

Effectiveness of AI Simulation as a Clinical Learning Tool for First-year MBBS Students: A Pilot Study

GOKUL SURESH REVATHY¹, ABHIMA MALLIKA BHASURANGAN²

ABSTRACT

Introduction: Traditional methods of medical education, including lectures and textbook-based learning, often emphasise passive knowledge acquisition and may fall short in preparing students for real-world clinical reasoning. With the advent of Artificial Intelligence (AI) and Large Language Models (LLM), simulation-based learning has gained attention as a tool for enhancing engagement, knowledge retention, and decision-making skills in medical training.

Aim: To evaluate the effectiveness of AI-driven clinical simulations in improving knowledge acquisition and retention among first-year MBBS students, compared to conventional teaching methods. It also aimed to assess students' perception of AI simulations as a teaching-learning method using structured feedback.

Materials and Methods: The current pilot study was conducted in the Department of Physiology of Sree Narayana Institute of Medical Sciences, Ernakulam, Kerala, India for a period of 6 months from December 2024 to May 2025. Among 60 Phase 1 MBBS students, allocated into 2 groups. Utilising a crossover design, each group received both AI simulation

(ChatGPT-based) and traditional lecture-based teaching for different clinical cases. Six clinical cases were used for the intervention. Pre-test and post-test assessments were conducted to measure immediate learning gains. Retention was evaluated using two follow-up tests administered at one-month interval. Statistical analysis included paired and unpaired t-tests, Mann-Whitney U tests, and effect size calculations. Additionally, structured feedback was collected from students in the AI group to assess perceptions of the learning experience.

Results: AI group showed significantly higher improvement in Cases 1 and 4. Gains were comparable in other cases, with trends favouring AI. Retention Test 1: AI group outperformed traditional group ($p=0.0323$, $r=0.272$). Retention Test 2: AI group retained a small advantage ($p=0.0678$, $r=0.240$). Students reported positive perceptions across domains, with mean Likert scores above 4.0.

Conclusion: AI simulation is an effective pedagogical tool for improving clinical knowledge and fostering deeper understanding in early medical learners. It promotes sustained engagement and meaningful learning, with measurable improvement in retention.

Keywords: Artificial intelligence, Bachelor of medicine, Bachelor of surgery, Computer-assisted instructions, Knowledge retention, Medical education, Teaching methods

INTRODUCTION

Medical education has traditionally depended on a combination of didactic lectures, textbooks, and clinical rotations to build foundational knowledge and practical competencies in students. These methods often emphasise passive learning and may not fully equip learners to navigate the complexities of real-world clinical decision-making [1,2]. In the early years of medical training, opportunities for clinical engagement are limited, and there is a growing need for active, experiential learning tools that can bridge this gap [3]. Recent advancements in AI, particularly in generative AI and Natural Language Processing (NLP), are opening new frontiers in medical education [4,5]. Tools like ChatGPT, powered by LLMs, have demonstrated the ability to simulate dynamic, context-aware patient interactions [6,7]. These AI-driven simulations mimic clinical reasoning pathways, and adapt to student inputs in real-time. As a result, learners are not only consuming information but also practising diagnostic thinking, decision-making, and communication in a safe, scalable, and consistent environment [5,7].

Simulation-based learning has already been validated as an effective strategy in medical education, improving knowledge retention, clinical reasoning, and learner confidence [1,2]. AI-enabled simulations reduce the dependence on physical infrastructure and faculty presence while increasing accessibility and personalisation [3,8]. The interactive and immersive nature of these tools fosters deeper cognitive engagement, which is crucial for early medical students forming their conceptual frameworks [9,10].

Loubbairi S et al., conducted a comprehensive systematic review on the use of AI simulations in health professions education [11]. Their findings indicated enhanced clinical reasoning abilities. AI simulations replicated the complexity of clinical scenarios, while adjusting in real-time to learner responses, facilitating higher-order thinking. Sriram A et al., through their study revealed that AI tools could tailor learning content to individual student performance, addressing cognitive gaps and reinforcing strengths, which increased learner motivation and self-efficacy [12]. Fazlollahi AM et al., quantitatively assessed academic outcomes in medical students using AI-based learning tools [13]. They found that students exposed to AI-assisted learning scored higher in assessments. This was attributed to the ability of AI tools to simulate clinical reasoning steps and provide contextual data, promoting deeper understanding.

Negi R et al., examined the practical barriers and enablers of AI simulation adoption in health professions education [14]. Their qualitative analysis highlighted the importance of institutional support, faculty training, and technological infrastructure. They concluded that while the pedagogical benefits are well-recognised, operational challenges must be addressed for sustainable integration. Collectively, the literature affirms that AI-based simulations not only improve immediate learning outcomes, but also foster learner engagement. While prior studies have demonstrated the benefits of AI-based tools and simulations in medical education, there is limited evidence regarding their effectiveness in early-phase (Phase 1 MBBS) learners, particularly in the context of foundational physiology integrated with clinical reasoning. Thus, the current project seeks to determine

whether AI simulations can serve as an impactful educational innovation in undergraduate medical training.

Study objectives:

- To compare the effectiveness of AI simulations with traditional lecture-based teaching for clinical knowledge acquisition among Phase 1 medical students;
- To analyse knowledge retention over time through follow-up assessments;
- To assess students' perception of AI simulations as a teaching-learning method using structured feedback.

MATERIALS AND METHODS

The present pilot study with a crossover design comparing AI-based simulation and traditional teaching was conducted in the Department of Physiology, Sree Narayana Institute of Medical Sciences, Ernakulam, Kerala, India from December 2024 to May 2025. Institutional Ethics Committee (IEC) approval (IEC No: APPL/IEC/117/88 dt 23.09.2024 SNIMS) was obtained before commencement.

Inclusion criteria: Phase 1 MBBS students of the academic year 2024-2025 who provided written informed consent.

Exclusion criteria: Students who did not give consent, students who were absent for one or more key components of the study (sessions or assessments).

Sample size calculation: A total of 60 Phase 1 MBBS students were included in the study using convenience sampling, based on the available cohort during the study period. The study was designed as a pilot study to evaluate the feasibility, acceptability, and preliminary effectiveness of AI-based clinical simulations in early medical education. Given the novelty of integrating AI-driven simulations into Phase 1 training and the limited existing data in this specific context, a pilot approach was considered appropriate to generate initial evidence and inform larger, adequately powered future studies.

Instruments:

- Six AI-based case simulations developed using ChatGPT;
- Validated pretest and post-test MCQs for each case;
- Retention tests 1 and 2;
- Feedback questionnaire.

The instruments were pilot tested on a group of 10 students prior to the study.

Study Procedure

The study commenced after getting written informed consent from the students. Students were allocated into two groups based on prior academic performance, as assessed using Class 12 examination scores and National Eligibility cum Entrance Test-Undergraduate (NEET-UG) ranks, to ensure a comparable distribution of baseline academic ability between the groups.

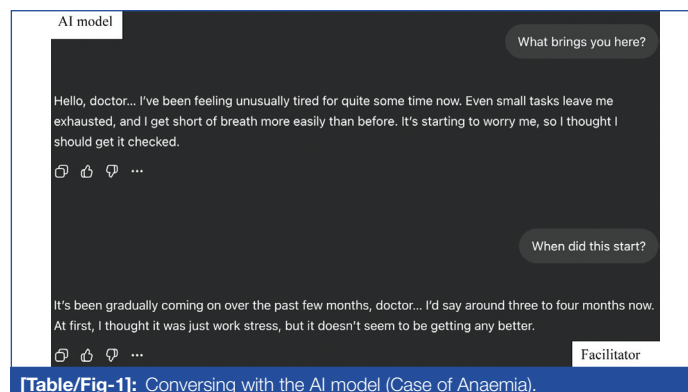
A total of 6 AI-based case simulations were developed using ChatGPT's customGPT function by configuring the model to emulate a virtual patient with predefined clinical conditions.

Selected cases:

1. Iron deficiency anaemia;
2. Haemophilia;
3. Myasthenia gravis;
4. Myocardial infarction;
5. Graves' disease;
6. Diabetes mellitus.

The study was conducted in two phases. During phase 1 sessions on each clinical condition, Group A received AI-based simulations while Group B received traditional teaching for the first three cases.

During the AI-based simulation sessions, the teacher acted as a facilitator, interacting with the virtual patient by eliciting history and guiding the discussion. Students were encouraged to actively participate by asking questions and applying their clinical reasoning. In contrast, the traditional lecture sessions incorporated clinical case vignettes with a comparable set of symptoms. For the subsequent three cases in phase 2, the intervention was reversed, with Group A receiving traditional teaching and Group B undergoing AI-based simulation. A representation of how the facilitator converses with the AI simulation is given in [Table/Fig-1].



[Table/Fig-1]: Conversing with the AI model (Case of Anaemia).

For each clinical case, a pretest was administered to assess baseline knowledge, followed by an immediate post-test using the same set of questions to evaluate learning gains. Each pretest and post-test consisted of 10 clinically oriented MCQs based on the respective case. Each correct response was awarded one mark, with no negative marking, resulting in a maximum possible score of 10 and minimum possible score of zero.

To assess knowledge retention, two follow-up tests were conducted at one-month intervals after the completion of each phase of the study. Each retention test consisted of 10 MCQs, distinct from those used in the pretest and post-test, derived from the previously covered three clinical cases, with scores ranging from 0 to 10.

Student perceptions regarding the AI-based simulation were collected using a structured and validated feedback questionnaire administered anonymously. The questionnaire consisting of 11 questions was based on a 5-point Likert scale and assessed domains such as understanding, engagement, critical thinking, and recall of foundational knowledge. The MCQs and questionnaire were developed by the investigators and validated through subject expert review for content validity, clarity and relevance, prior to administration.

STATISTICAL ANALYSIS

Data were entered in Microsoft Excel and analysed using Statistical Package for Social Sciences (SPSS) Statistics. The normality of data distribution was assessed using the Shapiro-Wilk test. For within-group comparisons, paired t-tests were used for normally distributed data, while the Wilcoxon signed-rank test was applied for non-normally distributed data. For between-group comparisons, independent t-tests were used for normally distributed data, and the Mann-Whitney U test was used for non-normally distributed data. Effect size (r) was calculated for significant comparisons to assess the magnitude of the intervention effect. A p-value of <0.05 was considered statistically significant.

RESULTS

The mean age of the population was 19.5±1.55 years. The study included 44 females and 16 males.

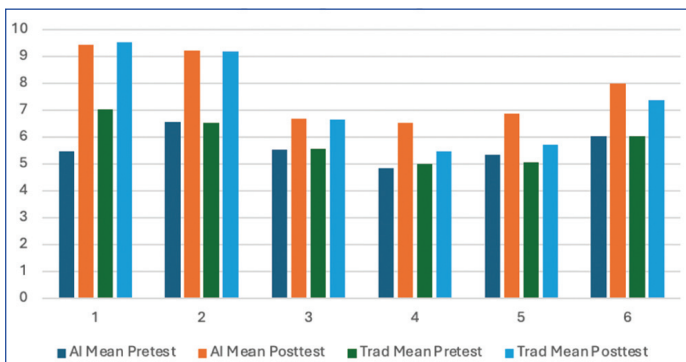
Knowledge Gain Analysis

The AI simulation group demonstrated consistently higher knowledge gains across all six clinical cases when compared to the traditional group [Table/Fig-2]. The most notable improvements

Case	AI mean pretest±SD	AI mean post-test±SD	p-value	Trad mean pretest±SD	Trad mean post-test±SD	p-value	AI mean gain±SD	Trad mean gain±SD	p-value	Effect size (r)	Interpretation
1	5.47±1.74	9.43±0.73	<0.0001	7.03±1.52	9.53±1.07	<0.0001	3.97±1.67	2.5±1.76	0.0028	0.382	Significant gain with AI
2	6.57±1.92	9.23±0.86	<0.0001	6.53±2.03	9.2±1.16	<0.0001	2.67±1.71	2.67±2.29	>0.99	0.001	No difference
3	5.53±1.7	6.7±1.73	0.0039	5.57±1.45	6.67±1.52	0.0045	1.17±2.04	1.1±1.95	0.8975	0.017	No meaningful difference
4	4.83±1.66	6.53±1.83	<0.0001	5±2.02	5.47±1.74	0.1946	1.7±1.7	0.47±1.93	0.0110	0.339	AI significantly better
5	5.33±2.31	6.87±1.87	0.0003	5.07±1.89	5.73±2.27	0.0597	1.53±2.03	0.67±1.86	0.0903	0.222	Trend favouring AI, not significant
6	6.03±2.08	8±1.62	<0.0001	6.03±2.27	7.37±2.2	0.0069	1.97±2.04	1.33±2.51	0.2883	0.138	Small improvement with AI, not significant

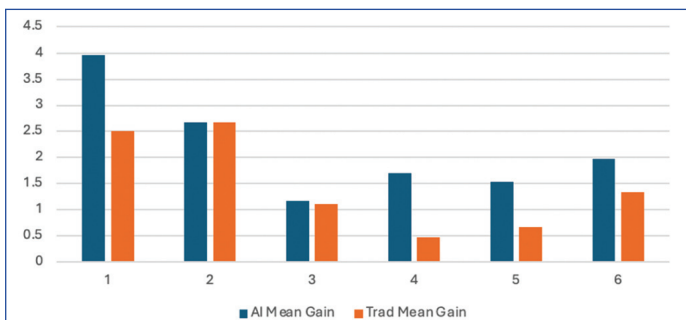
[Table/Fig-2]: Comparison of mean pretest, post-test scores and gain between AI case facilitated learning and traditional lectures.

were observed in Case 1 and Case 4. In Case 1, the AI group had a low pretest score but surpassed the traditional group in the post-test, indicating a marked improvement in conceptual understanding ($p=0.0028$, $r=0.382$). Similarly in Case 4, a statistically significant difference in gain scores favoured the AI group, demonstrating its ability to enhance learning, particularly in clinically applied physiology ($p=0.0110$, $r=0.339$) [Table/Fig-2,3].



[Table/Fig-3]: AI vs Traditional - Mean scores comparison pretest and post-test.

While Cases 2, 3, 5, and 6 showed no statistically significant differences between the two groups, all displayed positive trends favouring the AI group, with higher mean gain scores and effect sizes ranging from negligible to small ($r=0.138$ to 0.240) [Table/Fig-4]. This suggests that even in cases where significance wasn't achieved, the AI intervention still conferred educational advantages.



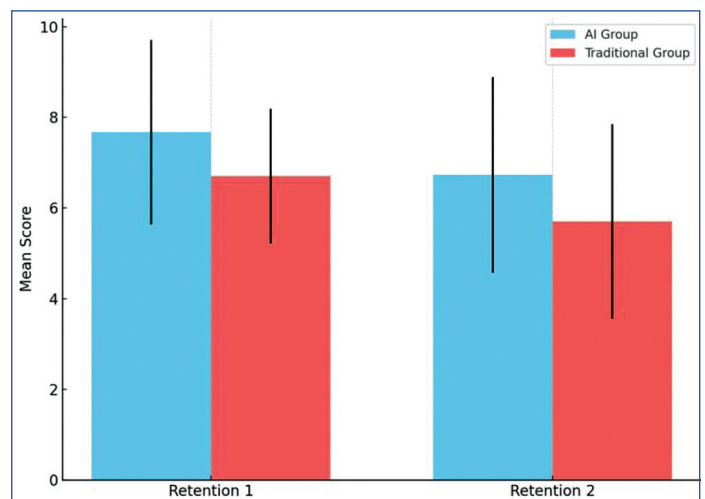
[Table/Fig-4]: AI vs Traditional-Mean gain comparison.

Retention Test Performance

The AI group showed a statistically significant advantage over the traditional group ($p=0.0323$) in retention test 1, with a mean score of $7.67±2.04$ compared to $6.70±1.49$ in the traditional group. The effect size ($r=0.272$) indicates a small to moderate educational impact, reinforcing the durability of learning through simulation. In retention Test 2, although not statistically significant ($p=0.0678$), the AI group maintained a slight edge in performance (mean: $6.73±2.16$) compared to the traditional group (mean: $5.70±2.14$), suggesting a continued trend of better knowledge retention even after a prolonged interval [Table/Fig-5]. These findings highlight the potential of AI simulations to promote not just immediate understanding but also long-term retention, a critical goal in foundational medical education [Table/Fig-6].

Retention test	AI Mean±SD	Traditional Mean±SD	p-value	Effect size (r)	Interpretation
Retention 1	7.67±2.04	6.70±1.49	0.0323	0.272	Significant advantage for AI
Retention 2	6.73±2.16	5.70±2.14	0.0678	0.240	Trend favouring AI, not statistically significant

[Table/Fig-5]: Comparison of mean scores of retention tests between AI case facilitated learning and traditional lectures.



[Table/Fig-6]: AI vs traditional - retention test performance.

Student Feedback

Feedback collected from students in the AI group revealed overwhelmingly positive perceptions. Quantitative scores averaged above 4.0 out of 5 across all domains, including critical thinking and recall of physiology. Standard deviations were relatively low, indicating consistent positive sentiment across the group [Table/Fig-7].

Question	Mean score±Standard deviation
Helped me understand clinical decision-making better.	4.2±0.7
Encouraged me to apply theoretical knowledge in a practical context.	4.2±0.8
Encouraged me to think critically and prioritise actions.	4.2±0.8
Recalled relevant physiology concepts during the simulation.	4.0±0.8
Debrief and reflection at the end of the case was useful for my learning.	4.1±0.9
Engaging and immersive.	4.2±0.9
Interface and instructions were easy to follow.	4.2±0.8
More involved compared to traditional lecture-based teaching.	4.3±0.9
Responded well to clinical inputs.	4.1±0.8
Interested in using similar simulations for other topics.	4.4±0.8
Retained information better through this interactive method.	4.2±0.8

[Table/Fig-7]: AI vs traditional - student feedback and perception.

These findings support the integration of AI as a learner-centred tool that fosters motivation, critical reasoning, and deeper conceptual grasp.

DISCUSSION

The present study evaluated the effectiveness of AI-based clinical simulation as a teaching-learning tool among Phase 1 MBBS students. As the present study was a pilot study, the findings should be interpreted as preliminary evidence regarding the educational utility of AI-based simulation in early medical training. The findings demonstrated that AI-based simulation resulted in greater or comparable knowledge gains across most clinical cases, with statistically significant improvement observed in select cases. In addition, students exposed to AI-based learning showed better performance in retention tests and reported high levels of engagement and perceived understanding.

The improvement in knowledge acquisition with AI-based simulations in this study is consistent with existing literature. Loubbairi S et al., in their systematic review reported that AI-driven simulations enhance clinical reasoning and promote active engagement by enabling learners to interact dynamically with clinical scenarios [11]. Fazlollahi AM et al., demonstrated that students exposed to AI-assisted learning tools like virtual patients scored better in competency scores compared to those receiving traditional teaching [13]. The findings related to knowledge retention also support the educational value of AI-based simulation. The AI group demonstrated significantly better performance in the first retention test and maintained a trend toward improved retention in the second test.

Student feedback in this study revealed high levels of satisfaction with AI-based simulation, particularly in domains such as engagement, critical thinking, and understanding. These findings are consistent with those reported by Sriram A et al., where AI-based simulations enhance learner motivation, engagement, and self-directed learning [12].

The ability of AI simulations to engage students in active decision-making via dynamic, dialogue-based format and respond contextually to user input created a personalised learning experience, allowing students to explore scenarios at their own pace and uncover gaps in their understanding in real-time [12]. Furthermore, the ability of AI-based tools to reinforce learning through repeated, risk-free interactive engagement may facilitate deeper cognitive processing and better retention [1,2]. The AI simulations emphasised experiential learning, probably the main reason for the improved knowledge acquisition, retention and feedback seen in our study. However, the successful implementation of AI-based tools in medical education depends on factors such as institutional support, faculty training, and technological infrastructure, as highlighted by Negi R et al., [14].

Overall, the findings of the current study suggest that AI-based clinical simulation is a promising educational tool for early medical training.

Limitation(s)

The study was limited to a single Institution and subject cohort. Post-test used the same set of questions as pre-test. Hence, the knowledge gain could be as a result of practice effect. The study had a relatively small sample size which may limit statistical power in some comparisons.

CONCLUSION(S)

The AI simulation is an effective pedagogical tool that enhances learner engagement, supports knowledge acquisition, and contributes to improved retention. Future studies should involve larger, multi-center cohorts and explore long-term outcomes, including the impact on clinical reasoning skills and performance in real clinical settings.

Acknowledgement

The authors extend their sincere gratitude to all student participants, their co-investigator, faculty of the Department of Physiology, and the Institutional Ethics Committee at SNIMS for their invaluable support.

REFERENCES

- [1] Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. *JAMA*. 2011;306(9):978-88.
- [2] Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Med Teach*. 2005;27(1):10-28.
- [3] Wartman SA, Combs CD. Reimagining medical education in the age of AI. *AMA J Ethics*. 2019;21(2):E146-52.
- [4] Masters K, Ellaway RH. e-learning in medical education guide 32 part 2: Technology, management and design. *Med Teach*. 2008;30(5):474-89.
- [5] Imran N, Jawaid M. Artificial intelligence in medical education: Are we ready for it? *Pak J Med Sci*. 2020;36(5):857-59. Doi: 10.12669/pjms.36.5.3042.
- [6] Popenici SAD, Kerr S. Exploring the impact of artificial intelligence on teaching and learning in higher education. *Res Pract Technol Enhanc Learn*. 2017;12(1):22. Doi: 10.1186/s41039-017-0062-8.
- [7] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-98.
- [8] Hasan S, Ahmed A, Ismail FW. The evolving role of AI in simulation-based medical education: A narrative review. *Adv Med Educ Pract*. 2026;17:58191. Doi: 10.2147/AMEPS581691.
- [9] Goh PS, Sandars J. A vision of the use of technology in medical education after the COVID-19 pandemic. *MedEdPublish* (2016). 2020;9:49. Doi: 10.15694/mep.2020.000049.1.
- [10] Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. Doi: 10.1038/s41591-018-0300-7.
- [11] Loubbairi S, Moussaoui YE, Lahlou L, Chakri I, Nassik H. The impact of artificial intelligence-driven simulation on the development of non-technical skills in medical education: A systematic review. *J Educ Eval Health Prof*. 2025;22:37.
- [12] Sriram A, Ramachandran K, Krishnamoorthy S. Artificial intelligence in medical education: Transforming learning and practice. *Cureus*. 2025;17(3):e80852. Doi: 10.7759/cureus.80852.
- [13] Fazlollahi AM, Bakhaider M, Alsayegh A, Yilmaz R, Winkler-Schwartz A, Mirchi N, et al. Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: A randomized clinical trial. *JAMA Netw Open*. 2022;5(2):e2149008.
- [14] Negi R, Chopra D, Maheshwari K, Mahajan A, Badyal D, Venkataramani P. Artificial intelligence in simulation-based training for health professions education: Navigating the rabbit hole. *Med J Armed Forces India*. 2025;81(6):637-43. Doi: 10.1016/j.mjafi.2025.08.001.

PARTICULARS OF CONTRIBUTORS:

1. Associate Professor, Department of Physiology, Sree Narayana Institute of Medical Sciences, Ernakulam, Kerala, India.
2. Professor, Department of Pharmacology, Sree Narayana Institute of Medical Sciences, Ernakulam, Kerala, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:

Gokul Suresh Revathy,
Revathy House, Karattupallikkara Perumbavoor, Ernakulam-683542, Kerala, India.
E-mail: gkr9900@gmail.com

AUTHOR DECLARATION:

- Financial or Other Competing Interests: None
- Was Ethics Committee Approval obtained for this study? Yes
- Was informed consent obtained from the subjects involved in the study? Yes
- For any images presented appropriate consent has been obtained from the subjects. No

PLAGIARISM CHECKING METHODS: [Jain H et al.]

- Plagiarism X-checker: Feb 19, 2026
- Manual Googling: May 12, 2026
- iThenticate Software: May 15, 2026 (1%)

ETYMOLOGY: Author Origin

EMENDATIONS: 7

Date of Submission: Feb 03, 2026

Date of Peer Review: Apr 11, 2026

Date of Acceptance: May 18, 2026

Date of Publishing: Jul 01, 2026