The Role of Generative AI Tools in Case-Based Learning and Teaching Evaluation of Medical Biochemistry

Weiwei Zhang, Liang Li, Kejian Pan, Lan Wang, Luo Zuo, Yiran Sun, Quekun Peng

Submitted to: JMIR Medical Education
on: October 14, 2024

Table of Contents

The Role of Generative AI Tools in Case-Based Learning and Teaching Evaluation of Medical Biochemistry

Weiwei Zhang[1]; Liang Li[2]; Kejian Pan[3]; Lan Wang[3]; Luo Zuo[4]; Yiran Sun[5]; Quekun Peng[3]

[1]Academic Affairs Division, Chengdu Medical College Chengdu CN
[2]Teaching Center of General Courses, Chengdu Medical College Chengdu CN
[3]School of Biosciences and Technology, Chengdu Medical College Chengdu CN
[4]Department of Gastroenterology, the Second Affiliated Hospital of Chengdu Medical College Chengdu CN
[5]School of Pharmacy, Chengdu Medical College Chengdu CN

Corresponding Author:
Quekun Peng
School of Biosciences and Technology, Chengdu Medical College
783# Xindu Avenue, Chengdu 610500, China
Chengdu
CN

Abstract

Background: With the continuous progress and development of the medical field, medical biochemistry, as an important basic course in medical education, has an increasingly large and complex knowledge system. The traditional teaching mode is often difficult to meet students' in-depth understanding of knowledge and individualized learning needs. Students may feel confused and overwhelmed when facing a large number of biochemical concepts, chemical reactions, molecular structures and other contents.

Objective: Generative AI tools are expected to be a powerful assistant in course teaching, while such tools are still rarely applied in biochemistry courses. This study attempts to investigate the important contribution of AI tools to students' knowledge expansion and teaching evaluation in medical biochemistry CBL teaching through a comparative study.

Methods: In this study, we assessed the important role of AI tools in assisting students' CBL learning by analyzing the specific learning performances of students in the control group (N=40) and the experimental group (N=39) in CBL learning. We also compared and analyzed the data on the evaluation scores of students' work by teachers and AI tools at the same time in an attempt to identify whether AI tools can replace teachers' comprehensive evaluation. In addition, a questionnaire was used to investigate the positive and negative impacts of AI tools on case study to identify whether the AI tool can be used as an important adjunct to medical biochemistry classroom teaching.

Results: The results of the study showed that AI tools in the experimental group significantly shorten the time students spend on assignments in case study and assist students to obtain higher scores in quizzes ($P<0.05$). In addition, the questionnaire data showed no significant difference between the AI evaluation and the instructor's evaluation scores ($P>0.05$), but the group using the AI tool scored higher than the control group ($P<0.05$). And while the majority of students (M=8.5) agreed with the idea that AI tools can provide basic information retrieval and knowledge expansion, more students (M=4.2) also believed that AI tools have flaws in answering complex and logical questions. In addition, students (M=7.17) generally have concerns that the use of AI tools reduces the communication between teachers and students and the tools have limitations in offering innovative and creative knowledge (M=5.5). As a result, such tools are not a complete replacement for existing digital resources and can lead to homogenized learning (M=8.12).

Conclusions: GAI tools as “smart teaching assistants” can significantly improve students' learning efficiency in medical biochemistry CBL learning and result in better performance (assignments and exam scores).

(JMIR Preprints 14/10/2024:67526)
DOI: https://doi.org/10.2196/preprints.67526

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

The Role of Generative AI Tools in Case-Based Learning and Teaching Evaluation of Medical Biochemistry

Weiwei Zhang[a#] , Liang Li[b#], Kejian Pan[c], Lan Wang[c] Luo Zuo[d], Yiran Sun[e], Quekun Peng[c]**

[a]*Academic Affairs Division, Chengdu Medical College, Chengdu 610500, China* [b]*Teaching Center of General Courses, Chengdu Medical College, Chengdu 610500, China*

[c]*School of Biosciences and Technology, Chengdu Medical College, Chengdu 610500, China*

[d]*Department of Gastroenterology, the Second Affiliated Hospital of Chengdu Medical College, Chengdu 610057, China*

[e]*School of Pharmacy, Chengdu Medical College, Chengdu 610500, China*

#These authors contributed equally to the work.

* Corresponding author:

Quekun Peng, PhD, *Department of Biochemistry and Molecular Biology, School of Biosciences and Technology, Chengdu Medical College, 783# Xindu Avenue, Chengdu 610500, China; Email: pengquekun@cmc.edu.com.*

Yiran Sun, PhD *Department of pharmacology, School of Pharmacy, Chengdu Medical College, Chengdu 610500, China 783# Xindu Avenue, Chengdu 610500, China, Email:18202863885@163.com.*

Abstracts:

Background: With the continuous progress and development of the medical field, medical biochemistry, as an important basic course in medical education, has an increasingly large and complex knowledge system. The traditional teaching mode is often difficult to meet students' in-depth understanding of knowledge and individualized learning needs. Students may feel confused and overwhelmed when facing a large number of biochemical concepts, chemical reactions, molecular structures and other contents.

Objective: Generative AI tools are expected to be a powerful assistant in course teaching, while such

tools are still rarely applied in biochemistry courses. This study attempts to investigate the important contribution of AI tools to students' knowledge expansion and teaching evaluation in medical biochemistry CBL teaching through a comparative study.

Methods: In this study, we assessed the important role of AI tools in assisting students' CBL learning by analyzing the specific learning performances of students in the control group (N=40) and the experimental group (N=39) in CBL learning. We also compared and analyzed the data on the evaluation scores of students' work by teachers and AI tools at the same time in an attempt to identify whether AI tools can replace teachers' comprehensive evaluation. In addition, a questionnaire was used to investigate the positive and negative impacts of AI tools on case study to identify whether the AI tool can be used as an important adjunct to medical biochemistry classroom teaching.

Results: The results of the study showed that AI tools in the experimental group significantly shorten the time students spend on assignments in case study and assist students to obtain higher scores in quizzes (P<0.05). In addition, the questionnaire data showed no significant difference between the AI evaluation and the instructor's evaluation scores (P>0.05), but the group using the AI tool scored higher than the control group (P<0.05). And while the majority of students (M=8.5) agreed with the idea that AI tools can provide basic information retrieval and knowledge expansion, more students (M=4.2) also believed that AI tools have flaws in answering complex and logical questions. In addition, students (M=7.17) generally have concerns that the use of AI tools reduces the communication between teachers and students and the tools have limitations in offering innovative and creative knowledge (M=5.5). As a result, such tools are not a complete replacement for existing digital resources and can lead to homogenized learning (M=8.12).

Conclusion: GAI tools as “smart teaching assistants” can significantly improve students' learning efficiency in medical biochemistry CBL learning and result in better performance (assignments and exam scores).

Keywords: generative AI tools, case-based teaching, teaching evaluation, medical biochemistry

Introduction

Medical education in China is currently undergoing rapid reform and innovation to meet the needs of the current society. In particular, some medical schools have widely implemented the Excellence in Physician Education Program that requires medical schools to integrate the basic medical curriculum(Lihong Liu, 2017). However, the length of basic medical curriculum in China is 2.5-3 years, shorter than that in medical schools in Europe and the United States. Meanwhile, at this basic stage, knowledge acquisition and testing are the core tasks of medical education, because knowledge mastery and passing diverse progress tests are the core embodiment of the future doctor's

literacy(Cleland, 2017; Malau-Aduli et al., 2019). Therefore, how to master more solid knowledge in a shorter time is an important starting point for the reform and innovation of basic medicine education in China. In addition, the rapid development of medical knowledge and the exponential growth of knowledge is a challenge for learners, who are required to study harder to meet the knowledge needed for success. In addition, the way physicians handle knowledge resources (e.g., literature searches) greatly affects their success in their specialty including the success in patient safety, diagnosis, and quality assurance, to name a few(Lim, Saffari, & Neo, 2022). Medical education requires medical students to learn these knowledge processing techniques at an early stage in order to integrate them into their academic and clinical practice.

The use of artificial intelligence in medical education is growing at an unprecedented pace and depth, significantly contributing to the transformation of traditional medical education models(Jowsey, Stokes-Parish, Singleton, & Todorovic, 2023). Generated artificial intelligence (GAI) is a new type of production method that automatically generates content using artificial intelligence technology. Based on big models, big data, and high computing power, GAI undertakes basic mechanical labor such as information mining, material invocation, reproduction editing, etc. with an efficiency significantly better than that of the human brain, and can even generate ideas, designs, and suggestions. In recent years, GAI has been widely used in various teaching scenarios, and its rapid iterative updating and functional enhancement will inevitably bring a great impact to the education system, and will also bring new opportunities and challenges to medical education(Boscardin, Gin, Golde, & Hauer, 2024). How to combine GAI with medical education and make GAI help medical education has become a hot research topic in the field of medical education nowadays.

GAI has many advantages in medical education(Moritz, Romeike, Stosch, & Tolks, 2023). Users can have no professional computer knowledge, do not need related professional skills nor to learn the use of related software, but only need to log in the computer or cell phone to the relevant Web site, voice input or text input their own interest in the issue. GAI can output the answers and give contextualized and reasonable answers, easily realizing the deep interaction between users and computers. Secondly, the answers provided by GAI can be in various forms, such as text, pictures, PPT, audio, video, program code, excel tool sheet, analysis and correction assignments, etc. GAI provides the answers to the questions in a very short time, which greatly reduces the time cost that the user has to spend on completing the work in the past.

Clinical biochemistry, as an important discipline in the field of medicine, has an urgent need for knowledge updating. With the deepening of medical research and the rapid development of technology, new disease markers and diagnostic indicators continue to emerge(Chang, 2021).

Clinical biochemistry courses must incorporate these latest research results in a timely manner to ensure that students master the most cutting-edge diagnostic methods. For example, rapid advances in genetic testing technology have led to a deeper understanding of the genetic factors of disease, and the curriculum needs to be updated with relevant content to enable students to understand how to apply genetic testing for early screening and diagnosis of disease. At the same time, drug development and therapeutic innovations are changing clinical practice (Berillo, Yeskendir, Zharkinbekov, Raziyeva, & Saparov, 2021). Courses need to reflect the mechanisms of action of new drugs and their metabolism in the body, so that students can provide patients with more precise drug treatment options(Dube, 2018). For example, the widespread use of immunotherapeutic drugs in tumor therapy requires a curriculum that details the biochemical basis of their action. In addition, with the application of big data and artificial intelligence in medicine, the clinical biochemistry curriculum should also cover related knowledge and develop students' ability to utilize these technologies for data analysis and disease prediction(Gupta et al., 2021). In conclusion, the clinical biochemistry curriculum must keep pace with the times, deeply integrate AI technologies, and continuously expand the scope of knowledge and update the field of knowledge, in order to cultivate professionals who can adapt to the development of modern medicine and provide high-quality healthcare services to patients.

Therefore, this research implemented a comparative study to analyze and compare the learning effects of students in two clinical medicine classes after using AI teaching aids for case-based learning and evaluated them. The impact of AI tools on students' learning behaviors was assessed through surveys of students' satisfaction with the ease of use of AI tools and the ability of such tools to expand their knowledge and aid their case studies. In addition, this study also compared and analyzed the AI tools and teachers' judgments on the quality of assignments and students' performance of grades after using the AI tools, and analyzed the evaluation of AI tools on students' case-based learning effects from multiple perspectives.

Methods

1.1 Participant recruitment and study background

A total of 89 students from two clinical medicine classes in the class of 2023 were randomly selected for the study. Class 1 had 45 students and Class 2 had 44 students. All students were randomly assigned at enrollment, and the students' college entrance scores were within the same range. In order to avoid influencing the results of the study due to students' personal preferences to AI, we set up a simple question “Please evaluate the use of AI tools in the medical biochemistry course” for initial screening of the sample. Students responded according to a scale of 0 to 10 points,

with higher ratings indicating greater agreement with the use of AI tools and lower ratings indicating less agreement. Samples of students with ratings lower than 2 or higher than 8 were excluded to minimize the interference of extreme attitudes with the results of the study. A total of 79 participants in the experiment were obtained by screening through questionnaire analysis (Fig 1). The control group (Class 1) had 40 students, 20 males and 20 females, with an age range of 18-22 years old and a mean age of 19±0.9. The experimental group (Class 2) had 39 students, 20 males and 19 females, with a mean age of 18±0.7. There were no significant differences in other characteristics such as age, gender, or English level between the two classes. The study was agreed and supported by the Education and Teaching Research Committee of Chengdu Medical College, and there were no teaching ethical issues.

1.2 Selection of GAI teaching tools

Before this study was conducted, six widely used software in China were recommended as candidate research tools for students, which were Kimi Chat (https://kimi.moonshot.cn/), Dou bao (https://www.doubao.com/chat/), Wen xin yi yan (https://yiyan.baidu.com/), Tong yi qian wen (https://tongyi.aliyun.com/), Chat GPT4.0 (https://chatgpt.com/), Zhi pu qing yan (https://chatglm.cn/). We distributed questionnaires in the experimental group to evaluate several AI software. Each student was asked to rate each software (on a scale of 0-10 points), and then the average score of each software and the proportion of selected students were counted. Finally, we identified the AI tools used in this study based on the software with the highest percentage of selections.

1.3 CBL learning

The clinical case was collected from the faculty of the First People's Hospital affiliated to our university and organized and edited by the biochemistry teaching team. The clinical case focuses on the physical examination, consultation, treatment and recovery process of a liver cancer patient (Fig S1). The questions discussed cover two segments: biochemistry and cell biology. The questions in the biochemistry section mainly cover five chapters: glucose metabolism, lipid metabolism, protein metabolism, blood biochemistry and liver biochemistry. The cell biology section mainly covers liver cancer indicators, oncogenes and oncogenes, tumor characteristics and mechanisms, and treatment options for liver cancer. Twelve scientific questions (Table 1) were set for the case. Both the control and experimental groups were divided into five groups of eight students each. The instructor distributed the case and questions online, and the students in both groups studied the case in terms of basic concepts, extended knowledge, and treatment strategy development, and prepared a report-back assignment. The offline discussion session consisted of “question answering, group presentation, and inter-group error correction”. The control group used traditional methods to complete the

assignments, while the experimental group used AI tools (Fig.1). Except for not using the AI tool, the other processes of the control group are exactly the same as those of the experimental group.

1.4 Teaching Evaluation

The evaluation in this study is mainly divided into homework evaluation and final exam evaluation. The homework evaluation method is further divided into AI evaluation and teacher's comprehensive evaluation. For the questions in Table 1, we require each group to complete four assignments, i.e., to answer the corresponding questions in the four classifications of C1, C2, C3 and C4, and make the answers to the questions into the corresponding PPT. The control group prepares the corresponding answers through traditional methods, such as academic websites, textbooks, literature, etc. The experimental group must use AI to prepare answers to the questions in the traditional methods. The experimental group was required to complete the PPT task with the help of AI tools. Subsequently, the PPT assignments of the control group and the experimental group were scored by the teacher and the AI tools according to the assignment rubric (Table S1). The evaluation scores of the two groups were then analyzed. The final exam evaluation mainly focuses on the comparative analysis of the final exam clinical case study questions and the total score of the test paper of the two groups of students, comparing the differences between the two.

1.5 Questionnaire survey and data analysis

We set up a questionnaire with two dimensions of positive and negative impacts, and the positive feedback questionnaire included 12 questions (Q1-Q12) and was divided into four levels: 1) the difficulty of using AI tools (Q1-Q3); 2) the impact of AI tools on the utilization of teaching resources (Q4-Q6); 3) the impact of AI tools on the mastery of knowledge (Q7-Q9); 4) the effect of AI tools on the enhancement of the clinical case analysis ability (Q10-12). A score of 0-10 was set for each question. The scores reflect the degree of student agreement with each point, ranging from 0 to 10. Higher scores indicate stronger agreement. Specifically, a score of less than 5 indicates a low level of agreement with the viewpoint; a score of 5-8 indicates a moderate level of agreement; and a score of more than 8 indicates a high level of agreement with the viewpoint.

The negative feedback questionnaire includes six questions, which are mainly divided into three levels: 1) the impact of AI tools on teacher-student interaction (Q1-2); 2) the impact of AI tools on active learning (Q3-4): 3) the impact of AI tools on competitive learning (Q5-6). The questionnaires were released through the Study Pass APP and collected online. In addition, we also researched online the time students spent completing case assignments and the number of times they sought help from teachers online.

All data were automatically downloaded by the Study Pass software and analyzed using SPSS software (version 22.0). Continuous variables were expressed as (' $x \pm s$) and two-sample t-tests were

used for comparisons between groups. Categorical data were expressed as percentages and tested with Pearson's chi-square test. $p< 0.05$ was considered statistically significant difference.

Results

1.1 Study subjects and GAI tool evaluation

In this study, six generative AI tools that are commonly used or have high visibility in China were recommended. The appropriate software was selected as the experimental tool through the students' pre-instruction independent evaluation. After the research Kimi Chat 2.0 tool received the highest rating of 7.6, which was chosen by 40% of the students (Fig 2). Wen xin yi yan and Dou bao received average scores of 6.2 and 5.8 respectively, while Tong yi qian wen and Zhi pu qing yan received scores of 4.3 and 3.6 respectively. While Chat GPT4.0 obtained a pairwise low score of 2.2. The reason for the high score of Kimi Chat may be that Kimi Chat is the AI tool with the highest recommendation value on the web at present, and it has an advantage in the field of scientific research. Chat GPT4.0 has the lowest score, mainly because the open registration of the tool is not supported in China for the time being, and it is more difficult to use in China. Therefore, we finally chose Kimi Chat 2.0 software as the tool for this study.

1.2 Consistency between the AI tool and the teachers' comprehensive evaluation on the assignment evaluation results

By analyzing the homework scores of the two groups (Fig 3A-D), we found that in the control group, there was no significant difference in homework scores between the comprehensive evaluation and the AI tool evaluation ($P>0.05$). The same results existed for the experimental group. However, the task completion scores of the experimental group that used the GAI tool were significantly higher than those of the control group ($P<0.05$). This result indicates that the quality of the assignments with the AI-assisted tool was significantly higher than without the AI tool. However, based on the same grading rules, there is no statistically significant difference between the evaluation of the assignments by AI tools and the manual comprehensive evaluation, indicating that AI tools have the potential to replace the teacher's evaluation of the assignments, which will also greatly reduce working load of teachers.

1.3 The AI tool provides more detailed answers to case studies

We compared and analyzed the answers given by the control group and the experimental group to the two main clinical questions in the case, as shown in Table 2, and found that the experimental group gave more comprehensive and detailed answers. For example, for question 1(What are the pathological and molecular mechanisms of hepatocellular carcinoma?), the AI tool cannot only give conventional answers, such as viral infection, cirrhosis, genetic factors, signaling pathway

abnormalities, inflammatory microenvironment, etc., but also give more cutting-edge answers. For example, tumor microenvironmental alterations, angiogenesis, epigenetic alterations, metabolic abnormalities and genomic instability. And in question 2 (What are the clinical and biological therapeutic options for the treatment of hepatocellular carcinoma?) the AI tool also gave more specialized answers than the control group students. For example, in addition to the answers of surgical resection, liver transplantation, radiotherapy and chemotherapy, tumor vaccine, and immunosuppressant given by the control group, the AI group also gave more novel treatments, such as immune-combination therapy, neoadjuvant therapy, stem cell therapy, and multidisciplinary integrated therapy and postoperative follow-up. It can be seen that the AI tool cannot only provide more comprehensive answers, but also provide students with integrated basic and clinical knowledge by analyzing and integrating information through multiple data centers. This is conducive to cultivating students' basic and clinical integrated thinking.

1.4 AI Tool reduces case study time and improve student test scores

AI software can quickly filter out content relevant to students' needs from a large number of learning resources, helping students acquire the knowledge they need more quickly and reducing the time they spend sifting through massive amounts of information. When students encounter problems, AI can quickly give accurate answers and detailed explanations, saving students the time to find information and think on their own(Alkhalaf, Yu, Yin, & Deng, 2024). We investigated the students' time in completing the case discussion assignments through a questionnaire, and found that the average completion time of students in the experimental group was greatly reduced after the use of the AI tool, while the time used by the control group (t=5.5h) was significantly more than that of the experimental group (t=2.6h) (Fig 4A). In addition, we also counted the number of times the control and experimental groups sought help from the instructor in completing their assignments, and found that the number of questions asked by the students in the experimental group (N=10) decreased significantly compared to that of the control group (N=35) after the use of the AI tool (Fig 4B). All these results show that using AI tools speeds up students' completion of assignments and saves more learning time. In addition, the AI teaching tool acted as a teaching assistant to help teachers share part of the task of answering questions, which reduced the teaching burden of teachers. Meanwhile, we also analyzed the examination scores of the control group and the experimental group in the final examination (Fig 4C). The results showed that the experimental group's case study question scores (8.7±2.1) and final exam scores (77.3±4.3) were significantly higher than those of the control group ((6.4±3.5; 66.5±5.4) (P<0.05). This indicates that AI tools can significantly enhance students' understanding of clinical cases and improve their learning performance.

1.5 AI tool has advantages in knowledge expansion of clinical cases

In this study, we investigated and analyzed the attitudes of students in the experimental group towards the role of AI tools in case study in biochemistry courses from these two levels (positive feedback and negative feedback) respectively (Fig 5). The positive feedback questionnaire contained 12 questions, and all 39 questionnaires were collected through an online questionnaire, and the mean score (Mean ± SD) of the scores for each question was analyzed. By analyzing the questionnaire data, it was found that in the first dimension (Q1-Q3), Q1M=9.50; Q2M=7.14; Q3M=7.49, indicating that the vast majority of students believe that AI tools are easier to acquisition and use. It can also greatly save learning time and reduce the time spent on reviewing and translating literature, and it is easier to set the search content, which can easily realize human-machine dialogue. In the second dimension (Q4-Q6), we found that students highly agreed that AI tools play a complementary role to the textbook (Q4M=9.18). Students also tend to agree that AI tools can provide them with a diversity of learning materials (Q5M=6.47), whereas students did not agree that AI tools replace existing e-learning resources (Q6M=4.12). In the third dimension (Q7-Q9), the vast majority of students (Q7M=8.48) were positive about the view that AI tools provide a more comprehensive explanation of basic knowledge. Most students also agreed that AI tools can expand biochemistry knowledge beyond the textbook (Q8M=6.18). However, students' agreement with the role of AI tools in the explanation of cutting-edge knowledge points was low (Q9M=4.50). This also suggests that AI tools may have shortcomings in the tracking of new knowledge. In the fourth dimension (Q10-Q12), the view that AI tools are more comprehensive in analyzing cases was affirmed by most students (Q10M=5.8), while more students believed that AI tools have limited role in deep understanding of clinical biochemistry in clinical cases (Q11M=4.20). Most of the students did not think that the AI tool could generate similar clinical cases (Q12M=3.90). This indicates that this AI tool has some limitations in generating complex concepts and cutting-edge knowledge.

1.6 Limitations of the AI tool in medical biochemistry case studies

Although AI tools have great potential in classroom teaching, we also need to be fully aware of their limitations and consider them carefully in their application to ensure that they are used sensibly, safely, and effectively. In this study, we also set six questions to investigate students' attitudes toward the limitations of AI in biochemistry classroom teaching (Fig 6). Most of the students believed that the use of AI tools reduced the interaction between students and teachers inside and outside the classroom (Q1M=7.17), which was evidenced by the decreased frequency of students’ questions to the teacher. However, students were neutral in their views on whether AI tools reduced classroom attention (Q2M=5.24). In addition, most students disagreed with the view that the use of AI tools reduced active thinking (Q3M=4.31), and students were also neutral on the view that the use of AI tools reduced the training of creative thinking (Q4M=5.5). All of these data suggest that AI tools

reduce students' ability to learn and think actively to some extent, and are likely to weaken students' learning ability. Secondly, we also investigated the effect of AI tools on learning pressure, and most students agreed that using AI tools rather increases the competitive pressure of learning (Q5M=8.10) and leads to homogenization of learning (Q6M=6.56). This can be explained by the fact that the homogeneous answers provided by AI tools make the variability of scores among students not obvious. This is corroborated by the data from Fig 2 of the experimental group's homework scores.

Discussion

Currently, the use of GAI in medical courses is increasing, and the willingness of students to use AI is getting stronger. Li et al. analyzed data from 1,243 undergraduate and graduate students from 13 universities and 33 hospitals, and 54.3% of the students demonstrated prior experience with medical AI. The results indicated that graduate medical students had a higher level of awareness of using medical AI than undergraduate students. Intention to use medical AI was positively correlated with factors such as performance expectations, habitus, hedonic motivation, and trust. In addition, some researchers have also found GAI to have some credibility in basic medical courses, showing its strong problem-solving ability(Li & Qin, 2023). Ghosh et al. used Chat GPT to analyze 200 medical biochemistry reasoning questions, and by analyzing the ability of ChatGPT to correctly answer queries in medical biochemistry that require higher-level thinking, researchers found that ChatGPT has the potential to be a useful tool for answering successful tool for answering questions in medical biochemistry that require higher-order thinking(Ghosh & Bir, 2023). The GPT-4 correctly answered 265 of 360 questions (73.6%) on the orthopaedic examination for training (OITE), which is equivalent to the national average PGY-5 mean score for postgraduate year 5 (PGY5) level orthopaedic residents and exceeds the corresponding pass score of 67% on the American Board of Orthopaedic Surgery Part I examination. ChatGPT performs at a high PGY1 level and GPT-4 performance better than the average PGY-5 level, suggesting that AI has the potential to aid in medical education and the delivery of healthcare (Kung, Marshall, Gauthier, Gonzalez, & Jackson, 2023). In addition, by analyzing ChatGPT's performance on the United States Medical Licensing Examination (USMLE) questions, the researchers found that ChatGPT achieved a critical value of more than 60% on the National Board of Medical Examiners-Free-Step-1 dataset, suggesting that the model reached the equivalent of a third-year medical student passing scores of third-year medical students (Gilson et al., 2023). The AI tools demonstrated in these studies show a strong ability to enhance teaching and learning as a potential “enabler” for medical education. The questionnaires in this study showed that most students were satisfied with both the interface and the personalized teaching resources provided by the AI tool.

With the rapid progress of biotechnology and the deepening of medical research, new knowledge, theories, and technologies are constantly emerging in the field of biochemistry. In particular, there are significant breakthroughs every year in the fields of gene editing, histology research, and biological drug development(Bauer et al., 2024). This requires students and professionals in clinical medicine to maintain their enthusiasm and sensitivity for learning, and to understand and apply the latest research results to clinical practice in a timely manner(Kocemba-Pilarczyk, Bentke-Imiolek, & Dudzik, 2023). GAI technology can quickly collect and organize the latest medical research results, clinical guidelines and industry trends, and ensure that the content of medical education is current and cutting-edge. At present, ChatGPT is widely used in medical education. However, there are few researches of its application in medical course teaching. He et al. selected ChatGPT, an advanced natural language model, as a tool and applied ChatGPT to the design of biochemistry and molecular biology experimental teaching program, and the results showed that GAI can assist teachers in the material update and assignment evaluation completed by teachers, which improves the efficiency of the teachers' lesson preparation(Xiaojin Zhang, 2024). Surapaneni et al. utilized 10 clinical cases to evaluate the performance of ChatGPT in medical biochemistry. In the first test ChatGPT generated correct answers for 4 questions, and in the second attempt ChatGPT provided correct answers for 6 questions and incorrect answers for 4 questions. But surprisingly, Case 3 got different answers for several attempts. They attribute this situation to the complexity of the cases(Surapaneni, 2023). In addition, Ghosh et al. conducted an investigation into the ability of ChatGPT to solve higher-order problems related to medical biochemistry. 200 medical biochemistry reasoning questions that required higher-order thinking to solve were randomly selected from the question pool. Finally, the AI software answered all of the questions, but the median score of 4.0 was less than the hypothesized maximum of 5 and significantly different(Ghosh & Bir, 2023). Although ChatGPT scores were not as high as expected, this is still an effective attempt to use ChatGPT in teaching clinical biochemistry.

Due to the limitations of ChatGPT's use in China, we took a comparative approach in this study to explore the role of Kimi Chat, a domestic Chinese AI tool, in biochemistry clinical case study and evaluate the impact of this tool on teaching and learning in the medical classroom. The results of the study show that the use of AI tools can provide more cutting-edge and comprehensive answers to case studies. For example, the AI tool provides students with novel biologic treatments for hepatocellular carcinoma, such as immune-combination therapy, targeted therapies, stem cell therapy, and other techniques. These detailed answers greatly expanded the scope of students' knowledge, constructed the connection between biochemistry and clinical diseases, and laid the foundation for cultivating students' systematic clinical thinking (Table 2). In addition, students in the experimental group can greatly save their study time and reduce their dependence on the instructor

with the help of AI tools. This can be clearly reflected in the students' completion time of homework and the frequency of questions. This shows that AI tools, with their easy operation and practicality, can provide students with personalized learning materials or answers regardless of time and space constraints that reduces students' access to traditional resource retrieval. Through the time saved, students can engage in broader and deeper learning. Secondly, AI tools act as teaching assistants, which can answer students' basic questions anytime and anywhere, which also reduces the transitional dependence on teachers, as well as the teaching pressure on teachers, and saves more time for teachers to engage in more complex teaching tasks.

The application of GAI technology significantly improves the objectivity and efficiency of teaching evaluation. Through automated data analysis and processing, the technology can generate detailed evaluation reports in a short period of time, which include not only students' test scores, but also their performance in class discussions, team projects, and clinical skills. This comprehensive, evidence-based evaluation method can help teachers more accurately identify students' learning strengths and weaknesses. At the same time, the evaluation results of AI technology are highly consistent and comparable, reducing human bias and increasing the credibility of the evaluation. Through these advanced technological means, the evaluation and teaching process of medical education has become more efficient, precise and fair(Lee, 2024). In this study, we set uniform standards for four assignments (PPTs) in CBL for two groups of students with dual evaluation by AI tools and teachers. The results show that there is no significant difference between the manual composite scores and AI scores for all four groups of homework grades, and the AI can accurately give detailed textual evaluations. It shows that the evaluation of AI tools in CBL is completely objective and fair, and it also suggests that we can use AI tools for homework correction and evaluation, which will greatly reduce the teaching burden of teachers. In addition, we also found that the experimental group's learning performance with the aid of the AI tool was significantly better than that of the control group, as reflected by the experimental group's scores on the case study questions and the overall performance. This also indicates that the AI tool can provide students with more detailed test answers.

The incorporation of generative AI technology brings innovation to medical education evaluation, but also poses challenges for teachers and students to adapt to the new technology. For teachers, they need to master the operation of more and more complete AI tools and effectively integrate them into teaching practice to improve teaching quality. Currently, AI tools are springing up, and how to choose the right tool from a large number of AI tools is also one of the challenges faced by teachers. This study provides over a variety of more widely used AI software on the market, adopts the form of questionnaire research according to the students' preference, so that the students can independently

choose the AI tools and reduces the problem of irrational data due to the bias of learning tools. Through preliminary comparison, it was found that the software Kimi Chat has more advantages in terms of the comprehensiveness of specialized knowledge points and the comparison of large texts compared with other tools. This may also be the main reason for the students' preference to Kimi Chat.

For students, generative AI technology provides a wealth of learning resources and aids, but over-reliance may lead them to neglect the interpersonal emotional communication between teachers and students, and reduce the desire for active thinking and the importance of in-depth learning and understanding of medical knowledge(Pagliari, Chambon, & Berberian, 2022). For example, data from Fig5-Q1 showed that students perceived less interaction with faculty and students in the classroom after using AI tools. We also found in our classroom observations that some students would focus on the use of AI tools and spend less time listening to lectures. Secondly, students also tend to think that the use of AI tools reduces creative thinking; the emergence of AI tools makes students think of using tools whenever they encounter difficult problems, instead of actively recalling knowledge and seeking for answers by themselves. Over time this may lead to dependence on the tools, harming the logical thinking and creative thinking of students. Furthermore, the analysis of the questionnaire data shows that although most students believe that AI tools have a more comprehensive display of basic and expanded knowledge of biochemistry, the students' scores on the questionnaire show that they believe that AI tools are not comprehensive in explaining complex concepts. In particular, more students felt that AI tools were unable to analogize and analyze similar cases. This also indicates that AI tools still have shortcomings in solving complex clinical medicine problems. From the questionnaire results, we also found that students prefer to think that learning through AI will cause learning pressure because AI has limitations in aiding students in innovative learning where it is easy for students to get similar answers through the aid of AI tools, thus not able to show the difference of learning effect. This is another major challenge currently faced by AI tools. Therefore, the issues of how to ensure the novelty and value of knowledge creation by AI tools, how to avoid over-imitation and truly realize innovation, and how to evaluate and understand the results of AI creation are still worthy of in-depth exploration by AI tool developers.

In addition, after the experiment, we investigated again the experimental group students' views on whether they agreed with the use of AI, and found that the experimental group students' ratings of their views on agreeing with the use of AI increased from 5.5 before the experiment to 7.6 (Fig 6), which suggests that more students benefited from the use of AI and had a willingness to use AI tools in their future studies. It shows that although the AI tool is still there with some disadvantages, its advantages in classroom teaching are still obvious. the use of AI tools changed the students' view of

traditional classroom lectures, and benefit the students. As the development of smart teaching tools improves, we believe that more and more students will try to use AI teaching aids in their classroom teaching.

Conclusion

This study analyzed the impact of AI tools on medical biochemistry course learning by comparing students' learning outcomes and performance before and after the use of AI tools. The results found that AI tools had a positive impact on the learning behaviors of students in the experimental group in many ways. First, AI tools reduce the time students spend on consulting materials and integrating resources, saving their learning time and allowing them to spend their time on more important learning tasks. At the same time, AI tools can also reduce the frequency of students asking teachers, help teachers evaluate students' work objectively and fairly, and reduce teachers' teaching burden. Secondly, AI tools can collect and integrate the latest medical research results, clinical guidelines, and industry news in a timely manner, ensuring the timeliness and cutting-edge of medical course content and expanding students' knowledge boundaries. However, in our research, we also found that there are certain drawbacks to the application of AI software in classroom teaching. For example, AI wisdom tools are still defective in terms of knowledge innovation and creativity; they cannot replace the logical thinking of the human brain, and are biased in their understanding of complex scientific issues. This is also a place where AI tools are in dire need of enhancement. In conclusion, this study shows that AI tools have a positive impact on promoting classroom teaching and learning, and are worthy of being widely promoted.

Acknowledgments

We would like to thank Ni, S from Jincheng College of Chengdu and Peng, T from Sichuan University for their help in revising the manuscript and the other researchers for their support in this study.

Authors’ contributions

Zhang, W and Pan, K conceptualized and designed the study, collected the data, and analyzed the data. Li, L revised the manuscript. Zuo, L provided the case from the affiliated hospital. Wang, L made and collected questionnaires. Sun, Y and Peng, Q prepared the main manuscript text. The authors have met the criteria for authorship and had a role in preparing the manuscript. Also, all authors approved the final manuscript.

Data availability statement
Data will be supplied upon request.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National First-class Undergraduate Course Construction Project (No. 2023232793); 2023 Teaching Reform Research Foundation of Chengdu Medical College (Key Project) (No. JG2022005); Higher Education Talent Training Quality and Teaching Reform Project of Sichuan Province in 2022 (No. JG2021-1210).

ORICD:

Yiran Sun: https://orcid.org/0000-0002-5574-6518

Quekun Peng: https://orcid.org/0000-0002-4065-1137

Figure legends:

Fig 1 Teaching design of biochemistry and molecular biology course based on GAI tools.

Fig 2 Selection of GAI tools.

Fig 3 A comparison of teacher and GAI tools in grading student assignments. * P<0.05; ** P<0.01; *** P<0.001; ns, no significance.

Fig 4 The influence of Kimi Chat on students' learning behavior. * P<0.05; ** P<0.01; *** P<0.001; ns, no significance.

Fig 5 The positive impact of Kimi Chat on students' learning outcomes.

Fig 6 The negative effects of Kimi Chat on students' learning.

Fig 7 Comparison of students' ratings of AI tools before and after the experiment.

Fig S1 Clinical case data

Reference

Alkhalaf, M., Yu, P., Yin, M., & Deng, C. (2024). Applying generative AI with retrieval augmented generation to summarize and extract key clinical information from electronic health records. *J Biomed Inform, 156*, 104662. doi:10.1016/j.jbi.2024.104662

Bauer, M. E., Akbar, S., Bauler, T. J., Chacon, J., McClelland, E. E., Staudaher, S., & Zhao, Y. (2024). Exploration of the integration of microbiology and immunology emerging topics into undergraduate medical education. *Med Educ Online, 29*(1), 2336331. doi:10.1080/10872981.2024.2336331

Berillo, D., Yeskendir, A., Zharkinbekov, Z., Raziyeva, K., & Saparov, A. (2021). Peptide-Based Drug Delivery Systems. *Medicina (Kaunas), 57*(11). doi:10.3390/medicina57111209

Boscardin, C. K., Gin, B., Golde, P. B., & Hauer, K. E. (2024). ChatGPT and Generative Artificial Intelligence for

Medical Education: Potential Impact and Opportunity. *Acad Med, 99*(1), 22-27. doi:10.1097/acm.0000000000005439

Chang, Z. (2021). Grasp the trends of the discipline and raise Biochemistry education to a higher level. *41*(7), 1357-1361.

Cleland, J. A. (2017). The qualitative orientation in medical education research. *Korean J Med Educ, 29*(2), 61-71. doi:10.3946/kjme.2017.53

Dube, D. H. (2018). Design of a drug discovery course for non-science majors. *Biochem Mol Biol Educ, 46*(4), 327-335. doi:10.1002/bmb.21121

Ghosh, A., & Bir, A. (2023). Evaluating ChatGPT's Ability to Solve Higher-Order Questions on the Competency-Based Medical Education Curriculum in Medical Biochemistry. *Cureus, 15*(4), e37023. doi:10.7759/cureus.37023

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Med Educ, 9*, e45312. doi:10.2196/45312

Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers, 25*(3), 1315-1360. doi:10.1007/s11030-021-10217-3

Jowsey, T., Stokes-Parish, J., Singleton, R., & Todorovic, M. (2023). Medical education empowered by generative artificial intelligence large language models. *Trends Mol Med, 29*(12), 971-973. doi:10.1016/j.molmed.2023.08.012

Kocemba-Pilarczyk, K. A., Bentke-Imiolek, A., & Dudzik, P. (2023). Molecular biology techniques in real life - case study introducing medical students to QF-PCR technique used in medical diagnostics. *Folia Med Cracov, 63*(2), 135-148. doi:10.24425/fmc.2023.145918

Kung, J. E., Marshall, C., Gauthier, C., Gonzalez, T. A., & Jackson, J. B., 3rd. (2023). Evaluating ChatGPT Performance on the Orthopaedic In-Training Examination. *JB JS Open Access, 8*(3). doi:10.2106/jbjs.Oa.23.00056

Lee, H. (2024). The rise of ChatGPT: Exploring its potential in medical education. *Anat Sci Educ, 17*(5), 926-931. doi:10.1002/ase.2270

Li, Q., & Qin, Y. (2023). AI in medical education: medical student perception, curriculum recommendations and design suggestions. *BMC Med Educ, 23*(1), 852. doi:10.1186/s12909-023-04700-8

Lihong Liu, H. B., Xiaowei Wang, Yuan Zou. (2017). A study of innovative curriculum application to clinical medicine in the excellent doctors educational training program. *Higher medical education in China, 2017*(11), 73-74.

Lim, I., Saffari, S. E., & Neo, S. (2022). A cross-sectional study of knowledge and practices in the management of patients with Parkinson's disease amongst public practice-based general practitioners and geriatricians. *BMC Health Serv Res, 22*(1), 91. doi:10.1186/s12913-022-07503-7

Malau-Aduli, B. S., Alele, F. O., Heggarty, P., Teague, P. A., Sen Gupta, T., & Hays, R. (2019). Perceived clinical relevance and retention of basic sciences across the medical education continuum. *Adv Physiol Educ, 43*(3), 293-299. doi:10.1152/advan.00012.2019

Moritz, S., Romeike, B., Stosch, C., & Tolks, D. (2023). Generative AI (gAI) in medical education: Chat-GPT and co. *GMS J Med Educ, 40*(4), Doc54. doi:10.3205/zma001636

Pagliari, M., Chambon, V., & Berberian, B. (2022). What is new with Artificial Intelligence? Human-agent interactions through the lens of social agency. *Front Psychol, 13*, 954444. doi:10.3389/fpsyg.2022.954444

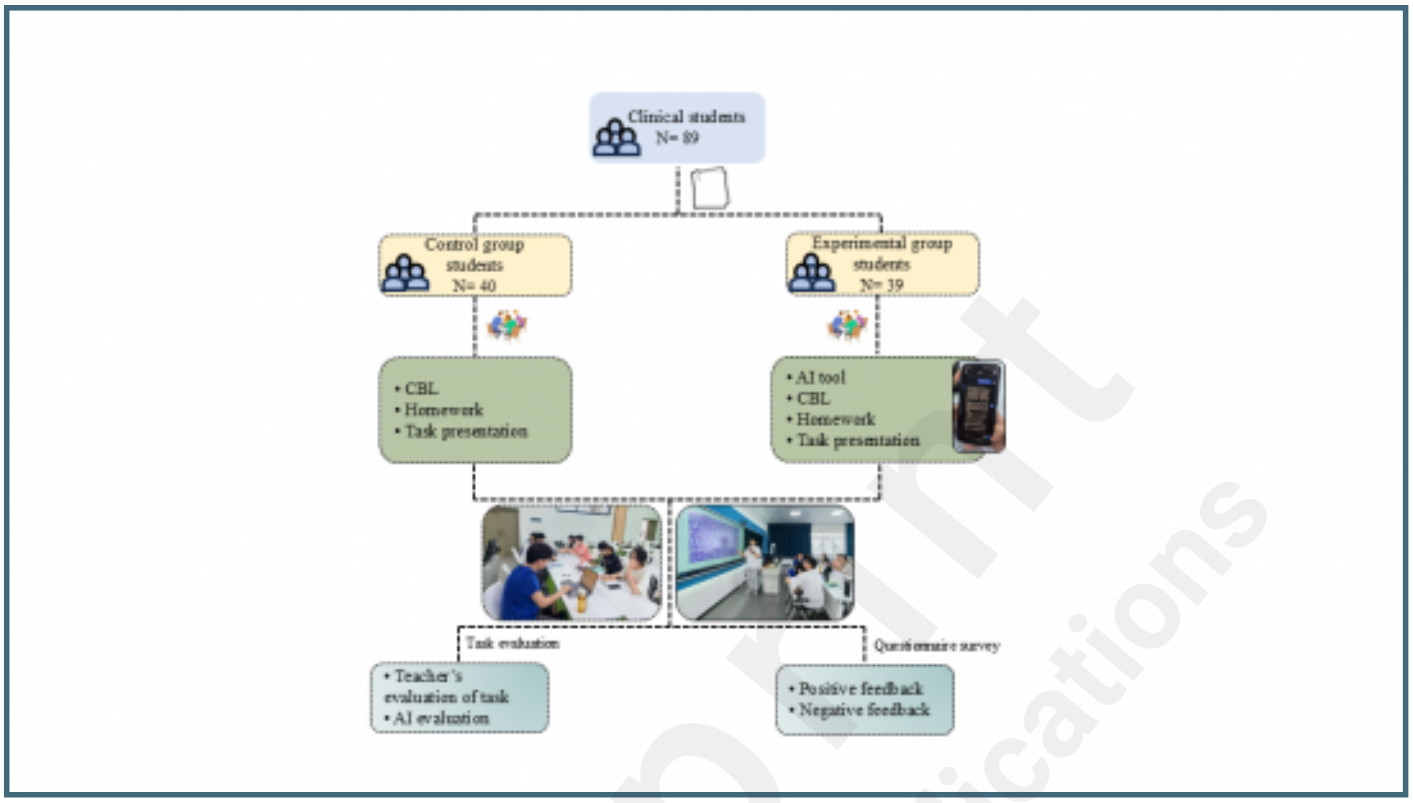
Surapaneni, K. M. (2023). Assessing the Performance of ChatGPT in Medical Biochemistry Using Clinical Case Vignettes: Observational Study. *JMIR Med Educ, 9*, e47191. doi:10.2196/47191

Xiaojin Zhang, Y. G., Sisi Cao, Junfei Tao, Xuehan He. (2024). application of artificial intelligence genereated content in experimental teaching of biochemistry and molecular biology. *Basic Medical Education, 26*(5), 406-411.

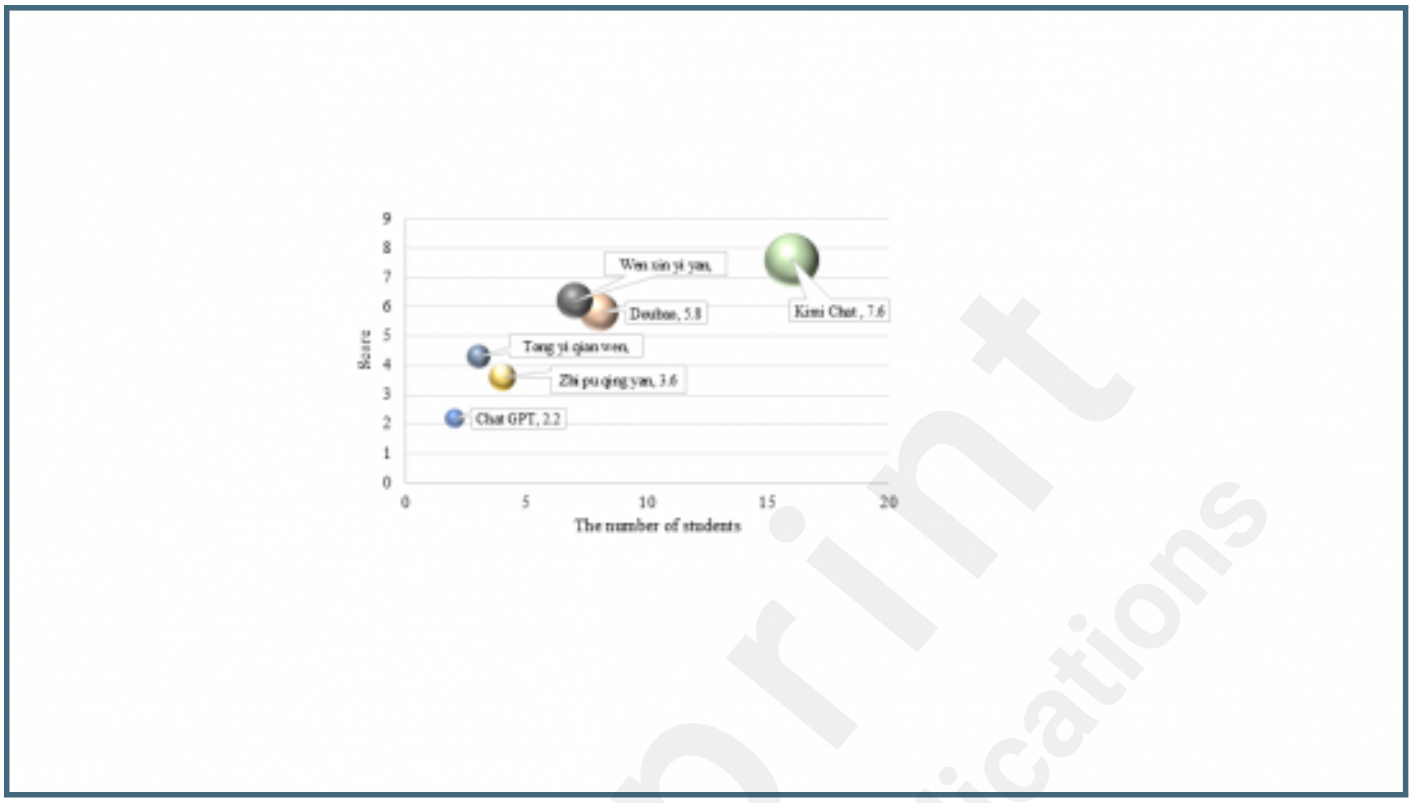
Supplementary Files

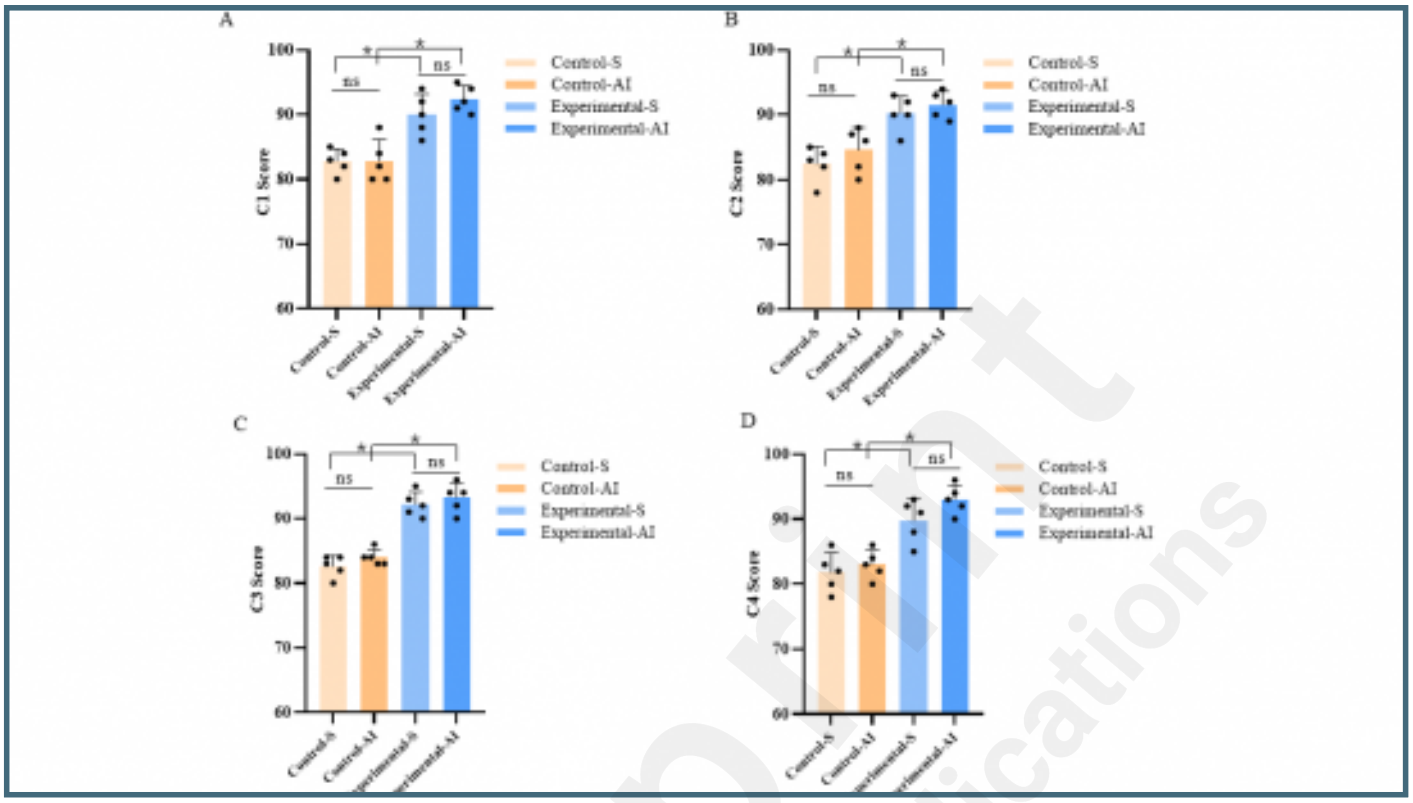
Figures

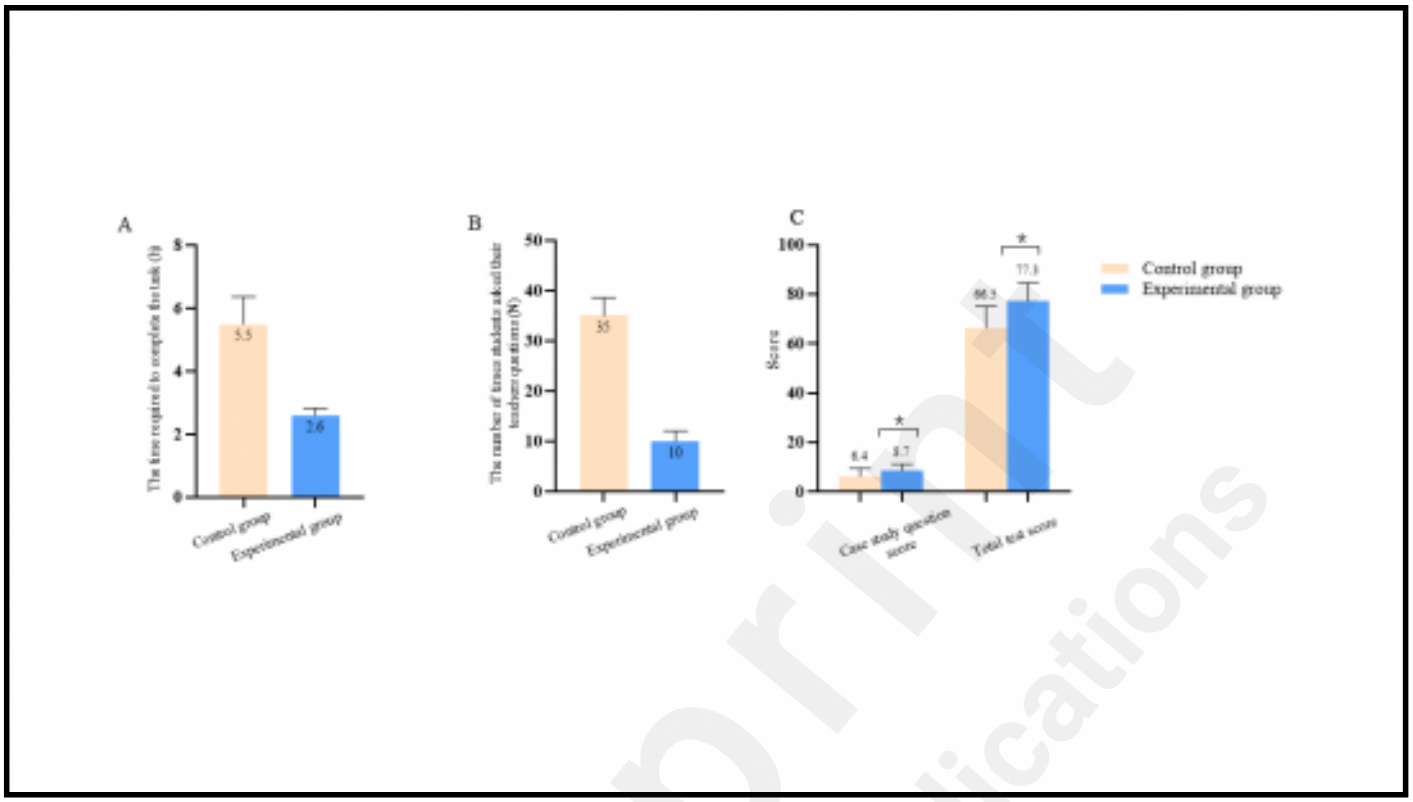
Teaching design of biochemistry and molecular biology course based on GAI tools.



Selection of GAI tools.

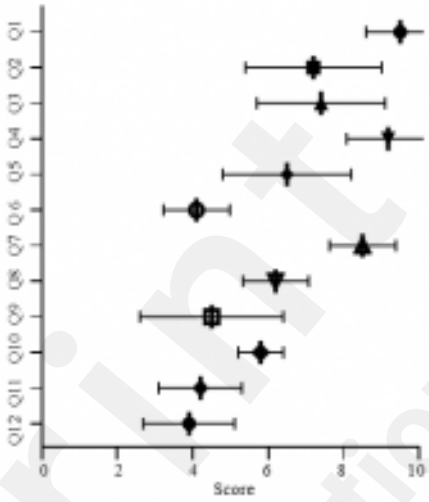






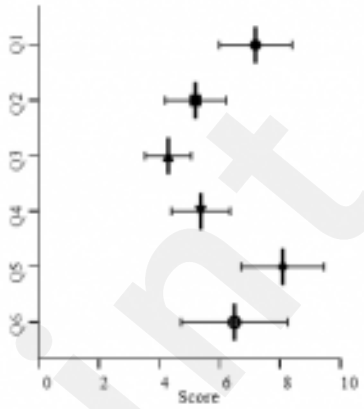
The positive impact of Kimi Chat on students' learning outcomes.

Questions
Q1. Ease of AI tool acquisition
Q2. Ease of interface operation of AI tools.
Q3. The role of the AI tool in saving learning time.
Q4. The role of AI tools in supplementing the knowledge system of teaching materials.
Q5. The role of AI tools in diversity learning resources.
Q6. Potential of AI tools to replace other electronic resources.
Q7. The role of AI tools in basic concept retrieval.
Q8. The role of AI tools in complex concept retrieval.
Q9. The role of AI tools for innovative concept retrieval.
Q10. The role of AI tools for comprehensive analysis of clinical cases
Q11. The role of AI tools for in-depth understanding of clinical cases
Q12. The role of AI tools for generating innovative cases.



The negative effects of Kimi Chat on students' learning.

Questions
Q1. AI tool reduces classroom interaction
Q2. AI tool reduces attention span
Q3. AI tool reduces active thinking
Q4. AI tool reduces innovative thinking
Q5. AI tool increases the competitive pressure to learn
Q6. AI tool increases the homogeneity of learning



Comparison of students' ratings of AI tools before and after the experiment.