
WHITE PAPER
Beyond legacy care:
How safe medical AI is reshaping the front door for health plans
Counsel

Beyond legacy care: How safe medical AI is reshaping the front door for health plans

According to physician-led evaluations of publicly available AI chatbots on common medical questions, unsafe or potentially harmful responses occurred in up to ~43% of cases for some models.[1] Considering that 8 out of 10 of Americans reportedly turn to search engines or general-purpose AI for medical questions, there is a significant risk that patients are receiving unsafe guidance.[2]

This emerging reality exposes a critical fault line in how patients access care today: an increasingly large proportion of them are turning to AI for medical guidance, even though current systems, mostly consumer tools, can provide inaccurate or unsafe advice.

For decades, healthcare was designed around the assumption that demand would be episodic, predictable, and manageable by human workflows. The reality is much different. Patients seek care that is continuous, personalized, and available when they need it most.

Today, health plans face a growing gap between these realities, leading to increased utilization of care, avoidable emergency department visits, accelerated disease progression, and higher total cost of care. Patients wait weeks for appointments, default to inappropriate care settings when guidance is unclear, and bounce between disconnected touchpoints. The result is not just a poor experience, but a systemic failure in healthcare.

While virtual care offers a narrow-scope solution, it still over-escalates and under-triages, and provides care experiences that feel transactional and reactive as they lack full context on each patient. These legacy front doors to care were designed for a different era, and remain constrained by a significant limiting factor: the growing physician shortage.

For health plans, reimagining the front door through the adoption of AI offers a new paradigm. AI-native care models, particularly those that pair medical AI with physicians to augment clinical efficiency, enable payers to safely expand access at scale.

In this white paper, we'll examine

- Limitations of legacy care models
- External factors driving the adoption of healthcare AI
- Considerations for adopting medical AI to improve access
- Defining the path to a modern, AI-native front door to healthcare

Limitations of legacy care models

Legacy care models, such as telehealth, nurse lines, or concierge primary or urgent care services, often leave patients with a fragmented experience and lead to clinical decisions that lack context, resulting in higher costs for payers.

Fragment patient experience

Legacy care models assume that episodes of illness are discrete events that can be resolved in isolation. In reality, patients' needs are continuous, context-dependent, and nonlinear. When care is deployed around standalone encounters, clinical context is lost, leading to repeated assessments, redundant diagnostics, and inconsistent clinical decisions. Patients are left bouncing between care settings, from urgent care to specialists to emergency departments (EDs).

In the U.S. alone, non-urgent and potentially unnecessary ED visits remain high, with many studies estimating that up to 60% of them fall into this category.[3] This misalignment is expensive, costing the healthcare system $80 billion in 2021 from treat-and-release ED visits alone, an average of approximately $750 per visit.[4]

These figures represent more than inefficiency. They reflect missed opportunities for appropriate upstream care that could be delivered in outpatient or managed settings.

Interoperability gaps degrade clinical efficiency

Fragmentation of clinical data across health systems, health plans, providers, and digital tools limits the longitudinal view of a patient's health. When a care setting lacks real-time access to medical history, medications, prior diagnoses, lab testing, or care plans, clinical decision-making becomes less effective due to inherent blind spots.

While stakeholders across the healthcare ecosystem have developed integrations with HIEs and EHRs, or adopted interoperability tools, care delivery remains bound by human physicians who are limited in their ability to retain context across an ever-changing clinical landscape.

Recent research shows that triage accuracy varies widely even in structured settings. Many digital symptom checkers achieve an average sensitivity of approximately 50%, compared to 10–19% for over-triage, and 13–19% for under-triage in primary care settings.[5]

The direct cost of legacy models

The financial impact of not modernizing a payer's front door to care is significant. Studies suggest that diverting non-urgent visits to appropriate outpatient care could yield savings in the billions annually.[6]

The average cost of an emergency department visit far exceeds comparable care settings. Treating the same clinical issues in urgent care or outpatient settings can be 3x to 10x cheaper than in the ED.[7] When multiplied across millions of avoidable encounters, the economic case for a more effective front door becomes self-evident.

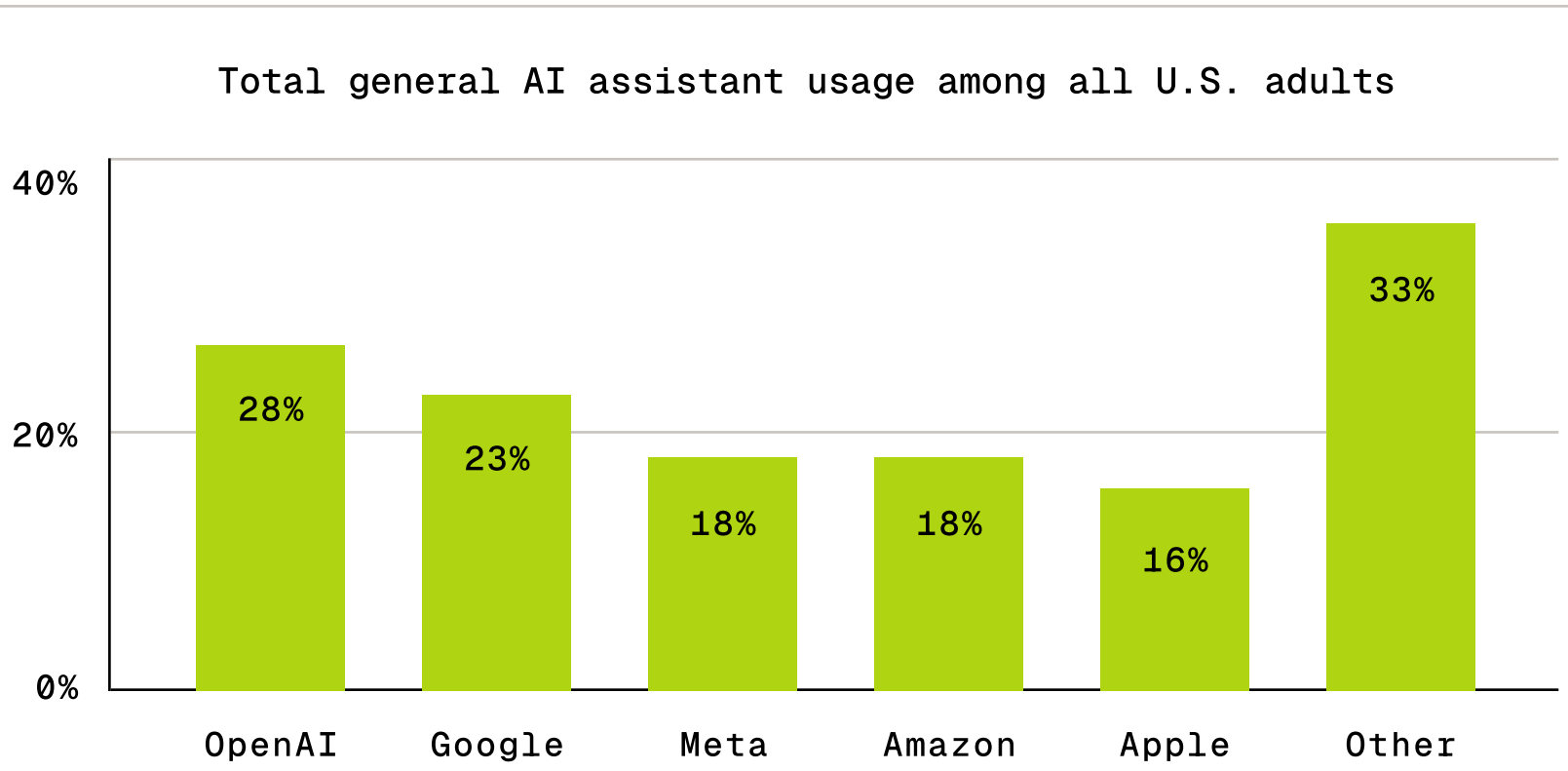
External factors driving the adoption of healthcare AI

Consumer behavior has shifted rapidly toward the use of general-purpose AI tools. A recent report from *OpenAI* showed more than 230 million people turn to ChatGPT weekly for health-related information,[8] accelerating expectations around speed, availability, and digital access across industries, including healthcare.

At the same time, access to care is becoming more constrained. Growing physician shortages are widening the gap between demand and available clinical capacity, placing additional pressure on already strained care models. Adding to these challenges is the emerging evidence that shows that general-purpose AI, while widely adopted, exhibits meaningful accuracy and reliability limitations when applied to medical decision-making.

Consumer adoption of general-purpose AI

Consumers have rapidly adopted AI tools. A recent report from *Menlo Ventures* found that 61% of American adults have used AI in the past six months. General usage by large language model (LLM) is as follows[9]:



While this data is not specific to health information, the scale of overall adoption, combined with how consumers currently seek information (71% via search engines vs. 20% via AI), underscores the growing risk of unsupervised AI being used as a front door to care.

How patients seek information today

71%

Percentage of patients that turn to search engines

20%

Percentage of patients that turn to AI

Growing physician shortages
According to a report from the *Association of American Medical Colleges (AAMC)*, the U.S. is projected to face a physician shortage of up to 86,000 by 2036.[10]

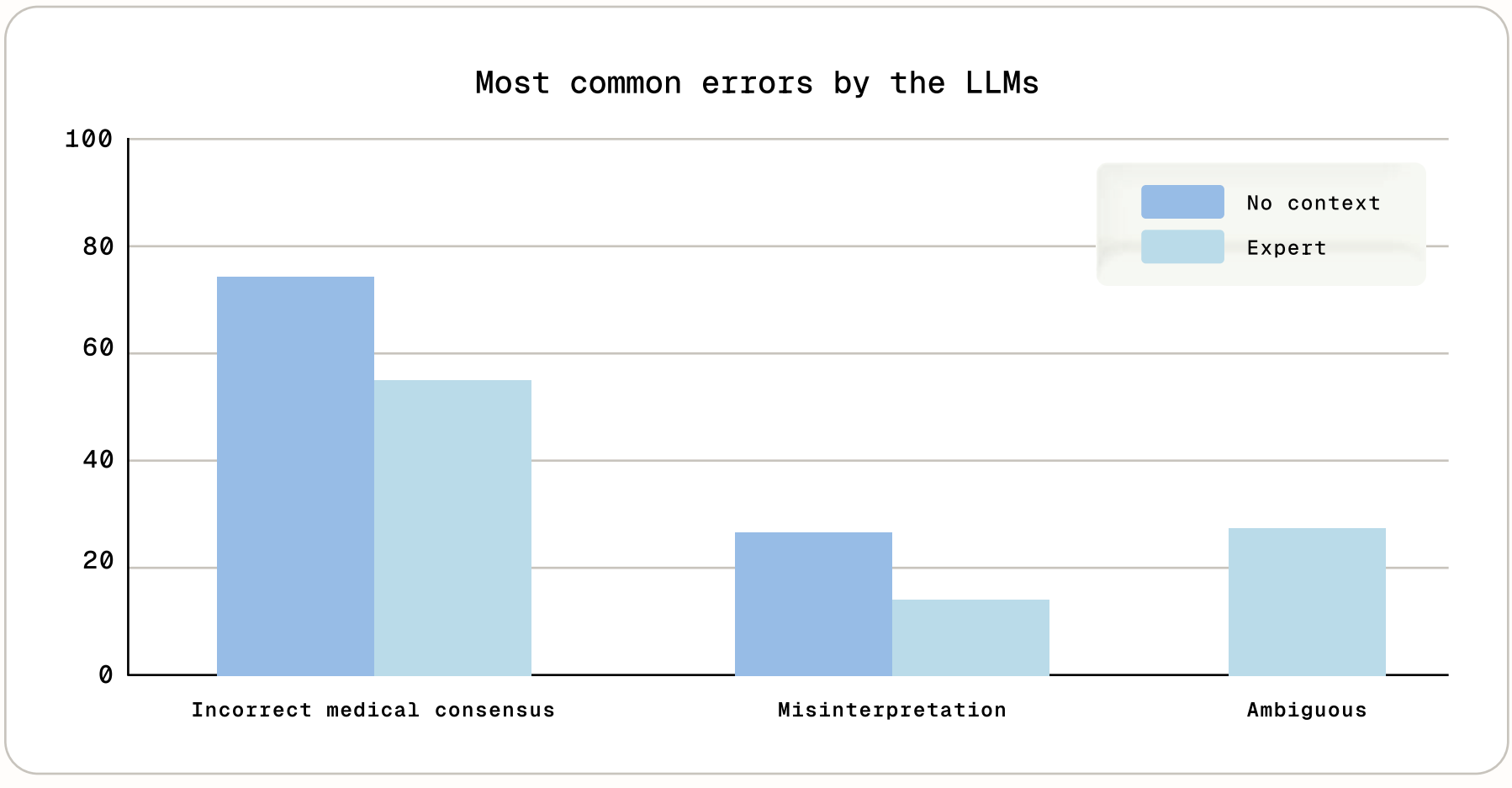
The report also outlines that as many as 74 million Americans lack access to basic healthcare. For health plans, this widening supply–demand gap makes it increasingly difficult to rely on traditional, physician-led care models to meet patient needs, accelerating the urgency for scalable approaches that expand clinical capacity without compromising care quality.

86K+
Projected physician shortage by 2036

74M+
Americans who lack access to basic healthcare

Accuracy of general-purpose AI in healthcare
According to a study published in *Nature*, LLMs provided incorrect answers in more than 30% of cases, with response quality shown to be highly sensitive to how prompts were constructed.[11]

In this study, prompts were categorized as either "no context" or "expert." "No context" prompts consisted solely of a health question, while "expert" prompts included contextual instructions intended to guide LLMs toward trusted clinical sources. Depending on the prompt category, the distribution of common response errors varied significantly, as shown below:



For health plans, this variability highlights why general-purpose AI, regardless of prompt quality, cannot be trusted as the first point of care for patients, reinforcing the need for purpose-built medical AI designed for safety, consistency, and accountability.

Considerations for adopting medical AI to improve access

As health plans accelerate adoption of AI-native care models, the outstanding question is what constitutes “safe” and “responsible” use of AI. Regulators, policymakers, and stakeholders across the ecosystem are increasingly aligned on one principle: AI systems involved in medical decision-making must be governed as clinical infrastructure, not consumer technology.

Healthcare AI expert Saurabh Gombar recently shared that “chatbots rely heavily on disclaimers that say their outputs should not be considered medical advice.”[12] More should be demanded from patient-facing medical AI than the consumer models accept, similar to how traditional providers are highly-regulated and bound by accountability measures.

While progress on standardized policies to evaluate, deploy, and oversee AI as a front door to care has been slow, guidance and accountability frameworks have emerged from coalition groups.

To effectively evaluate AI-native care solutions, payers must keep these guidelines in mind while also considering a vendor’s approach to agentic frameworks.

Emerging regulatory and governance guidelines

Scrutiny of AI in healthcare has intensified, leading top healthcare organizations to begin publishing guidance that clarifies expectations for the responsible use of AI in clinical environments.

The Joint Commission

Notably, The Joint Commission (TJC) has expanded its focus on health IT and clinical decision support to include AI-enabled systems, emphasizing the need for governance, validation, transparency, and human oversight when technology influences patient care. While TJC does not certify algorithms themselves, its standards reinforce a critical expectation for health plans and providers alike: systems that shape clinical decisions must be auditable, monitored, and accountable.

Coalition for Health AI

Similarly, the Coalition for Health AI (CHAI), a multi-stakeholder consortium including health systems, payers, technology companies, and policymakers, has published principles and frameworks focused on trustworthy AI in healthcare. CHAI’s work emphasizes fairness, safety, transparency, and real-world performance monitoring, underscoring that responsible AI adoption is not solely a technical challenge, but an operational and governance one.

Due to unclear guidelines on who is accountable for the responsible use of medical AI, these two entities have partnered to publish guidelines, including[13]:

AI policies and governance

- Require expertise and cross-functional representation.
- Oversee selection, life cycle, compliance, and risk management of AI tools.
- Report AI activities and adverse events to the governing board.

Data security and data use protections

- Encrypt data in transit and at rest.
- Apply strict access controls and audit logs.
- Perform regular security assessments and maintain an incident response plan.

Ongoing quality monitoring

- Require vendors to supply validation + bias-testing data.
- Continuously test, validate, and monitor AI performance and bias after deployment.
- Scale monitoring frequency of risk proximity to patient care.

These guidelines indicate a critical takeaway

Medical AI solutions will increasingly be evaluated through the same lens as other clinical infrastructure, with expectations around oversight, risk management, and continuous performance evaluation.

To function as a reliable front door to care, medical AI must meet a higher standard than consumer tools. Based on emerging regulatory expectations and industry guidance, including frameworks advanced by *TJC* and *CHAI*, safe medical AI requires several foundational components:

- **Physician-supervised architectures:** AI systems must operate within a framework where licensed clinicians retain oversight, authority, and accountability. This includes clearly defined escalation thresholds, review mechanisms, and the ability for clinicians to intervene when needed.
- **Constrained clinical domains:** Safe systems are explicit about what they can and cannot do. Narrowly defined scopes of practice reduce hallucinations, improve reliability, and ensure alignment with clinical protocols and standards of care.
- **Evidence-based reasoning and traceability:** AI outputs should be grounded in established clinical guidelines and supported by transparent reasoning paths. Traceability is essential, not only for regulatory audits, but also for clinician trust and continuous improvement.
- **Deterministic escalation logic:** Uncertainty should not default to blanket reassurance or excessive escalation. Safe medical AI must define when confidence thresholds are met, when additional information is required, and when escalation to a human clinician is mandatory.
- **Continuous monitoring and governance:** Safety does not end at deployment. Ongoing performance monitoring, clinician feedback loops, and outcome tracking are required to detect drift, bias, or emerging failure modes over time, an expectation increasingly reflected in both accreditation standards and industry guidance.

As these expectations become table stakes, health plans will need to prioritize medical AI solutions that adhere to these governance and safety principles.

The role of agentic frameworks in safe care delivery

One of the most important architectural advances enabling safe medical AI is the shift toward agentic frameworks. Rather than relying on a single, monolithic model, agentic systems decompose clinical workflows into specialized agents, each designed for a specific task such as intake, risk stratification, triage, follow-up, or care coordination.

This design closely aligns with emerging expectations emphasized by *CHAI*, particularly around transparency, controllability, and task-specific accountability. For health plans, agentic architectures offer several advantages:

- **Modularity:** Individual agents can be evaluated, constrained, or updated independently
- **Controllability:** Guardrails can be tailored to each agent's role and risk profile
- **Transparency:** Decision pathways are easier to audit and explain
- **Scalability with safety:** Complex workflows can be orchestrated without relying on a single model to reason across all dimensions of care

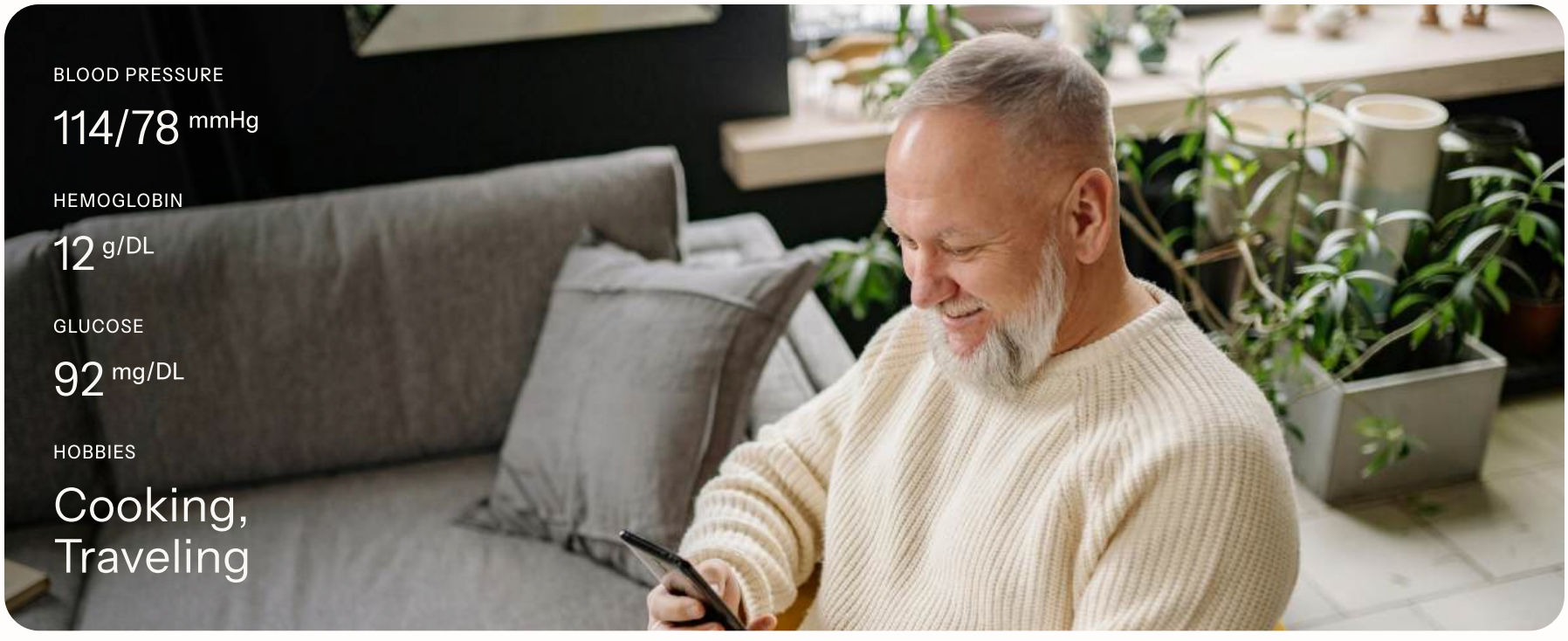
Agentic frameworks transform AI from a black box into a governed clinical system, making them far better suited for payer environments.

Guardrails are infrastructure, not features

A common misconception is that safety can be layered on after deployment. In practice, organizations like *TJC* and *CHAI* reinforce that safety must be embedded directly into system design. Effective guardrails include:

- Protocol-driven constraints aligned with clinical guidelines
- Hard stops for high-risk scenarios
- Mandatory escalation pathways
- Comprehensive logging and audit trails
- Clear delineation of AI versus clinician responsibility

When guardrails are treated as infrastructure rather than features, AI systems become predictable, inspectable, and trustworthy, qualities essential for any front door guiding patient care.



Defining the path to a modern, AI-native front door to healthcare

Meet Counsel, an AI-native primary care platform delivering care through a messaging-based chat experience. Designed to become the safe, modern front door to healthcare, Counsel helps payers expand access while meeting high clinical, regulatory, and operational standards.

Responsible by design

Every interaction is safeguarded by physician supervision and audit-ready guardrails that align with payer compliance, clinical risk, and governance standards.

Member-preferred access

With an average two-minute physician response time and always-on chat, Counsel resolves most concerns at first contact and appropriately steers members to in-network care and plan-aligned resources.

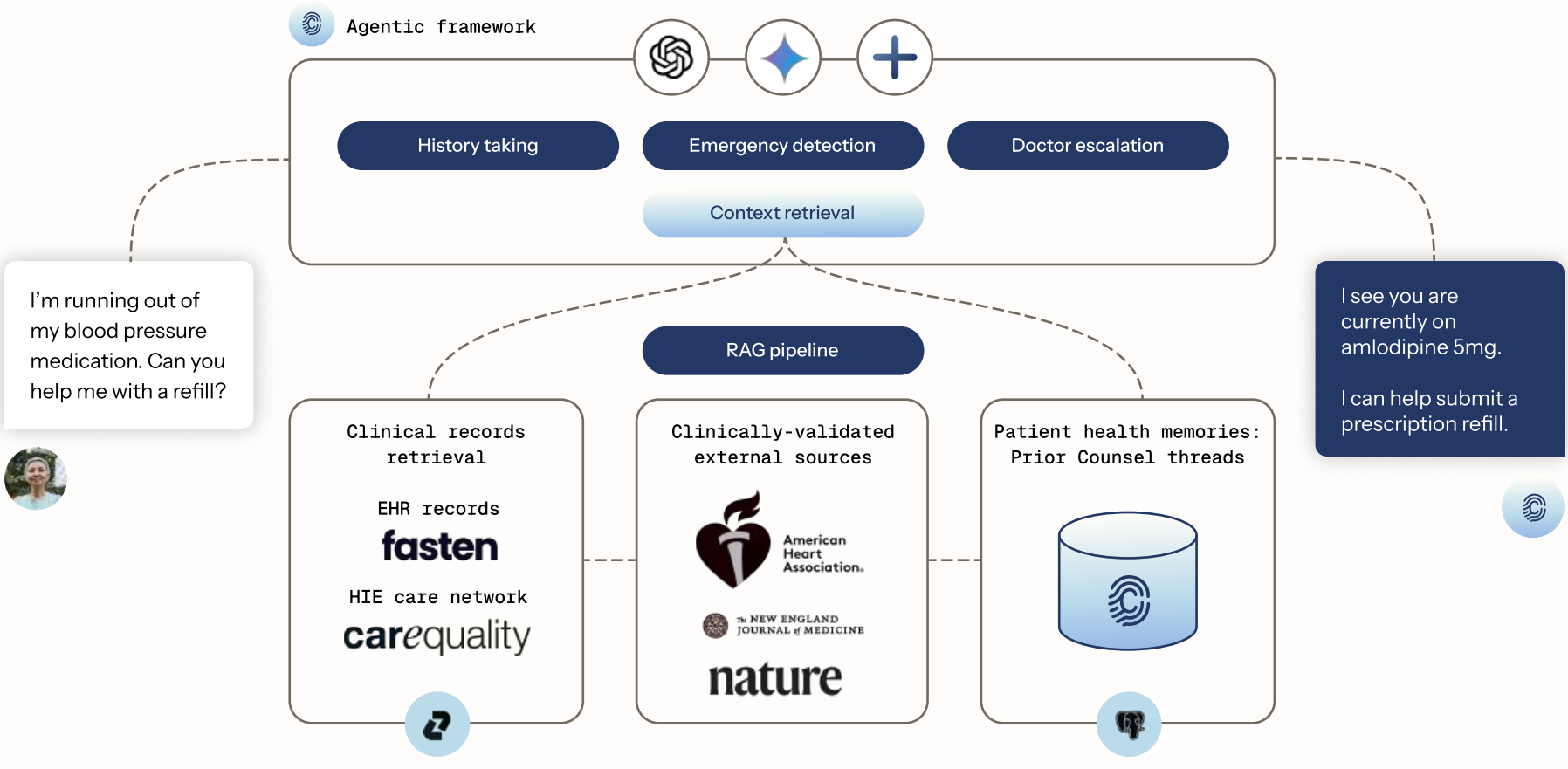
Continuous & context-aware

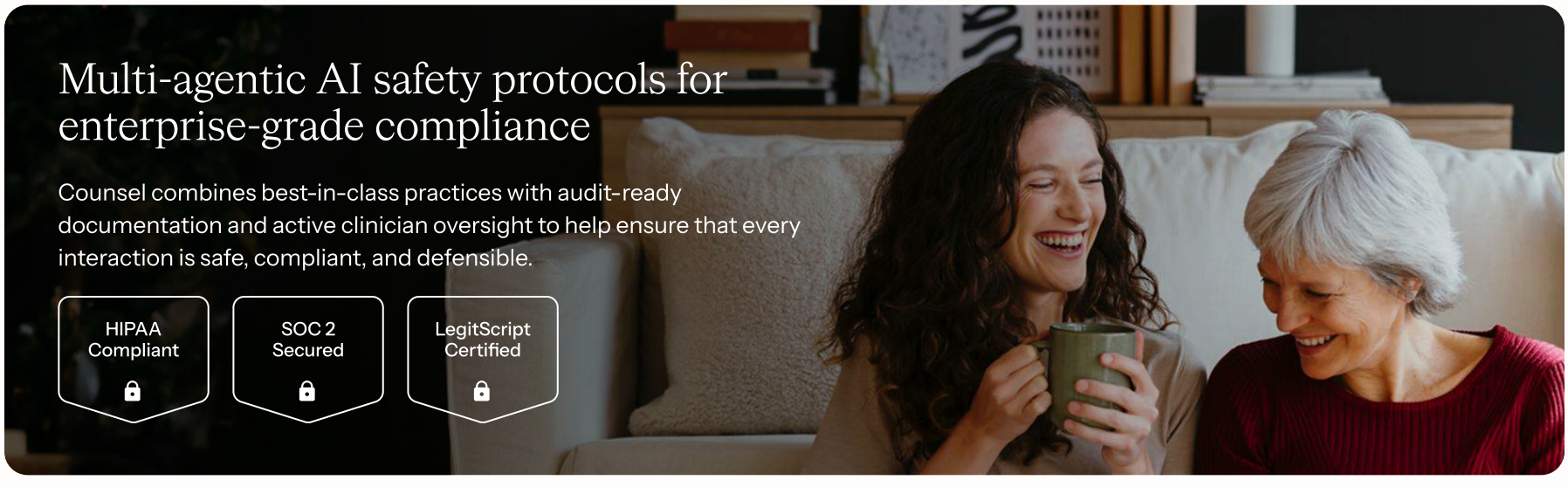
Each interaction builds on the last. By maintaining longitudinal clinical context, Counsel supports ongoing relationships that improve member health, reduce fragmentation, and lower avoidable downstream spend.

Advanced agentic framework and RAG pipeline

Counsel's platform operates with multiple independent agents in every interaction. This agentic framework is built upon third-party models trained individually by their provider. Counsel does not contribute any additional training data to these base models.

To provide patients with highly personalized care, Counsel's context-retrieval agent serves as the orchestrator of our RAG pipeline. This pipeline pulls relevant medical records, the latest evidence-based research from vetted clinical sources, and health memories.





The new standard for safe, connected care

Counsel operationalizes what health plans should demand from a modern AI front door:

- **HIE and EHR connectivity:** Counsel integrates health information exchanges and electronic health records to ensure every encounter reflects a complete clinical history, across care settings and time.
- **Plan-aligned navigation:** Members are steered to preferred networks and point solutions. Counsel intelligently guides members toward covered benefits, in-network providers, and partner programs to amplify engagement and maximize total cost-of-care performance for payer organizations.
- **White-labeled embedded deployment:** Counsel fits into any portal or member apps, enabling brand continuity while extending clinical workflows and enterprise healthcare platform capabilities to any organization.
- **Evidence-based intelligence:** Counsel's AI is trained and reinforced exclusively on peer-reviewed guidelines, green-listed sources, and clinically validated literature, ensuring accuracy, relevance, and adherence to regulatory compliance and ethical standards.
- **Continuous AI improvement:** Through physician feedback and few-shot prompting, Counsel's medical AI gets validated for accuracy, safety, and alignment with clinical best practices.

Safer care. Real health outcomes. Impactful cost savings.

96%

Resolution rate

Nearly all member concerns resolved without escalation

4.9/5

Member satisfaction

Personalized experience secures long-term engagement

$381

Annual savings

Per-engaged member based on care avoided

24%

Reduction in ER visits

Compared with consumer-facing AI tools

~2 min

Response time

When physician assistance is required

Reimagining healthcare for the modern era

Legacy care models were not designed for today's realities. Transactional interactions, fragmented data, and limited clinical context continue to drive over-escalation, under-triage, and avoidable downstream spend. At the same time, the rapid adoption of consumer AI tools has made it clear that AI is shaping how patients seek care, making payer adoption critical to ensuring its safe use.

Medical AI solutions like Counsel offer a path forward. When designed with a robust clinical infrastructure, including physician supervision and a secure agentic framework that preserves context, access to high-quality care can be delivered at scale.

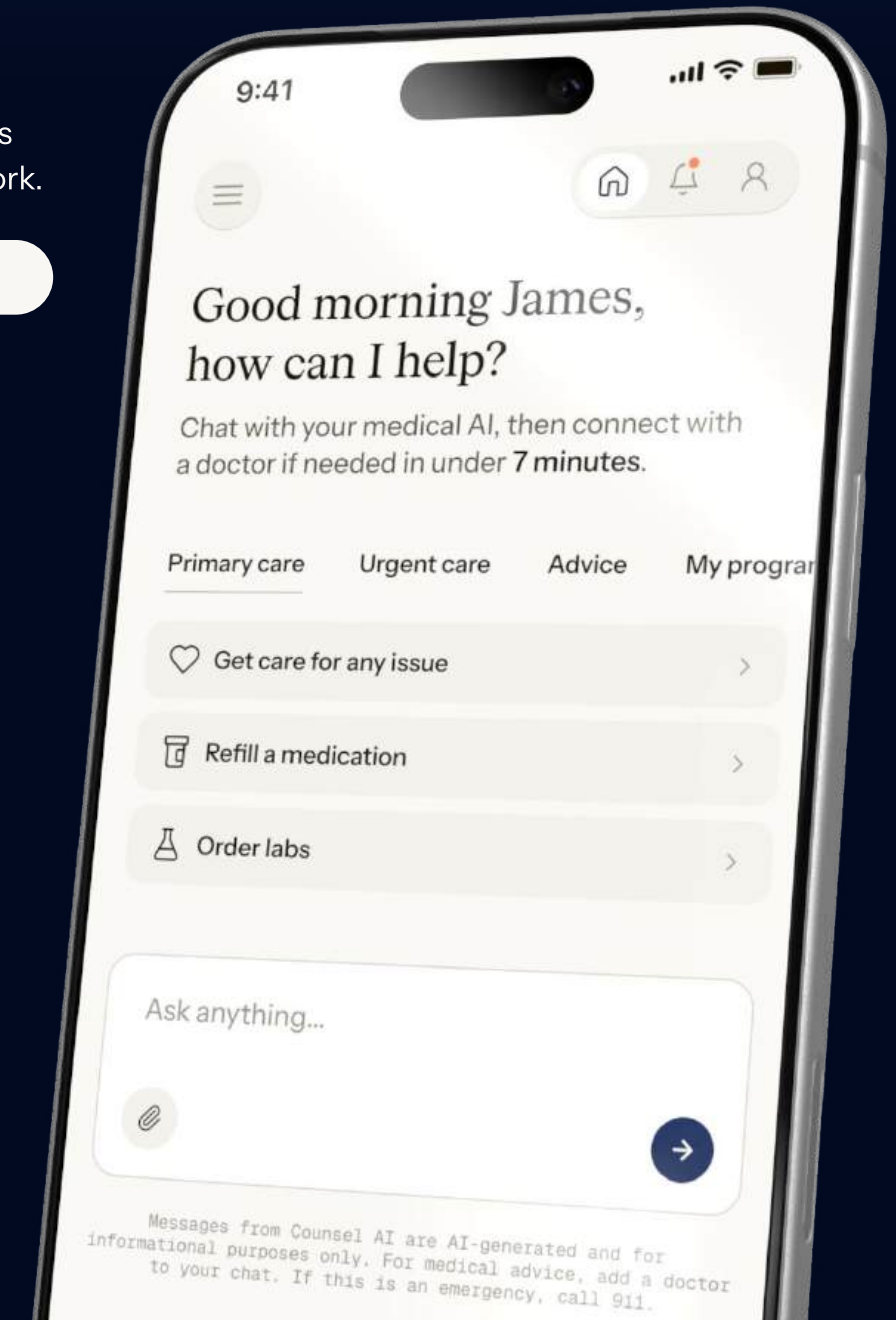
Counsel demonstrates that the future is possible today. Let's get started.

Request a demo

See how Counsel's responsible medical AI can improve access and lower costs, creating measurable value across your network.

counselhealth.com/contact-sales | business@counselhealth.com

Try Counsel now
Scan the QR code to download the app and start your first chat.



Sources

[1] Draelos, R. L., Afreen, S., Blasko, B., Brazile, T. L., Chase, N., Desai, D. P., Evert, J., Gardner, H. L., Herrmann, L., Vaikom House, A., Kass, S., Kavan, M., Khemani, K., Koire, A., McDonald, L. M., Rabeeah, Z., & Shah, A. (2025). *Large language models provide unsafe answers to patient-posed medical questions* (arXiv:2507.18905). arXiv. https://arxiv.org/abs/2507.18905

[2] Annenberg Public Policy Center. (2025). *Many in U.S. consider AI-generated health information useful and reliable*. https://www.annenbergpublicpolicycenter.org/many-in-u-s-consider-ai-generated-health-information-useful-and-reliable/

[3] National Committee for Quality Assurance. (2024). *Emergency department utilization (EDU)*. https://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality-report/emergency-department-utilization-edu/

[4] Agency for Healthcare Research and Quality. (2024). *Emergency department visit costs in 2021 (Statistical Brief No. 311).* U.S. Department of Health and Human Services. https://hcup-us.ahrq.gov/reports/statbriefs/sb311-ED-visit-costs-2021.pdf

[5] Nguyen, H., Meczner, A., Burslam-Dawe, K., & Hayhoe, B. (2022). *Triage errors in primary and pre–primary care.* Journal of Medical Internet Research, 24(6), e37209. https://www.jmir.org/2022/6/e37209/

[6] Connect America. (2024). *Preventing unnecessary ED utilization and hospitalizations with PERS CareSage analytics.* https://www.connectamerica.com/insights/preventing-unnecessary-ed-utilization-and-hospitalizations-with-pers-caresage-analytics/

[7] Illustra Health. (2024). *Reducing avoidable emergency department utilization in value-based care.* https://illustra.health/blog/reducing-avoidable-emergency-department-utilization-in-value-based-care

[8] TechCrunch. (2026, January 7). *OpenAI unveils ChatGPT Health, says 230 million users ask about health each week.* https://techcrunch.com/2026/01/07/openai-unveils-chatgpt-health-says-230-million-users-ask-about-health-each-week/

[9] Menlo Ventures. (2025). *The state of consumer AI: 2025.* https://menlovc.com/perspective/2025-the-state-of-consumer-ai/

[10] American Association of Medical Colleges. (2024). *The complexities of physician supply and demand: Projections from 2022 to 2037.* https://www.aamc.org/media/75236/download?attachment

[11] npj Digital Medicine. (2025). *Evaluating search engines and large language models for answering health questions.* npj Digital Medicine, 8, Article 153. https://www.nature.com/articles/s41746-025-01546-w

[12] MedCity News. (2026, January). *OpenAI, Anthropic expand into healthcare.* https://medcitynews.com/2026/01/openai-anthropic-healthcare/

[13] JAMA Network. (2025). *New Guidance on Responsible Use of AI.* JAMA. https://jamanetwork.com/journals/jama/fullarticle/2842278