Nicolas Lescano

Review of article by Partha Pratim Ray

The letter is refreshing in how it cuts through the hype of ChatGPT and gets to the real problems. The model can produce helpful suggestions for clinical decision support, sure, but that usefulness collapses the moment you realize it only knows medicine up to 2021. Healthcare knowledge does not sit still. If your CDS tool cannot keep pace with new guidelines or therapies, what exactly are you supporting? This is less of a technical curiosity and more like selling a GPS that refuses to acknowledge new roads after a certain year.

Evaluation is another weak spot. Having experts rate ChatGPT’s outputs is better than nothing, but it is also subjective and does not touch the real question: do patients benefit? Until studies measure outcomes, satisfaction, or cost-effectiveness, we are essentially grading AI on style points. And then there is the issue of hallucinations. In everyday conversation, a made-up fact is an annoyance; in healthcare, it can be dangerous. If the tool is to be taken seriously, validation against guidelines and expert knowledge cannot be optional.

Ray also raises the subtle but important point of bias. If evaluators know they are reading AI-generated text, their ratings will be skewed, whether consciously or not. A double-blind setup, where the reviewers cannot tell human from AI, is common sense. Without it, the results say more about human perception of AI than about the tool’s actual value.

One thing I would add is that most health systems are already drowning in clinical decision support that is either ignored, clicked through, or outright resented. In that context, the smarter first use of ChatGPT might not be to churn out more alerts but to help trim and simplify the bloated ones we already have. Reducing noise before adding new signals would probably have more impact on safety and usability than layering on yet another round of CDS interventions.

The most useful framing here is that AI is not a replacement for human clinicians but a complement. ChatGPT can throw ideas onto the table, but it is the clinicians who must decide whether those ideas are worth acting on. That relationship is where the real potential lies. The risk is pretending the system is ready to shoulder responsibility without the hard work of integration, updating, and validation.

In other words, this article is not a takedown but a reminder. ChatGPT for CDS might work, but only if paired with constant updating, proper study designs, and a healthy respect for its flaws. Without that, we are just adding another layer of complexity to an already overcomplicated system.