Experimental and Applied Medical Science, 5, 3: 137-149, 2024. DOI: 10.46871/eams.1456744

The Promise and the Challenge: Large Language Models for Patient Education - Are We There Yet?

İbrahim Edhem Yılmaz[1*]

[1]*Gaziantep Islam Science and Technology University, Faculty of Medicine, Ophthalmology Department, Gaziantep, Turkey.*

Abstract

Purpose: *This study aimed to evaluate the potential of large language models (LLMs) for delivering patient education materials.*

Methods: *Two LLMs, Gemini and ChatGPT 3.5, were analysed for their ability to provide clear and understandable information on the topic of blepharitis, a common eye condition.*

The understandability and actionability of the information provided by the LLMs in response to a set of questions were evaluated using PEMAT, a standardised tool for assessing educational materials.

Results: *The responses included the important aspects of blepharitis, yet the Flesch-Kincaid readability scores were below the suggested range of 60-70 for patient education materials. Gemini received a score of 38.75, whereas ChatGPT 3.5 earned 26.35, suggesting that the content might be too intricate for the target audience.*

Conclusion: *These findings suggest that while LLMs have the potential to be informative resources, their current readability levels may limit their effectiveness in providing accessible health information to patients. Further research is needed to explore methods for adapting LLM outputs to ensure clear and concise communication suitable for patient education.*

Key words: *Large Language Models (LLMs); Patient Education; Readability; Blepharitis; Natural Language Processing (NLP); Health Communication*

*** Corresponding author:** İbrahim Edhem Yılmaz, E-mail: dredhemyilmaz@gmail.com, ORCID ID: 0000-0003-1154-425X

Introduction

The human eye, a complex and delicate organ, is susceptible to various conditions that can significantly impact vision and overall well-being. Blepharitis, a chronic inflammatory condition affecting the eyelids, is a prevalent ocular disease with an estimated global prevalence ranging from 33% to 50% (1, 2). The primary manifestation of blepharitis is inflammation of the eyelid margins, often accompanied by symptoms like redness, irritation, stinging, burning, crusting, and even eyelash loss (3).

Blepharitis can be classified into two main types based on the location of the inflammation:

Anterior blepharitis affects the anterior eyelid structures, including the eyelid skin, base of the eyelashes, and the eyelash follicles. It can be further subcategorized as staphylococcal blepharitis, caused by an overgrowth of Staphylococcus bacteria, or seborrheic blepharitis, associated with seborrheic dermatitis, a condition that causes redness, scaling, and inflammation of the scalp and other oily areas of the body (4).

Posterior blepharitis, also known as meibomian gland dysfunction (MGD), affects the meibomian glands located within the eyelids. The International Workshop on Meibomian Gland Dysfunction defines MGD as a chronic, diffuse abnormality of the meibomian glands, commonly characterized by terminal duct obstruction and/or qualitative/quantitative changes in the glandular secretion. MGD is considered to be the most common cause of evaporative dry eye disease, a condition in which the tears evaporate too quickly due to insufficient oil production by the meibomian glands. MGD can also occur in conjunction with anterior blepharitis (5, 6).

The Significance of Patient Education in Blepharitis Management

Effective blepharitis management relies on a multifaceted approach that includes proper lid hygiene, warm compresses, and sometimes topical medications or oral antibiotics (4). However, the cornerstone of successful treatment lies in patient education. Empowering patients with accurate and comprehensive knowledge about their condition equips them to actively participate in their own care and adhere to prescribed treatment regimens (7).

Limitations of Traditional Patient Education Methods

Traditionally, patient education on blepharitis has relied on methods such as brochures, websites, and in-person consultations with ophthalmologists. While these methods can provide valuable information, they often lack accessibility, engagement, and personalization (8). Busy schedules, complex medical jargon, and information overload can prevent effective knowledge transfer. Brochures and websites may not be readily available to patients, particularly those in underserved communities or with limited technological access. Additionally, the information presented in these materials may be written in complex medical language that is difficult for patients to understand. In-person consultations, while offering the opportunity for personalized interaction, can be time-consuming for both patients and healthcare professionals, limiting the frequency and depth of education provided.

The Rise of Chatbots and their Potential for Patient Education

The emergence of chatbots, conversational AI (artificial intelligence) programs designed to simulate human conversation, presents a novel and potentially transformative approach to patient education. Chatbots can offer readily accessible, interactive, and on-demand information delivery, potentially overcoming some of the limitations of traditional methods (9).

Within the healthcare domain, chatbots hold immense promises for disseminating health information, promoting patient engagement, and even offering basic symptom assessment and triage. Studies have explored the application of chatbots in various areas. This list is not exhaustive, but it effectively highlights the diverse applications of chatbots in healthcare education and support (9, 10).

Research Gap and Rationale for the Study

While research has begun to explore the potential of chatbots in healthcare education, their application to educating patients about blepharitis remains relatively unexplored. Given the prevalence of blepharitis and the crucial role of patient education in its management, investigating the efficacy of chatbots in this context can be highly valuable.

This study aims to address this gap in knowledge by assessing the quality and reliability of information delivered by chatbots on blepharitis compared to a trusted source (the American Academy of Ophthalmology website). By evaluating chatbot responses using established scoring tools employed by ophthalmologists, this research will offer insights into the potential of chatbots to serve as a complementary tool for patients.

Materials and Methods

This study used a descriptive comparative design to evaluate the quality and reliability of information delivered by chatbots on blepharitis compared to a trusted source, the American Academy of Ophthalmology (AAO) website.

Selection of Frequently Asked Questions (FAQs) on Blepharitis

To ensure a comprehensive and representative sample of patient concerns, a two-step process was used to select the FAQs on blepharitis.

Step 1: The AAO website's "Eye Conditions" and “Ask an ophthalmologist” sections were visited. From this webpage, the twenty (20) most frequently asked questions (FAQs) were identified and documented. This initial selection ensures the chosen questions reflect the most common patient inquiries about blepharitis.

Step 2: Two ophthalmologists, independent of the study team, reviewed the twenty (20) initial FAQs.

This two-step approach ensures the selected FAQs represent a combination of the most commonly viewed questions on the AAO website and additional questions deemed crucial by ophthalmologists for comprehensive patient education.

Selection of Chatbots

Two large language models (LLMs) will be chosen for this study: Gemini and ChatGPT 3.5.

Gemini, developed by Google, is a generative AI that can generate text, answer questions, and perform tasks similar to ChatGPT. It’s known for its multimodal capabilities, allowing it to produce not only text but also images, video, audio, and code. Initially launched as Bard, it was later upgraded to the Gemini large language model and rebranded, offering various tiers for different user needs.

ChatGPT 3.5 is a conversational AI developed by OpenAI, capable of understanding and generating human-like text. It supports a wide range of languages and is fine-tuned for conversational applications, though it performs best in English. The model is built on a transformer architecture and can be used for various tasks, from answering questions to creative writing and code generation.

The data sources were responses to 20 frequently asked questions about blepharitis on the AAO website and outputs from ChatGPT 3.5 and Gemini to these questions as prompts.

As of March 3, 2024, we accessed the latest versions of chatbots on the websites of ChatGPT free version (which uses the GPT-3.5 engine) and Gemini of Google. The chat history and cache of each chatbot was cleared before the next question. This was done to ensure that each chatbot did not remember or use any previous information or conversation in answering the questions.

Expertise of Participating Ophthalmologists

Two ophthalmologists, board-certified in ophthalmology and with experience managing blepharitis, will be recruited to participate in this study. Their expertise in diagnosing and treating blepharitis is crucial for evaluating the accuracy, comprehensiveness, and appropriateness of the information delivered by the chatbots.

Prior to commencing the evaluation, the ophthalmologists were briefed on the study objectives and methodology. They were provided with a standardized scoring sheet outlining the specific criteria for each assessment tool. The responses were blinded to evaluators before scoring.

Scoring Tools for Chatbot Response Evaluation

Three established scoring tools will be employed to evaluate the quality and reliability of the chatbot responses:

Patient Education Material Assessment Tool (PEMAT): The PEMAT is a standardized tool used to assess the understandability and actionability of educational materials for patients (11). It consists of two subscales:

PEMAT-U (Understandability) assesses factors like clarity, organization, layout, and use of language (17 items). Scores are calculated as a percentage (0-100%) with higher scores indicating better understandability for patients.

PEMAT-A (Actionability) assesses the specificity, practicability, and effectiveness of suggested actions (7 items). Scores are also calculated as a percentage (0-100%) with higher scores indicating clearer and more actionable information for patients.

SOLO Taxonomy (1-5): Developed by Biggs and Collis, the SOLO Taxonomy is a framework that categorizes the level of cognitive demand required to understand information (12). This tool will be used to assess the complexity and depth of knowledge conveyed in the chatbot responses. The SOLO Taxonomy categorizes learning outcomes into five levels:

Prestructural: Responses at this level demonstrate a lack of understanding of the question or task.

Unistructural: Responses focus on a single, isolated fact or aspect of the topic.

Multistructural: Responses include multiple facts or details but lack connection or integration.

Relational: Responses demonstrate a relational understanding by connecting different aspects of the topic or explaining cause-and-effect relationships.

Extended Abstract: Responses at this highest level demonstrate a deep and comprehensive understanding of the topic. They may involve applying knowledge to new situations, making generalizations, or drawing critical judgments.

By analysing the SOLO Taxonomy level of the chatbot responses, the ophthalmologists can assess whether the information provided is appropriate for patient education and fosters a deeper understanding of blepharitis.

Flesch-Kincaid Readability Ease Score (0-100): The Flesch-Kincaid readability ease score is a metric used to assess the reading difficulty of a text. It is calculated based on the average number of words per sentence and the average number of syllables per word. Higher scores indicate easier readability, with scores ranging from 0 (most difficult) to 100 (easiest). A score of 60-70 corresponds to a typical eighth or ninth grade reading level, which is considered a suitable target for patient education materials (13). This ensures the information is comprehensible for a broad audience of patients with varying levels of education.

The Flesch-Kincaid score was used to evaluate the readability of the chatbot responses. Scores within the 60-70 range will be considered optimal for patient education materials on blepharitis.

Data Analysis

Following the scoring, the collected data was analysed using descriptive statistics. This involved calculating measures of central tendency (mean, median) and dispersion (standard deviation) for the scores obtained from each evaluation tool (PEMAT-U, PEMAT-A, SOLO Taxonomy, and Flesch-Kincaid).

Additionally, a comparative analysis (independent sample t-test or Mann Whitney U) was performed to assess any significant differences between the information quality provided by the two chatbots (Gemini and ChatGPT 3.5) and the information from the trusted source (AAO website).

Ethical Considerations

Since this project does not involve human subjects, approval from the Institutional Review Board was not required.

Results

The data presented in Table 1 allows for a comparative analysis of the information quality, reliability, and readability provided by the chatbots (Gemini and ChatGPT 3.5.)

Understandability (PEMAT-U): Both chatbots achieved scores exceeding 78% on the PEMAT-U understandability subscale, indicating that the information they delivered was generally clear and easy for patients to comprehend. A statistical analysis using an independent samples t-test revealed a trend towards higher scores for Gemini compared to ChatGPT 3.5. However, this difference did not reach statistical significance at the alpha level of 0.05 (p=0.152).

Table 1. Scores from Ophthalmologist Evaluations.

Scoring Tool	Chatbot	Ophthalmologist 1 (Mean)	Ophthalmologist 2 (Mean)	Mean Score (SD)
PEMAT-U (Understandability)	Gemini	82%	85%	83.5% (9.42)
	ChatGPT 3.5	78%	80%	79.0% (8.91)
PEMAT-A (Actionability)	Gemini	49%	50%	49.5% (10.21)
	ChatGPT 3.5	45%	48%	46.5% (8.89)
SOLO Taxonomy	Gemini	3.5	3.5	3.5 Relational (0.51)
	ChatGPT 3.5	3.5	3.4	3.45 Relational (0.6)
Flesch-Kincaid Readability*	Gemini			38.75 (8.3)
	ChatGPT 3.5			26.35 (8.73)

*p<0.001

Actionability (PEMAT-A): Scores on the PEMAT-A actionability subscale were lower than understandability scores for both chatbots. While the information provided was understandable, it could be further improved in terms of offering specific and actionable steps for patients to manage their blepharitis. Gemini again received slightly higher scores, indicating a relatively clearer presentation of recommended actions for patients. Mann Whitney U test analysis showed a non-significant difference between Gemini and ChatGPT 3.5 (p=0.28462).

Depth of Knowledge (SOLO Taxonomy): Both ophthalmologists categorized the responses from both chatbots as almost reaching the "Relational" level within the SOLO Taxonomy. This indicates that the chatbots were able to connect different aspects of blepharitis and explain cause-and-effect relationships, demonstrating a more comprehensive understanding of the condition than isolated facts. Mann Whitney U test analysis showed a non-significant difference between Gemini and ChatGPT 3.5 (p=0.90448).

Readability (Flesch-Kincaid ease): The Flesch-Kincaid scores for both chatbots fell below the 60-70 range (Gemini:38.75, ChatGPT 3.5: 26.35), which is considered poor score for patient education materials. A statistically significant difference was found between the chatbots (t(2)=4.60, p=0.00004557). Gemini scored higher than ChatGPT 3.5 suggesting advantage in readability for Gemini's responses.

Table 2. Comparison of Information Depth (SOLO Taxonomy) of chatbots and the AAO website.

FAQ Topic	Chatbot Response (Level)	AAO Website FAQ (Level)
Symptoms of Blepharitis	Relational (Both Chatbots)	Relational
Causes of Blepharitis	Relational (Both Chatbots)	Extended Abstract
Treatment Options for Blepharitis	Multistructural (Both Chatbots)	Relational
Preventing Blepharitis	Multistructural (Both Chatbots)	Relational

Table 2 allows for a more granular comparison of how each FAQ topic is addressed by the chatbots and the AAO website.

It highlights that while the chatbots can demonstrate a relational understanding by connecting different aspects of blepharitis, the AAO website, likely due to its creation by medical professionals, may delve deeper into certain topics, reaching the "Extended Abstract" level.

Table 3. Inter-Rater Reliability for Ophthalmologist Evaluations.

Scoring Tool	Kappa Statistic
PEMAT-U (Understandability)	0.85
PEMAT-A (Actionability)	0.80
SOLO Taxonomy	0.75

Table 3 presents the Kappa statistic, a measure of inter-rater reliability between the two ophthalmologists who evaluated the chatbot responses. All Kappa values fall within the "substantial" agreement range according to benchmarks (14). This indicates a good level of consistency between the ophthalmologists in their scoring of the information provided by the chatbots using the PEMAT and SOLO Taxonomy tools. This strengthens confidence in the overall findings of the study.

Discussion

Leveraging Chatbots for Enhanced Patient Education in Ophthalmology

This study investigated the potential of chatbots as a tool for patient education on blepharitis. By analysing information quality, readability, and actionability, the research yielded valuable insights into both the strengths and weaknesses of this emerging technology in healthcare.

Strengths of Chatbot Information for Patient Education

The findings revealed that both Gemini and ChatGPT 3.5 delivered information on blepharitis that was generally clear and understandable (PEMAT-U scores exceeding 78%). This aligns with existing research highlighting the potential of chatbots to simplify complex medical topics for patients (15-17). Conversational agents, with their ability to access and process vast amounts of text data and generate human-like conversation, empowers patients by providing them with on-demand access to a wealth of medical knowledge (18).

Furthermore, both chatbots demonstrated a close to relational understanding of blepharitis, connecting various aspects of the condition and explaining cause-and-effect relationships (SOLO Taxonomy). This aligns with studies suggesting that chatbots can go beyond isolated facts, presenting a more comprehensive picture of health topics (19, 20). This can be particularly beneficial for patients seeking to understand the underlying mechanisms of their condition and how their behaviours or lifestyle choices might influence it.

Beyond basic information delivery, the chatbots also offered some actionable recommendations, such as maintaining lid hygiene and using warm compresses. While there is room for improvement (lower PEMAT-A scores suggesting a need for more specific action steps), these initial steps could empower patients to take a more active role in managing their blepharitis. This aligns with the growing emphasis on patient-centred care and self-management of chronic conditions (21). However, it is important to acknowledge that, chatbots in their current form are not designed to provide definitive medical advice. Their reliance on training data, which may include outdated information, necessitates human oversight and fact-checking before delivering chatbot-generated content to patients.

Despite these strengths, the study also revealed areas for improvement in chatbot-delivered patient education. Analysis of readability scores revealed that the Flesch-Kincaid ease score for both chatbots fell below the recommended range for patient education materials (60-70). Specifically, the scores for Gemini and ChatGPT 3.5 were 38.75 and 26.35, respectively. This suggests that the information presented by the chatbots may be too complex for the intended audience of patients seeking health information.

One key limitation was the lack of in-depth information compared to the AAO website (SOLO Taxonomy). While chatbots reached close to relational level, the website potentially reached an extended abstract level, providing a more nuanced and detailed understanding. This highlights the current limitations of chatbots in replicating the comprehensiveness of established medical resources (22).

Chatbots medical knowledge bases could be enhanced by integrating with established databases and collaborating with ophthalmologists to curate information. This provides a more nuanced understanding of eye conditions, including potential variations and treatment considerations. Additionally, chatbots can personalize recommendations, empowering patients to take an active role in managing their condition by suggesting specific treatment approaches or medication use (23).

Another area for improvement is the provision of actionable recommendations. The chatbots primarily focused on general hygiene practices, and specific details regarding treatment options or medication use were often lacking. This could be due to limitations in the chatbot's programming or the need for further development to tailor recommendations to individual patients' needs, as discussed earlier.

Developing chatbots with multilingual functionalities can improve accessibility for patients with diverse linguistic backgrounds. This is crucial for ensuring equitable access to health information and empowering patients from all communities to actively participate in their healthcare (24). Imagine a chatbot that can seamlessly switch between languages based on the patient's preference, ensuring they receive clear and accurate information regardless of their native tongue (25).

Ethical Considerations and Human Oversight

Furthermore, the role of human oversight and ethical considerations should remain paramount in the development and implementation of chatbots for patient education. Integrating chatbots into ophthalmic care should be done in a way that complements, not replaces, the expertise of healthcare professionals. Chatbots can act as a first point of contact, providing basic information and education while directing patients to seek professional medical advice for diagnosis and treatment decisions. Additionally, ensuring data privacy and security is crucial when collecting and storing patient information through chatbots (26).

The current study acknowledges several limitations. Firstly, it focused on evaluating only two large language models (Gemini and ChatGPT 3.5). The evolving landscape of chatbot development suggests that a wider range of chatbots with potentially different functionalities and information delivery approaches might yield varying results. Future research should explore a more diverse selection of chatbots to provide a more comprehensive understanding of their capabilities in patient education.

Secondly, using convenience sampling method, the study evaluated a limited set of FAQs on blepharitis. While these FAQs covered core aspects of the condition, a broader range of topics could reveal additional strengths or weaknesses in the chatbots' information delivery. Additionally, the information provided by the AAO website, used as a benchmark, might not encompass the full spectrum of details a medical professional would address during a consultation. Expanding the scope of evaluated topics and incorporating consultations with ophthalmologists as a reference point could provide a more holistic understanding of how chatbots compare to traditional patient education methods.

Conclusion

The findings of this study suggest that chatbots hold promise as valuable tools for patient education on blepharitis. They can deliver clear, understandable, and relatable information, potentially empowering patients and improving their health literacy. However, there is room for improvement in terms of comprehensiveness, personalization of recommendations, and interactivity. Future research should explore these areas while emphasizing the importance of human oversight, ethical considerations, and data privacy. As chatbot technology continues to evolve, ophthalmologists and healthcare professionals have the opportunity to leverage these advancements to enhance patient education, promote self-management of eye conditions, and ultimately improve patient care outcomes.

Acknowledgements

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used QuillBot tool to improve language and readability, with caution. After using this tool/service, the author(s) reviewed and edited the content as needed and took full responsibility for the content of the publication.

Conflict of Interest

The authors declare no conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Eberhardt M, Rammohan G. Blepharitis. In: StatPearls. Treasure Island (FL): StatPearls Publishing; January 23, 2023.
2. Trattler W, Karpecki P, Rapoport Y, et al. The Prevalence of Demodex Blepharitis in US Eye Care Clinic Patients as Determined by Collarettes: A Pathognomonic Sign. Clinical Ophthalmology 2022;16; doi: 10.2147/OPTH.S354692.
3. Pflugfelder SC, Karpecki PM, Perez VL. Treatment of blepharitis: Recent clinical trials. Ocular Surface 2014;12(4):273–284; doi: 10.1016/j.jtos.2014.05.005.
4. Duncan K, Jeng BH. Medical Management of Blepharitis. Curr Opin Ophthalmol 2015;26(4); doi: 10.1097/ICU.0000000000000164.
5. Shah PP, Stein RL, Perry HD. Update on the Management of Demodex Blepharitis. Cornea 2022;41(8); doi: 10.1097/ICO.0000000000002911.
6. Nichols KK, Foulks GN, Bron AJ, et al. The International Workshop on Meibomian Gland Dysfunction: Executive Summary. Investigative Opthalmology & Visual Science 2011;52(4):1922; doi: 10.1167/iovs.10-6997a.
7. Rhodes L, Huisingh C, McGwin G, et al. Eye Care Quality and Accessibility Improvement in the Community (EQUALITY): impact of an eye health education program on patient knowledge about glaucoma and attitudes about eye care. Patient Relat Outcome Meas 2016; doi: 10.2147/prom.s98686.
8. Görtz M, Baumgärtner K, Schmid T, et al. An artificial intelligence-based chatbot for prostate cancer education: Design and patient evaluation study. Digit Health 2023;9. doi: 10.1177/20552076231173304.

9. Frangoudes F, Hadjiaros M, Schiza EC, et al. An Overview of the Use of Chatbots in Medical and Healthcare Education. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2021;12785 LNCS:170–184; doi: 10.1007/978-3-030-77943-6_11/cover.
10. Chaix B, Bibault JE, Pienkowski A, et al. When chatbots meet patients: One-year prospective study of conversations between patients with breast cancer and a chatbot. JMIR Cancer 2019;5(1); doi: 10.2196/12856.
11. PEMAT for Printable Materials (PEMAT-P) | Agency for Healthcare Research and Quality. Available from: https://www.ahrq.gov/health-literacy/patient-education/pemat-p.html.
12. Biggs J, Collis K. Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome). 2014.
13. Kincaid PJ. Derivation of New Readability Formulas. Naval Technical Training Command Millington TN Research Branch 1975;(February).
14. Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas 1960;20(1); doi: 10.1177/001316446002000104.
15. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Switzerland) 2023;11(6); doi: 10.3390/healthcare11060887.
16. Zheng Y, Wang L, Feng B, et al. Innovating Healthcare: The Role of ChatGPT in Streamlining Hospital Workflow in the Future. Ann Biomed Eng 2023; doi: 10.1007/S10439-023-03323-W.
17. Cascella M, Montomoli J, Bellini V, et al. Evaluating the Feasibility of ChatGPT in Healthcare: An Analysis of Multiple Clinical and Research Scenarios. J Med Syst 2023;47(1); doi: 10.1007/S10916-023-01925-4.
18. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: A systematic review. Journal of the American Medical Informatics Association 2018;25(9):1248–1258; doi: 10.1093/jamia/ocy072.
19. Car LT, Dhinagaran DA, Kyaw BM, et al. Conversational agents in health care: Scoping review and conceptual analysis. J Med Internet Res 2020;22(8); doi: 10.2196/17158.
20. Ayers JW, Zhu Z, Poliak A, et al. Evaluating Artificial Intelligence Responses to Public Health Questions. JAMA Netw Open 2023;6(6); doi: 10.1001/jamanetworkopen.2023.17517.
21. Wu Q, Jiang S. The Effects of Patient-Centered Communication on Emotional Health: Examining the Roles of Self-Efficacy, Information Seeking Frustration, and Social Media Use. J Health Commun 2023;28(6):349–359; doi: 10.1080/10810730.2023.2208537.
22. Bhirud N, Tatale S, Randive S, et al. A Literature Review On Chatbots In Healthcare Domain. 2019.
23. Lin J, Joseph T, Parga-Belinkie JJ, et al. Development of a practical training method for a healthcare artificial intelligence (AI) chatbot. BMJ Innov 2021;7(2):441–444; doi: 10.1136/bmjinnov-2020-000530.
24. Multilingualism and WHO. n.d. Available from: https://www.who.int/about/policies/multilingualism.
25. Crossing the Quality Chasm: A New Health System for the 21st Century. 2001; doi: 10.17226/10027.
26. Li H, Moon JT, Purkayastha S, et al. Ethics of large language models in medicine and medical research. Lancet Digit Health 2023;5(6):e333–e335; doi: 10.1016/S2589-7500(23)00083-3.