Letter to Editor

# Utility of ChatGPT in neuroanesthesia: Still long way to go

**Chandrika Balasubramanian[1], Santha Prem Kumar Edison[1], Krishna Mohan Surapaneni [1]\***

[1]*Panimalar Medical College Hospital & Research Institute, Banaveduthottam, Tamil Nadu, India*

**ABSTRACT**

With the evolution of Artificial Intelligence (AI), complex and intricate medical decisions are expected to benefit from the insights of such AI-powered large language models like ChatGPT. This study was conducted to examine ChatGPT 3.5's performance in neuroanesthesia through ten clinical cases. While the model shows moderate competency in some respects, it could not generate correct responses for over half of the scenarios underscoring the need for specific training and refinement before practical use in specialized clinical decision support.

## 1. Introduction

Dear Editor,

Neuro anesthesiology plays a crucial role in ensuring the safe and effective management of patients undergoing neurosurgical procedures.[1] With the advent of artificial intelligence (AI), particularly language models like ChatGPT 3.5, there is a growing interest in evaluating their potential applications in the medical field.[2,3] This study aims to assess the performance of ChatGPT 3.5 in the context of neuro anesthesia by analyzing its responses to a set of clinical cases sourced from a publicly accessible website.

For this study, ten clinical cases as Multiple Choice Questions (MCQ) were selected from the Neuroanesthesia Quiz section of the Society for Neuroscience in Anesthesiology and Critical Care (SNACC) website.[4] The cases were diverse, covering a range of neurosurgical scenarios to test the versatility and accuracy of ChatGPT. The methodology involved presenting each case to

ChatGPT 3.5 in MCQ format with five options. ChatGPT's responses were then compared to the correct answers provided in the answer key.[4] All responses were generated twice to confirm the answers generated by ChatGPT.

ChatGPT 3.5 demonstrated a moderate level of performance in responding to the neuroanesthesia clinical cases. Out of the ten cases presented, ChatGPT provided correct answers for only four, while the remaining six responses did not match with the key. This suggests a limitation in the model's ability to consistently generate accurate and contextually relevant answers in the domain of neuroanesthesia. Table 1 shows the topics of neuro anesthesiology and the questions in brief which were used to converse with ChatGPT. ChatGPT demonstrated accuracy in topics related to neuroanesthesiology such as position-related complications, and induction agent considerations, postoperative diuresis, fluid and electrolyte management and topical anesthetics, complications in airway management. There were notable inaccuracies in identifying complications during pediatric anesthesia, autonomic responses, cardiac imaging interpretation, air embolism, postoperative complication management,

\* *Corresponding author.*
*E-mail address*: krishnamohan.surapaneni@gmail.com (K. M. Surapaneni).

**Table 1:** Case scenarios inneuro anesthesiology and ChatGPT responses

| S. No | Topic | Scenario on | ChatGPT response |
|---|---|---|---|
| 1. | Neuroanesthesiology, Position-related Complications. | Quadriplegia after suboccipital craniotomy in the sitting position. | Correct |
| 2. | Pediatric Anesthesia, Surgical Complications. | Severe hypotension and hypoxemia during augmentation cystoplasty. | Incorrect |
| 3. | Autonomic Responses, Vagal Stimulation. | Bradycardia and facial flushing during nephrectomy under general anesthesia. | Incorrect |
| 4. | Pediatric Anesthesia, Intraoperative Crisis Management. | Systolic BP drop and EtCO2 decrease during craniectomy in a 4-month-old child. | Incorrect |
| 5. | Cardiac Imaging, Myocardial Perfusion. | Persistent myocardial filling defect at three hours on a dipyridamole-thallium scan. | Incorrect |
| 6. | Induction Agents, Anesthesia in High-Risk Patients. | Induction agent disadvantage in a patient with carotid artery stenosis. | Correct |
| 7. | Postoperative Diuresis, Fluid and Electrolyte Management. | Increased urine output, serum osmolarity, and hematocrit after occipital astrocytoma resection. | Correct |
| 8. | Air Embolism, Postoperative Complications. | Sudden CVP increase and premature ventricular contractions after craniotomy in the sitting position. | Incorrect |
| 9. | Neurosurgical Procedures, Ventriculostomy Care. | Management of ventriculostomy system during transport after aneurysm coiling. | Incorrect |
| 10. | Topical Anesthetics, Complications in Airway Management. | Cyanosis and decreased SpO2 after topical benzocaine for fiberoptic intubation | Correct |

neurosurgical procedures, and ventriculostomy care.

The findings of this study carry significant implications for the potential integration of language models like ChatGPT 3.5 into the complex landscape of medical decision-making, specifically within specialized fields such as neuroanesthesia. While ChatGPT demonstrated a moderate level of competency by providing accurate responses for a few of the cases, the notable limitation of being unable to deliver accurate answers in more than half of the scenarios raises substantial concerns about its reliability, particularly in critical medical situations where precision is paramount.

The inability of ChatGPT 3.5 to consistently generate accurate responses suggests a need for further research and refinement before considering its practical deployment in neuroanesthetic decision-making processes. Neuro anesthesiology involves intricate scenarios with nuances that demand a high level of contextual understanding, and the model's limitations in this regard underscore the complexity of translating language models into reliable tools for specialized medical domains.[5,6]

To enhance ChatGPT's overall performance, future research efforts should focus on targeted training and refinement tailored to the specific challenges posed by neuro anesthesiology cases. This might involve incorporating more recent and diverse medical data, as well as collaborating with domain experts to fine-tune ChatGPT's responses to the intricacies of real-world scenarios. Additionally, exploring methods to provide real-time interaction and feedback mechanisms could contribute to refining the model's accuracy and addressing its limitations in critical medical contexts.

## 2. Conclusion

This study provides an initial assessment of ChatGPT 3.5 in the field of neuro anesthesiology, revealing both its strengths and limitations. While the model demonstrated some capability in answering clinical cases, its accuracy fell short in a significant portion of scenarios. However, this study is subjected to limitations as only 10 clinical cases were used. More comprehensive assessments are required to generalize the findings. Future iterations and improvements in training methodologies may be required to enhance the model's performance and reliability for practical clinical applications in neuro anesthesiology. This research contributes to the ongoing dialogue surrounding the integration of AI in healthcare and underscores the importance of rigorous evaluation before widespread implementation in specialized medical domains.

## 3. Sources of Funding

## 4. Conflict of Interest

None to be declared.

## Acknowledgements

## References

1. Longhitano Y, Zanza C. The Route of Neuro-Critical Care. *Rev Recent Clin Trials*. 2022;17(4):225–6.
2. Liu J, Wang C, Liu S. Utility of ChatGPT in Clinical Practice. *J Med Internet Res*. 2023;25:e48568.
3. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT - Reshaping medical education and clinical management. *Pak J Med Sci*. 2023;39(2):605–7.
4. Society for Neuroscience in Anesthesiology and Critical Care (SNACC). Last accessed on 30-11-23. Available from: https://snacc.org/education-corner/neuroanthesia-quiz/.
5. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare (Basel)*. 2023;11(6):887.
6. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the Feasibility of ChatGPT in Healthcare: An Analysis of Multiple Clinical and Research Scenarios. *J Med Syst*. 2023;47(1):33.

## Author biography

**Chandrika Balasubramanian,** Professor & HoD

**Santha Prem Kumar Edison,** Assistant Professor

**Krishna Mohan Surapaneni,** Vice Principal & Professor
https://orcid.org/0000-0002-5204-5708

**Cite this article:** Balasubramanian C, Edison SPK, Surapaneni KM. Utility of ChatGPT in neuroanesthesia: Still long way to go. *Indian J Clin Anaesth* 2024;11(3):442-444.