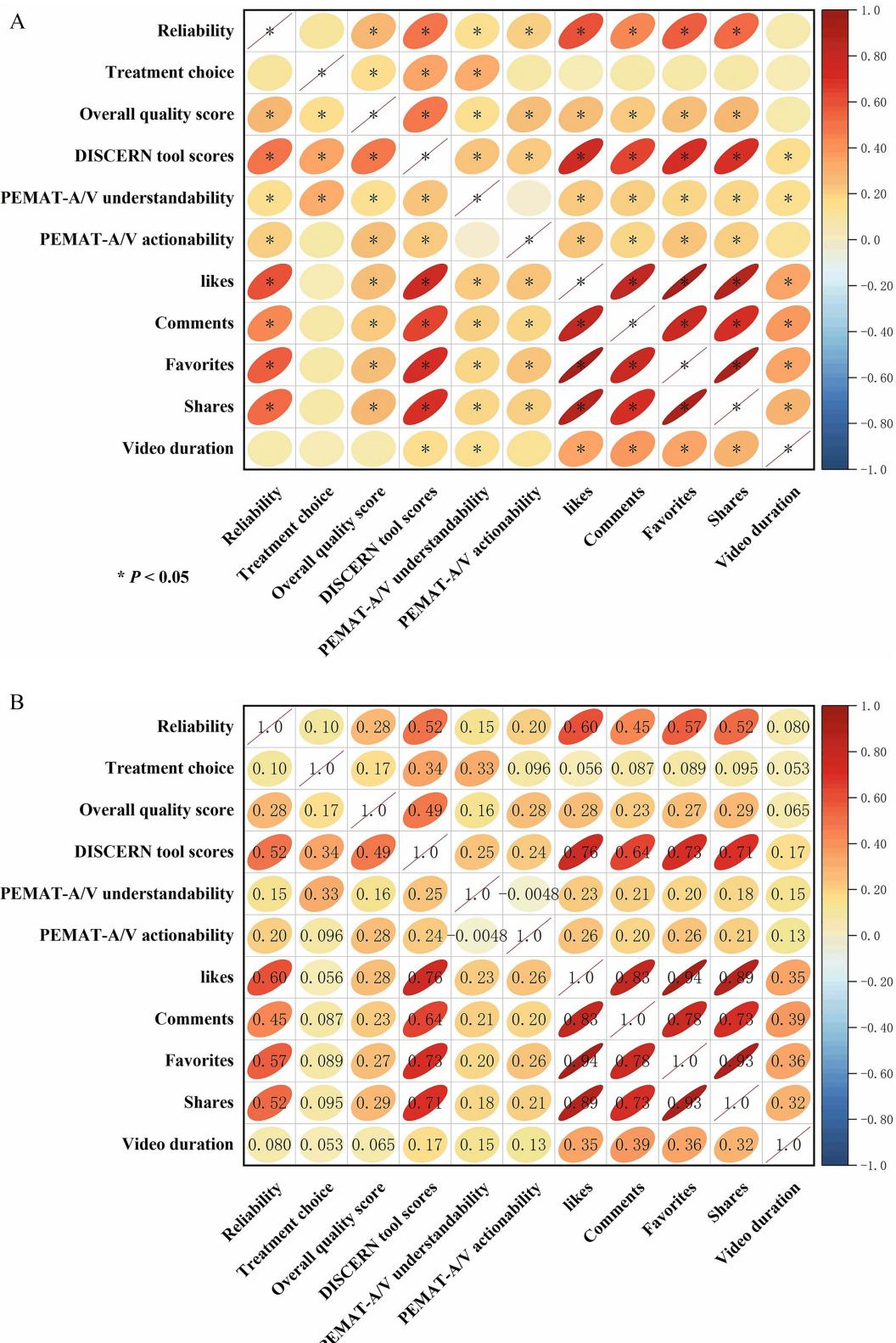
FIGURE 5
Correlation matrix illustrating the relationships between video quality assessment scores and user engagement metrics. **(A)** Visualization of the correlation coefficients. The color intensity and the shape of the ellipses represent the strength of the correlation; red indicates a positive correlation. Asterisks (*) denote a statistically significant correlation ($p < 0.05$). **(B)** Detailed matrix showing the specific correlation coefficients (r-values) in the upper triangle and the corresponding visual representation in the lower triangle.

quality, indicating that the scientific rigor and the method of presentation are more crucial factors than duration alone. Consequently, for both platform promotion efforts and the health education practices of physicians, it is essential to focus on organically combining scientifically rigorous content with effective communication strategies (38, 39).

Notably, DR is a chronic, progressive microvascular complication of diabetes that requires life-long management and, in many cases, multidisciplinary care (40). Because early and even sight-threatening stages may be asymptomatic, patients are often unaware of the need for regular ocular screening and tight systemic control until irreversible damage has occurred (41). High-quality DR-related educational short videos should therefore do more than simply list symptoms or provide brief "management tips." They should clearly explain what DR is and how it progresses, highlight key systemic and ocular risk factors, emphasize the importance of regular dilated fundus examinations and timely referral, and offer balanced information on evidence-based treatment options such as laser photocoagulation, intravitreal anti-VEGF therapy, and vitrectomy, together with realistic expectations regarding visual outcomes and the need for long-term follow-up (42, 43). In our sample, however, videos from medical, non-medical, and for-profit users tended to prioritize practical themes such as "management" and "symptoms," whereas risk-factor education, screening recommendations, and long-term prognosis were relatively underrepresented (44). This pattern suggests substantial room for improvement in the completeness of DR-related health education on TikTok. From a communication perspective, it is equally important how these core messages are presented. Effective DR popular-science videos should use plain, patient-centered language, avoid excessive jargon, and employ intuitive visual aids—such as simple diagrams, fundus photographs, or animations—to illustrate retinal damage and treatment procedures in a way that is accessible to viewers with varying levels of health literacy. Information can be structured in a clear sequence, end with a specific call-to-action encouraging people with diabetes to undergo regular eye examinations, and explicitly remind viewers that online content cannot replace professional medical consultation. At the same time, creators should refrain from sensational titles, exaggerated promises of "cure," or promotion of unproven remedies (45, 46), and should transparently disclose any commercial intent or sponsorship. Taken together, our findings imply that partnering with non-profit organizations and professional bodies to produce and promote videos that meet these content and presentation criteria may be a practical strategy to improve the quality and public-health impact of DR-related educational short videos on TikTok (47).

The present study is subject to certain limitations. First, our analysis was restricted to Mandarin-language DR videos on TikTok and relied on a single disease-specific keyword (Diabetic Retinopathy) for video retrieval. Although this term is the standard clinical designation used in Chinese guidelines and patient-education materials, this keyword-based strategy may have failed to capture videos that discussed DR using alternative lay expressions or more generic phrases related to diabetic eye disease. Moreover, TikTok is currently one of the most widely used short-video platforms in mainland China but represents only a portion of the broader Chinese short-video ecosystem; other popular platforms, such as Xiaohongshu and WeChat video channels, also host DR-related educational content, and the quality, framing, and user reach of the videos posted there may differ from those observed on TikTok. Our findings should therefore be interpreted as platform- and keyword-specific and may not be directly generalizable to other platforms or non-Mandarin language settings. Second, the cross-sectional design of this study captures video quality at only one specific time point. Given the extremely fast pace of content renewal and algorithm-driven content distribution on short-video platforms, both the visibility and composition of DR-related videos may change rapidly over time, and the applicability of our results is consequently constrained by their temporal nature. Furthermore, while internationally validated instruments such as DISCERN and PEMAT-A/V were employed to evaluate video quality, the effectiveness of health communication is also shaped by numerous audience-related factors, including educational attainment, media literacy, and cultural context. These audience-level variables were not incorporated into the present analysis and may modify how information of similar "objective" quality is actually perceived and acted upon. Subsequent studies could integrate user surveys and experimental designs to assess how various demographic groups comprehend, trust, and are behaviorally influenced by health-related short videos. Such work would provide a more nuanced and robust evidence base for developing an evidence-informed framework to guide the design and dissemination of DR-related public health communication on short-video platforms.

5 Conclusion

In conclusion, this study demonstrates that there is considerable variation in the overall quality of educational short videos about DR on the TikTok platform, and the source of the video is a key factor influencing its quality. Videos produced by non-profit organizations and medical professional users are of significantly higher quality than those from for-profit accounts and general users. In the future, promoting the standardized and scientific application of short videos in DR health education will require enhancing the participation of professionals, improving platform content review and regulatory mechanisms, and optimizing video content formats and presentation strategies, combined with long-term follow-up and effectiveness evaluation studies, to better improve patient health literacy and benefit public health.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the study involving human data in accordance with the local legislation and institutional requirements. Written informed consent was not required, for either participation in the study or for the publication of potentially/indirectly identifying information, in accordance with the local legislation and institutional requirements. The social media data was accessed and analyzed in accordance with the platform's terms of use and all relevant institutional/national regulations.

Author contributions

LJ: Data curation, Funding acquisition, Investigation, Resources, Writing – original draft, Writing – review & editing. WW: Conceptualization, Methodology, Writing – review & editing. XJ: Conceptualization, Formal analysis, Visualization, Writing – review & editing. FJ: Conceptualization, Validation, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Teo, ZL, Tham, YC, Yu, M, Chee, ML, Rim, TH, Cheung, N, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology*. (2021) 128:1580–91. doi: 10.1016/j.ophtha.2021.04.027

2. Modjtahedi, BS, Wu, J, Luong, TQ, Gandhi, NK, Fong, DS, and Chen, W. Severity of diabetic retinopathy and the risk of future cerebrovascular disease, cardiovascular disease, and all-cause mortality. *Ophthalmology*. (2020) 128:1169–79. doi: 10.1016/j.ophtha.2020.12.019

3. VanderBeek, BL, Yu, Y, Cardillo, S, and Hubbard, R. Twenty-year trends in prevalence and incidence of diabetic retinal disease. *Ophthalmology*. (2025) 132:767–74. doi: 10.1016/j.ophtha.2025.01.022

4. Yu, CW, Nanji, K, Garg, A, Zaman, M, Jomy, J, El-Sayes, A, et al. The prevalence of diabetic retinopathy in American Indians/Alaska natives and non-indigenous Americans: a systematic review and meta-analysis. *Ophthalmology*. [Ahead of Print] (2025). doi: 10.1016/j.ophtha.2025.08.004

5. Nwanyanwu, K, Andoh, J, Chen, E, Xu, Y, Granados, A, Deng, Y, et al. Social determinants associated with diabetic retinopathy awareness: national health and nutrition examination survey (2005-2008). *Am J Ophthalmol*. (2025) 278:389. doi: 10.1016/j.ajo.2025.06.028

6. Prasad, M, Agrón, E, Vitale, S, Arunachalam, T, Duic, C, Siddig, F, et al. Diabetic retinopathy incidence, progression, and health-related quality of life from the ACCORD trial. *Am J Ophthalmol*. (2025) 279. doi: 10.1016/j.ajo.2025.07.018

7. Liu, D, Chen, X, Wang, Q, Pan, X, Wei, X, Wang, B, et al. Evidence-based pharmacological investigation of the clinical practice and rational use of metformin in prediabetes mellitus. *Precis Med*. (2025) 2:100035. doi: 10.1016/j.prmedi.2025.100035

8. Zhang, K, Zhang, B, Wu, Q, Zhu, S, Wang, D, and Zhang, C. Exploring the mechanism of Simiao Yong'an decoction in treating diabetic foot based on metabolomics. *Precis Med*. (2025) 2:100028. doi: 10.1016/j.prmedi.2025.100028

9. Lum-Wang, S, Ingledew, P, and Chai, B. YouTube videos as a tool to educate medical students about careers in radiation oncology. *Int J Radiat Oncol Biol Phys*. (2025) 123:e15. doi: 10.1016/j.ijrobp.2025.05.055

10. Pham, H, Finedore, M, Skovira, C, Dean, L, Vu, A, Sargent, J, et al. 227 assessing the quality of YouTube videos on cannabinoid hyperemesis syndrome. *Ann Emerg Med*. (2024) 84:S104–5. doi: 10.1016/j.annemergmed.2024.08.229

11. Waidyaratne, G, Daboul, J, Liyanarachchi, S, and Chakraborty, S. The evaluation and analysis of irritable bowel syndrome-related short videos on social media (TikTok). *Aliment Pharmacol Ther*. (2024) 60:350–6. doi: 10.1111/apt.18096

12. Chang, IA, Wells, MW, Slack, GC, Rampazzo, A, and Gatherwright, JR. Analysis of TikTok's most viewed #PlasticSurgery content: an opportunity for educational outreach. *Plast Reconstr Surg*. (2022) 149:1271e–3e. doi: 10.1097/PRS.0000000000009113

13. Quartey, A, and Lipoff, JB. Is TikTok the new Google? Content and source analysis of Google and TikTok as search engines for dermatologic content in 2024. *J Am Acad Dermatol*. (2025) 93:e25–6. doi: 10.1016/j.jaad.2024.10.131

14. Zhang, X, Zhou, L, Wang, S, Fan, C, and Huang, D. Facilitating patient adoption of online medical advice through team-based online consultation. *JTAER*. (2025) 20:231. doi: 10.3390/jtaer20030231

15. Zhang, X, Wu, Y, Valacich, JS, and Jenkins, JL. Transforming patient-physician interaction through asynchronous online health interaction: a relational communication perspective. *J Assoc Inf Syst*. (2025) 26:1681–702. doi: 10.17705/1jais.00952

16. Javidan, A, Nelms, MW, Li, A, Lee, Y, Zhou, F, Kayssi, A, et al. Evaluating YouTube as a source of education for patients undergoing surgery: a systematic review. *Ann Surg*. (2023) 278:e712–8. doi: 10.1097/SLA.0000000000005892

17. Dand, H, Ranasinghe, U, Macarulay, C, Dand, HA, Ranasinghe, UA, Mokhtur, S, et al. S364 validity, reliability, and accuracy of YouTube content on colorectal cancer compared to trusted sources. *Am J Gastroenterol*. (2023) 118:S263. doi: 10.14309/01.ajg.0000951096.34967.96

18. Karass, S, Chahal, R, Singh, N, Karass, SH, Kutheala, N, Sukaina, M, et al. YouTube as a reliable source of information on antibody drug conjugates: a qualitative analysis. *J Clin Oncol*. (2024) 42:9050. doi: 10.1200/jco.2024.42.16_suppl.9050

19. Zhou, X, Ma, G, Su, X, Li, X, Wang, W, Xia, L, et al. The reliability and quality of short videos as health information of guidance for lymphedema: a cross-sectional study. *Front Public Health*. (2025) 12:1472583. doi: 10.3389/fpubh.2024.1472583

20. Jiang, S, Zhou, Y, Qiu, J, and Gou, X. Search engines and short video apps as sources of information on acute pancreatitis in China: quality assessment and content assessment. *Front Public Health*. (2025) 13:1578076. doi: 10.3389/fpubh.2025.1578076

21. Sun, F, Zheng, S, and Wu, J. Quality of information in gallstone disease videos on TikTok: cross-sectional study. *J Med Internet Res*. (2023) 25:e39162. doi: 10.2196/39162

22. Abbas, S, Abbas, A, Bozeman, L, Wood, L, and Chu, DI. Mo2093: assessing the quality of YouTube videos on celiac disease delivered by experts vs. non-experts. *Gastroenterology*. (2025) 169:S-1287. doi: 10.1016/s0016-5085(25)03845-4

23. Zhang, X, Xie, K, Gu, B, and Guo, X. Engaging physicians with introductory incentives: references to online and offline income. *MIS Q*. (2025) 49:917–52. doi: 10.25300/MISQ/2024/18172

24. Chen, S, Lai, K, Guo, X, Lai, K-h, and Zhang, X. The influence of digital health technology on the allocation of regional medical resources in China. *Health Policy Technol.* (2025) 14:101013. doi: 10.1016/j.hlpt.2025.101013

25. Wu, M, Shin, D, Waitayangkoon, P, Wu, MY-C, and Konter, J. Quality matters: assessing bronchiectasis videos on YouTube for patient education. *Chest.* (2024) 166:A169. doi: 10.1016/j.chest.2024.06.158

26. Lim, JI, Kim, SJ, Bailey, ST, Kovach, JL, Vemulakonda, GA, Ying, GS, et al. Diabetic retinopathy preferred practice pattern®. *Ophthalmology.* (2025) 132:P75–P162. doi: 10.1016/j.ophtha.2024.12.020

27. Goobie, GC, Guler, SA, Johannson, KA, Fisher, JH, and Ryerson, CJ. YouTube videos as a source of misinformation on idiopathic pulmonary fibrosis. *Ann Am Thorac Soc.* (2019) 16:572–9. doi: 10.1513/AnnalsATS.201809-644OC

28. Demircan, B, Kiyak Yorulmaz, Y, and Adigüzel Yaşar, Y. Pressurized metered-dose inhaler practices on YouTube: usefulness, reliability, and quality analysis. *J Allergy Clin Immunol Pract.* (2025) 13:2692–2699.e2. doi: 10.1016/j.jaip.2025.07.005

29. Gaikwad, A, Rachh, P, Raut, K, and Nadgere, J. Critical appraisal of YouTube videos regarding peri-implant diseases and conditions: a cross-sectional analysis. *Clin Oral Implants Res.* (2024) 35:498–509. doi: 10.1111/clr.14246

30. Musbahi, A, and Immanuel, A. P-OGC50 an objective evaluation of YouTube videos on oesophageal cancer using the PEMAT score. *Br J Surg.* (2021) 108:znab430.178. doi: 10.1093/bjs/znab430.178

31. Kong, W, Song, S, Zhao, YC, Zhu, Q, and Sha, L. TikTok as a health information source: assessment of the quality of information in diabetes-related videos. *J Med Internet Res.* (2021) 23:e30409. doi: 10.2196/30409

32. Wu, M, Yang, Y, and Chen, Y. Evaluation of treatment information quality on hypertension and diabetes on WeChat and TikTok: a cross-sectional content analysis. *Health Informatics J.* (2024) 30:14604582241275824. doi: 10.1177/14604582241275824

33. Chen, J, and Wang, Y. Social media use for health purposes: systematic review. *J Med Internet Res.* (2021) 23:e17917. doi: 10.2196/17917

34. Suarez-Lledo, V, and Alvarez-Galvez, J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res.* (2021) 23:e17187. doi: 10.2196/17187

35. Huang, M, Wang, J, Wei, J, Zhang, R, Wang, X, Gan, J, et al. Assessing the quality of educational short videos on dry eye care: a cross-sectional study. *Front Public Health.* (2025) 13:1542278. doi: 10.3389/fpubh.2025.1542278

36. Liu, Z, Chen, Y, Lin, Y, Ai, MM, Lian, DL, Zhang, YH, et al. YouTube/ Bilibili/ TikTok videos as sources of medical information on laryngeal carcinoma: cross-sectional content analysis study. *BMC Public Health.* (2024) 24:1594. doi: 10.1186/s12889-024-19077-6

37. Metzler, H, and Garcia, D. Social drivers and algorithmic mechanisms on digital media. *Perspect Psychol Sci.* (2024) 19:735–748. doi: 10.1177/17456916231185057

38. Cui, N, Lu, Y, Cao, Y, Chen, X, Fu, S, and Su, Q. Quality assessment of TikTok as a source of information about mitral valve regurgitation in China: cross-sectional study. *J Med Internet Res.* (2024) 26:e55403. doi: 10.2196/55403

39. Lee, S, Narula, N, Weglarz, M, Lee, J, and Kim, D. Evaluation of social media videos on weight-loss surgery and GLP-1 agonists: a content quality and reliability analysis. *Surg Endosc.* (2025) 39:6032–6. doi: 10.1007/s00464-025-11938-4

40. Sivaprasad, S, Wong, TY, Gardner, TW, Sun, JK, and Bressler, NM. Diabetic retinal disease. *Nat Rev Dis Primers.* (2025) 11:62. doi: 10.1038/s41572-025-00646-x

41. Wong, TY, Cheung, CM, Larsen, M, Cheung, CMG, Sharma, S, and Simó, R. Diabetic retinopathy. *Nat Rev Dis Primers.* (2016) 2:16012. doi: 10.1038/nrdp.2016.12

42. Kirkpatrick, CE, and Lawrie, LL. TikTok as a source of health information and misinformation for young women in the United States: survey study. *JMIR Infodemiol.* (2024) 4:e54663. doi: 10.2196/54663

43. Cheung, N, Mitchell, P, and Wong, TY. Diabetic retinopathy. *Lancet.* (2010) 376:124–36. doi: 10.1016/S0140-6736(09)62124-3

44. Basch, CH, Fung, IC, Berdnik, A, and Basch, CE. Widely viewed English language YouTube videos relating to diabetic retinopathy: a cross-sectional study. *JMIR Diabetes.* (2016) 1:e6. doi: 10.2196/diabetes.6450

45. AlBloushi, AF, and Abouammoh, MA. YouTube videos related to diabetic retinopathy: are they good enough? *J Fr Ophtalmol.* (2022) 46:223–30. doi: 10.1016/j.jfo.2022.07.010

46. Sule, S, DaCosta, MC, DeCou, E, Gilson, C, Wallace, K, and Goff, SL. Communication of COVID-19 misinformation on social media by physicians in the US. *JAMA Netw Open.* (2023) 6:e2328928. doi: 10.1001/jamanetworkopen.2023.28928

47. Sampige, R, Rodgers, EG, Huang, A, and Zhu, D. Education and misinformation: exploring ophthalmology content on TikTok. *Ophthalmol Ther.* (2023) 13:97–112. doi: 10.1007/s40123-023-00834-6

Changes in parental knowledge and concerns regarding pediatric fever from 2017 to 2024: repeated cross-sectional surveys on the association of a smartphone application

Masahiko Sakamoto[1,2]*, Asuka Suzuki[2] and Hirono Ishikawa[2]

[1]Department of Pediatrics, Saku Central Hospital Advanced Care Center, Saku, Japan, [2]Teikyo University Graduate School of Public Health, Itabashi, Japan

Introduction: Fever in children is typically not an indication of severe illness; however, parental anxiety remains high. Recently, factors such as declining birth rates and the COVID-19 pandemic have influenced parental health awareness and behavior. This study aimed to evaluate the concerns and knowledge about fever among parents whose children visited the emergency department in 2017 and 2024, and to evaluate the changes over the 7 years and the association with the download of the smartphone app "Oshiete! Doctor," which provides child healthcare information.

Methods: These repeated cross-sectional surveys were conducted at the Holiday Pediatric Medical Center in Saku in 2017 and 2024. A questionnaire was administered to caregivers of children aged ≤8 years to assess their anxiety about fever (brain damage, seizures, and dehydration) and knowledge (need for antibiotics for fever, body temperature as a criterion for consultation, and the criteria for attending daycare). The awareness and utilization of the application were also investigated. Data from 2017 and 2024 were compared using the chi-square and t-tests.

Results: A comparison between the 2017 survey (n = 224) and the 2024 survey (n = 261) revealed a significant increase in concerns about brain damage and seizures in response to fever (+14.4% and +14.9%, respectively). In contrast, substantial improvements were observed in the recognition of the necessity of antibiotics for fever, the recognition of body temperature as a criterion for consultation, and the criteria for attending daycare. When participants were stratified according to the application download status, anxiety levels increased in both groups; however, the magnitude of increase was slightly smaller in the download group. Nevertheless, statistical significance was not confirmed in the multivariable analysis.

Discussion: Over the past 7 years, an increase in parents' knowledge of fever and an escalation in their anxiety were observed. However, the dissemination of the application alone did not demonstrate an effect strong enough to offset the overall increase in societal anxiety. Future studies may be able to more clearly verify a suppressive effect on this increase by using more refined exposure measurements based on behavioral indicators, such as actual frequency of use, including by expanding the sample size.

Introduction

Although fever in children rarely indicates severe illness, parental anxiety about fever is high (1). Parents often consulted a doctor when their child had a fever (2). Approximately two-thirds of parents typically visit a medical facility within the first 12 h after their child develops a fever (3). This strong anxiety about fever has been widely known since it was defined as "fever phobia" by Schmitt et al. in the 1980s (4). Fever phobia is characterized by parental anxiety and exaggerated, unrealistic misconceptions about fever, including concerns about brain damage, impaired consciousness, and hearing loss (5). Despite years of education, many parents still have misconceptions about fever.

In recent years, following the declining birth rates and the COVID-19 pandemic, parental awareness of health has changed. During the COVID-19 pandemic, a significant amount of medical information was disseminated; however, this did not necessarily result in effective health behaviors. Ishikawa et al. reported that health literacy declined during the pandemic (6). Given this context, awareness and behaviors associated with fever may also have changed. Kupcova et al. reported that anxiety worsened during the pandemic, particularly among women and young adults (7). During the pandemic, many public health institutions recommended treating fever with acetaminophen or ibuprofen (8). However, this may have resulted in a heightened sensitivity to fever and exacerbated anxiety related to fever. However, to our knowledge, no previous studies have examined the changes in fever-related anxiety or the behavior of parents before and after the pandemic.

To reduce parental anxiety, it is necessary to provide accurate knowledge and guidance on fever management. Parents seek more accessible and consistent information regarding fever management (9). Providing parents with appropriate health information about their children can increase their knowledge and decrease anxiety. Previous randomized controlled trials have reported that educational interventions, such as informational leaflets, significantly improved parents' knowledge regarding fever, including its definition (temperature above 38 °C) and appropriate management strategies (10). Furthermore, a Randomized control trial showed that video tutorials guided managing common childhood symptoms and deciding when to seek medical help, significantly and safely increasing parents' self-efficacy when used by people calling out-of-hour medical helplines (11). Recently, research on digital health has also been conducted. For example, FeverApp, an app that provides fever management protocols for parents in Germany, has been introduced. Beerenbrock et al. emphasize that structured guidance based on this app has the potential to reduce parents' concerns about fever (12). Conversely, there is limited evidence on evaluating the effectiveness of such apps.

In 2015, a parent education initiative, "Oshiete! Doctor" ("Teach Me, Doctor"), was launched in Saku City, Nagano Prefecture, and a mobile application was developed to provide health information aimed at reducing parents' anxiety. The application was designed to be free of charge and user-friendly, with a simple interface to facilitate ease of use and provide reassurance to parents during emergencies.

In 2022, this study surveyed to investigate the association between application use and parental health literacy (13). This study suggests that if parents have lower health literacy, they are less likely to use the app and may not be able to obtain sufficient benefits. Many questions remain regarding how providing information through applications can help reduce parental anxiety.

This study aimed to evaluate the fever-related anxiety and knowledge among parents of infants and young children who visited the emergency department with fever as the primary complaint in 2024, and to analyze the changes over the 7 year period by comparing the data with that from a similar survey conducted in 2017. Furthermore, this study evaluated the association between fever-related anxiety and behaviors and the download of mobile applications.

Materials and methods

Study participants and data collection

These repeated cross-sectional surveys were conducted at the Holiday Pediatric Medical Center in Saku City, Nagano Prefecture, in 2017 and 2023. The center is the only medical institution responsible for morning holiday outpatient services (no ambulances accepted) in the Saku region, which has a target population of 230,000. The survey period was from October 1, 2017, to January 30, 2018, and from December 20, 2024, to March 31, 2024. The study targeted guardians of children aged ≤8 years who visited the center during these periods, with 229 participants in the first period and 425 in the second period. The parents received an explanation of the survey from the nurses and participated in the survey with their consent. Participants who selected "agree to participate" in the questionnaire were regarded as having consented to the study. The questionnaire was collected by the nursing staff after consultation (Supplementary File S1). Respondents were excluded if they were not a child's mother or father, did not own a smartphone, or visited the center more than once during the study period. The total number of respondents in the 2017 survey was 224 (224/229; response rate 98%). All of these respondents were individuals who had visited the medical facility because their child presented with fever symptoms. In the 2024 survey, the number of respondents was 386 (386/425; response rate 91%). Among these, 261 respondents who visited the medical facility due to their child's fever symptoms were included as the study participants. This study was approved by the Ethics Review Committee of the Saku Central Hospital (Approval Number: R202310-05).

Overview of the application

The "Oshiete! Doctor" application was launched in 2016 and is offered free of charge, with support from Saku City, to improve guardians' childcare skills. The application provides information on seven topics: guidelines for seeking medical care for sick children,

explanations of childhood illnesses, vaccination information, details of childcare support organizations, and disaster preparedness for children (See Supplementary File S2). This app has a section on fever management, which provides guidelines on how to determine whether to seek medical consultation and knowledge about home care for fever.

Survey questions

Concerns about fever

Concerns about fever were assessed by asking whether participants agreed with the following three statements based on previous studies (2) (14): "Brain damage may occur," "Convulsions may occur," and "Dehydration may occur." Responses were categorized as "concerned" if participants selected "strongly agree" or "somewhat agree" and as "not concerned" if they selected "somewhat disagree" or "strongly disagree."

Knowledge about fever

Knowledge about fever was assessed using the following statement: "Antibiotics are necessary for the treatment." "Body temperature serves as a criterion for seeking medical attention," and "If there was no fever the previous day and the child is afebrile in the morning, they can attend school." Responses regarding antibiotic necessity were categorized as "Necessary" if participants selected "Strongly agree" or "Somewhat agree" and as "Not necessary" if they selected "Somewhat disagree" or "Strongly disagree." Responses regarding body temperature as a criterion for seeking medical attention were categorized as "Serves as a criterion" or "Does not serve as a criterion." Responses regarding school attendance were classified as "Can attend school" or "Cannot attend school" after excluding the "Cannot decide" response. These questions were selected based on the following clinical facts: "Most cases of fever in children are viral infections and do not require antibiotics" (15); "The judgment to consult a doctor for a child should be based on an evaluation of the child's overall condition and cannot be made based on body temperature alone" (16); and "It is not uncommon for children to have no fever in the morning and then develop a fever in the afternoon" (17). Although a scale to assess parents' fever management practices has been developed overseas (18), it had not been validated in Japanese at the time of this study. Therefore, instead of that scale, this study employed these three questions commonly used to advise patients with fever in pediatric clinical practice.

Recognition and download of the "Oshiete! Doctor" application

The participants were asked about their recognition and download of the application. In the present study, only the download status was available; therefore, it was used as a surrogate measure for actual app use. Responses were categorized as "not downloaded" if they selected "do not know the application" or "know the application but have not downloaded it," and as "downloaded" if they selected "have downloaded it."

Socio-demographic data

The analysis included the respondent's relationship with the child, age group, the child's age and sex, and birth order. Respondent-child relationships were categorized as either "Father" or "Mother." Age groups were classified into four categories: 20s, 30s, 40s, 50s, or older. Birth order was categorized as either "first child" or "second child and later."

Data analysis and statistics

Continuous variables are summarized using descriptive statistics (mean and standard deviation or median and quartiles), and categorical variables are expressed as frequencies and percentages. Socio-demographic characteristics, application recognition and usage, knowledge and attitudes toward fever, and related behaviors were compared between the 2017 and 2024 surveys using chi-square or t-tests. This study also constructed a multivariable logistic regression model in which the survey year (2017 vs. 2024), application download status (no vs. yes), and their interaction (year × app) were the main explanatory variables, adjusted for the child's age, birth order, and the parent's age group. With regard to the handling of missing values, a complete-case analysis was conducted.

Data analysis was performed using the Stata 17 software (StataCorp LP, College Station, Texas, United States).

Results

In the 2017 survey, 224 out of 229 parents responded to the questionnaire (response rate of 98%). All of their children presented with fever. In the 2024 survey, 386 of the 425 parents responded (response rate of 91%). Of these respondents, 261 were parents of children with fever. Table 1 presents the respondents' background characteristics. The proportion of fathers among respondents showed a slight increase from 19.6% in 2017 to 24.1% in 2024; however, this difference was not statistically significant. The most common age group among parents in both surveys was the ≥30s. However, the proportion of parents in this age group decreased from 66.1% in 2017 to 51% in 2024, whereas the proportion in other age groups increased. No significant differences were observed in the age and sex of the children or the proportion of first-born children between 2017 and 2024.

Regarding awareness of the application, the proportion of parents who knew the application increased from 52% in 2017 to 66.5% in 2024. However, the proportion of parents who reported downloading the application decreased from 32% in 2017 to 25% in 2024.

Table 2 presents a stratified comparison of concerns regarding fever, categorized by the download status of the application between 2017 and 2024. Overall, concerns about brain damage and seizures increased in 2024 compared to those in 2017. Particularly, the proportion of parents who expressed concerns about brain damage and seizures was significantly higher in the non-download group between 2017 and 2024. However, in the multivariable logistic regression analyses regarding concerns about brain damage, seizures, and dehydration (not shown in the table), none of the year × application interactions were statistically significant ($p = 0.956$, $p = 0.306$, and $p = 0.903$, respectively). In other words, no evidence was obtained to suggest that the magnitude of change from 2017 to 2024 differed according to application use. The child's age showed an independent positive association with anxiety about brain damage and seizures (Adjusted odds ratio [aOR] = 1.1 and 1.2, respectively). No clear associations were observed with birth order or the parents' age group.

Table 3 presents a stratified comparison of parental knowledge regarding fever, categorized by application download status, between 2017 and 2024. The proportion of respondents who believed that antibiotics were necessary for fever, those who considered body temperature to be a criterion for seeking medical care, and those who allowed children to attend day care if they were afebrile on the day of attendance despite having a fever the previous day all demonstrated a substantial decrease. This trend was observed regardless of the download status of the application. Similarly, in the multivariable logistic regression analysis, knowledge regarding fever improved from 2017 to 2024 in all items; however, none of the year × application interactions were statistically significant ($p = 0.238$, $p = 0.165$, $p = 0.998$), and no evidence was obtained that "the degree of change from 2017 to 2024 differed according to application use.

Discussion

This study aimed to compare changes in anxiety and knowledge about fever between 2017 and 2024 among parents of children aged ≤8 years who visited a holiday medical center with a chief complaint of fever, and to assess the association between these factors and the download of the Japanese parent-oriented app "Oshiete! Doctor."

The importance of health literacy has been emphasized during the pandemic [19]. Particularly, digital health literacy has been reported to influence individuals' ability to access, understand, and evaluate information and services that support healthy behaviors during a pandemic [20]. In the present study, the proportion of parents who had downloaded the application in 2024 among those visiting a holiday medical care center was 25%. In contrast, a 2023 study targeting parents who attended infant health checkups in the same region found that 43.4% of them had downloaded and utilized the application [13]. In Japan, infant health checkups are nearly universal, with a very high participation rate. Therefore, this result reflects the overall utilization rate of the app by parents in this region. Consequently, it is suggested that parents who seek emergency outpatient care are less likely to download the application than the general parent population. It is also possible that parents who downloaded the application may have avoided emergency outpatient visits.

Furthermore, it was found that among parents visiting holiday medical care centers, awareness of the application increased between 2017 and 2024. However, the proportion of those who downloaded the application decreased. After the COVID-19 pandemic, parents who actively used the application may have become more inclined to refrain from emergency outpatient visits. The comparison of these survey data before and after the pandemic suggests an association between app downloads and reduced emergency outpatient visits, indicating that digital health literacy may be associated with preventive behavior during the pandemic.

The present study demonstrated a substantial decline in 2024 compared to 2017 in the proportion of respondents who acknowledged fever as a valid indication for seeking medical intervention, including those who regarded fever as a suitable reason to skip school if it had subsided the previous day. An elevated body temperature does not

TABLE 1 Background characteristics of parents and children and awareness and utilization of the application.

Characteristics	2017 n = 224		2024 n = 261		p-value*
Relationship					0.234
Father	44	19.6%	63	24.1%	
Mother	180	80.4%	198	75.9%	
Age group of parents					**0.026**
20s	25	11.2%	39	15.4%	
30s	148	66.1%	133	52.6%	
40s	49	21.9%	76	30.0%	
50s and older	2	0.9%	5	2.0%	
Age of child (years)	3.7 ± 2.6		4.0 ± 2.5		
Sex of child					0.243
Male	123	54.9%	132	50.6%	
Female	97	43.3%	129	49.4%	
Birth order					0.827
First	104	49.3%	126	48.3%	
Second or later	107	50.7%	135	51.7%	
Recognition and download of the App					**<0.001**
Do not know the app	107	47.8%	88	34.0%	
Know the App, but have not downloaded it	45	20.1%	108	41.5%	
Have downloaded the App	72	32.1%	64	25.0%	

*P-values were generated using a t-test or χ² test. P values indicating statistically significant differences ($p < 0.05$) were shown in bold.

TABLE 2 Comparison of concerns related to fever by application download (2017 and 2024).

Concerns related to fever	Group	2017	2024	p-value*
Brain damage	Total	115/217 (53.0%)	174/259 (67.2%)	0.001
	Non-download group	79/150 (52.7%)	131/194 (67.5%)	**0.005**
	Download group	36/67 (53.7%)	43/64 (67.2%)	0.116
Seizure	Total	142/216 (65.7%)	209/259 (80.7%)	<0.001
	Non-download group	95/150 (63.3%)	158/194 (81.4%)	**<0.001**
	Download group	47/66 (71.2%)	50/64 (78.1%)	0.365
Dehydration	Total	191/216 (88.4%)	227/261 (87.0%)	0.71
	Non-download group	133/150 (88.7%)	173/196 (88.3%)	0.908
	Download group	58/66 (87.9%)	54/64 (84.4%)	0.563

*P-values were generated using a t-test or χ² test. P values indicating statistically significant differences (p < 0.05) were shown in bold.

TABLE 3 Comparison of knowledge related to fever by application usage (2017 and 2024).

Knowledge related to fever	Group	2017	2024	p-value*
Antibiotics were necessary for the fever	Total	117/217 (53.9%)	93/258 (36.1%)	**<0.001**
	Non-download group	89/151 (58.9%)	73/194 (37.6%)	**<0.001**
	Download group	28/66 (42.4%)	20/63 (31.8%)	0.21
Consider body temperature as a criterion for seeking medical care	Total	191/215 (88.8%)	99/229 (43.2%)	**<0.001**
	Non-download group	134/149 (89.9%)	71/173 (41.0%)	**<0.001**
	Download group	57/66 (86.4%)	28/56 (50.0%)	**<0.001**
If the fever is gone in the morning, the child can attend school.	Total	40/183 (21.9%)	11/254 (4.3%)	**<0.001**
	Non-download group	28/124 (22.6%)	9/192 (4.7%)	**<0.001**
	Download group	12/59 (20.3%)	2/62 (3.2%)	**0.003**

*P-values were generated using a t-test or χ² test. P values indicating statistically significant differences (p < 0.05) were shown in bold.

necessarily indicate severe infectious disease; therefore, it should not be used as a criterion for seeking medical care (21). Furthermore, owing to the influence of the circadian rhythm, body temperature tends to be higher in the afternoon than in the morning (22). Therefore, if a fever is present on a given day, there is a possibility of recurrence in the afternoon, even if the fever subsides in the morning. This underscores the importance of rest and the necessity of refraining from attending school. Consequently, these changes in knowledge are considered favorable, suggesting that parental understanding of fever has improved over the past 7 years. This may be due to the increase in public messages during COVID-19, such as appropriate fever management and the limited role of antibiotics, which may have influenced parents' knowledge.

With respect to the proportion of caregivers who believed that antibiotics were necessary in the presence of fever, in 2017 the figures were 89/151 (non-download group, 58.9%) versus 28/66 (download group, 42.2%), indicating that the download group provided fewer incorrect responses (p = 0.025), suggesting an association between app download and higher health literacy. However, in 2024, the proportion of incorrect responses decreased in both groups, with

73/194 (37.6%) in the non-download group and 20/63 (31.8%) in the download group, and the difference between the groups had narrowed. In Japan, nationwide campaigns promoting the appropriate use of antibiotics have been implemented to address antimicrobial resistance as the Action Plan for Countermeasures against Antimicrobial Resistance (AMR) since 2016. Initiatives, such as monthly antimicrobial resistance awareness, have been implemented to promote public education on the proper use of antibiotics. Furthermore, educational programs for healthcare professionals have been implemented to support the appropriate use of antibiotics. Some reports indicated a decrease in antibiotic consumption in Japan following the launch of these campaigns (23). This study's findings may also suggest the effectiveness of the nationwide campaign promoting the appropriate use of antibiotics.

Conversely, the proportion of individuals expressing concerns about brain damage and seizures related to fever increased over the past 7 years. This may be attributed to the influence of media coverage related to the COVID-19 pandemic and the heightened uncertainty, which increased parental vigilance and anxiety. Previous studies have also reported that health-related anxiety worsens during pandemics (7).

When participants were stratified according to the presence or absence of application download, anxiety levels increased in both groups; however, the magnitude of increase was slightly smaller in the group that had downloaded the application, suggesting a potential mitigating effect on the rise in anxiety. We have previously reported that individuals unfamiliar with the application exhibited comparatively lower health literacy (13). Parents with lower digital health literacy may have been less likely to download the application and, as a result, may have demonstrated higher levels of anxiety. However, in multivariable analysis, the statistical significance of the interaction was not confirmed, which may be attributable to insufficient statistical power or to the fact that the exposure was measured as "download or not," rather than capturing the actual intensity of use. This may have introduced non-differential exposure misclassification (download ≠ actual use). With more refined exposure measurements based on behavioral indicators such as actual frequency of use and type of content accessed, including an expanded sample size, it may be possible to more clearly verify the potential attenuating effect on the increase in anxiety.

If the use of the smartphone application is associated with better knowledge and lower unnecessary fear of fever, this could lead to more appropriate home management of fevers and possibly fewer unnecessary emergency department visits or calls to pediatricians. This aligns with public health goals of focusing medical attention where it is truly needed and empowering parents to care for mild illnesses at home. Therefore, the use of such apps and improved health literacy may also lead to a reduction in the burden on the healthcare system.

However, these studies are repeated cross-sectional surveys, and the inference of a causal relationship between application download and the reduction of anxiety is limited; these mechanisms remain at the hypothetical stage. Further research is needed to demonstrate these hypotheses.

This study has several limitations. First, the study period was limited to the winter season from December to March. Although numerous diseases cause fevers, pediatric illnesses often exhibit seasonality. Thus, conducting the study exclusively during winter may have influenced the results. However, winter is the peak season for emergency department overcrowding due to influenza outbreaks and other factors. This study aimed to explore measures to alleviate the burden of emergency services. Therefore, evaluating this specific period was of utmost importance, and seasonal selection was deemed appropriate.

Second, the participants included in the studies conducted in 2017 and 2024 were not from the same cohort. Consequently, individual-level changes could not be assessed. However, the participants were parents of children aged ≤8 years or younger who sought medical attention at the same holiday medical center during winter because fever was the primary complaint. Given the similarities in their backgrounds, valid comparisons could be made at the group level.

Third, although the present study examined the proportion of parents who downloaded the mobile application, it did not evaluate the extent to which they utilized it after downloading. Therefore, differences were observed by download status, but the extent of app use was not measured. App download coarsely measures the true exposure (app use), and because participants who downloaded but did not use the app were included in the use group. If reanalysis were to be conducted using app use, the effect estimate would likely appear stronger, and we believe this would affect the results of multivariable analysis.

Fourth, since there is only one holiday medical center for children in this area, this study is considered to be representative of the emergency consultation behavior of parents in this area. However, since this study was conducted in a single region, region-specific factors may influence the results, which limits the generalizability of the findings to other regions.

Fifth, in the 2017 survey, the questionnaire was administered to parents whose children visited the medical facility with fever, whereas in 2024, the questionnaire was administered to all parents whose children visited the medical facility, and from among them, those with children presenting with fever were identified. Because the recruitment methods in the 2 years were not completely identical, there is a possibility that the comparison may have been rendered less precise. However, we consider that this methodological difference was minimal and likely had little actual impact on the findings.

Furthermore, as this study is based on self-reported data, it may be subject to response bias. Conversely, respondents may have overestimated their knowledge or underreported their concerns due to social desirability, which may have affected the results.

Finally, this study comprises two repeated cross-sectional surveys. Therefore, it is necessary to consider the insufficient control of potential confounding factors and the possibility of reverse causation between application use and anxiety, and it should be noted that causal relationships cannot be inferred from this study.

Parental knowledge of fever may have improved following the COVID-19 pandemic. In particular, there is a suggestion that awareness of refraining from antibiotic use has increased, indicating that campaigns promoting the appropriate use of antibiotics may have been effective. In contrast, anxiety related to fever tended to worsen, with this trend being more pronounced among parents who had not downloaded the application. In the future, longitudinal research targeting multiple regions that directly measures broad contextual factors such as application use, parental health literacy, and.

pandemic-related stress is needed to study the causal relationship between measures to improve digital health literacy and the reduction of parents' concerns.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Review Committee of the Saku Central Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MS: Funding acquisition, Writing – original draft, Project administration, Formal analysis, Software, Resources, Visualization, Conceptualization, Methodology, Investigation, Data curation, Writing – review & editing. AS: Visualization, Software, Formal analysis, Investigation, Writing – review & editing, Validation,

Supervision, Methodology. HI: Formal analysis, Visualization, Writing – review & editing, Validation, Conceptualization, Investigation, Supervision, Methodology.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2025.1619134/full#supplementary-material

References

1. de Bont, EG, Francis, NA, Dinant, GJ, and Cals, JW. Parents' knowledge, attitudes, and practice in childhood fever: an internet-based survey. *Br J Gen Pract*. (2014) 64:e10–6. doi: 10.3399/bjgp14X676401

2. Hamideh Kerdar, S, Himbert, C, Martin, DD, and Jenetzky, E. Cross-sectional study of parental knowledge, behaviour and anxiety in management of paediatric fever among German parents. *BMJ Open*. (2021) 11:e054742. doi: 10.1136/bmjopen-2021-054742

3. Çelik, T, and Güzel, Y. Parents' knowledge and Management of Fever: "parents versus fever!". *Turk Arch Pediatr*. (2024) 59:179–84. doi: 10.5152/TurkArchPediatr.2024.23152

4. Schmitt, BD. Fever phobia: misconceptions of parents about fevers. *Am J Dis Child*. (1980) 134:176–81. doi: 10.1001/archpedi.1980.02130140050015

5. Kwak, YH, Kim, DK, Jang, HY, Kim, JJ, Ryu, JM, Oh, SB, et al. Fever phobia in Korean caregivers and its clinical implications. *J Korean Med Sci*. (2013) 28:1639–44. doi: 10.3346/jkms.2013.28.11.1639

6. Ishikawa, H, Kato, M, and Kiuchi, T. Declines in health literacy and health-related quality of life during the COVID-19 pandemic: a longitudinal study of the Japanese general population. *BMC Public Health*. (2021) 21:2180. doi: 10.1186/s12889-021-12092-x

7. Kupcova, I, Danisovic, L, Klein, M, and Harsanyi, S. Effects of the COVID-19 pandemic on mental health, anxiety, and depression. *BMC Psychol*. (2023) 11:108. doi: 10.1186/s40359-023-01130-5

8. Wrotek, S, LeGrand, EK, Dzialuk, A, and Alcock, J. Let fever do its job: the meaning of fever in the pandemic era. *Evol Med Public Health*. (2021) 9:26–35. doi: 10.1093/emph/eoaa044

9. Kelly, M, Sahm, LJ, Shiely, F, O'Sullivan, R, McGillicuddy, A, and McCarthy, S. Parental knowledge, attitudes and beliefs regarding fever in children: an interview study. *BMC Public Health*. (2016) 16:540. doi: 10.1186/s12889-016-3224-5

10. Kelly, M, Sahm, L, McCarthy, S, O'Sullivan, R, Mc Gillicuddy, A, and Shiely, F. Randomised controlled trial of an intervention to improve parental knowledge and management practices of fever. *BMC Pediatr*. (2019) 19:447. doi: 10.1186/s12887-019-1808-9

11. Borch-Johnsen, L, Gren, C, Lund, S, Folke, F, Schrøder, M, Frederiksen, MS, et al. Video tutorials to empower caregivers of ill children and reduce health care utilization: a randomized clinical trial. *JAMA Netw Open*. (2023) 6:e2336836. doi: 10.1001/jamanetworkopen.2023.36836

12. Beerenbrock, Y, Jenetzky, E, and Martin, DD. FeverApp for parents: a multilingual and socially accountable approach to Paediatric fever Management in Germany. *J cme*. (2024) 13:2352940. doi: 10.1080/28338073.2024.2352940

13. Sakamoto, M, Ishikawa, H, and Suzuki, A. Evaluation of parents' use of a child health care information app and their health literacy: cross-sectional study. *JMIR Pediatr Parent*. (2024) 7:e48478. doi: 10.2196/48478

14. Arias López, MDP, Ong, BA, Borrat Frigola, X, Fernández, AL, Hicklent, RS, Obeles, AJT, et al. Digital literacy as a new determinant of health: a scoping review. *PLoS Digit Health*. (2023) 2:e0000279. doi: 10.1371/journal.pdig.0000279

15. Karageorgos, S, Hibberd, O, Mullally, PJW, Segura-Retana, R, Soyer, S, and Hall, D. Antibiotic use for common infections in pediatric emergency departments: a narrative review. *Antibiotics*. (2023) 12:1092. doi: 10.3390/antibiotics12071092

16. National Institute for Health and Care Excellence Fever in under 5s: assessment and initial management [Internet]. London: National Institute for Health and Care Excellence; 2019 Nov 7 [updated 2021 Nov]. (NICE guideline [NG143]). (2021). Available online at: https://www.nice.org.uk/guidance/ng143

17. Harding, C, Pompei, F, Bordonaro, SF, McGillicuddy, DC, Burmistrov, D, and Sanchez, LD. Fever incidence is much lower in the morning than the evening: Boston and US National Triage Data. *West J Emerg Med*. (2020) 21:909–17. doi: 10.5811/westjem.2020.3.45215

18. Walsh, A, Edwards, H, and Fraser, J. Parents' childhood fever management: community survey and instrument development. *J Adv Nurs*. (2008) 63:376–88. doi: 10.1111/j.1365-2648.2008.04721.x

19. Rohwer, E, Mojtahedzadeh, N, Neumann, FA, Nienhaus, A, Augustin, M, Harth, V, et al. The role of health literacy among outpatient caregivers during the COVID-19 pandemic. *Int J Environ Res Public Health*. (2021) 18:11743. doi: 10.3390/ijerph182211743

20. Hange, N, Agoli, AM, Pormento, MKL, Sharma, A, Somagutta, MR, Paikkattil, N, et al. Impact of COVID-19 response on public health literacy and communication. *Health Promot Perspect*. (2022) 12:1–9. doi: 10.34172/hpp.2022.01

21. De, S, Williams, GJ, Teixeira-Pinto, A, Macaskill, P, McCaskill, M, Isaacs, D, et al. Lack of accuracy of body temperature for detecting serious bacterial infection in febrile episodes. *Pediatr Infect Dis J*. (2015) 34:940–4. doi: 10.1097/inf.0000000000000771

22. Baehr, EK, Revelle, W, and Eastman, CI. Individual differences in the phase and amplitude of the human circadian temperature rhythm: with an emphasis on morningness-eveningness. *J Sleep Res*. (2000) 9:117–27. doi: 10.1046/j.1365-2869.2000.00196.x

23. Honda, H, Goto, T, Uehara, Y, and Takamatsu, A. Promotion of antimicrobial stewardship following issuance of the antimicrobial resistance national action plan in Japan: a systematic review of 2016-2020. *Int J Antimicrob Agents*. (2023) 62:106829. doi: 10.1016/j.ijantimicag.2023.106829

Frontiers in
Public Health

Explores and addresses today's fast-moving healthcare challenges

One of the most cited journals in its field, which promotes discussion around inter-sectoral public health challenges spanning health promotion to climate change, transportation, environmental change and even species diversity.

Discover the latest Research Topics

See more →

frontiers | Research Topics