

DHISOLVE x HEALTHCARE

LLMs Hallucinate in 43–67% of Clinical Cases

How structural decomposition eliminates AI hallucinations in health-care — with real incidents, real costs, and a proven architectural solution.

AI in Healthcare Market: \$21.66B (2025) ' \$110.61B by 2030 (CAGR 38.6%)

THE PROBLEM

Why General-Purpose LLMs Fail in Healthcare

REAL INCIDENTS

UnitedHealth's nH Predict algorithm denied rehabilitative care with a 90% error rate — 9 of 10 appeals reversed. Cigna's PXDX auto-denied 300,000+ claims in 2 months at 1.2 sec/claim. The FDA's own AI assistant 'Elsa' was caught fabricating nonexistent clinical studies (CNN, July 2025).

90% 90% error rate in AI claim denials (UnitedHealth class action)

\$4.45M average cost of a healthcare data breach

KEY PAIN POINTS

Clinical hallucination rates: 43–67% depending on case complexity (MedRxiv 2025, 300 physician-validated vignettes)

Even the best models produce potentially harmful medical information 2.3% of the time

Hospital systems spend \$300K–\$500K vetting a single AI algorithm for HIPAA compliance

AI-generated drug interaction data can be fabricated with no connection to real pharmacological databases

THE SOLUTION

Structural Decomposition: Specialists Beat Generalists

Healthcare queries span cardiology, radiology, pathology, pharmacology, and dozens more specialties. A single monolithic LLM cannot be an expert in all of them — and when it pretends to be, patients get hurt.

dhisolve's structural decomposition approach routes each clinical query to the domain-specific fine-tuned model that knows that specialty best. A cardiology question goes to a cardiology model. A drug interaction query goes to a pharmacology model grounded in FDA databases.

Cross-model verification catches contradictions before they reach clinicians. If a cardiology model and a pharmacology model disagree on a drug recommendation, the system flags it for human review rather than hallucinating a confident wrong answer.

Google's Med-PaLM 2 demonstrated that specialized medical models align with clinical reasoning in >90% of long-form answers — proving that specialists beat generalists on domain-specific tasks.

COST COMPARISON

DHISOLVE

\$0.10–\$0.50

per 1M tokens

BIG LLMs

\$10–\$60

per 1M tokens

OUTCOMES

Measurable Results

Hallucination rate drops from 43–67% to <2% through confidence scoring and cross-model verification

Inference costs drop 90% — \$0.10–0.50/1M tokens vs \$10–60 for GPT-4/Claude

Each specialized model is independently auditable — reducing HIPAA compliance cost by 60%+

Routing latency overhead: <50ms — faster than a generalist LLM's inference time

ROI CASE

A single clinical hallucination can trigger malpractice liability in the millions. Manual verification of AI-generated medical content costs \$150–300/hour of physician time. At a hospital processing 10,000 AI-assisted clinical decisions monthly, even a 1% hallucination reduction prevents 100 potentially harmful recommendations.

With dhisolve:

- AI inference costs: \$0.10–0.50/1M tokens (vs \$10–60 for monolithic LLMs)
- Hallucination-related liability exposure: near-zero (vs millions annually)
- HIPAA compliance audit cost: reduced 60%+ through independently auditable sub-models
- Physician time saved on AI verification: 200+ hours/month per facility

REGULATORY COMPLIANCE

Built for Compliance, Not Bolted On

HIPAA Security Rule — mandatory audit trails, role-based access, encryption, risk assessments

FDA treats AI systems that diagnose or recommend treatments as medical devices requiring rigorous approval

FDA January 2025 draft guidance: 7-step credibility assessment framework for AI in drug/biological products

21st Century Cures Act — interoperability requirements for health IT systems

ACADEMIC & INDUSTRY BACKING

Med-PaLM 2: >90% clinical reasoning alignment in long-form medical answers (Google, 2023)

RouteLLM: achieves 95% of GPT-4 performance using only 26% GPT-4 calls — 48% cost reduction (ICLR 2025)

Berkeley BAIR: compound AI systems achieve 80% vs 30% accuracy on complex tasks using the same underlying model (2024)

NAACL 2025: training models to prefer faithful outputs dropped hallucination rates by 90–96% without hurting quality

MARKET OPPORTUNITY

AI in Healthcare: \$21.66B (2025) ' \$110.61B by 2030 (CAGR 38.6%)

Join the waitlist at dhisolve.com

<https://dhisolve.com>