

# AI Clinical Safety Audit

*Prompt-battery testing for health-AI products, before you ship*

## Dr. Tolulope Ajidahun

Licensed Medical Doctor | USMLE Qualified | AI Evaluation & LLM Safety Testing

### What I Do

I run a structured, category-based prompt battery against health-AI products to catch what generic QA misses: unsafe omissions, hallucinated clinical facts, and missed escalation to emergency or professional care. The battery spans drug interactions, red-flag symptoms, mental-health crisis response, pregnancy and pediatric contraindications, and vague-symptom triage, scored against a fixed four-label rubric.

### What a Sample Run Found

A 25-prompt condensed battery run directly against a leading general-purpose model returned 25 of 25 correct or appropriately escalated responses, with 0 unsafe omissions and 0 hallucinations. Independent published research shows this result varies meaningfully across models and configurations, with hallucination rates ranging from about 5% to over 11% in one comparative study, which is exactly what a pre-launch audit of a specific product is built to catch. The full methodology and worked examples are available in the accompanying Portfolio Sample.

Study: [Comparative accuracy and hallucination rates in patient-facing information, PMC](#)

## Engagement Options

Engagement	Scope & Deliverables	Timeline	Investment
<b>Free Look</b>	Send five safety-critical prompt and output pairs. Anything concerning gets flagged at no cost, so the caliber of review can be evaluated before committing to anything.	1-2 Days	<b>\$0</b>
<b>Focused Audit</b>	A 25-prompt battery run directly against a live product, scoped to one risk area (for example, medication safety or crisis response), with a written findings report and risk rating.	5 Days	<b>\$1,000</b>
<b>Full Audit</b>	The complete 50+ prompt battery across all five categories, run directly against the deployed product configuration, with a written report, flagged outputs, severity ratings, and a 30-minute debrief call.	10 Days	<b>\$2,000</b>
<b>Ongoing Monitoring</b>	<b>Safety</b> Quarterly re-runs of the battery as the model, prompts, or retrieval layer change, with a regression report flagging anything that got worse.	Ongoing	<b>\$500 / mo</b>

*Pricing reflects a single physician-credentialed reviewer with direct clinical and AI-evaluation background, positioned against specialist AI-consulting rates rather than enterprise governance-consulting rates.*

### Get Started

Send five prompt and output pairs to [contact@drtolu.com](mailto:contact@drtolu.com) for a free look, or connect on LinkedIn to scope a Focused or Full Audit.