QUICK READ · EXIA BRIEF 03 ~4 min read · Version 1.0

Same blood test. Different AI prompt. Different answer?

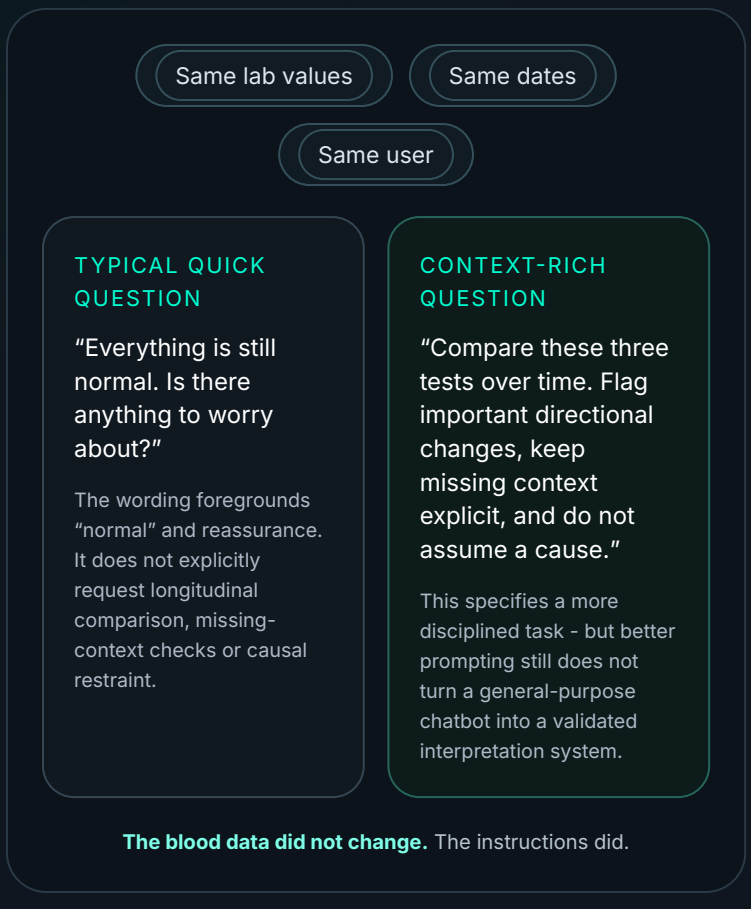
A general-purpose chatbot does not see only your laboratory values. It also sees how you ask, what context you include and - in an ongoing chat - what came before.

The prompts below are illustrative. They are not benchmark outputs from a named AI product, and this Brief does not claim that every model will respond in the same way.

THE SIMPLE IDEA

The data can stay the same while the task changes.

Prompt wording can foreground reassurance, trajectory, missing context or causal restraint. A knowledgeable user may ask for those safeguards explicitly. Most consumers should not have to know the perfect scientific prompt before a health-AI system behaves responsibly. [1-5]



1 · WHAT IS AI ACTUALLY RESPONDING TO?

Your prompt is part of the input.

In patient-facing use, questions are often incomplete, colloquial or framed around a prior belief. Research also shows that multi-turn conversation can influence model behavior in ways that are not obvious from the current question alone. [1-4]

CURRENT PROMPT

What did you ask the model to pay attention to?

"Is this normal?" and "compare the trajectory and tell me what is missing" are different task specifications even when the laboratory data are identical.

Wording shapes the task.

CONVERSATION HISTORY

What assumptions are already sitting in the chat?

A user's belief, an earlier AI hypothesis or an irrelevant prior turn can become additional context for the next response unless its status stays explicit.

History can become hidden input.

Short is not itself the failure mode. A short prompt can be adequate for a bounded question, and a long prompt can still contain wrong assumptions. The reliability issue is whether the system recognizes which context and safeguards the task requires.

Conversation context can help - or contaminate.

Relevant history can make an answer more useful. But multi-turn medical studies have also observed models shifting toward a user's incorrect position, degrading across sequential alternatives, or propagating earlier model errors into later turns. [2-4]

That does not make every history-dependent answer a "hallucination." A broader problem is **provenance collapse**: measured facts, user beliefs and generated hypotheses can start to look like the same kind of truth.

More words ≠ more governance.

A detailed prompt can improve task specification. It cannot, by itself, guarantee source fidelity, longitudinal consistency, causal restraint or safe handling of missing information.

2 · WHERE STATEMENTS COME FROM

Not every statement in a chat has the same status.

Laboratory facts, user-reported context, model hypotheses and external evidence can appear in one conversation. Their provenance should stay visible.

1

LAB FACT

Measured or reported data

Values, units, dates and other recorded laboratory information. These still need provenance and comparability checks.

2

USER CONTEXT

Relevant, but not automatically verified

Symptoms, medication use, lifestyle changes or beliefs supplied by the person can matter while remaining user-reported information.

3

MODEL HYPOTHESIS

An inference is not a patient fact

An earlier AI sentence such as “this may be related to X” should not silently become established history in the next turn.

4

EXTERNAL EVIDENCE

Source-bound support

Scientific evidence should be distinguishable from conversational claims and should fit the specific conclusion being made.

These inputs are not interchangeable. A previous AI suggestion should not silently become a laboratory fact.

3 · WHAT THIS DOES NOT MEAN

Prompt sensitivity is not proof that every AI answer is unreliable.

- It does not mean every short prompt is unsafe.
- It does not mean longer prompts are automatically more accurate.
- It does not mean conversation history is always harmful.
- It does not mean every context-dependent change is a hallucination.
- It does not prove that Sentinel eliminates hallucination or other AI failure modes.

ONE TAKEAWAY

A consumer should not need to know the perfect prompt before a health-AI system behaves responsibly.

Prompting can improve task specification. Governance should decide what the system must check even when the user does not know to ask.

EVIDENCE · KEPT COMPACT

Sources behind this Brief

These sources support the prompt/context and multi-turn reliability propositions used here. They do not show that every chatbot, model version or conversation will fail in the same way.

1 Draelos RL, et al. Large language models provide unsafe answers to patient-posed medical questions. *npj Digital Medicine.* 2026;9:241. Source

2 Kim TM, Luo L, Kim SE, Manrai AK, Topol E, Rajpurkar P. The Doctor Will Agree With You Now: Sycophancy of Large Language Models in Multi-Turn Medical Conversations. *HeaLing 2026.* 2026:19-34. Source

3 Guo KH, et al. Stop Listening to Me! How Multi-turn Conversations Can Degrade LLM Reliability. arXiv:2603.11394v3. 2026. Preprint. Source

4 Munnangi M, Savage S. Evaluating Large Language Models on Misconceptions in Multi-Turn Medical Conversations. arXiv:2607.12884. 2026; accepted to MLHC 2026. Source

5 National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile.* NIST AI 600-1. 2024. Source

Scientific and claims review · 16 Aug 2026 · Version 1.0. Educational and non-diagnostic. Derived from the AI Guide v1.1 prompt/context-provenance section; principal sources rechecked for production. External clinical/laboratory review has not been performed or represented. This Brief is not a benchmark of a named chatbot and does not establish that Sentinel is clinically validated or hallucination-free.