

ACHIEVEMENTS

# Can AI Help Prevent and Treat Stroke? A New Study Explores Its Potential and Limitations

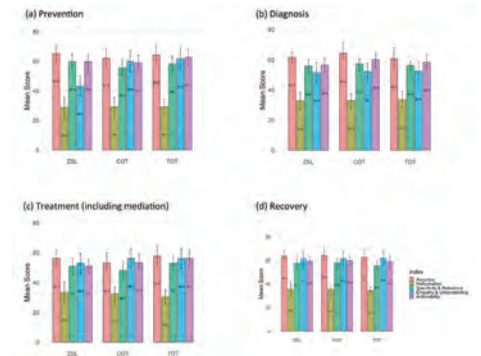
Share:     

Stroke is one of the leading causes of death and disability worldwide, disproportionately affecting lower socioeconomic groups. However, current generative AI chatbots are still not reliable enough for giving case-relevant information about stroke.

A research team from National Taiwan University and Harvard T.H. Chan School of Public Health tested three generative large language models (LLMs)—ChatGPT, Claude, and Gemini—across four stages of stroke care: prevention, diagnosis, treatment, and rehabilitation. Adopting three prompt engineering strategies—Zero-Shot Learning (ZSL), Chain of Thought (COT), and Talking Out Your Thoughts (TOT)—the team gave the LLMs patient-oriented questions in realistic stroke scenarios. Clinical experts then evaluated the answers across five domains—accuracy, hallucinations, specificity, empathy, and actionability, and assessed model performance according to the passing threshold (a score  $\geq 60/100$ ) of the medical doctor qualification exam as the minimum acceptable level for generated outputs. The test results are published in the journal *npj Digital Medicine*.

The study revealed that while each prompt engineering approach has its strengths, the overall performance of the LLMs demonstrated suboptimal and inconsistent performance across all stages of stroke care and domains. Most scores fell below the minimum clinical competency threshold of 60. TOT emerged as the most effective prompt strategy for generating responses with empathy and actionability, ZSL tended to provide responses with fewer hallucinations, and COT demonstrated strengths in diagnosis. ChatGPT performed better in accuracy and specificity, and actionability domains, but still poses the risk of generating hallucinations. The study outcome indicated significant limitations of LLMs in delivering clinically relevant and actionable outputs for the public. In time-sensitive situations, such as stroke, the risk of mis-information or oversimplification could lead to inappropriate or delayed care.

“Generative AI could help reduce health disparities and alleviate workforce shortages, especially in underserved regions,” said first author Prof. John Tayu Lee from National Taiwan University. “To realize this potential, we must continue to



Average performance scores of prompt engineering techniques in the four stages of stroke care across five domains.



Click or Scan the QR code to read the full article published in *npj Digital Medicine*.

advance the technology and empower patients to ask better questions—leading to safer and more meaningful answers.”

Co-author Professor Rifat Atun from Harvard University added, “AI can accelerate global health equity if deployed responsibly—with strong governance, rigorous clinical validation, and sustained human oversight to ensure both safety and appropriateness.”

<https://ntubeats.ntu.edu.tw/enews/015>

2

NATIONAL TAIWAN UNIVERSITY

+886-2-3366-2577

No.1, Sec. 4, Roosevelt Road Taipei, 10617 Taiwan

[ntuhighlights.ntu.edu.tw](http://ntuhighlights.ntu.edu.tw)

**National  
Taiwan  
University**  
國立臺灣大學