

Clinical Use of AI in Medicine

Week 4 Lead Team

Research Source

- Published February 2025- it's been a year!
- Nature Medicine
 - "Monthly journal, publishing original peer-reviewed research"
 - "Nature Medicine also publishes commissioned content, including News, Reviews and Perspectives"
- Online Observatory of International Research (OOIR)
 - Ranks Nature medicine as 1st in Medicine, Research & Experimental publications
- Google Scholar
 - Rated Nature Medicine the top 4 medical publication in the Health and Medical Sciences (general) category

Author Background

- Led by Ethan Goh, MD and Robert J. Gallo, MD (Stanford Medicine)
- Senior supervision by Jonathan H. Chen, MD, PhD (Stanford) and Adam Rodman, MD (Harvard/BIDMC)
- Large interdisciplinary team, including:
 - Practicing physicians (internal, emergency, family medicine)
 - Experts in clinical reasoning & medical education
 - Biomedical informatics and AI researchers
 - Industry AI expertise (Microsoft)

Summary

- Randomized controlled trial comparing GPT-4 to standard resources for clinical trial experiments
- No increase in harmful decisions



92 physicians completed
expert-designed
management cases



GPT-4 assistance
improved scores
by 6.5%



GPT-4 users spent
more time per
case

Harm Analysis

- Likelihood of Harm: LLM assisted and control physicians groups had similar rates of answers based on the medium or high likelihood of harm.
- Severity of Harm: Mild-to-moderate potential harm occurred at roughly similar rates between the groups. In the case of severe potential harm cases, they presented similar numbers.
- Conclusion: GPT 4 assisted physicians improved scores by 6.5 points compared to conventional resources. This suggests GPT 4 helped make doctors better at case management decisions, improving scores.

Changes Since Publishing

- No published studies have contradicted core finding
 - Growing alignment with later research
- Subsequent studies show LLMs improve:
 - Complex clinical reasoning
 - Decision structuring
- AI increasing reach in health space
 - ChatGPT Health
- The study is now viewed as an early, credible baseline for human–AI collaboration in medicine

Discussion

- Clinical safety, accuracy, and reliability
- Data privacy/security/integrity
- Ethical considerations and bias
- Regulatory & legal liability
- Implementation
- Patient experience
- Declining Clinician Intuition
- Professional Identify

