AI Engineering
https://www.sciltp.com/journals/aieng

Article

Knowledge of Chatbots on Oriental Medicine-Based Theory and Practice of Acupuncture Is Inferior to That on Western Medicine-Based Acupuncture-Transmitted Infections

Hsi Liang [1,2], Zi-Jie Lee [3], Chi-Ching Tsang [4], Susanna K. P. Lau [5] and Patrick C. Y. Woo [1,5,6,*]

[1] Doctoral Program in Translational Medicine and Department of Life Sciences, National Chung Hsing University, Taichung 402, Taiwan
[2] Fuke Fengze Chinese Medicine Clinic, Taichung 407, Taiwan
[3] Institute of Molecular Biology, National Chung Hsing University, Taichung 402, Taiwan
[4] School of Medical and Health Sciences, Tung Wah College, Homantin, Hong Kong
[5] Department of Microbiology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong
[6] The iEGG and Animal Biotechnology Research Center, National Chung Hsing University, Taichung 402, Taiwan
* Correspondence: pcywoo@hku.hk; Tel.: +886-4-22840370 (ext. 19); Fax: +886-4-22860164

How To Cite: Liang, H.; Lee, Z.-J.; Tsang, C.-C.; et al. Knowledge of Chatbots on Oriental Medicine-Based Theory and Practice of Acupuncture Is Inferior to That on Western Medicine-Based Acupuncture-Transmitted Infections. *AI Engineering* **2026**, *2*(1), 4. https://doi.org/10.53941/aieng.2026.100004

Received: 23 January 2026
Revised: 11 April 2026
Accepted: 9 June 2026
Published: 15 June 2026

Abstract: Background—Chatbots are increasingly used in clinical medicine. However, all studies that examined the applications of chatbots so far were on their usefulness in western medicine. There has been no study that explored their knowledge on alternative medical practices. Objectives—To examine the performance of five free chatbots (ChatGPT, Claude, Copilot, Gemini and Perplexity) as well as the subscribed versions of ChatGPT, Claude and Perplexity in answering questions on oriental medicine-based theory and practice of acupuncture and western medicine-based acupuncture-transmitted infections. Methods—A total of 2528 multiple choices and true/false questions from four books on theory and practice of acupuncture, and 100 questions on acupuncture-transmitted infections set by two microbiology and infectious disease professors, were used for evaluation of the chatbots. Results—Overall, the median score for the eight free/subscribed chatbots in answering questions on theory and practice of acupuncture (75%) was significantly lower than that on acupuncture-transmitted infections (86%) ($p < 0.001$). Further analysis also showed significantly lower median score for all the five free chatbots and all the three subscribed chatbots in answering the questions on theory and practice of acupuncture (75% for both) than those on acupuncture-transmitted infections (86% and 90% respectively) ($p = 0.014$ and $p = 0.003$ respectively). For the three subscribed chatbots, GPT-4o achieved the highest median score (84%), followed by Perplexity Pro (79%) and Claude 3 Opus (66%) for the questions on theory and practice of acupuncture ($p = 0.036$ by Kruskal Wallis test). Post-hoc Dunn's test revealed that the median score for GPT-4o was significantly higher than Claude 3 Opus ($p = 0.032$). Conclusions—Performance of the chatbots on oriental medicine-based theory and practice of acupuncture was inferior to that on western medicine-based acupuncture-transmitted infections. In order to further improve their usefulness, efforts should be spent on improving the chatbots' knowledge on the theory and practice of acupuncture.

Keywords: chatbots; oriental medicine; acupuncture; western medicine

Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Artificial intelligence (AI) has become an increasingly important tool in clinical medicine, supporting clinicians across diagnosis, treatment and healthcare delivery. One of the most established uses of AI is in medical imaging and diagnostics. Machine-learning algorithms, particularly deep learning can analyze radiological images to detect abnormalities, such as cancers, fractures, and strokes [1]. In pathology, AI systems assist with digital slide analysis by identifying malignant cells and grading tumor [2,3]. In microbiology, we have recently shown in a proof-of-concept study that it was useful for identification of Aspergillus species, and other groups have also used similar approaches for parasite classification and diagnosis of bacterial vaginosis [4–8]. In ophthalmology and dermatology, validated AI tools have been used to screen for conditions such as diabetic retinopathy and skin cancer [9]. These applications improve diagnostic accuracy, reduce reporting times and help address workforce shortages.

ChatGPT and other AI-powered chatbots are rooted in the need to make large amounts of information more accessible, interactive and useful to humans. As digital data has grown exponentially, traditional interfaces have become increasingly inefficient for complex, conversational or context-dependent tasks. Chatbots based on large language models aim to bridge this gap by enabling natural, human-like interaction with computers. At their core, chatbots are designed to understand and generate natural language. They are trained on vast datasets containing text from books, articles, websites, and other sources, allowing them to learn patterns in language, reasoning and communication. This enables users to ask questions, give instructions, or explore ideas in everyday language. Chatbots are increasingly used in clinical medicine as tools that support patients, clinicians and healthcare systems by improving access to information, efficiency and continuity of care. Rather than replacing clinicians, they function as adjuncts that enhance communication, decision-making, and workflow. Recently, we have demonstrated that chatbots were not inferior to professors in their infectious disease knowledge and data analysis [10]. However, all the chatbots that were evaluated (i.e., ChatGPT, Claude, Copilot, Gemini and Perplexity) were products from the western world. Moreover, all the studies that have examined the applications of these chatbots focused on their performance in answering questions on western medicine. So far, no study has examined the performance of these chatbots on alternative medical practices that are widely employed by both westerners and non-westerners for the prevention and treatment of medical diseases.

Among all modalities of alternative medicine, acupuncture is one of the most widely recognized and evidence-based (Figure 1A). In traditional Chinese medicine, acupuncture is based on the concept of qi (vital energy) flowing through pathways called meridians. Illness and pain are thought to arise when the flow of qi is disrupted or imbalanced. Acupuncture aims to restore balance by stimulating specific acupuncture points along the meridians, thereby normalizing qi flow and supporting the body's natural healing processes. One of its most commonly used areas in clinical practice is pain and related disorders. Chronic pain conditions are the most frequent reason for referral, particularly low back pain, neck pain, knee osteoarthritis. In addition to pain disorders, acupuncture is also well-established in the treatment of nausea and vomiting, including post-operative nausea, chemotherapy-induced nausea and pregnancy-induced nausea. In this study, we examined the performance of five of these software applications, including ChatGPT, Claude, Copilot, Gemini and Perplexity; as well as the free and subscribed versions of ChatGPT, Claude and Perplexity; in answering multiple choice questions (MCQs) and true/false (T/F) questions obtained from four books on the theory and practice of acupuncture as well as T/F questions on acupuncture transmitted infections.

Figure 1. Panel **A**, example of a 23-year-old woman presented with right forearm pain undergoing acupuncture treatment. The acupuncture points LI4 (Hegu, or joining valley) (blue arrow) and LI10 (Shousanli) (yellow arrow) on the large intestine meridian were punctured. Panel **B**, the four books where questions on the theory and practice of acupuncture were extracted in this study.

2. Methods

2.1. Chatbots and Questions

Five chatbots and three subscribed versions (Supplementary Table S1) were examined using MCQs and T/F questions extracted from four books on the theory and practice of acupuncture (Figure 1B). The first book, *750 Questions and Answers about Acupuncture: Exam Preparation and Study Guide* (Book 1), contains five sections of questions [11]. The second one, *Tests: Chinese Acupuncture and Moxibustion* (Book 2), contains five sections of questions [12]. The third one, *TCM Study Guide, Acupuncture Board Exam Practice Questions* (Book 3), contains 8 chapters of questions [13]. The fourth one, *Acupuncture Board Exam Study Questions—DIFFERENTIAL DIAGNOSIS* (Book 4), contains 12 chapters of questions [14]. All questions that required interpretation of pictures or images were excluded. Therefore, a total of 2528 questions on the theory and practice of acupuncture, which included 650, 543, 1017 and 318 questions extracted from Book 1, Book 2, Book 3 and Book 4 respectively, were used in the study. All the questions were examined by a practicing doctor on Chinese medicine and acupuncture to ensure that they were of good quality and the answers provided by the books reasonable. In addition, 100 T/F questions related to acupuncture transmitted infections [15–28], set and cross-validated by two microbiology and infectious disease professors, were used as controls (Supplementary Table S2). These 100 questions were divided into four subsets, including (1) pyogenic bacterial infections associated with acupuncture, (2) mycobacterial infections associated with acupuncture, (3) viral infections associated with acupuncture, and (4) infection control for acupuncture.

To ensure methodological consistency and reproducibility, a standardized prompting protocol was implemented for all chatbot evaluations. Each question was presented using a uniform template: “Please answer the following multiple-choice question by selecting the single best answer. Respond with the letter corresponding to your choice only (e.g., A, B, C, or D). Do not provide any explanation.” This instruction was followed immediately by the question and its answer options. No additional contextual cues or leading information were provided. Moreover, each question was asked in a new, independent session to prevent any carryover effects from prior conversations. This approach ensured that responses were not influenced by conversational memory. Furthermore, no prior dialogue or contextual history was included when presenting each question. This eliminated potential bias arising from accumulated context.

2.2. Scoring

For each question extracted from the four books as well as the 100 questions on acupuncture transmitted infections, it was used for testing the five free and three subscribed chatbots. Answers provided by the four books on the theory and practice of acupuncture were regarded as correct answers. The scoring scheme was as follows: one point given for each correct answer, no point deducted for wrong answers, and 0.5 and 0.2 points awarded for T/F questions and MCQs respectively that the chatbot gave a pass. The study was carried out from February to April 2025.

2.3. Runtime Calculation

A laptop computer with its networking specifications listed in Supplementary Table S3 was used for estimating the chatbots’ speed of performance. The first 90 questions from each of the four books were grouped into 3 sets (i.e., 30 questions per set) and each set of questions was fed to the chatbots and the time for running each set of question was recorded and analyzed.

2.4. Statistical Analysis

Comparison of the scores for questions on theory and practice of acupuncture and acupuncture transmitted infections, as well as those between the free and subscribed versions of five free chatbots (Claude 3.5 Sonnet, Copilot, Gemini, GPT-4o mini and Perplexity) and all three subscribed chatbots (Claude 3 Opus, GPT-4o and Perplexity Pro), was done by Mann-Whitney U-test. The Kruskal-Wallis test, with post-hoc Dunn’s test, was used to compare the scores for the five free chatbots and the three subscribed ones, as well as the chatbot runtime.

3. Results

Performance of chatbots for questions on theory and practice of acupuncture vs. questions on acupuncture transmitted infections. Overall, the median score for the eight free and subscribed chatbots for the questions on theory and practice of acupuncture (75%) was significantly lower than that on acupuncture transmitted infections

(86%) ($p < 0.001$) (Figure 2A). Further analysis also showed significantly lower median score for all five free chatbots (Claude 3.5 Sonnet, Copilot, Gemini, GPT-4o mini and Perplexity) and all three subscribed chatbots (Claude 3 Opus, GPT-4o and Perplexity Pro) in answering the questions on theory and practice of acupuncture (75% for both) than those on acupuncture transmitted infections (86% and 90% respectively) ($p = 0.014$ and $p = 0.003$ respectively) (Figure 2B).

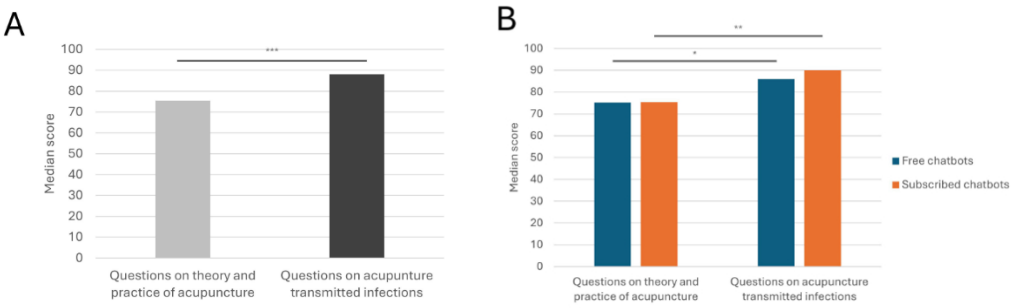

Figure 2. Panel **A**, Overall comparison of the scores for the free and subscribed chatbots on the theory and practice of acupuncture and questions on acupuncture-transmitted infections. Panel **B**, Comparison of the five free chatbots and the three subscribed ones for questions on the theory and practice of acupuncture and questions on acupuncture-transmitted infections. *, $p \leq 0.05$; **, $p \leq 0.005$; ***, $p \leq 0.001$.

3.1. Performance of Free Chatbots

As a whole, no significant difference among the median scores obtained by the five free chatbots for both the questions on theory and practice of acupuncture in the four books and the questions on acupuncture transmitted infections was observed.

3.2. Performance of Subscribed Chatbots

As a whole, GPT-4o achieved the highest median score (84%), followed by Perplexity Pro (79%) and Claude 3 Opus (66%) for the questions on theory and practice of acupuncture in the four books (Figure 3). Statistically, significant difference was observed among the median scores of these three subscribed chatbots in answering the questions on theory and practice of acupuncture in the four books ($p = 0.036$ by Kruskal Wallis test). Post-hoc Dunn's test revealed that the score of GPT-4o was significantly higher than Claude 3 Opus ($p = 0.032$) (Figure 3). As for the questions on acupuncture transmitted infections, no significant difference was detected among the scores of the three subscribed chatbots.

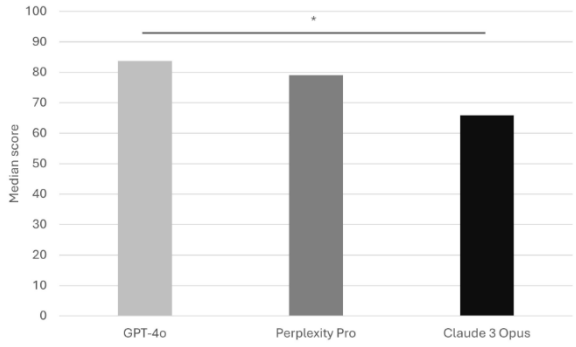

Figure 3. Comparison of the three subscribed chatbots in answering the questions on theory and practice of acupuncture. *, $p \leq 0.05$.

3.3. Free vs. Subscribed Versions

For both the four books of questions on theory and practice of acupuncture as well as the questions on acupuncture transmitted infections, no significant difference was detected between the median scores of GPT-4o and GPT-4o mini, as well as between Perplexity and Perplexity Pro and between Claude 3.5 Sonnet and Claude 3 Opus (Figure 4).

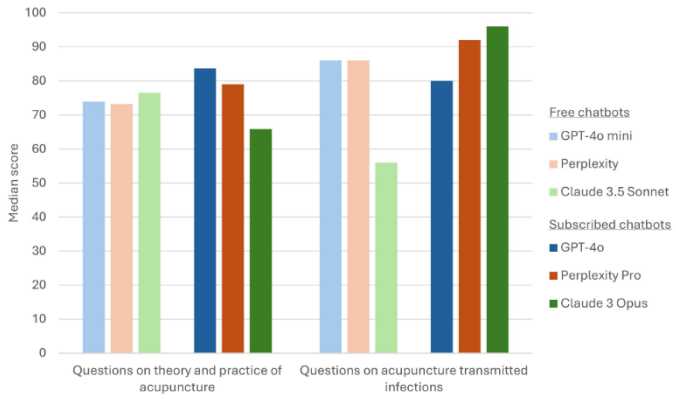

Figure 4. Comparison of Claude 3.5 Sonnet, GPT-4o mini and Perplexity and their subscribed counterparts (Claude 3 Opus, GPT-4o and Perplexity Pro) in answering questions on the theory and practice of acupuncture and questions on acupuncture-transmitted infections.

3.4. Runtime Calculation

For the five free chatbots, it was found that in general the median runtime was the longest for Claude, followed by Perplexity and the other three chatbots (Figure 5). As for the comparison between the free and subscribed versions, the median runtime was significantly longer for GPT-4o and Perplexity Pro when compared to their corresponding free counterparts [GPT-4o mini (27 s) vs. GPT-4o (14 s), $p = 0.019$ and Perplexity Pro (93 s) vs. Perplexity (15 s), $p = 1.86 \times 10^{-5}$ by Mann-Whitney U-test] (Figure 5).

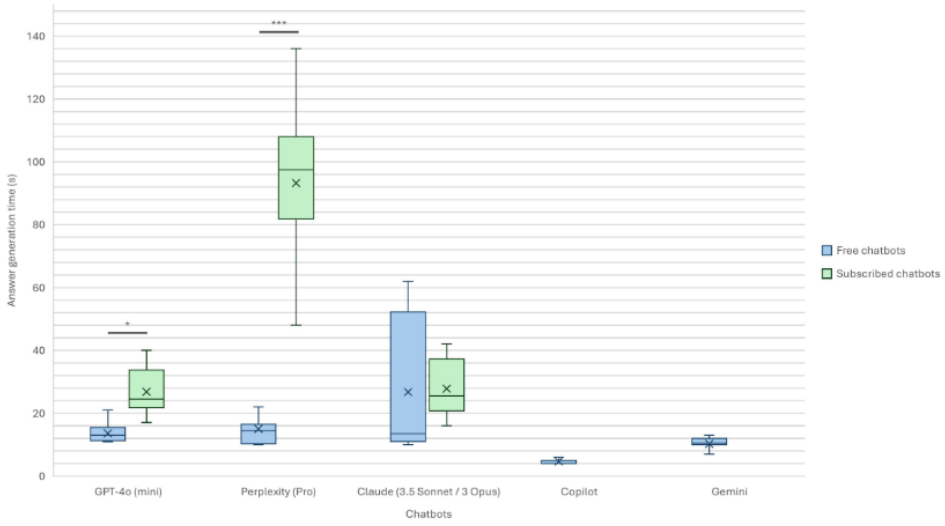

Figure 5. Answer generating time for the chatbots for questions on the theory and practice of acupuncture. *, $p \leq 0.05$; ***, $p \leq 0.001$.

4. Discussion

Among all forms of alternative medicine, acupuncture is both one of the most evidence based as well as widely accepted in a global scale; while at the same time, it is also closely monitored and tightly regulated in the western world. Back in the last millennium, Australia has published its first infection control guideline for acupuncture [29]. Currently, in the USA, one of the requirements of the common acupuncture licensure is to pass the National Certification Commission for Acupuncture and Oriental Medicine (NCCAOM) Board Examination; in the United Kingdom, acupuncturists have to obtain the British Acupuncture Council Certification in order to practice acupuncture; and in Germany, medical doctors can obtain acupuncture certification through the German Medical Acupuncture Association whereas non-medical practitioners have to obtain the Heilpraktiker (Naturopath) License in order to practice acupuncture. In this study, we compared the performance of five different chatbots, three of which included both free and subscribed versions, in answering more than 2500 questions on the theory and practice of acupuncture obtained from four books that were mostly used by candidates preparing for the NCCAOM Board Examination of the USA. The questions covered diverse subjects in the discipline of acupuncture

and moxibustion, such as meridians and acupuncture points, acupuncture techniques and modalities, safety and professionalism, as well as some short mini-case studies. Although some of the questions may not perfect and the correct answers arguable, such a problem was diluted by the large number of questions used and would not have significant effect on the overall results. From the results of this study, it was shown that for the free versions of the five chatbots themselves, no significant difference on the scores achieved was observed. However, for the subscribed versions of the chatbots, GPT-4o performed significantly better than Claude 3 Opus. As a whole, it seems that GPT-4o appeared to be the best chatbot on acupuncture related knowledge, followed by Copilot, Perplexity Pro, etc. In general, the different chatbots performed differently because they are built with varying training data, model architectures, tuning methods, and safety rules which affect how well they understand questions, reason, and generate accurate responses. More specifically, the relatively superior performance of GPT-4o could be because it has integrated strong reasoning, higher accuracy in clinical decision-making and adaptability [30–33], allowing it to achieve high performance even in specialized tasks even without explicit cultural or domain-specific optimization.

The most attention-grabbing result revealed in the present study was that the performance of the chatbots on the theory and practice of acupuncture was inferior to that on acupuncture transmitted infections. In our previous study on the performance of chatbots in answering more than 2500 questions from three books on microbiology and infectious disease, the overall median score for all the eight free and subscribed chatbots was 85% [10], which is comparable to the median score (86%) for the same eight chatbots for the 100 questions on acupuncture transmitted infections in the present study (Figure 2). These 100 questions on acupuncture transmitted infections that we set were based on the knowledge we acquired through our own research in the last 25 years as well as literature review on four different areas, including acupuncture transmitted bacterial infections, acupuncture transmitted mycobacterial infections, acupuncture transmitted viral infections, and infection control [15–19]. Although these four categories of knowledge-based questions were related to acupuncture, the approach of building up the knowledge was based on the scientific approach to microbiology and infectious disease. Therefore, these questions were used as controls for testing the chatbots on their performance in answering the questions on the theory and practice of acupuncture, which was primarily based on the oriental medicine approach. In contrast to a median score of 85–86% for the western medicine-based questions on clinical microbiology and infectious disease as well as acupuncture transmitted infections, the performance of the chatbots on the four books of questions on the theory and practice of acupuncture was significantly poorer, with an overall median score of only 75% (Figure 2). Such an overall inferior performance of the eight chatbots for questions on the theory and practice of acupuncture was also observed when further similar analysis on the five free chatbots or the three subscribed ones was performed on their scorings for questions on the theory and practice of acupuncture and those on acupuncture transmitted infections were compared (Figure 2). We speculate that such an inferior knowledge of the chatbots on acupuncture could be due to the less well-established databases of the chatbots on oriental medicine as compared to those on western medicine; or alternatively, specific terms and concepts in oriental medicine could be translated differently in different English textbooks and journals, which may lead to misunderstanding and misinterpretation. In fact, it is quite well-acknowledged that due to the inherent difference in linguistic complexity and conceptual framing between the two languages, English is more precise than Chinese for specific scientific and medical concepts and terms.

In order to further improve their usefulness, efforts should be spent on improving the chatbots' knowledge on the theory and practice of acupuncture. Acupuncture is not only practiced by fully trained licensed acupuncturists and oriental medicine practitioners, but also other health care professionals, such as physiotherapists, chiropractors, and nurse practitioners, who have completed the recognized training and use acupuncture as a means to relieve pain due to muscle spasm and menstruation, improve sleep pattern, etc. From the angle of health care professionals who perform acupuncture treatment on patients, they may consult chatbots for some ready-to-use knowledge, such as refreshing some concepts and hardcore knowledge in a timely manner, checking for the possibility of an alternative treatment approach when the original one has failed, and looking for a second or third opinion based on the information recorded in a few different ancient Chinese medicine books. Furthermore, from another perspective, as acupuncture is so widely used, it is not uncommon for clinicians from different specialties to encounter patients who have recent histories of acupuncture treatment. For example, orthopedic surgeons may see patients with low back pain, gynecologists may have consultations from women with dysmenorrhea or other menstrual symptoms, and psychiatrists may encounter clients with sleep disorders. Since not all clinicians have abundant knowledge of acupuncture, they may consult chatbots because they would like to understand in greater depth about the acupuncture treatment that their patients have gone through, which will facilitate their own management of the corresponding patients. Since the chatbots are frequently upgraded through ongoing cycles of improved training data, algorithm refinement, user feedback and software fixes to enhance

accuracy and performance and new chatbots are being developed, regular evaluation of the chatbots could be performed accordingly so that end-users will be updated on their usefulness.

Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/26061516 13159707/AIEng-26010121-SI.pdf. Table S1: Five chatbots and three subscribed versions evaluated in this study, Table S2: True/false questions on acupuncture transmitted infection and answers provided by chatbots, Table S3: Computer and networking for runtime calculation.

Author Contributions

Conceptualization, P.C.Y.W.; Data curation, H.L., Z.-J.L.; Formal analysis, H.L., C.-C.T., P.C.Y.W.; Funding acquisition, S.K.P.L., P.C.Y.W.; Investigation, S.K.P.L., P.C.Y.W.; Methodology, H.L., Z.-J.L., P.C.Y.W.; Project administration, H.L., Z.-J.L.; Supervision, S.K.P.L., P.C.Y.W.; Visualization, C.-C.T.; Writing—original draft, H.L., P.C.Y.W.; Writing—review & editing, H.L., Z.-J.L., C.-C.T., S.K.P.L., P.C.Y.W. All authors have read and agreed to the published version of the manuscript.

Funding

This work was partly supported by the Feature Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE-114-S-0023-A) in Taiwan.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

The patient has provided written informed consent for her photographs to be taken and used for publication and academic purposes. Images are presented with respect to the patient's privacy and confidentiality.

Data Availability Statement

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No additional AI tools except those mentioned in the manuscript were utilized for this paper.

References

1. Bera, K.; Braman, N.; Gupta, A.; et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat. Rev. Clin. Oncol.* **2022**, *19*, 132–146.
2. Försch, S.; Klauschen, F.; Hufnagl, P.; et al. Artificial intelligence in pathology. *Dtsch. Ärztebl. Int.* **2021**, *118*, 199.
3. Shafi, S.; Parwani, A.V. Artificial intelligence in diagnostic pathology. *Diagn. Pathol.* **2023**, *18*, 109.
4. Tsang, C.-C.; Zhao, C.; Liu, Y.; et al. Automatic identification of clinically important *Aspergillus* species by artificial intelligence-based image recognition: Proof-of-concept study. *Emerg. Microbes Infect.* **2024**, *14*, 2434573.
5. Van Noorden, R.; Webb, R. ChatGPT and science: The AI system was a force in 2023—for good and bad. *Nature* **2023**, *624*, 509.
6. Kumar, S.; Arif, T.; Alotaibi, A.S.; et al. Advances towards automatic detection and classification of parasites microscopic images using deep convolutional neural network: Methods, models and research directions. *Arch. Comput. Methods Eng.* **2023**, *30*, 2013–2039.
7. Wang, Z.; Zhang, L.; Zhao, M.; et al. Deep neural networks offer morphologic classification and diagnosis of bacterial vaginosis. *J. Clin. Microbiol.* **2021**, *59*, e02236-20. https://doi.org/10.1128/jcm.02236-20.
8. Song, Y.; He, L.; Zhou, F.; et al. Segmentation, splitting, and classification of overlapping bacteria in microscope images for automatic bacterial vaginosis diagnosis. *IEEE J. Biomed. Health Inform.* **2016**, *21*, 1095–1104.

9. Goyal, M.; Knackstedt, T.; Yan, S.; et al. Artificial intelligence-based image classification methods for diagnosis of skin cancer: Challenges and opportunities. *Comput. Biol. Med.* **2020**, *127*, 104065.
10. Wang, C.-S.; Hsiao, Y.; Tsou, C.-H.; et al. Chatbots are just as good as professors in both factual recall and clinical scenario analysis: Emergence of a new tool in clinical microbiology and infectious disease. *J. Infect.* **2024**, *89*, 106274.
11. Jennes, F. *750 Questions and Answers about Acupuncture: Exam Preparation and Study Guide*; Blue Poppy Enterprises, Inc.: Portland, OR, USA, 2003.
12. Yongqiang, C. *Tests: Chinese Acupuncture and Moxibustion*; Foreign Language Press: Beijing, China, 2001.
13. Luecke, K. *Acupuncture Board Exam Practice Questions*; Independently Published: Seattle, WA, USA, 2024.
14. Luecke, K. *Acupuncture Board Exam Study Questions—Differential Diagnosis*; Independently Published: Seattle, WA, USA, 2024.
15. Woo, P.C.; Li, J.H.; Tang, W.; et al. Acupuncture mycobacteriosis. *N. Engl. J. Med.* **2001**, *345*, 842–843.
16. Woo, P.C.; Leung, K.W.; Wong, S.S.; et al. Relatively alcohol-resistant *mycobacteria* are emerging pathogens in patients receiving acupuncture treatment. *J. Clin. Microbiol.* **2002**, *40*, 1219–1224.
17. Woo, P.C.; Lau, S.K.; Wong, S.S.; et al. *Staphylococcus aureus* subcutaneous abscess complicating acupuncture: Need for implementation of proper infection control guidelines. *New Microbiol.* **2003**, *26*, 169–174.
18. Woo, P.C.; Lau, S.K.; Yuen, K.Y. First report of methicillin-resistant *Staphylococcus aureus* septic arthritis complicating acupuncture: Simple procedure resulting in most devastating outcome. *Diagn. Microbiol. Infect. Dis.* **2009**, *63*, 92–95.
19. Woo, P.C.; Lin, A.W.; Lau, S.K.; et al. Acupuncture transmitted infections. *BMJ.* **2010**, *340*, c1268.
20. Walsh, B. Control of infection in acupuncture. *Acupunct. Med.* **2001**, *19*, 109–111.
21. Song, J.Y.; Sohn, J.W.; Jeong, H.W.; et al. An outbreak of post-acupuncture cutaneous infection due to *Mycobacterium abscessus*. *BMC Infect. Dis.* **2006**, *6*, 6.
22. Huang, C.-C.; Kotha, P.; Tu, C.-H.; et al. Acupuncture: a review of the safety and adverse events and the strategy of potential risk prevention. *Am. J. Chin. Med.* **2024**, *52*, 1555–1587.
23. Koh, S.J.; Song, T.; Kang, Y.; et al. An outbreak of skin and soft tissue infection caused by *Mycobacterium abscessus* following acupuncture. *Clin. Microbiol. Infect.* **2010**, *16*, 895–901.
24. Murray, R.; Pearson, J.; Coombs, G.; et al. Outbreak of invasive methicillin-resistant *Staphylococcus aureus* infection associated with acupuncture and joint injection. *Infect. Control Hosp. Epidemiol.* **2008**, *29*, 859–865.
25. Kim, Y.-J.; Kim, S.-H.; Lee, H.J.; et al. Infectious adverse events following acupuncture: clinical progress and microbiological etiology. *J. Korean Med. Sci.* **2018**, *33*, e164.
26. Tang, P.; Walsh, S.; Murray, C.; et al. Outbreak of acupuncture-associated cutaneous *Mycobacterium abscessus* infections. *J. Cutan. Med. Surg.* **2006**, *10*, 166–169.
27. Ryu, H.J.; Kim, W.J.; Oh, C.H.; et al. Iatrogenic *Mycobacterium abscessus* infection associated with acupuncture: clinical manifestations and its treatment. *Int. J. Dermatol.* **2005**, *44*, 846–850.
28. Chung, A.; Bui, L.; Mills, E. Adverse effects of acupuncture. Which are clinically significant? *Can. Fam. Physician* **2003**, *49*, 985–989.
29. Limited, A.A.A. *Infection Control Guidelines for Acupuncture*; Australian Acupuncture Association: West End, QLD, Australia, 1997.
30. Sanderson, K. GPT-4 is here: What scientists think. *Nature* **2023**, *615*, 773.
31. Abosi, O.J.; Kobayashi, T.; Ross, N.; et al. A head-to-head comparison of the accuracy of commercially available large language models for infection prevention and control inquiries, 2024. *Infect. Control Hosp. Epidemiol.* **2025**, *46*, 309–311.
32. Tao, K.; Osman, Z.A.; Tzou, P.L.; et al. GPT-4 performance on querying scientific publications: reproducibility, accuracy, and impact of an instruction sheet. *BMC Med. Res. Methodol.* **2024**, *24*, 139.
33. Lahat, A.; Sharif, K.; Zoabi, N.; et al. Assessing generative pretrained transformers (GPT) in clinical decision-making: comparative analysis of GPT-3.5 and GPT-4. *J. Med. Internet Res.* **2024**, *26*, e54571.