

OpenAI Brings Improved Health Responses To Free ChatGPT via @sejournal, @MattGSouthern

OpenAI says GPT-5.5 Instant, the default model for free ChatGPT users, now performs comparably to its frontier Thinking models on health questions. The claim is based on the company's own health evaluations.

Health is one of the categories drawing the most scrutiny over AI-generated answers. For example, a Guardian investigation reported that some Google AI Overviews provided inaccurate medical guidance, and Google later removed AI Overviews for certain medical queries. OpenAI's update lands in that same high-risk category, but with a claim of improvement rather than a retreat.

For publishers and SEOs in health, that means a large, free audience can get medical answers in ChatGPT instead of clicking through to a source.

What OpenAI Reported

OpenAI points to gains on HealthBench and HealthBench Professional, the clinical version. It says GPT-5.5 Instant scores higher than GPT-5.3 Instant, the model it replaced.

The company also reported a drop in factuality problems on live traffic. It says the rate of health responses flagged for at least one possible factuality issue fell 71% over two months. That figure comes from monitors OpenAI runs on production traffic.

OpenAI ran a third comparison against physicians. It asked doctors to write responses to representative health conversations, then had a separate panel of physicians compare those with model responses. In that comparison, the panel rated GPT-5.5 Instant's responses higher than the physician-written ones on criteria including accuracy, communication, and completeness, across 3,500 reviewed responses.

OpenAI says the model showed fewer failure modes than both older models and the physicians. It pointed to fewer cases of missing a red flag or failing to ask the user for more context.

How OpenAI Measured It

HealthBench is a benchmark the company built with its physician network, using doctor-written rubrics rather than exam-style questions.

OpenAI says it works with more than 260 physicians across 60 countries and that doctors have reviewed more than 700,000 example responses to date. The company has cited the 260-physician figure since it launched ChatGPT Health in January. None of the results have been published for outside review.

Health Is Already One Of ChatGPT's Biggest Use Cases

OpenAI has said more than 230 million people ask ChatGPT health and wellness questions each week, one of the most common reasons people use the chatbot.

Health also sits in a protected category in OpenAI's policies. When the company began testing ads in ChatGPT, it said it would not run them in conversations about health, mental health, or politics.

Why This Matters

Medical queries already draw heavy AI-answer exposure, with the highest rate of any category in a recent Ahrefs analysis of Google's AI Overviews. More of that demand moving into ChatGPT's free tier could increase the zero-click pressure on publishers.

The accuracy claims are harder to act on. OpenAI ran the tests in-house, so you face the same measurement gap as with other AI answers in health. The company says its health responses improved, but the claims aren't verified by an independent third-party.

Looking Ahead

The post doesn't specify how changes impact citations. If more platforms shift health answers to free tiers, verifying answers and handling traffic loss become the practitioners' responsibility.

<https://www.searchenginejournal.com/openai-brings-improved-health-responses-to-free-chatgpt/579919/>