EXECUTIVE BRIEFING | January 2026

# Beyond the Algorithm: Ensuring Clinical Validity at Scale

Why Continuous Evidence is the New Standard of Care.

# The Execution Gap

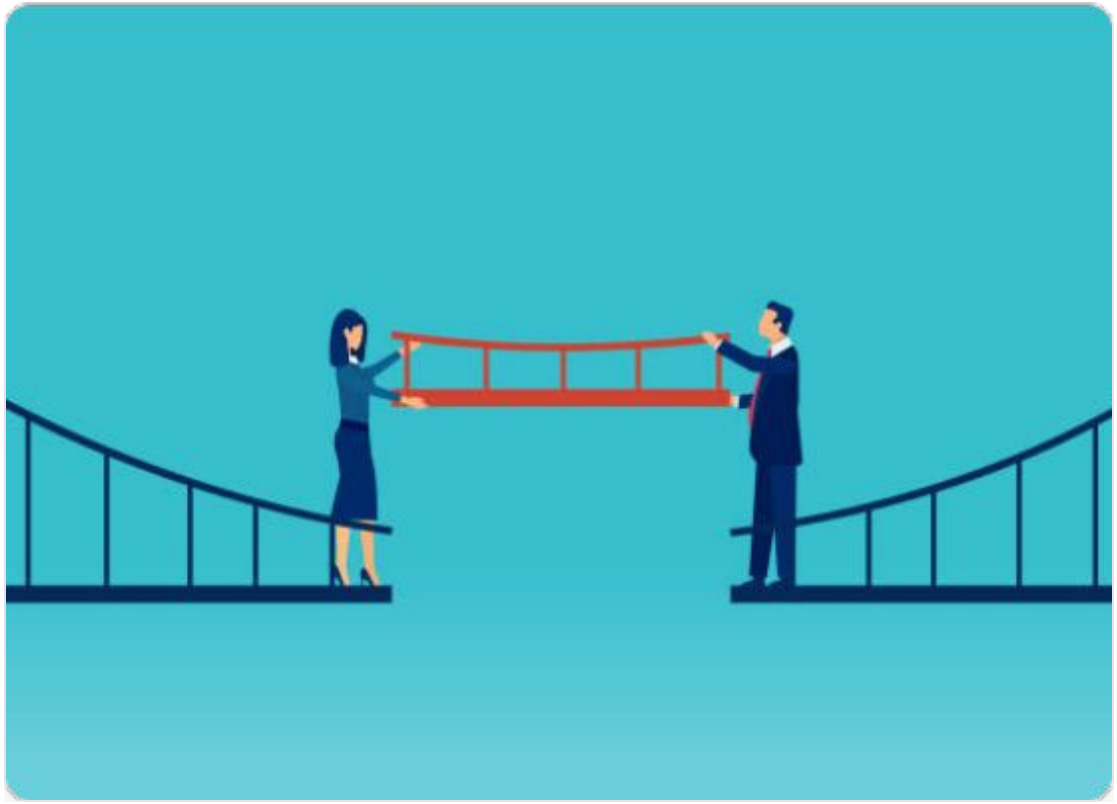
## The Current State

Clinical AI governance has reached conceptual maturity. Regulators (FDA), standards bodies (NIST), and clinical associations (AMA) broadly agree on the requirements: **Transparency**, **Oversight**, and **Equity**.

## The Problem

However, frameworks define *intent*, not *execution*. Most health systems lack the technical infrastructure to generate continuous evidence of AI performance. Instead, they rely on episodic, manual validation that fails to capture real-world shifts in data.

> **"Only 61% of hospitals validated AI tools on local data prior to deployment, and fewer than half tested for bias."**



*Source: American Heart Association (AHA), "Science Advisory on Predictive AI in Cardiovascular Care," November 2025.*

# A Unified Standard of Care

A consensus has emerged across clinical, technical, and federal bodies. It is no longer sufficient to validate a model once; systems must demonstrate **Continuous Evidence** of safety and efficacy.

## Clinical & Ethical

AMA CHAI

**Focus: Patient Safety & Equity**
Physicians must remain the "Human-in-the-loop." Governance must use "Assurance Labs" to validate fairness on local populations before and during deployment.

## Technical Standards

NIST ISO 42001

**Focus: Reliability & Measurement**
Reliability is a statistical requirement. Systems must demonstrate "continuous monitoring" of validity, ensuring models do not degrade over time.

## Federal Regulation

FDA ONC

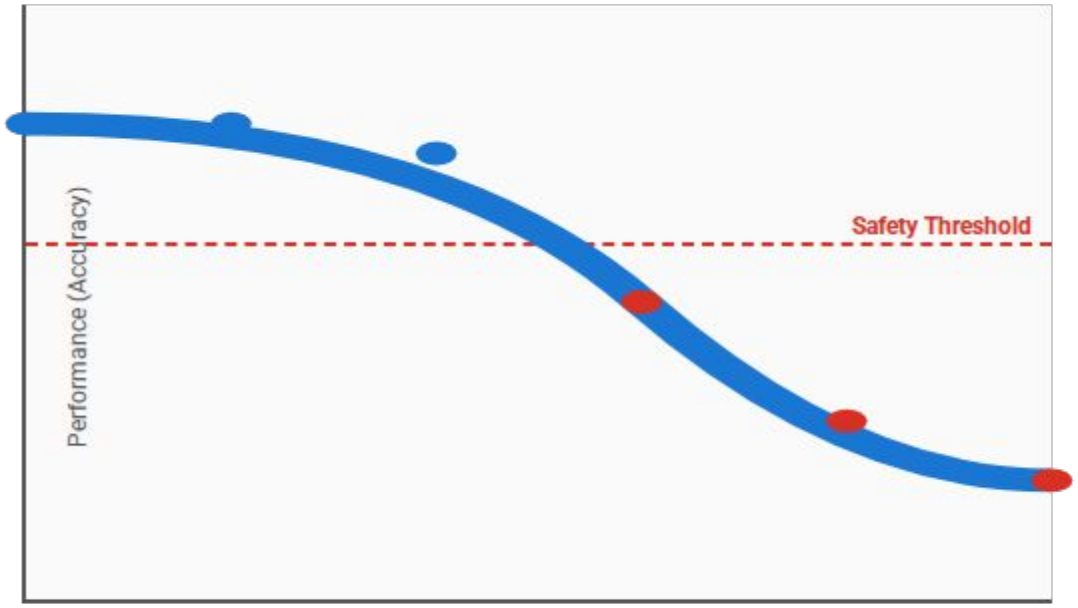
**Focus: Surveillance & Transparency**
Pre-market clearance is not enough. "Total Product Lifecycle" (TPLC) rules mandate post-market surveillance for data drift.

# The Risk: Silent Failure (Model Drift)

## Why Continuous Monitoring Matters

Unlike traditional software, AI models are probabilistic. They degrade over time as patient populations, scanner protocols, and operational workflows shift.

- **Silent Failure:** Models don't crash; they simply start issuing incorrect predictions with high confidence.
- **The Liability:** A model validated last year may be unsafe today. Without monitoring, this risk accumulates invisibly.




**The Drift Zone:** Where liability accumulates.

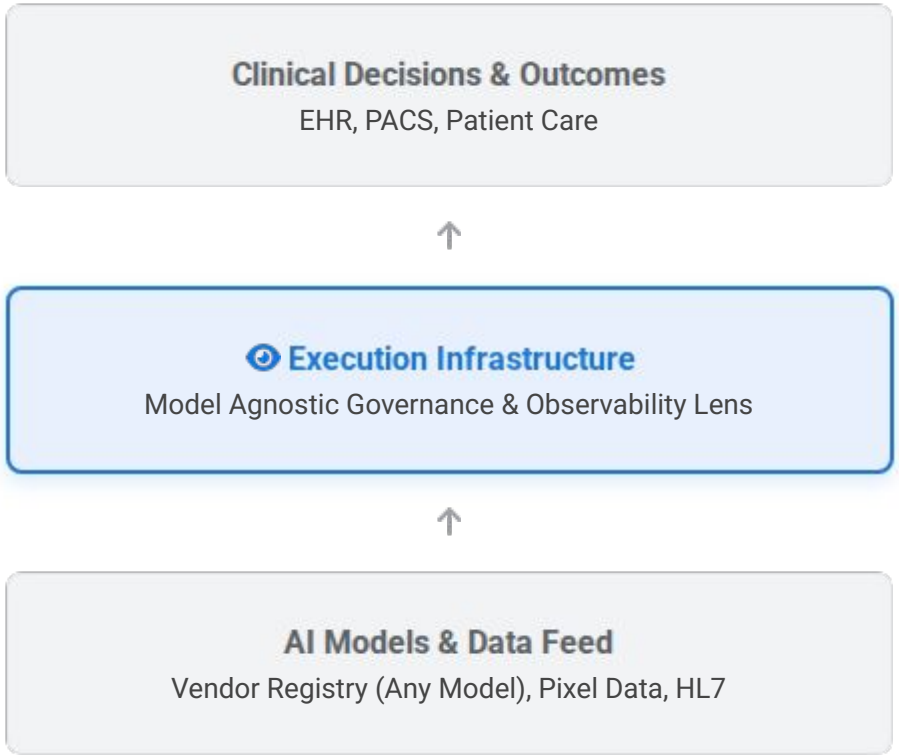
*Source: JAMA Health Forum, "Accountability and Governance Challenges in Clinical AI," 2025; West Monroe, "Applications of AI in Healthcare".*

# The Missing Infrastructure

## Bridging Policy and Practice

The gap described by Harvard and the AMA is not a lack of rules; it is a lack of tooling. You cannot manage 21st-century AI with 20th-century spreadsheets.

We require an **Execution Infrastructure**—a translation layer that sits between the AI models and clinical care. This layer automatically translates high-level regulatory mandates into hard, daily evidence without burdening clinical staff.



# The Observability Lens Delivers Automated Ground Truth

How do we technically satisfy the mandates for surveillance, fairness, and oversight? By reconciling the **AI Prediction** with the **Clinical Outcome** in a continuous loop.

### 1. Capture

Intercept raw AI model outputs in real-time across all vendors.

**Meets FDA TPLC**

### 2. Extract

Use Clinical LLM & NLP to parse unstructured reports for the "True" diagnosis.

**Meets AMA Oversight**

### 3. Match

Reconcile prediction vs. ground truth at the patient level to validate accuracy.

**Meets NIST Validity**

### 4. Evidence

Generate continuous Sensitivity & Specificity metrics for audit.

**Meets ISO Audit**

# The Strategic Imperative

## Why Infrastructure is Critical

Legacy systems (PACS/EHR) were designed for storage, not surveillance. As AI scales, the friction of manual validation becomes a safety risk.

We built Observability Lens because **Vendor Neutrality** is essential for trust. You cannot ask a model to grade its own homework.

> **"You cannot govern what you cannot measure. A neutral verification layer is the only way to ensure clinical validity at scale."**

### The Trust Barrier

Adoption fails without independent verification.

# Meeting the Standard

Execution Infrastructure directly satisfies the core requirements of the modern regulatory stack, moving compliance from "episodic" to "continuous."

**FDA TPLC (Surveillance)**
Longitudinal drift monitoring alerts immediately on performance degradation, satisfying post-market rules.

**NIST (Fairness)**
Subgroup stratification detects hidden bias across demographics, ensuring equitable performance.

**AMA (Oversight)**
Outcome-linked verification keeps humans in the loop by validating AI against physician consensus.

**ISO 42001 (Audit)**
Automated logging creates a defensible audit trail for every prediction, timestamp, and outcome.

*Sources: FDA TPLC Guidance; NIST AI RMF; AMA Principles for Augmented Intelligence; Joint Commission Sentinel Event Alert.*

# From Liability to Asset

## The Liability Trap

*"Hoping it works"*

- ✕ Episodic Validation
- ✕ Unknown Performance
- ✕ Silent Risk Accumulation

## The Strategic Asset

*"Proving it works"*

- ✓ Continuous Monitoring
- ✓ Defensible Evidence
- ✓ Trust & Adoption

# Recommended Path Forward

Moving from intent to execution requires a phased operational roadmap. We recommend starting with a baseline audit to understand current exposure.

1

## Retrospective Audit (Gap Analysis)

**Goal: Baseline Exposure**

Retrospectively audit the top 3 deployed algorithms on local historical data to identify hidden drift and demographic bias.

2

## Automate (Governance)

**Goal: Enable Surveillance**

Deploy the execution infrastructure to automate FDA TPLC monitoring and create real-time dashboards for leadership.

3

## Scale (Trust)

**Goal: Enterprise Adoption**

Expand to the full enterprise portfolio. Use validated performance data to negotiate value-based contracts with AI vendors.

# References


- **FDA:** *Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device: Lifecycle Management.*
- **Joint Commission:** *Sentinel Event Alert: Guidance on Responsible AI in Healthcare.*
- **American Heart Association:** *Science Advisory on Predictive AI in Cardiovascular Care.*
- **AMA (H-480.939):** *Principles for Augmented Intelligence in Health Care.*
- **NIST:** *AI Risk Management Framework (AI RMF 1.0).*
- **ISO/IEC 42001:** *Artificial Intelligence Management System Standard.*
- **JAMA Health Forum:** *Accountability and Governance Challenges in Clinical AI.*
- **Harvard Gazette:** *AI is speeding into healthcare. Who should regulate it?*
- **West Monroe:** *Applications of AI in Healthcare and Governance Implications.*