Application of ChatGPT-assisted problem-based learning teaching method in clinical medical education

Zeng Hui[1], Zhu Zewu[2], Hu Jiao[2] and Cui Yu[2]*

Abstract

Introduction Artificial intelligence technology has a wide range of application prospects in the field of medical education. The aim of the study was to measure the effectiveness of ChatGPT-assisted problem-based learning (PBL) teaching for urology medical interns in comparison with traditional teaching.

Methods A cohort of urology interns was randomly assigned to two groups; one underwent ChatGPT-assisted PBL teaching, while the other received traditional teaching over a period of two weeks. Performance was assessed using theoretical knowledge exams and Mini-Clinical Evaluation Exercises. Students' acceptance and satisfaction with the AI-assisted method were evaluated through a survey.

Results The scores of the two groups of students who took exams three days after the course ended were significantly higher than their scores before the course. The scores of the PBL-ChatGPT assisted group were significantly higher than those of the traditional teaching group three days after the course ended. The PBL-ChatGPT group showed statistically significant improvements in medical interviewing skills, clinical judgment and overall clinical competence compared to the traditional teaching group. The students gave highly positive feedback on the PBL-ChatGPT teaching method.

Conclusion The study suggests that ChatGPT-assisted PBL teaching method can improve the results of theoretical knowledge assessment, and play an important role in improving clinical skills. However, further research is needed to examine the validity and reliability of the information provided by different chat AI systems, and its impact on a larger sample size.

Keywords Problem-based learning, ChatGPT, Medical education

*Correspondence:
Cui Yu
cuiyeyu@126.com
[1]Nursing Department, The Third Xiangya Hospital, Central South University, Changsha, Hunan 410000, China
[2]Department of Urology, Xiangya Hospital, Central South University, Changsha, Hunan 410000, China

Hui *et al. BMC Medical Education*   (2025) 25:50

Page 2 of 7

The rapid development of artificial intelligence technology is leading to a profound change in people's daily lives and work. ChatGPT, as a type of natural language processing technology, can automatically process and generate natural language. The technology has attracted widespread attention and has been referred to as having great significance, and has been widely used in multiple fields. In the field of medical education, ChatGPT has a wide range of application prospects [1–3].

In contrast to other professional students, medical students are not only required to complete the basic knowledge of specialized fields but also need to acquire clinical practice capabilities. The problem-based learning (PBL) teaching method places emphasis on student-centered learning and problem-oriented learning, guiding students to explore and solve problems independently [4, 5]. This approach avoids the limitations of traditional lecture-based teaching, where students are not able to fully exert their exploratory and active nature. Previous research has shown that PBL teaching methods can effectively improve the understanding of theoretical knowledge and clinical practice skills of medical students [6–8].

However, the combination of PBL and ChatGPT-assisted teaching methods in medical education is still in the exploratory stage [9]. To improve the quality and effectiveness of education, to enhance students' self-study and team cooperation capabilities, we have introduced the teaching model of PBL and ChatGPT-assisted learning in the field of urology, and conducted a research on medical undergraduates as the research object, through evaluating the results of students' work, exploring its feasibility and application in medical education.

Materials and methods
Participants
We evaluated the impact of PBL ChatGPT-assisted teaching on medical education of 42 medical interns in Xiangya Medical College of Medicine's 5th-year medical curriculum. The study population consisted of 23 males and 19 females, who were randomly assigned to either the PBL ChatGPT-assisted ($n = 21$) or traditional teaching group ($n = 21$) for a two-week rotation in urology. No participants had previously completed PBL teaching mode. All students participated in a skills training program organized by the same instructor. Randomization was stratified by baseline clinical assessment scores to ensure comparability of the groups.

Study design
This is a prospective, randomized, controlled study. Students were randomly assigned to two groups. The experimental group received PBL and ChatGPT-assisted teaching methods, while the control group received traditional teaching methods. Both groups had the same

textbooks, teaching content, instructor, and total teaching time. The study design process flowchart is illustrated in Fig. 1. Ethical approval was obtained from the institutional review board, and informed consent was secured, with special attention to privacy concerns due to the involvement of patient data.

Instructional implementation
The complexity and individualized treatment needs of renal tumors require students to have a comprehensive understanding of kidney diseases and related clinical skills. In this study, we selected to teach students about the clinical manifestations, history taking, physical examination, diagnostic strategies, differential diagnosis, and treatment options of renal tumors. The curriculum included two parts: theoretical learning and practical skills training. We developed the teaching objectives based on the characteristics of renal tumors and helped students master the relevant content.

PBL ChatGPT-assisted teaching group
Study group formation
The students were divided into four groups, three groups consisting of 5 students and one group consisting of 6 students. Within each group, a communication platform was established, and the group leader was appointed.

Pre-class knowledge transfer
The students were familiarized with the features and potential educational applications of the ChatGPT 4.0 version. Prior to the commencement of the class, the instructor dedicated one session to educating students on the utilization of ChatGPT 4.0 version. Subsequently, the students' mastery of ChatGPT was assessed through a classroom-based test.

Prior to the class, the teaching objectives were sent to the students. Each group of students communicated with the ChatGPT, followed the teaching objectives, and engaged in the preliminary exploration of relevant content. The students interacted with the artificial intelligence in an interactive way, exploring the dynamic generation of clinical cases based on renal tumors, which included clinical manifestations, history taking, physical examination, diagnostic strategies, differential diagnosis, and treatment options. This provided students with the opportunity to enhance their understanding of various clinical scenarios through the use of the artificial intelligence.

Asking questions
Before the class, the students were given several open-ended questions generated by ChatGPT, which they discussed and explored independently. We have also used ChatGPT to generate simulated patients and conducted
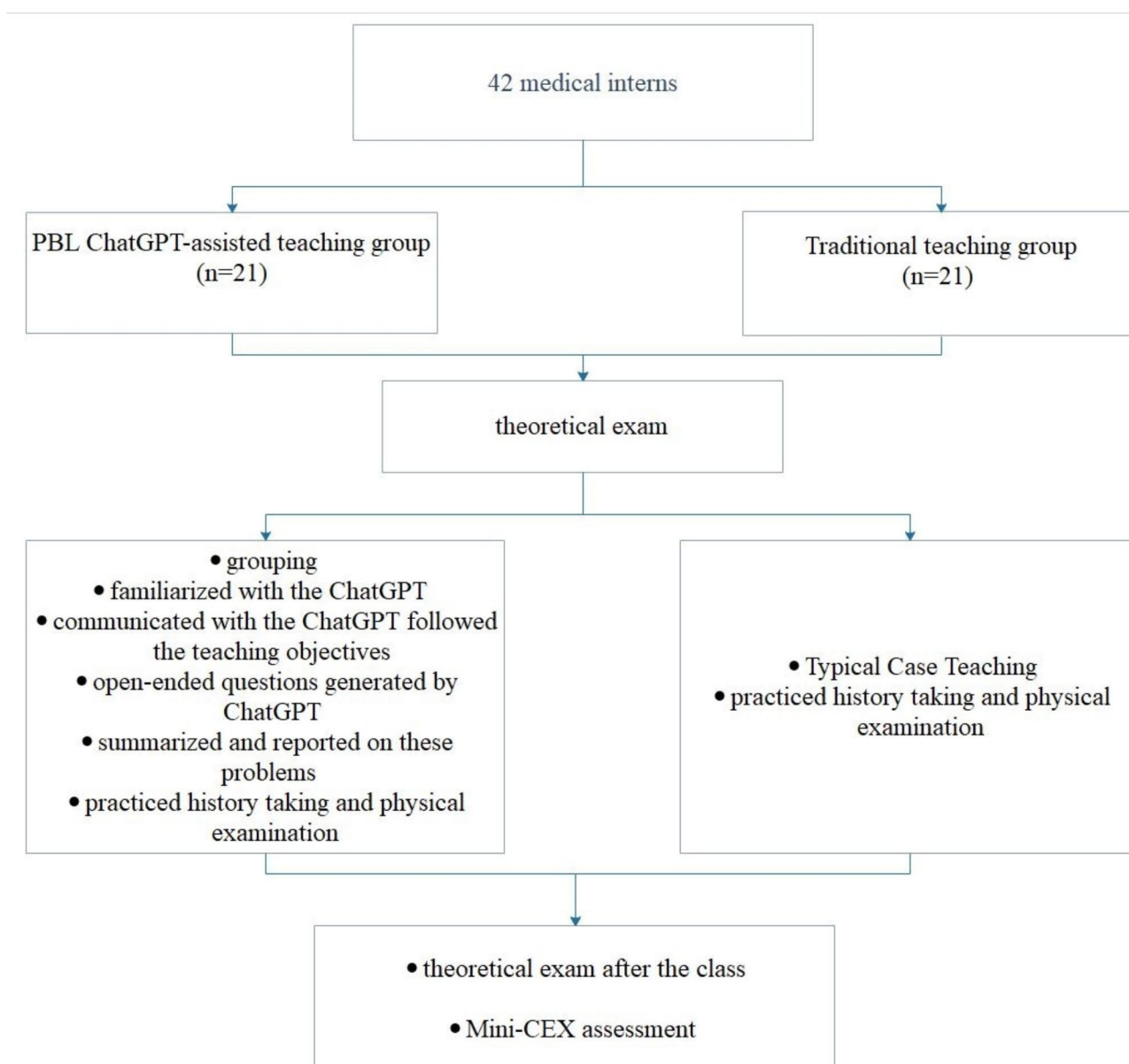
Fig. 1 Study design and flow chart

simulated interviews with the virtual patients to practice the process of history-taking. This approach enables students to practice their skills in collecting patient histories while engaging in interactive and dynamic learning experiences with ChatGPT. The purpose of this process was to enable students to receive appropriate diagnoses, select appropriate treatments, and apply the knowledge they had learned to analyze and solve clinical problems.

Teaching process

a)Classroom Discussions: Each group of students summarized and reported on problems generated by Chat-GPT before, proposing answers in the classroom. Finally, the instructor summarized and reviewed all the problems discussed.

b)Clinical Skills Training: After the discussions, students practiced history taking and physical examination in the patients' bedsides.

During the entire process, instructor critically reviewed the content provided by ChatGPT and provided instant, personalized feedback to ensure its accuracy. This direct involvement of AI and educators aims to improve learning outcomes by leveraging the scalability of AI and the expertise of educators.

Hui *et al. BMC Medical Education*   (2025) 25:50

Page 4 of 7

Traditional teaching group

Pre-class knowledge transfer

The instructor designed a typical case of renal tumors and students actively preview relevant knowledge to achieve the learning objectives by studying the case.

Teaching process

The instructor first introduced the clinical cases in detail, explaining the diagnostic reasoning process and emphasizing the importance of history taking and physical examination techniques. We developed multimedia presentations (text, audio, video, images, etc.) that detailed each renal tumor's clinical manifestations, history taking, physical examination, diagnostic strategies, differential diagnosis, and treatment options in detail. Afterward, students independently conducted patient interviews and physical assessments under the observer's guidance of the instructor. The teaching time of the traditional teaching group is the same as that of the PBL-ChatGPT group.

Theoretical knowledge exam of two groups

Before the class began, all students took a theoretical exam. Then, the third day after the class ended, they took another exam to evaluate their knowledge of the theoretical material. The maximum score for the exam was 100 points, and a score below 80 was considered a failing grade.

Mini-CEX assessment of two groups

Mini-CEX has been widely accepted as an effective and reliable method for evaluating clinical skills [10, 11]. It includes students' ability to obtain a patient's history from the patient and family and to perform physical exams under the guidance of the instructor. The Mini-CEX scoring system uses a nine-point scale and includes seven standards: medical interviewing skills, physical examination skills, humanistic qualities/professionalism, clinical judgment, counseling skills, organization efficiency and overall clinical competence. The scale ranged from unsatisfactory (1–3 points), satisfactory (4–6 points), superior (7–9 points). To maintain consistency, all Mini-CEX evaluations are performed by a single evaluator.

ChatGPT satisfaction survey questionnaire

We evaluated students' satisfaction and impact of using PBL-ChatGPT on their learning experience. The study used a survey with six questions.(Supplementary Table 1) The survey was completed by students who participated in the PBL ChatGPT-assisted teaching group.

Statistical analysis

Descriptive statistics are represented by the mean ± standard deviation (x ± s) and are compared using independent t-tests. Categorical data is represented by frequency and percentage (n,%) and is analyzed using the chi-squared test when appropriate. All statistical analyses are performed using SPSS (version 23.0, Chicago, IL, USA). A P-value < 0.05 is considered statistically significant.

Results

Theoretical knowledge exam scores

The scores of the two groups of students who took exams three days after the course ended were significantly higher than their scores before the course (PBL-ChatGPT group: $p < 0.01$;Traditional teaching group: $p < 0.01$). the scores of the PBL-ChatGPT assisted group were significantly higher than those of the traditional teaching group three days after the course ended (PBL-ChatGPT group: 93.90 ± 3.65;Traditional teaching group: 90.33 ± 4.08; $p < 0.01$) (Table 1).

Mini-CEX evaluation results

All participants completed the Mini-CEX assessment on average within 37 ± 1.3 min and the average time for immediate feedback for each student was 4.3 ± 0.8 min. The PBL-ChatGPT group showed statistically significant improvements in medical interviewing skills, clinical judgment and overall clinical competence compared to the Traditional Teaching group. The Table 2 summarizes the detailed comparison of Mini-CEX scores between the two groups of students.

Satisfaction survey

The students gave highly positive feedback on the PBL-ChatGPT teaching method. The satisfaction rate was very

Table 1 Distribution of theoretical knowledge exam scores

Exam scores	≤79	80–85	86–90	91–95	96–100	Mean scores ± SD	P
Pre class scores							
PBL ChatGPT-assisted group	4	8	6	2	1	83.21 ± 6.24	
Traditional teaching group	3	10	5	1	2	84.16 ± 4.35	
Post class scores							< 0.01*
PBL ChatGPT-assisted group	0	2	5	9	5	93.90 ± 3.65	< 0.01**
Traditional teaching group	0	4	8	6	3	90.33 ± 4.08	< 0.01***

* The *P*-value for the comparison of post-class scores between PBL ChatGPT-assisted group and traditional teaching group

** The *P*-value for the comparison between pre-class scores and post-class scores in PBL ChatGPT-assisted group

*** The *P*-value for the comparison between pre-class scores and post-class scores in traditional teaching group

Table 2 Mini-CEX scores between the two groups of students

	Mean score (± SD)		
	PBL ChatGPT-assisted group	traditional teaching group	*P*
N	21	21	
Medical interviewing skills	6.29 ± 1.10	5.33 ± 1.39	0.02
Physical examination skills	5.76 ± 1.30	5.43 ± 1.29	
Humanistic qualities/professionalism	5.09 ± 1.67	4.62 ± 1.39	
Clinical judgment	6.48 ± 1.33	5.24 ± 1.55	< 0.01
Counseling skills	4.91 ± 1.55	5.24 ± 0.83	
Organization efficiency	5.29 ± 1.49	5.57 ± 1.12	
Overall clinical competence	6.14 ± 1.15	5.14 ± 1.06	< 0.01

Table 3 Result of ChatGPT satisfaction survey questionnaire

Items	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
ChatGPT can effectively assist with answering questions in PBL courses.				16	5
ChatGPT can deepen your understanding of the knowledge points in PBL courses.			1	11	9
ChatGPT can enhance communication between groups in PBL courses.			3	14	4
ChatGPT can help you train and improve your clinical skills.			2	12	7
Your communication with ChatGPT is effective and efficient.			4	13	4
In general, ChatGPT has been beneficial for your learning process.				5	16

high, and there were no dissatisfied situations reported. The Table 3 provides a detailed overview of the feedback items and results specific to the PBL-ChatGPT teaching method.

Discussion

The teaching method of PBL encourages students to take an active role in learning. Several studies have confirmed that the appropriate application of PBL teaching methods can overcome the limitations of traditional teaching methods and improve students' clinical thinking ability [12–14]. It enables students to better understand and master theoretical knowledge, and enhance their professional skills [15, 16]. ChatGPT, as a new generation of artificial intelligence language processing models, has brought about a profound change in medical education [17, 18]. It provides extensive medical information, clinical case analysis, and medical skills simulation, helping medical students better understand and master medical knowledge. At the same time, ChatGPT offers more flexible and personalized learning ways, catering to the diversified learning needs of students [19–21]. By combining PBL teaching methods with ChatGPT, we can leverage their advantages and enhance the quality of teaching and the training of medical professionals.

Our research shows that the use of ChatGPT as a teaching aid in the PBL method can improve the results of theoretical knowledge assessment, and play an important role in improving clinical skills. The high level of satisfaction among students towards the use of Chat-GPT indicates that artificial intelligence has the ability to increase students' engagement. This positive response can be attributed to the personalized and interactive

nature of the AI experience, which meets the diverse learning styles of students. However, some students have also reported that the AI system sometimes struggles to understand the specific language context, leading to misunderstandings and inaccurate answers. Students also noticed that ChatGPT has limitations in understanding technical terms in surgery. In the process of teaching, we have observed that the probability of errors occurring in AI is approximately 15%. Some of these errors can be resolved by adding qualifying words. However, another portion of the errors arises from AI's inability to filter out accurate answers, necessitating teachers to review the answers provided by AI during the instructional process.

ChatGPT can personalize education content based on individual performance indicators, thereby promoting more precise and effective learning experiences [22–24]. The advantages of PBL-ChatGPT in theory learning and clinical skills prove that artificial intelligence technology can assist in improving students' practical operation skills and clinical thinking ability. ChatGPT can provide extensive medical information, but its answers depend on pre-trained data [25]. If ChatGPT's training data has biases and is incomplete, the algorithm may have biases, leading to information incomplete, outdated or inaccurate. Therefore, before PBL teaching begins, it is crucial for students to acquire the accurate way to communicate with ChatGPT. It is essential for students to develop skills in constructing high-quality prompts, including proficiently selecting keywords, so as to effectively convey tasks or problems to the artificial intelligence model. As students often fail to create effective prompts, this can lead to answers with biases. Therefore, during PBL teaching, teachers should closely review the accuracy of

information provided by ChatGPT. Medical education is not just a process of transmitting knowledge, but also emphasizes cultivating students' interpersonal communication and communication skills, as well as emphasizing cultivating students' emotional intelligence and empathy ability. Although ChatGPT has the ability to process natural language, it cannot provide the ability of emotional communication and understanding. When medical students rely too heavily on ChatGPT, they will lose trust in human relationships, and the relationship between people and people will be limited to "human- machine" level [26]. Over-reliance on ChatGPT can lead to the lack of interconnectedness of information, resulting in students' thinking becoming fixed, lacking innovative thinking, and further affecting the effectiveness of medical education [27]. Therefore, ChatGPT should be viewed as a supplement to teaching methods, rather than a substitute. During PBL teaching, the advantages of ChatGPT should be utilized as a teaching aid, while teachers should pay attention to providing appropriate guidance and supervision to ensure that students receive reliable relevant information, while maintaining the personalized learning experience of individualization and interaction.

This research has some limitations. Firstly, the sample size was small, which may limit the generalizability of the results. Additionally, the survey used in this study may have limitations in terms of accuracy and validity. Self-reported data may be prone to bias, and the design of the survey may not have effectively captured the experiences and feelings of the participants. The current study did not analyze the individual effects of PBL and Chat-GPT-assisted learning on the outcomes. Future research is needed to separately compare the results of these two learning methods in order to determine the impact of ChatGPT on learning effectiveness. Moreover, further research is needed to assess the effectiveness of other chat AI systems in enhancing PBL education outcomes, and to explore the unique features and limitations of each system, in order to determine their potential advantages and disadvantages in improving the PBL learning experience.

Abbreviations
PBL   Problem based learning
AI   Artificial intelligence

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12909-024-06321-1.

Supplementary Material 1

Author contributions
HJ.ZZ.ZH and CY. involved in the study conceptualization. HJ.collected and preprocessed the data. ZH. conducted data analysis, results interpretation and manuscript preparation. ZZ. and CY. contributed to the review and editing of the manuscript. CY. supervised the study. All authors read and approved the final manuscript.

Data availability
The datasets during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The Ethics Committee at Xiangya Hospital approved the study. All participants were conveyed about the study by email, written informed consent was obtained once they filled the questionnaire. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Garg N, Campbell DJ, Yang A, McCann A, Moroco AE, Estephan LE, Palmer WJ, Krein H, Heffelfinger R. Chatbots as Patient Education resources for aesthetic facial plastic surgery: evaluation of ChatGPT and Google Bard responses. Facial Plast Surg Aesthet Med. 2024 Jul 1.
2. Rotem R, Zamstein O, Rottenstreich M, O'Sullivan OE, O'reilly BA, Weintraub AY. The future of patient education: a study on AI-driven responses to urinary incontinence inquiries. Int J Gynaecol Obstet. 2024 Jun 30.
3. Hoang T, Liou L, Rosenberg AM, Zaidat B, Duey AH, Shrestha N, Ahmed W, Tang J, Kim JS, Cho SK. An analysis of ChatGPT recommendations for the diagnosis and treatment of cervical radiculopathy. J Neurosurg Spine 2024;28:1–11.
4. Azzahrani M. Problem-based learning for Interprofessional Education: a review of the Concept and its application in a geriatric team. Cureus. 2024;16(6):e63055.
5. de Andrade Gomes J, Braga LAM, Cabral BP, Lopes RM, Mota FB. Problem-based learning in Medical Education: A Global Research Landscape of the last ten years (2013–2022). Med Sci Educ. 2024;34(3):551–60.
6. Jaganathan S, Bhuminathan S, Ramesh M. Problem-based learning - an overview. J Pharm Bioallied Sci. 2024;16(Suppl 2):S1435–7. https://doi.org/10.4103/jpbs.jpbs_820_23.
7. Cheng Y, Sun S, Hu Y, Wang J, Chen W, Miao Y, Wang H. Effects of different geriatric nursing teaching methods on nursing students' knowledge and attitude: systematic review and network meta-analysis. PLoS ONE. 2024;19(5):e0300618.
8. Arruzza E, Chau M, Kilgour A. Problem-based learning in medical radiation science education: a scoping review. Radiography (Lond). 2023;29(3):564–72.
9. Divito CB, Katchikian BM, Gruenwald JE, Burgoon JM. The tools of the future are the challenges of today: the use of ChatGPT in problem-based learning medical education. Med Teach. 2024;46(3):320–2.
10. Mortaz Hejri S, Jalili M, Masoomi R, Shirazi M, Nedjat S, Norcini J. The utility of mini-clinical evaluation Exercise in undergraduate and postgraduate medical education: a BEME review: BEME guide 59. Med Teach. 2020;42(2):125–42. https://doi.org/10.1080/0142159X.2019.1652732
11. Motefakker S, Shirinabadi Farahani A, Nourian M, Nasiri M, Heydari F. The impact of the evaluations made by Mini-CEX on the clinical competency of nursing students. BMC Med Educ. 2022;22(1):634.

12. Chen D, Zhao W, Ren L, Tao K, Li M, Su B, Liu Y, Ban C, Wu Q. Digital PBL-CBL teaching method improves students' performance in learning complex implant cases in atrophic anterior maxilla. PeerJ. 2023;11:e16496.
13. Forbes HM, Syed MS, Flanagan OL. The role of Problem-based learning in Preparing Medical students to work as Community Service-oriented Primary Care Physicians: a systematic literature review. Cureus. 2023;15(9):e46074.
14. Fung JTC, Chan SL, Takemura N, Chiu HY, Huang HC, Lee JE, Preechawong S, Hyun MY, Sun M, Xia W, Xiao J, Lin CC. Virtual simulation and problem-based learning enhance perceived clinical and cultural competence of nursing students in Asia: a randomized controlled cross-over study. Nurse Educ Today. 2023;123:105721.
15. Sun M, Chu F, Gao C, Yuan F. Application of the combination of three-dimensional visualization with a problem-based learning mode of teaching to spinal surgery teaching. BMC Med Educ. 2022;22(1):840.
16. Wang A, Xiao R, Zhang C, Yuan L, Lin N, Yan L, Wang Y, Yu J, Huang Q, Gan P, Xiong C, Xu Q, Liao H. Effectiveness of a combined problem-based learning and flipped classroom teaching method in ophthalmic clinical skill training. BMC Med Educ. 2022;22(1):487.
17. Yang R, Tan TF, Lu W, Thirunavukarasu AJ, Ting DSW, Liu N. Large language models in health care: development, applications, and challenges. Health Care Sci. 2023;2(4):255–63.
18. Zhang F, Liu X, Wu W, Zhu S. Evolution of chatbots in nursing education: Narrative Review. JMIR Med Educ. 2024;10:e54987.
19. Solano C, Tarazona N, Angarita GP, Medina AA, Ruiz S, Pedroza VM, Traxer O. ChatGPT in Urology: Bridging Knowledge and Practice for Tomorrow's Healthcare, a Comprehensive Review. J Endourol. 2024 Jul 3.
20. Srinivasan M, Venugopal A, Venkatesan L, Kumar R. Navigating the Pedagogical Landscape: exploring the implications of AI and chatbots in nursing education. JMIR Nurs. 2024;7:e52105.
21. Sharma A, Medapalli T, Alexandrou M, Brilakis E, Prasad A. Exploring the role of ChatGPT in Cardiology: a systematic review of the current literature. Cureus. 2024;16(4):e58936.
22. Alhur A, Redefining Healthcare With Artificial Intelligence (AI). The contributions of ChatGPT, Gemini, and co-pilot. Cureus. 2024;16(4):e57795.
23. Wu J, Ma Y, Wang J, Xiao M. The application of ChatGPT in Medicine: a scoping review and bibliometric analysis. J Multidiscip Healthc. 2024;17:1681–92.
24. Law S, Oldfield B, Yang W. Global Obesity Collaborative. ChatGPT/GPT-4 (large language models): opportunities and challenges of perspective in bariatric healthcare professionals. Obes Rev. 2024;25(7):e13746.
25. Tan S, Xin X, Wu D. ChatGPT in medicine: prospects and challenges: a review article. Int J Surg. 2024;110(6):3701–6.
26. Stretton B, Kovoor J, Arnold M, Bacchi S. ChatGPT-Based learning: generative Artificial Intelligence in Medical Education. Med Sci Educ. 2023;34(1):215–7.
27. Choi W. Assessment of the capacity of ChatGPT as a self-learning tool in medical pharmacology: a study using MCQs. BMC Med Educ. 2023;23(1):864.

Publisher's note