

SIGNAL 1

Building a Robust AI Monitoring Program in Healthcare

The Promise and the Problem

Artificial intelligence is no longer confined to pilot projects or innovation labs. Across health systems, AI is woven into everyday workflows—screening patients, triaging cases, drafting documentation, predicting discharges, and optimizing resources. The promise is enormous: better outcomes, faster workflows, lower costs.

But with adoption comes risk. Unlike traditional software, AI does not fail in linear, predictable ways. Rules-based programs either work or they don't. AI models, by contrast, can drift from their original accuracy, degrade silently, or generate errors no one anticipated. A deterioration model may miss certain cases when patient demographics shift. A generative tool may confidently insert lab results that were never ordered. Unlike a coding error, which can be patched and resolved, these risks evolve over time and demand continuous oversight.

This is why health systems must approach AI differently—and why monitoring is not a “nice to have” but the foundation of safe, scalable adoption.

Why AI Is Different

The defining feature of AI is that it is **probabilistic, not deterministic**. Every output carries a degree of uncertainty. Some level of error is not only possible, but inevitable. That reality makes ongoing scrutiny essential.

| Traditional Software | AI/ML Models |
|-----------------------------|----------------------------------|
| Deterministic – predictable | Probabilistic - uncertain |
| Fixed outputs | Variable outputs |
| Errors are static | Errors evolve over time |
| Test once and deploy | Must monitor continuously |

Compounding the challenge is **performance drift**. A model that worked well on last year’s data may falter when applied to this year’s patient population, workflows, or operational environment. Shifts in coding practices, new treatment protocols, or changes in case mix can all erode accuracy.

Finally, AI introduces **new classes of failure** that traditional software simply does not. Large language models may hallucinate plausible but false information. Agentic systems may misinterpret instructions or take unintended actions. These mistakes cannot always be anticipated or prevented at the design stage—they must be caught and corrected in the real world.

Taken together, these characteristics demand monitoring practices that are ongoing, adaptive, and deeply tied to governance.

The Limits of Today’s Monitoring

Health systems are accustomed to tracking adoption and impact when they deploy digital tools. They measure **usage**—who is logging in—and **impact**—whether a workflow appears faster. These measures are useful, but they are insufficient for AI.

Neither usage nor impact metrics reveal whether a model remains accurate, fair, and safe to use. An AI tool may be widely used and save time yet steadily produce errors or biases that put patients or the organization at risk. Unlike traditional digital solutions, AI requires performance-oriented monitoring that does not stop once the go-live date has passed.

Monitoring AI is not simply a matter of good IT hygiene. It is a matter of clinical safety, organizational risk management, user trust, and value realization.

AI requires performance-oriented monitoring that does not stop once the go-live date has passed.

This raises an important question: whose job is it to monitor AI—the vendor, the health system, or both?

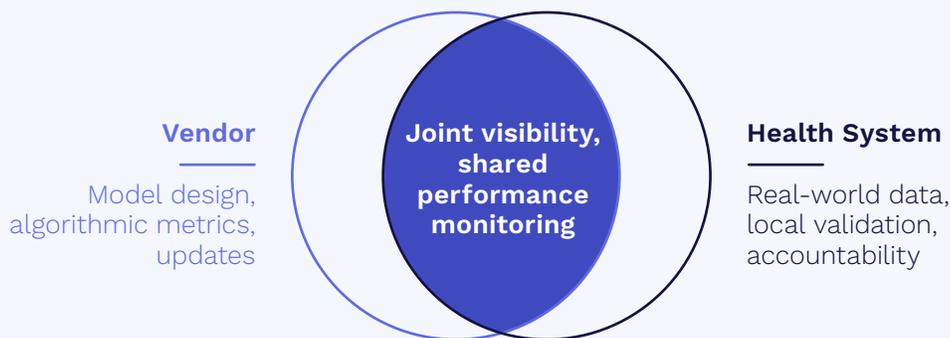
Monitoring Is a Shared Responsibility

Vendors play a critical role in monitoring. They often provide valuable dashboards, reports, and performance metrics that offer depth and technical detail. But no vendor can see the full picture. Real-world performance is shaped by local data and practices, specific workflow integrations, and governance standards. Robust monitoring may require access to downstream data that does not otherwise interact with the vendor’s solution.

That is why health systems cannot outsource monitoring entirely. Just as software vendors secure their code, but hospitals still run their own cybersecurity programs, AI vendors contribute important metrics while health systems supply the context and accountability. Vendor metrics are a starting point, not the finish line.

The most effective approach is partnership. Vendors and health systems each bring complementary insights: vendors know the model’s inner workings; health systems know the environment where it operates and the highest priority risks. Together, they create a more accurate and trustworthy picture. This collaboration not only makes AI safer; it builds the confidence among clinicians, patients, and boards that accelerates adoption and unlocks impact at scale.

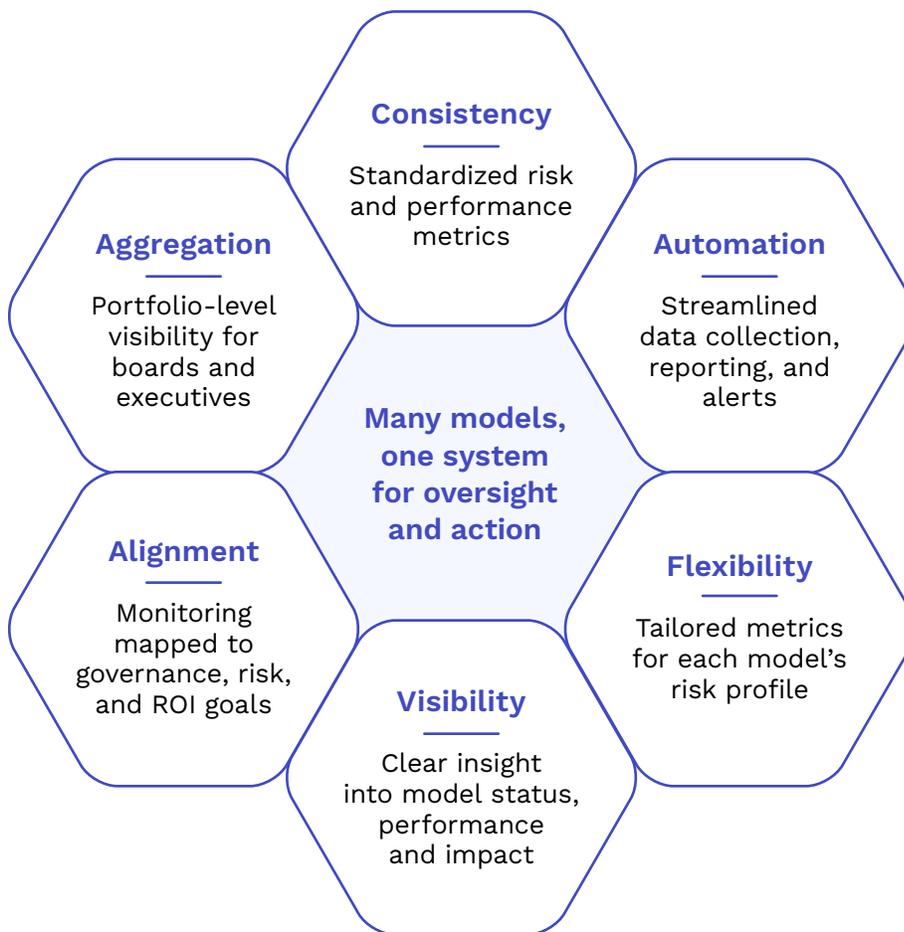
Effective oversight requires partnership—vendors provide model insight; health systems supply context.



Why Health Systems Need a Central Monitoring Program

As the number of AI tools multiplies, the challenge is not just monitoring each tool individually but managing them as a portfolio and enabling interventions when problems arise. A deterioration predictor, a discharge planner, a documentation assistant—all may require different oversight, but the health system must be able to see them together and operationalize a common set of workflows and integrations.

A central monitoring program delivers that system-wide view. It creates:



Without centralization, oversight fragments. With it, health systems gain the visibility and resilience needed to manage dozens, or soon hundreds, of models at scale.

What Strong AI Monitoring Looks Like

A robust program blends standardization with configurability. Categories of oversight must be consistent, but the techniques and metrics must adapt to each model. A predictive model for sepsis risk should be monitored differently than a generative model for claims denial responses. Even within a single class of models, context matters: a discharge summary tool carries different risks than a patient-facing chatbot.

| Model Type | Primary Risk | Monitoring Focus | Example |
|------------|--------------------|----------------------|----------------------|
| Predictive | Accuracy drift | Performance tracking | Sepsis model |
| Generative | Hallucination | Content validation | Discharge summaries |
| Agentic | Unintended actions | Behavior audit | Scheduling assistant |

At its core, strong monitoring includes:

- ◆ **Risk-based intensity.** Higher-risk models receive more frequent and detailed oversight.
- ◆ **Standardized categories.** Safety, accuracy, utilization, ROI, and fairness tracked in common terms.
- ◆ **Comparative intelligence.** Systems can identify which models deliver value and which should be retired.
- ◆ **Clear escalation.** Defined pathways ensure monitoring insights lead to timely, risk-appropriate action.
- ◆ **Governance integration.** Monitoring is not just a technical task—it's embedded within broader governance, informed by initial intake, evaluation, and ongoing oversight.

When done well, monitoring shifts from a compliance function into a strategic advantage.

From Pilots to Portfolio: Monitoring as the Enabler of Scale

The AI adoption story in healthcare is often one of frustration. AI investments are growing, yet, value realization feels slow and underwhelming. At the same time, the expanding portfolio of AI tools and types is creating anxiety at the executive level, largely due to the possibility of undetected, and therefore unmanaged, risk.

Robust monitoring changes the equation. It reduces concerns about hidden risk by building guardrails in from the start. It also makes pilots far more informative as leaders gain visibility into the performance, risk, and impact of new tools, enabling them to make future investments based on real data rather than abstract business cases or generic vendor claims.

In this way, monitoring is not just a guardrail. It is the enabler of value realization.

Signal 1's Role: Turning Monitoring Into Management

Monitoring and measurement are the foundation of safe and effective AI adoption. But health systems need more than a set of disconnected dashboards—they need the right underlying infrastructure and easy-to-use tools.

Signal 1's AI Management Platform provides that foundation. It standardizes monitoring across vendors and models, aligns oversight with governance frameworks, aggregates results into executive dashboards, and turns insights into action.

With our platform, monitoring becomes more than observation. It becomes management. Health systems gain the ability not only to see where AI stands today, but to guide where it goes tomorrow—safely, measurably, and at scale.

SIGNAL 1

signal1.ai