Evaluation of the Effectiveness of MED·SMART-AI in Undergraduate Clinical Anesthesia Teaching

Zhen Wu[1], Jie Hou[2], Lixin Yin[1], Wanchao Liu[1], Jun Geng[1], Laiyou Wen[1,*]
[1]*Department of Anesthesiology, Jiangyin Hospital Affiliated to Nantong University, Jiangyin, Wuxi, Jiangsu, China*
[2]*Department of Science and Education, Jiangyin Hospital Affiliated to Nantong University, Jiangyin, Wuxi, Jiangsu, China*
**Corresponding Author.*

Abstract: This study aimed to assess the impact of MED·SMART-AI on undergraduate clinical anesthesia internship teaching. Sixty undergraduate interns at Jiangyin Hospital Affiliated to Nantong University were divided into two groups. One group (n=30, February-April 2024) received traditional instruction, while the other (n=30, May-July 2024) was taught using the MED·SMART-AI method. The evaluation included theoretical knowledge, practical skills, case analysis performance, student learning initiative, and self-evaluation components, such as learning ability, communication skills, and self-confidence. Interns utilizing MED·SMART-AI demonstrated significantly higher scores in theoretical examinations, skills assessments, and case analysis compared to those receiving traditional instruction. Furthermore, the MED·SMART-AI group exhibited notable improvements in learning initiative (information, self-management, and learning cooperation abilities), alongside enhanced self-evaluations of learning ability, doctor-patient communication, literature retrieval, scientific research innovation, and self-confidence. These findings indicate that integrating MED·SMART-AI into undergraduate clinical anesthesia education effectively improves academic outcomes and fosters a more proactive and self-assured learning approach.

Keywords: MED·SMART-AI; Clinical-Anesthesia; Undergraduate teaching; Educational effectiveness; Learning initiative

1. Introduction

Anesthesiology is an important secondary clinical discipline. Undergraduate students systematically learn theoretical knowledge such as physiology, internal medicine, surgery, human anatomy, pain management, and pathophysiology. However, there is still a significant gap between them and becoming clinicians who can provide medical services to patients. Clinical internships are a bridge between theoretical medical knowledge and clinical medical knowledge, allowing interns to experience the process of diagnosing and treating patients, improving their clinical medical knowledge and overall quality. Clinical teaching is crucial to the effectiveness of internships. Conventional clinical anesthesia teaching can effectively impart anesthesiology knowledge, but interns learn passively and cannot establish a systematic, three-dimensional clinical anesthesia knowledge network. Teachers cannot provide personalized teaching, nor can they update the latest advances in anesthesiology in a timely manner to compensate for the shortcomings of the teaching staff, all of which help to improve the quality of clinical teaching in anesthesiology [1]. With the development of artificial intelligence (AI), OpenAI in the United States launched ChatGPT (Chat Generative Pre-Trained Transformer), a generative chatbot, in November 2022. This has revolutionized the traditional AI functions. It can not only answer continuous questions but also reject unscientific questions. It has the potential to innovate teaching and reshape the educational ecosystem [2]. Currently, medical applications based on ChatGPT are gradually being carried out in clinical practice, showing powerful medical assistance functions [3], and significantly improving learning efficiency and enhancing practical ability in medical education [4]. At present, there are few studies on the use of ChatGPT in undergraduate clinical anesthesia

Academic Education Publishing House

teaching. This study creates MED·SMART-AI based on ChatGPT for undergraduate clinical anesthesia teaching, and has achieved remarkable results, which are reported as follows.

2. Materials and Methods

2.1 General Information

Thirty undergraduate interns at Jiangyin Clinical Hospital of Xuzhou Medical University from February 2024 to April 2024 were selected as the control group and received traditional teaching methods. Thirty undergraduate interns at Jiangyin Clinical Hospital of Xuzhou Medical University from May 2024 to July 2024 were selected as the study group and received the MED·SMART-AI teaching method. In the control group, there were 17 males and 13 females, aged 21-24 years, with an average age of (22.38±0.78) years, and a pre-internship GPA of 3.1-3.9, with an average of (3.52±0.17). In the study group, there were 16 males and 14 females, aged 21-24 years, with an average age of (22.41±0.73) years, and a pre-internship GPA of 3.0-3.9, with an average of (3.51±0.18). There was no statistically significant difference between the two groups in terms of gender, age, and pre-internship GPA ($P>0.05$).

2.2 Methods

Both groups were taught according to the content of anesthesiology teaching in the internship syllabus. The teaching hours and content were exactly the same. The control group was given traditional teaching methods. According to the teaching syllabus, teaching PPTs were written, and classroom lectures were used to explain the theory and practical knowledge of clinical anesthesiology. Anesthesia techniques were explained and trained in a simulated classroom. Interns could ask the teacher questions they did not understand in class or after class. The teachers were all associate chief physicians and above. The study group implemented the MED·SMART-AI teaching method, as follows.

2.2.1 Knowledge framework and adaptive language model establishment

A knowledge framework covering basic medical theoretical knowledge and main clinical knowledge was established, covering main conceptual entities and entity relationships. The knowledge hierarchy transitions from general knowledge to professional knowledge. The general layer includes knowledge of basic human structure, physiological systems, and common diseases. The professional layer involves knowledge of various clinical specialties, fully considering the connections between different professional knowledge, forming a unified knowledge framework. A large-scale corpus of medical education texts was constructed, covering medical textbooks, exercises, handouts, examination papers, and literature databases (CNKI, Wanfang Database, etc.). Based on the general language model, this domain text was used for fine-tuning to obtain a professional language model. In addition, the corpus should continue to be expanded, and incremental learning should be used for model updates to adapt to new knowledge and new corpora at any time.

2.2.2 Intelligent question answering implementation and virtual environment construction

The system accurately understands the intention of the question, performs knowledge reasoning based on the knowledge graph, and gives a standard answer that meets the teaching requirements. The question-answering strategy is continuously optimized to improve the smoothness of interaction and support multi-turn dialogues. The virtual environment should be realistic and restore clinical work scenarios. Students can learn typical cases, operate and query patient information, and conduct discussion-based learning. The system can appropriately increase the complexity of cases according to the level of interns and exercise their adaptability.

2.2.3 Formulation of intelligent teaching strategies

The platform continuously tracks the learning status of each intern, forms learning trajectory data, intelligently analyzes and evaluates the intern's knowledge status, automatically formulates the best teaching strategy for each intern, and establishes an AI algorithm to evaluate the teaching effect to achieve strategy optimization.

2.2.4 MED·SMART-AI application guidance

Teachers who master the system explain the basic knowledge and application of MED·SMART-AI to ensure that interns can skillfully use the system, guide theoretical learning, and guide anesthesia technical operations according to system prompts, pictures, and videos. Teacher Yin Lixin is responsible for the maintenance and upgrade of the MED·SMART-AI system.

2.3 Observation Indicators

The theoretical scores, skill scores (use of monitors and ventilators, placement of spinal block positions, tracheal intubation, etc.), and case analysis scores of the two groups before and after learning were recorded. The theoretical scores and case analysis scores were assessed by the Science and Education Department, and the skill scores were assessed by chief physicians who did not know the grouping. The total score for each item was 100 points. The learning initiative and self-evaluation of the two groups before and after learning were compared. The learning initiative was evaluated by the "Active Learning Ability Evaluation Scale", with a Cronbach's α of 0.94. The Likert 5-level scoring method was adopted, including three dimensions: information ability (55 points), self-management ability (50 points), and learning cooperation ability (35 points), with a total score of 140 points. Learning ability is positively correlated with the score [5]. Self-evaluation was assessed by the teaching director based on the self-made self-evaluation questionnaire according to the internship syllabus. The content includes five dimensions: learning ability, doctor-patient communication ability, literature retrieval ability, scientific research innovation ability, and self-confidence. Each dimension contains 4 items, a total of 20 items. The Likert 5-level scoring method is used, with a total score of 100 points. The higher the score, the better the self-evaluation.

2.4 Statistical Processing

SPSS25.0 was used for processing. Measurement data conforming to normal distribution were expressed as mean ± standard deviation (±s). The t-test was used for comparison between groups. P<0.05 was considered statistically significant.

3. Results

3.1 Comparison of Theoretical Scores, Skill Scores, and Case Analysis Scores Between the Two Groups

Before learning, there was no statistically significant difference in theoretical scores, skill scores, and case analysis scores between the two groups (P>0.05). After learning, the theoretical scores, skill scores, and case analysis scores of the study group were higher than those of the control group (P<0.05). See Table 1.

Table 1. Comparison of Theoretical Scores, Skill Scores, and Case Analysis Scores Between the Two Groups (±s, points)

Group	Cases	Theoretical Score	Skill Score	Case Analysis Score
		Before	After	Before
Study Group	30	78.13±5.51	89.43±3.31	71.43±4.45
Control Group	30	79.17±6.09	84.57±4.81	70.67±4.40
t-value		0.689	4.565	0.670
p-value		0.493	0.000	0.505

3.2 Comparison of Learning Initiative Between the Two Groups

Before learning, there was no statistically significant difference in the scores of information ability, self-management ability, and learning cooperation ability between the two groups (P>0.05). After learning, the scores of information ability, self-management ability, and learning cooperation ability of the study group were higher than those of the control group (P<0.05). See Table 2.

Table 2. Comparison of Learning Initiative Between the Two Groups (±s, points)

Group	Cases	Information Ability	Self-Management Ability	Learning Cooperation Ability
		Before	After	Before
Study Group	30	22.23±2.57	28.83±2.12	21.37±2.34
Control Group	30	21.83±2.61	26.77±2.45	20.67±3.26
t-value		0.598	3.499	0.955
P-value		0.552	0.001	0.344

3.3 Comparison of Self-Evaluation Between the Two Groups

Before learning, there was no statistically significant difference in the scores of learning ability, doctor-patient communication ability, literature retrieval ability, scientific research innovation ability, and self-confidence between the two groups (P>0.05). After learning, the scores of learning ability, doctor-patient communication ability, literature retrieval ability, scientific research innovation ability, and self-confidence of the study group were

higher than those of the control group (P<0.05).
See Table 3.

Table 3. Comparison of Self-Evaluation Between the Two Groups (±s, points)

Group	Cases	Learning Ability	Doctor-Patient Communication Ability	Literature Retrieval Ability	Scientific Research Innovation Ability	Self-Confidence
		Before	After	Before	After	Before
Study Group	30	13.13±1.74	15.87±1.22	13.27±1.60	16.33±1.21	9.97±1.81
Control Group	30	12.93±1.70	14.77±1.43	13.47±1.22	15.10±1.37	9.77±1.55
t-value		0.451	3.200	0.545	3.687	0.460
P-value		0.654	0.002	0.588	0.001	0.647

4. Discussion

Anesthesiology is a clinical discipline involving internal medicine, surgery, obstetrics and gynecology, pediatrics, and other disciplines. It not only requires interns to be familiar with the causes, mechanisms, pathophysiological characteristics, surgical methods and steps, postoperative complications, and rehabilitation of surgical diseases, but also puts forward higher requirements for interns' anesthesia technical operation and anesthesia management. How to improve interns' theoretical knowledge and practical skills of anesthesia in a limited time is an important topic for teaching staff. Although optimizing the traditional teaching model can improve the quality of teaching, it still cannot keep up with the rapidly changing pace of medical development and meet the interns' desire for medical knowledge, nor can it effectively improve the interns' humanistic care ability, independent thinking ability, literature retrieval, and scientific research innovation ability [6]. With the application of AI in the medical field, especially the emergence of ChatGPT, medical students' access to medical knowledge and learning methods have undergone major changes, which also poses a major challenge to clinical medical teaching.

ChatGPT is a language processing model driven by AI technology. It has the characteristics of self-learning, iterative updating, and incremental training, and can provide massive medical knowledge. The key to the quality of the model is the knowledge information fed [7]. The MED·SMART-AI model in this study is based on the general ChatGPT and feeds knowledge such as medical textbooks, exercises, handouts, examination papers, and literature databases (CNKI, Wanfang Database). Through self-learning and pre-training, it provides anesthesia medical knowledge, assists interns in understanding and mastering theoretical knowledge, establishes a systematic anesthesiology knowledge system, interacts with interns, promotes knowledge absorption and memory, and improves learning effects. It guides interns' skill operations through videos, pictures, etc., and enhances practical operation capabilities. This study shows that before learning, there was no statistically significant difference in theoretical scores, skill scores, and case analysis scores between the two groups. After learning, the theoretical scores, skill scores, and case analysis scores of the study group were higher than those of the control group, indicating that MED·SMART-AI significantly improved the interns' academic performance, which is similar to the report of Aburumman [8]. In addition to a large amount of detailed clinical medical knowledge, the MED·SMART-AI model has also deeply learned anesthesiology professional knowledge, the latest medical literature knowledge, etc., which can provide interns with the ability to answer questions and solve doubts, and quickly improve the level of anesthesia medical theory. It also provides rich simulated anesthesia scenarios and skill operation videos, pictures, etc. [9] to help interns master the essentials of anesthesia skill operations and improve their skill operation level; it provides simulated surgical patients with different ASA classifications, guides interns on how to conduct preoperative anesthesia assessment and visits, do a good job in intraoperative anesthesia quality control, and deal with abnormal conditions and anesthesia complications in a targeted manner to improve case analysis capabilities. MED·SMART-AI forms learning trajectory data based on interaction with interns, intelligently assesses the status of the knowledge structure, automatically proposes personalized optimal learning strategies, and answers interns' corresponding questions in a targeted, comprehensive, and detailed manner [10], assisting interns in mastering the theoretical knowledge of anesthesia medicine, providing realistic anatomical maps and 3D videos, helping interns master skill operation knowledge, and improving learning effects. In addition, MED·SMART-AI continues to expand the corpus and uses incremental learning to update the model, adapting to new knowledge and new corpus at any time, expanding the scope of knowledge and extending the depth of cognition [11], which is conducive to improving the knowledge level of interns.

Learning initiative is an important driving factor for interns' clinical practice, prompting interns to formulate learning plans, establish learning goals, and choose scientific learning methods and learning resources to maximize the effect of the internship. This study shows that before learning, there was no statistically significant difference in the scores of information ability, self-management ability, and learning cooperation ability between the two groups. After learning, the scores of information ability, self-management ability, and learning cooperation ability of the study group were higher than those of the control group, suggesting that MED·SMART-AI improves the learning initiative of interns. MED·SMART-AI changes the previous "search questions" learning method and realizes the learning mode of real-time interactive question and answer with interns. It can also provide functions such as literature translation and auxiliary reading to improve interns' ability to obtain information; with the help of powerful language processing capabilities, it provides personalized answers according to questions, assists interns in formulating learning plans, optimizing learning programs, enhancing time management, improving decision-making capabilities, and clinical thinking skills [12]; it provides rich learning materials, such as instant videos, 3D pictures, virtual case materials, and the latest medical guidelines or expert consensus, to enhance learning interest and motivation [13]. MED·SMART-AI changes the learning style of interns, provides individualized and interesting learning paths, simulation technology and simulated scenarios quickly improve skill levels, and timely updated medical knowledge broadens the learning horizon, improving the learning initiative of interns in multiple aspects.

Self-evaluation is a subjective evaluation of interns' overall learning status during the learning process. It has the characteristics of immediacy, repetition, and flexibility, and is an effective indicator reflecting the overall learning performance of interns. This study shows that before learning, there was no statistically significant difference in the scores of learning ability, doctor-patient communication ability, literature retrieval ability, scientific research innovation ability, and self-confidence between the two groups. After learning, the scores of learning ability, doctor-patient communication ability, literature retrieval ability, scientific research innovation ability, and self-confidence of the study group were higher than those of the control group, indicating that MED·SMART-AI improved self-evaluation. The reason for this is that MED·SMART-AI has the following advantages: ① It provides intelligent test exercises and examinations. According to the learning characteristics of interns, it automatically generates practice questions and more difficult examination papers that meet the current scores, intelligently marks papers, generates score reports and evaluation results, and promotes interns to improve their learning in a targeted manner; ② Based on data analysis and sequence item execution programs, it intelligently generates classroom or stage learning evaluations, promptly discovers learning problems, and provides targeted training and exercises; ③ The knowledge structure contains a large amount of literature data content. Powerful search and literature knowledge acquisition capabilities enable interns to continuously improve their literature retrieval capabilities, and provide literature evaluation, subtly improving interns' literature evaluation level [14]; ④ AI technology and interdisciplinary thinking help interns analyze data, discover cutting-edge hotspots in professional development, critically think about clinical problems, etc., personalized improve scientific research and innovation capabilities, and improve communication skills with patients through virtual simulation of clinical scenarios, effectively handle doctor-patient relationships, and enhance interns' self-confidence.

Although MED·SMART-AI is intelligent and can assist junior anesthesiologists in preoperative visits [15], it lacks humanistic care, has ethical risks [16], and may be used dishonestly [17]. Interns may rely too much on MED·SMART-AI for learning, which affects the overall and systematic nature of medicine, and is not conducive to teacher-student communication, leading to teachers lacking accurate judgment of interns. In addition, this study has a small sample size and a short research period, and further research is needed for clinical teaching.

In summary, MED·SMART-AI for undergraduate clinical anesthesia teaching significantly improves interns' academic performance, enhances learning initiative, and improves self-evaluation. It is worthy of promotion and in-depth research in clinical teaching.

Acknowledgments

1.Nantong University Teaching Reform Research Project (Project Number: 2023B39)

2.2023 Wuxi Municipal Health Commission Youth Project:

Application of MED·SMART-AI Medical Artificial Intelligence Model in Clinical Practice and Teaching under Residency Standardization Training, Project Number 0202360.

3.2023 Nanjing Medical University Teaching Reform Project:

Research on the Use of MED·SMART-AI in Standardized Residency Training for Anesthesiology, Project Number 2023LX114.

References

[1] Sallam, M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare 11, 887 (2023).

[2] Kung, James, andy et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLoS Digit. Health 2, e0000198 (2023).

[3] Cascella, M, Joseph W Burns, Flavia B Consens et al. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. J. Med. Syst. 47, 33 (2023).

[4] Ali, Sat, Rralic. et al. Using ChatGPT to write patient clinic letters. Lancet Digit. Health 5, e231–e232 (2023).

[5] Mbakwe, Roderick J Little, Karen J Angell et al. ChatGPT passing USMLE shines a spotlight on the challenges of assessing clinical reasoning. PLoS Digit. Health 2, e0000229 (2023).

[6] Eysenbach, The role of ChatGPT, chatbots, and other large language models in health care and medical research. JMIR Med. Educ. 9, e46892 (2023).

[7] Lee, P., Bubeck, S. & Petro, J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. N. Engl. J. Med. 388, 1233–1239 (2023).

[8] Aburumman, The impact of artificial intelligence-powered ChatGPT on improving the learning experience of anesthesia students. Int. J. Adv. Comput. Sci. Appl. 14, (2023).

[9] Shorey, Shohreh Deldari, Dimitris Spathis et al. A virtual counseling application using artificial intelligence for communication skills training in nursing education: development and pilot feasibility study. JMIR Nurs. 2, e12657 (2019).

[10]Gilson, How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. JMIR Med. Educ. 9, e45312 (2023).

[11]Singh, N, Mann, D. M, Artificial intelligence in medical education: a scoping review and proposed research agenda. J. Med. Educ. Curric. Dev. 10, 23821205231171685 (2023).

[12]Biswas, S. Potential use of Chat GPT in global health. Ann. Biomed. Eng. 51, 1181–1182 (2023).

[13]Lubowitz, J. H. Editorial commentary: ChatGPT, artificial intelligence, and the future of orthopedic surgery education. Arthroscopy 39, 1151–1152 (2023).

[14]Temsah, Overview of early ChatGPT's impact on higher education: strengths, challenges, and future directions. J. Artif. Intell. Educ. (2023).

[15]Perera, S. N, Abeysekera, W. Y. & Perera, A. I. Usefulness of Chat GPT in clinical medicine. J. Med. Internet Res. 26, e52712 (2024).

[16]Patel, S. B. & Lam, K. ChatGPT: the future of discharge summaries? Lancet Digit. Health 5, e107–e108 (2023).

[17]Shen, Y ChatGPT and other large language models are double-edged swords. Radiology 307, e230163 (2023).