

ACCURACY AND TEACHABILITY OF ARTIFICIAL INTELLIGENCE CHATBOTS IN SOLVING PHARMACEUTICAL CALCULATIONS: A DESCRIPTIVE STUDY

K. G. SATHEESH KUMAR, UPPU BHARATHI*

Department of Pharmacology, Sri Venkateswara Medical College (SVMC), Tirupati, India
*Corresponding author: Uppu Bharathi; *Email: ubharathi1212@gmail.com

Received: 10 Nov 2025, Revised and Accepted: 03 Jan 2026

ABSTRACT

Objectives: This study evaluated the accuracy and teachability of widely available AI chatbots in performing pharmaceutical calculations, including adult and paediatric dosing, IV infusion rates, and dilution/concentration problems.

Methods: A descriptive study was conducted in the Department of Pharmacology, Sri Venkateswara Medical College, Tirupati. Five free-access AI chatbots (A–E) were given 35 questions each. Responses were compared with gold standard answers. Incorrect answers were corrected through feedback, and performance was reassessed. Data were analysed descriptively using Microsoft Excel.

Results: Accuracy varied across chatbots, with Chatbot A performing best (89%) and Chatbot D close behind (86%). All chatbots performed well in simpler tasks such as percentage and ratio calculations (86%), whereas paediatric dosing, IV infusion, and dilution/concentration problems were more error-prone. Following feedback, teachability was high: Chatbots A, B, and D corrected all errors (100%), while C and E improved to 91% and 82%, respectively.

Conclusion: AI chatbots show potential as educational and clinical support tools for pharmaceutical calculations. They handle simpler tasks reliably and can improve performance after feedback. Nevertheless, supervision remains crucial for complex calculations to ensure patient safety. When integrated carefully, AI chatbots can complement traditional learning and clinical practice.

Keywords: Artificial intelligence, Chatbots, Pharmaceutical calculations, Accuracy, Teachability, Paediatric dosing, IV infusion, Medication safety

© 2026 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>) DOI: <https://dx.doi.org/10.22159/ijcpr.2026v18i2.8068> Journal homepage: <https://innovareacademics.in/journals/index.php/ijcpr>

INTRODUCTION

Pharmaceutical calculations are a fundamental skill for medical and pharmacy students, clinicians, and researchers, ensuring the accurate dosing of medications and the safe preparation and administration of drugs. Proficiency in these calculations is vital to patient treatment, as healthcare professionals must dispense drugs in the correct concentrations and calculate appropriate doses to ensure that each particular drug is being administered correctly [1]. Medication errors, particularly those arising from incorrect pharmaceutical calculations, are a significant concern in healthcare settings. Such errors can lead to underdosing, toxicity, or therapeutic failure, posing serious risks to patient safety [2]. Traditionally, learners have relied on textbooks, manual calculations, calculators, and faculty guidance. However, the emergence of digital learning tools, particularly artificial intelligence (AI) chatbots, has introduced innovative methods to support education and clinical decision-making [3]. With the rise of AI, chatbots are increasingly being used in healthcare education and clinical decision support. These AI-powered tools have the potential to assist learners in solving pharmaceutical calculations, provide immediate feedback, and enhance understanding through interactive guidance [4]. Beyond education, AI chatbots may also serve as practical clinical tools. They have the potential to assist healthcare professionals with rapid dosage calculations in real-time, particularly in urgent or high-pressure situations, potentially reducing medication errors and enhancing patient safety [5]. However, the accuracy and teachability of AI chatbots in solving pharmaceutical calculations remain underexplored. A recent study found that while AI chatbots demonstrated limited accuracy for multi-step pharmaceutical calculations, they may be more reliable for low-complexity calculations. Additionally, many chatbots were found to be teachable on at least one question, suggesting potential for improvement through feedback [6]. In conclusion, while AI chatbots hold promise as educational and clinical support tools in healthcare, their accuracy and teachability in solving pharmaceutical calculations require further investigation. This descriptive study aims to evaluate

both the accuracy and teachability of AI chatbots in solving pharmaceutical calculations, providing insights into their utility and identifying areas for improvement.

MATERIALS AND METHODS

Study setting

This descriptive study was conducted in the Department of Pharmacology, Sri Venkateswara Medical College (SVMC), Tirupati. The department provides a structured academic environment for evaluating educational and clinical support tools such as artificial intelligence (AI) chatbots.

Study design

A descriptive study design was employed to assess the accuracy and teachability of AI chatbots in solving pharmaceutical calculations.

Selection of AI chatbots

Five AI chatbots were selected for evaluation. To maintain confidentiality and avoid ethical issues, the chatbots were labelled A–E for the purposes of this study. The labelling was arbitrary, meaning that “A” could be any of the five chatbots, “B” any other, and so on. This approach ensured that readers could understand the results without being able to identify the specific platforms. The chatbots were chosen because they are widely available and capable of providing educational support for pharmaceutical calculations.

Sample size calculation

The sample size was calculated using the standard formula for proportions:

$$n = \frac{Z^2 \cdot p \cdot q}{d^2}$$

Where:

- $Z = 1.96$ (for 95% confidence)

- $p = 0.90$ (anticipated accuracy based on Campbell *et al.*¹)
- $q = 1 - p = 0.10$
- $d = 0.05$ (allowable error)

Applying the values:

$$n = \frac{(1.96)^2 \cdot 0.90 \cdot 0.10}{(0.05)^2} = \frac{0.345744}{0.0025} \approx 138.3 \approx 139$$

To ensure sufficient data, 35 pharmaceutical calculation questions covering adult dosing, pediatric dosing, IV infusion rates, and dilution/concentration problems were selected. Each of the five chatbots was presented with all 35 questions, yielding $35 \times 5 = 175$ responses. This exceeds the calculated sample size, ensuring adequate statistical power.

Procedure

1. Each chatbot (labeled A–E) was individually presented with all 35 questions, and their responses were recorded.
2. A gold standard answer sheet, prepared and validated by a senior professor, was used to compare chatbot responses.
3. For every incorrect answer, the correct solution was provided as feedback to the chatbot.

4. The chatbot's responses were reassessed on the same or similar questions to evaluate teachability and improvement in accuracy.

Data analysis

All data were compiled and analyzed using Microsoft Excel (MS Excel). Accuracy for each chatbot was calculated as the proportion of correct answers compared to the gold standard. Teachability was assessed by evaluating improvements in performance after feedback. Descriptive statistics including percentages were used to summarize results.

RESULTS

In the present study, 35 pharmaceutical calculation questions were administered to five artificial intelligence chatbots to assess their accuracy and teachability. The responses were analyzed with respect to calculation type, error distribution, and post-feedback improvement, and the results are summarized below.

Fig. 1 presents the overall accuracy of the chatbots in solving 35 pharmaceutical calculation questions. Chatbot A achieved the highest accuracy (89%), followed closely by Chatbot D (86%). Chatbots B, C, and E showed moderate performance, with accuracies of 71%, 69%, and 69%, respectively.

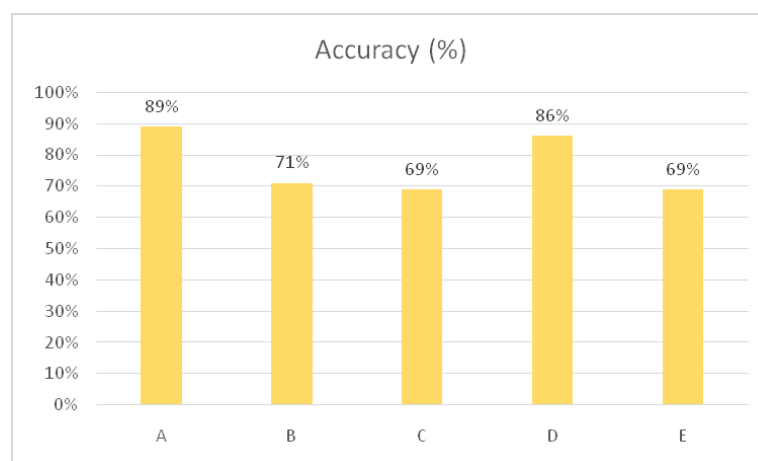


Fig. 1: Overall accuracy of chatbots in pharmaceutical calculations

Table 1: Accuracy by type of pharmaceutical calculation

Type of calculation	No. of questions	Chatbot A accuracy (%)	Chatbot B accuracy (%)	Chatbot C accuracy (%)	Chatbot D accuracy (%)	Chatbot E accuracy (%)
Adult dose calculation	7	7 (100%)	5 (71%)	5 (71%)	6 (86%)	5 (71%)
Paediatric dose calculation	7	6 (86%)	4 (57%)	4 (57%)	6 (86%)	4 (57%)
IV and infusion rate	7	6 (86%)	5 (71%)	5 (71%)	6 (86%)	5 (71%)
Dilution and Concentration	7	6 (86%)	5 (71%)	4 (57%)	6 (86%)	4 (57%)
%/Ratio strength	7	6 (86%)	6 (86%)	6 (86%)	6 (86%)	6 (86%)

Table 1 shows the accuracy of chatbots according to the type of pharmaceutical calculation. For adult dose calculations, Chatbots A and D scored 86%, while Chatbots B, C, and E were less accurate at 57%. In pediatric dosing, Chatbot A achieved 100% accuracy, Chatbot D scored 86%, and Chatbots B, C, and E each scored 71%.

For IV and infusion rate calculations, Chatbots A and D again performed best at 86%, whereas B, C, and E ranged from 57–71%. Similar trends were observed for dilution and concentration calculations, while all chatbots performed equally well at 86% for percentage and ratio strength calculations.

Table 2: ANOVA (Two-factor without replication)

Source of variation	SS	df	MS	F	P-value	F crit
Rows (Chatbots)	9.36	4	2.34	9.75	0.000342	3.006917
Columns (Types of calculation)	4.56	4	1.14	4.75	0.01019	3.006917

Table 2 shows a significant difference among chatbots ($F = 9.75$, $p = 0.000342$), indicating that chatbot accuracy varies meaningfully depending on which chatbot is used. Since the F-value exceeds the

critical value ($F_{crit} = 3.0069$), this effect is statistically significant. Similarly, there was also a significant difference across the different types of calculations ($F = 4.75$, $p = 0.01019$).

Table 3: Error pattern analysis

Chatbot	Total errors (n)	Adult dose errors (n, %)	Pediatric dose errors (n, %)	IV/Infusion errors (n, %)	Dilution and Conc errors (n, %)	Ratio strength errors (n, %)
A	4	0 (0%)	1 (25%)	1 (25%)	1 (25%)	1 (25%)
B	10	2 (20%)	3 (30%)	2 (20%)	2 (20%)	1 (10%)
C	11	2 (18%)	3 (27%)	2 (18%)	3 (27%)	1 (10%)
D	5	1 (20%)	1 (20%)	1 (20%)	1 (20%)	1 (20%)
E	11	2 (18%)	3 (27%)	2 (18%)	3 (27%)	1 (10%)

Table 3 shows the error patterns of the chatbots across different calculation types. Chatbot A made 4 errors, mostly in pediatric dosing, IV/infusion, dilution/concentration, and ratio calculations, with no adult dose errors. Chatbots B, C, and E made the most errors (10–11 each), primarily in pediatric dosing and dilution/concentration, while

adult dosing and IV/infusion errors were moderate. Chatbot D made 5 evenly distributed errors across all calculation types. Overall, pediatric dosing and dilution/concentration were the most error-prone areas, whereas ratio/strength calculations had the fewest errors. Table 4. Teachability Assessment

Table 4: Teachability assessment

Chatbot	Incorrect responses selected for feedback (n)	Corrected after feedback (n)	Teachability (%)
A	4	4	100%
B	10	10	100%
C	11	10	91%
D	5	5	100%
E	11	9	82%

Table 4 summarizes the teachability assessment. Chatbots A, B, and D demonstrated complete teachability, correcting 100% of previously incorrect responses after feedback. Chatbot C corrected 91% of errors, and Chatbot E corrected 82%. These findings indicate that AI chatbots can effectively improve performance when provided with corrective feedback, although the extent of improvement varies across different platforms.

DISCUSSION

This study evaluated the accuracy and teachability of five widely accessible AI chatbots in solving pharmaceutical calculations. The overall accuracy varied among chatbots, with Chatbot A demonstrating the highest accuracy (89%) and Chatbot D slightly lower at 86%, whereas Chatbots B, C, and E showed moderate performance (69–71%). These results indicate that while AI chatbots are generally capable of handling pharmaceutical calculations, performance can vary significantly depending on the platform used. Similar findings have been reported in previous studies, where AI-assisted tools exhibited variability in problem-solving accuracy across different clinical and educational tasks [5]. The analysis of accuracy by calculation type revealed that chatbots performed best in percentage and ratio/strength calculations, with all five chatbots achieving 86% accuracy [7]. In contrast, adult and pediatric dose calculations, as well as IV/infusion rate and dilution/concentration problems, demonstrated more variability. For instance, Chatbot A excelled in pediatric dosing with 100% accuracy, whereas B, C, and E scored 71%. These trends suggest that simpler calculations with fewer steps, such as ratio/percentage problems, are handled more reliably by AI chatbots, whereas calculations requiring multiple steps or context-specific interpretation, such as pediatric dosing and infusion rates, are more error-prone [3]. There is a significant difference between the chatbots and in the way they solve different types of calculations, suggesting that each system relies on different underlying models, computation methods, and reasoning processes. This finding aligns with the concept that AI tools, while capable of rapid computation, may be limited in handling multi-step problems that require judgment or contextual understanding [4]. The error pattern analysis further highlighted the specific areas where AI performance was limited. Pediatric dosing and dilution/concentration calculations were the most error-prone across multiple chatbots,

while ratio/strength calculations consistently had the fewest errors. Chatbots B, C, and E exhibited the highest total errors [10–11], particularly in pediatric dosing and dilution/concentration calculations, indicating that complex or multi-step calculations remain challenging for AI systems. These findings are consistent with previous literature suggesting that AI systems may misinterpret intermediate steps or fail to correctly apply unit conversions in pharmaceutical calculations⁽²⁾. Interestingly, Chatbots A and D made fewer errors overall, suggesting that the underlying algorithms or problem-solving models used by these platforms may be more robust or better trained for clinical and educational scenarios [5]. An important aspect of this study was the assessment of teachability. Chatbots demonstrated the ability to learn from feedback, with Chatbots A, B, and D achieving 100% correction of previously incorrect responses, and Chatbots C and E improving to 91% and 82%, respectively. This indicates that AI chatbots are capable of incremental learning and adapting to guidance, which has important implications for both educational and clinical applications. Real-time feedback can potentially enhance chatbot reliability and mitigate errors in future tasks [1]. Moreover, the teachability feature suggests that these systems could be incorporated into pharmacy education as interactive learning tools, providing students with immediate feedback while reinforcing correct problem-solving strategies [7]. Despite the promising findings, several limitations were noted. First, only five AI chatbots were evaluated, which may not fully represent the broader landscape of AI-assisted educational tools. Second, the study focused exclusively on specific types of pharmaceutical calculations; real-world clinical scenarios often require integration of multiple calculations along with patient-specific factors, which may further challenge AI performance [8]. Third, while teachability was assessed using corrective feedback on similar questions, it is unclear whether this learning translates across entirely new problem types or more complex clinical calculations. Future research should explore AI performance in dynamic clinical environments and assess long-term learning capabilities.

Overall, the study demonstrates that AI chatbots can be valuable educational and clinical support tools for pharmaceutical calculations. They are particularly effective for simpler or structured calculations, and their teachability enables progressive improvement. However, careful supervision remains essential,

particularly for complex, multi-step calculations, to prevent potential errors in clinical decision-making. The integration of AI tools in pharmacological education and practice should focus on complementing human judgment rather than replacing it, thereby enhancing learning, efficiency, and patient safety [9].

CONCLUSION

This study demonstrates that AI chatbots can serve as valuable tools for solving pharmaceutical calculations, with varying levels of accuracy depending on the platform and type of calculation. Chatbots performed best in simpler calculations, such as percentage and ratio/strength problems, while more complex tasks, including pediatric dosing, IV/infusion rates, and dilution/concentration calculations, were more prone to errors. Importantly, all chatbots showed the capacity for teachability, improving performance when provided with corrective feedback, highlighting their potential as adaptive educational tools. These findings suggest that AI chatbots can complement traditional pharmacy education and clinical practice, providing rapid, interactive support for learners and healthcare professionals. However, careful supervision remains essential for complex or high-stakes calculations to ensure patient safety. Future research should explore AI integration in dynamic clinical scenarios and assess long-term performance improvement.

FUNDING

Nil

AUTHORS CONTRIBUTIONS

All authors have contributed equally

CONFLICT OF INTERESTS

Declared none

REFERENCES

1. Spence C, Martin C, McCambridge A, McCague PJ. M Pharm students' views and experiences of the teaching of pharmaceutical calculations and their importance in pharmacy practice. *Int J Pharm Pract.* 2024 Nov 1;32(Suppl 2):ii1. doi: [10.1093/ijpp/riae058.011](https://doi.org/10.1093/ijpp/riae058.011).
2. Mulac A, Hagesaether E, Granas AG. Medication dose calculation errors and other numeracy mishaps in hospitals: analysis of the nature and enablers of incident reports. *J Adv Nurs.* 2022 Jan;78(1):224-38. doi: [10.1111/jan.15072](https://doi.org/10.1111/jan.15072), PMID [34632614](https://pubmed.ncbi.nlm.nih.gov/34632614/).
3. Locke M, Suen RM, Williamson AK, Nieto MJ. FIP1L1-PDGFR α clonal hypereosinophilic syndrome with eosinophilic myocarditis and intracardiac thrombus. *Cureus.* 2023;15(8):e43138. doi: [10.7759/cureus.43138](https://doi.org/10.7759/cureus.43138), PMID [37692703](https://pubmed.ncbi.nlm.nih.gov/37692703/).
4. Srinivasan M, Venugopal A, Venkatesan L, Kumar R. Navigating the pedagogical landscape: exploring the implications of AI and chatbots in nursing education. *JMIR Nurs.* 2024 Jun 13;7:e52105. doi: [10.2196/52105](https://doi.org/10.2196/52105), PMID [38870516](https://pubmed.ncbi.nlm.nih.gov/38870516/).
5. Shiferaw MW, Zheng T, Winter A, Mike LA, Chan LN. Assessing the accuracy and quality of artificial intelligence (AI) chatbot-generated responses in making patient-specific drug-therapy and healthcare-related decisions. *BMC Med Inform Decis Mak.* 2024 Dec 24;24(1):404. doi: [10.1186/s12911-024-02824-5](https://doi.org/10.1186/s12911-024-02824-5), PMID [39719573](https://pubmed.ncbi.nlm.nih.gov/39719573/).
6. Campbell N, Kalabalik Hoganson J. Accuracy and teachability of artificial intelligence chatbots in solving pharmaceutical calculations: a descriptive study. *Int J Clin Pharm.* 2025 Aug;47(4):1109-13. doi: [10.1007/s11096-025-01947-7](https://doi.org/10.1007/s11096-025-01947-7), PMID [40493330](https://pubmed.ncbi.nlm.nih.gov/40493330/).
7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019 Jan;25(1):44-56. doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7), PMID [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/).
8. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J.* 2019 Jun;6(2):94-8. doi: [10.7861/futurehosp.6-2-94](https://doi.org/10.7861/futurehosp.6-2-94), PMID [31363513](https://pubmed.ncbi.nlm.nih.gov/31363513/).
9. Vasudevan R, Alqahtani T, Alqahtani S, Devanandan P, Kandasamy G, Saad R. Integrating AI in healthcare education: attitudes of pharmacy students at King Khalid University towards using ChatGPT in clinical decision-making. *Healthcare (Basel).* 2025 May 27;13(11):1265. doi: [10.3390/healthcare13111265](https://doi.org/10.3390/healthcare13111265), PMID [40508877](https://pubmed.ncbi.nlm.nih.gov/40508877/).