Large language models as a diagnostic support tool in neuropathology

Info

PUBLISHED DATE
2024-11-01

JOURNAL
The Journal of Pathology: Clinical Research

LINK
https://doi.org/10.1002/2056-4538.70009

Authors

Katherine J Hewitt Isabella C Wiest

Zunamys I Carrero Laura Bejan Thomas O Millner

Sebastian Brandner Jakob Nikolas Kather

Abstract

The WHO guidelines for classifying central nervous system (CNS) tumours are changing considerably with each release. The classification of CNS tumours is uniquely complex among most other solid tumours as it incorporates not just morphology, but also genetic and epigenetic features. Keeping current with these changes across medical fields can be challenging, even for clinical specialists. Large language models (LLMs) have demonstrated their ability to parse and process complex medical text, but their utility in neuro-oncology has not been systematically tested. We hypothesised that LLMs can effectively diagnose neuro-oncology cases from free-text histopathology reports according to the latest WHO guidelines. To test this hypothesis, we evaluated the performance of ChatGPT-4o, Claude-3.5-sonnet, and Llama3 across 30 challenging neuropathology cases, which each presented a complex mix of morphological and genetic information relevant to the diagnosis. Furthermore, we integrated these models with the latest WHO guidelines through Retrieval-Augmented Generation (RAG) and again assessed their diagnostic accuracy. Our data show that LLMs equipped with RAG, but not without RAG, can accurately diagnose the neuropathological tumour subtype in 90% of the tested cases. This study lays the groundwork for a new generation of computational tools that can assist neuropathologists in their daily reporting practice.