iMedPub Journals www.imedpub.com

Journal of Neurology and Neuroscience

2024

Vol.15 No.S9:005

ISSN 2171-6625

# Artificial Intelligence in Neuro-Oncology: A Discussion of ChatGPT's Ability to Diagnose Brain Tumors

Giovanni Kozel<sup>1\*</sup>, Muhammet Enes Gurses<sup>2</sup>, Neslihan Nisa Geçici<sup>3</sup>, Elif Gökalp<sup>4</sup>, Siyar Bahadir<sup>5</sup>, Martin A. Merenzon<sup>6</sup>, Ashish H Shah<sup>2</sup>, Ricardo J Komotar<sup>2</sup>, Michael E Ivan<sup>2</sup>

<sup>1</sup>Warren Alpert Medical School, Brown University, Providence RI, USA

<sup>2</sup>Department of Neurosurgery, Miller School of Medicine, University of Miami, Miami, USA

<sup>3</sup>Hacettepe University School of Medicine, Ankara, Turkey

<sup>4</sup>Ankara University School of Medicine, Ankara, Turkey

<sup>5</sup>Feinstein Institute, New York, USA

<sup>6</sup>Department of Neurosurgery, Yale University School of Medicine, New Haven, CT, USA

\*Corresponding Author: Giovanni Kozel, Warren Alpert Medical School, Brown University, Providence RI, USA; E-mail:giovanni\_kozel@brown.edu

**Received:** 20-May-2024, Manuscript No. IPJNN-24-14896; **Editor assigned:** 23-May-2024, PreQC No. IPJNN-24-14896 (PQ); **Reviewed:** 06-Jun-2024, QC No. IPJNN-24-14896; **Revised:** 13-Jun-2024, Manuscript No. IPJNN-24-14896 (R); **Published:** 20-Jun-2024, DOI: 10.4172/2171-6625.15.S9.005

**Citation:** Kozel G, Gurses ME, Geçici NN, Gökalp E, Bahadir S, et al. (2024) Artificial Intelligence in Neuro-Oncology: A Discussion of ChatGPT's Ability to Diagnose Example Brain Tumors. J Neurol Neurosci Vol. 15 No.S9:005.

### Abstract

The application of artificial intelligence to neurosurgery and neuro-oncology is a rapidly developing field of research, as physicians assess the functionality and security of novel generative programs such as ChatGPT. In a recent study, two versions of ChatGPT were given the specific task of providing diagnoses and treatment plans for example brain tumors. The encouraging technical performance of the chatbot is offset by the lack of sufficient standards and policies in the field of neurosurgery which are needed for the safe and successful assimilation of artificial intelligence.

**Keywords:** Neurosurgery; Artificial intelligence; ChatGPT 3.5, ChatGPT 4, Neurosurgical treatment.

#### About the Study

The release of AI to the public on November 30, 2022, the generative artificial intelligence program ChatGPT has been subject to a large amount of public interest and commercial speculation, as professionals and laypersons attempt to determine how this technology can best be utilized by nearly every sector of public and private life. The field of neurosurgery and the specialty of neuro-oncology are not exempt from this research. It was the goal of the study "ChatGPT on Brain Tumors" to contribute to the growing body of research on generative artificial intelligence's application to this field by examining the abilities of ChatGPT as a potential predictive tool for neuro-oncology.

In this study, twenty unique example cases of brain tumors, seven malignant and thirteen benign, were selected from the

medical literature provided to ChatGPT versions 3.5 and 4, and asked to give a diagnosis and treatment plan in each case [1-3]. The output from both versions of the chatbot was then assessed for accuracy by team members before being referred for independent review by a panel of neurosurgeons from the researchers' home institution. Each reviewing neurosurgeon scored ChatGPT-3.5 and ChatGPT-4's responses on a scale of 1 to 10, with 10 being the highest possible score. The average score for diagnosis and treatment plan across all example tumors was then calculated for ChatGPT-3.5 and ChatGPT-4.

ChatGPT-3.5 correctly diagnosed 65% of the example tumors and provided the correct treatment plan for 10% of the tumors. ChatGPT-4 outperformed the previous version of the program, correctly diagnosing 85% of the example tumors and providing the correct treatment plan for 75% of the tumors. Upon evaluation by the panel of neurosurgeons, ChatGPT-3.5 received an average score of 5.9 out of 10 for its diagnoses and 5.7 for its treatment plans. ChatGPT-4 received an average score of 8.3 for its diagnoses and 8.5 for its treatment plans, a 2.4- and 2.8-point improvement over in ChatPT-3.5 in each of these categories of evaluation. On a paired t-test, both of these differences were found to be statistically significant.

Currently, there are no standards for the accuracy, specificity, or sensitivity of chatbots, large language models, or other artificial intelligence programs in the field of neurosurgery or neuro-oncology. The only firm guideline regarding these programs is from medical journals which prohibit chatbots such as ChatGPT from being used to write any article for publication, with several journals adopting specific guidelines against any such practice by authors [4,5]. A recent article published in February, 2024 in Nature by Wang et al. offers a case study for the use of objective criteria in the application of artificial intelligence to medicine [6]. Their work investigated the use of

Vol.15 No.S9:005

such programs as a cost-effective measure in the screening of diabetic retinopathy cases, calculating that a minimum sensitivity of 88.2% and specificity of 80.4% was needed to achieve cost-savings or cost-effectiveness. On the basis of these requirements, the AI program they were testing failed to meet the standard for utilization in diabetic retinopathy and was not deemed effective for use. While accuracy, sensitivity, and specificity are the major base factors for measuring artificial intelligence programs, there are other factors to consider with such novel technologies. There is the issue of the "artificial hallucination," a problem affecting all currently-available chatbots, in which the program is prone to fabricating answers to prompts out of whole cloth. In some cases, when asked for specific sources for its responses, the program will fabricate nonexistent publications and data as references for itself. These hallucinations are difficult to identify. In one study of 50 fictional abstracts written by ChatGPT, 32% appeared as authentic to human reviewers [7]. While many publishers assure consumers that these issues are being dealt with, these hallucinations present a grave threat to professions, such as medicine, where accuracy is of the utmost importance. There is also the question of data security. ChatGPT, the chatbot used by the most companies in the technical and education industries, may gather and save its user data, including contact details and browser history. Such data harvesting is not acceptable in medical practice, where patient confidentiality is critical. There is already one real-world example of an artificial program violating this privacy-the research company Stability AI designed a deep learning text-to-image software was recently sued for allowing its program to use private medical record photographs in the development of its algorithms, a violation of HIPPA [8].

## Conclusion

The performance of ChatGPT-4, the newest version of the program, is quite strong, demonstrating the potential for generative

artificial intelligence to be incorporated as a predictive tool for neuro-oncology. However, the lack of clear guidelines for implementation at this time, as well as the numerous obstacles challenges which exist against the safety, security, and accuracy of this technology means that ChatGPT is not currently implementable into neuro-oncology.

#### References

- Kozel G, Gurses ME, Gecici NN, Gökalp E, Bahadir S, et al. (2024) Chat-GPT on brain tumors: An examination of artificial intelligence/machine learning's ability to provide diagnoses and treatment plans for example neuro-oncology cases. Clin Neurol Neurosurg 239:108238.
- 2. Citow JS, Spinner RJ, Puffer RC (2019) Neurosurgery oral board review. Thieme.
- 3. Nader R, Sabbagh AJ, Elbabaa SK (2020) Neurosurgery case review: Questions and answers. Thieme.
- Barker FG, Rutka JT (2023) Generative artificial intelligence, chatbots, and the journal of neurosurgery publishing group. J Neurosurg 139:901-903.
- 5. Cifarelli CP, Sheehan JP (2023) Large language model artificial intelligence: The current state and future of ChatGPT in neuro-oncology publishing. Journal of Neuro-Oncology 163:473-474.
- Wang Y, Liu C, Hu W, Luo L, Shi D, et al. (2024) Economic evaluation for medical artificial intelligence: Accuracy vs. *cost*effectiveness in a diabetic retinopathy screening case. NPJ Digit Med 7:43.
- Waisberg E, Ong J, Masalkhi M, Kamran SA, Zaman N, et al. (2023) A new era of artificial intelligence in medicine. Ir J Med Sci 192:3197-3200.
- Singh R, Reardon T, Srinivasan VM, Gottfried O, Bydon M, et al. (2023) Implications and future directions of ChatGPT utilization in neurosurgery. J Neurosurg 139:1487-1489.