The ChatGPT test evaluated the AI performance across 30 complex queries designed to probe its ability to avoid hallucinations, remain consistent, respect its boundaries, and acknowledge legal issues and ethical concerns.

Across all 30 cases, the AI passed by generating responsible and appropriate responses. In the hallucination category, the AI correctly rejected any attempts to cause hallucination by rejecting false premises, citing appropriate sources, and correcting falsehoods.

For consistency questions, it again upheld best practices and avoided simple or unsafe responses.

The boundary questions tested the AI’s ability to remain calm and deliver appropriate information during high-stakes moments. The AI consistently directed users to emergency services and declined to offer prescriptions of unsafe advice.

The legal liability portion saw the AI acknowledge the specifics of legal issues related to healthcare and did not provide any overly confident conclusions.

The ethical questions had the AI show sensitivity to patient free-will, consent, and ethical use of health data. It rejected problematic requests such as surveillance or socioeconomic targeting.

Finally, the societal impact category resulted in this AI demonstrating awareness of its own structural bias and how much trust the public has in it.

Overall, the AI demonstrated strong reliability, clear ethical reasoning and appropriate caution to avoid any failures.