Dr. ChatGPT is Getting Remarkably Good at Diagnosing Health Problems - But Actual Doctors Are Still Better at Weighing Treatment Options

Though AI can present treatment options, it lacks the human touch needed to tailor care to patients.

June 4, 2026 By The Conversation and Andrew Parsons

A father is worried about his toddler, who has been running a fever for two days and pulling at one ear. A 65-year-old woman has been getting winded on her morning walks and feeling more fatigued than usual. Both reach for their phones and type their symptoms into an AI chatbot.

“Your child likely has an ear infection,” the father learns. “Your symptoms could indicate a cardiac condition,” the woman reads.

Those are helpful answers – and there’s a good chance they’re correct. Artificial intelligence is approaching, and in some cases exceeding, doctors’ ability to make accurate diagnoses. An April 2026 study found OpenAI’s o1 model had a 78% accuracy rate on complex diagnostic cases published in the New England Journal of Medicine and also outperformed experienced doctors when diagnosing actual emergency room patients. Similarly, ChatGPT, working on its own, outperformed physicians in diagnosing complex cases, a 2024 study found – even when the physicians were able to use ChatGPT themselves.

Making a correct diagnosis, though, is only half a doctor’s job. The other half is knowing what to do about it – in other words, deciding how to manage a patient’s health condition.

I am a doctor and medical educator studying how doctors make these decisions, a process known as management reasoning, and how doctors in training develop this ability. For clear-cut health concerns, an AI diagnosis may be enough for someone to get the care they need – a little numbing cream for a baby’s gums, say, or an appointment with a cardiologist.

But uncertainty is common in clinical practice. Often, knowing what ails a patient is necessary but not sufficient for determining how to care for them. And how to manage a patient, even after the diagnosis is settled, is a complex question.

Diagnosis categorizes, but management prioritizes

Experienced doctors do not assess each patient from scratch. Over years of practice, they build mental shortcuts called illness scripts.

Illness scripts are more than symptom checklists. They capture what a disease typically looks like, who tends to get it and how it most often progresses. When a doctor sees a new patient, they match what they observe against these mental scripts – a process of categorization and pattern recognition.

When a patient appears with a familiar pattern of signs and symptoms, a doctor calls up the matching mental script almost without thinking. This frees them to notice elements that don't quite align: a symptom that doesn't fit, or a detail in the patient's history – a recent trip abroad, an unusual exposure at work – that points toward a different diagnosis.

It's not surprising that AI is good at this pattern-matching process. Large language models like ChatGPT work in a similar way. They predict what word should come next in a sentence based on patterns learned from enormous amounts of text, including the medical literature. In that literature, the word "pneumonia" reliably follows certain symptom patterns: fever, say, combined with a cloudy patch on a chest X-ray. Pattern matching, at this level, is essentially the same thing a doctor does when fitting a patient's symptoms to an illness script.

But deciding what to do next – what tests to run, what treatments to try, what to monitor and what to follow up on – works differently. Instead of one right answer, a doctor faces multiple reasonable options. The art of medical management is prioritizing which among these options is best for the patient in front of you.

The human advantage

So how does a doctor go from diagnosing a patient to figuring out how best to care for them? The answer is almost always, "It depends."

Consider two men, Marcus and Tomás, both 68, both just diagnosed with early-stage prostate cancer. Their biopsies show the same thing: a slow-growing tumor confined to the prostate.

Both are offered the same two management options. Treat now, with surgery or radiation, accepting the risks of urinary incontinence and changes to sexual function. Or monitor closely with regular tests and biopsies, treating only if it grows. A study that followed more than 82,000 men with early-stage prostate cancer for 15 years found that fewer than 3 in 100 died of their prostate cancer regardless of which path they chose, though men who chose monitoring were about twice as likely to see their cancer spread.

AI can present both options alongside those statistics. What a doctor brings is knowledge of the person sitting across from them.

Marcus has no other significant health conditions. His doctor knows this, and knows Marcus well enough to know that uncertainty sits badly with him. For a patient without other pressing health concerns, a slow-growing tumor has time to progress and become something more serious. Both management paths are genuinely reasonable, but Marcus cannot live with waiting. Knowing cancer is in his body, watched but untreated, is not something he can set aside. He chooses treatment.

Tomás has advanced heart failure, something his doctor has been managing alongside him for years. She knows that his heart condition poses a more immediate threat to his health than this slow-growing tumor does. She knows, too, that he watched a friend go through radiation and come out diminished. Treating aggressively would mean bearing real costs for a benefit that may never arrive. She recommends active surveillance. For Tomás, it is the right answer and a relief.

Different management decisions are the norm in medicine. The right path for any patient depends on who that patient is and what they value, and on a doctor's judgment about where the evidence is reliable and where genuine uncertainty remains.

Judging risk and uncertainty

To decide how to manage a patient's condition, a doctor first considers evidence from the medical literature and then applies the available management options to the patient's particular circumstances. This requires honest communication, shared decision-making, jointly navigating risk and acknowledging uncertainty.

Some risk can be measured. For chest pain, doctors use scoring tools that estimate a patient's short-term likelihood of a heart attack based on their symptoms and test results. AI can likely work through those numbers faster than most doctors.

But risk and uncertainty at the bedside or in the clinic are difficult to measure. Scoring systems and practice guidelines are designed for the average patient – an idealized person, who does not exist. And both doctors' and patients' sense of risk and uncertainty are shaped by their experience. For many patients, this includes a long and justified history of mistrust in the healthcare system.

AI does not know what you have been through or what risk trade-offs you are willing to accept. It cannot acknowledge uncertainty the way a good doctor can, returning to it with you as your circumstances change.

This is where diagnosis and management part ways. The father of the feverish toddler probably got a useful answer: AI has seen enough feverish toddlers in the medical literature to make a reasonable call. But knowing what to do next, including when to stop watching and start worrying, is a conversation best had with your doctor.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

© 2026 Smart + Strong All Rights Reserved.
http://beta.docker.cancerhealth.com/article/dr-chatgpt-getting-remarkably-good-diagnosing-health-problems-actual-doctors-still-better-weighing-treatment-options