

PUBLIC REFERENCE BRIEF

HEALTHCARE AI EVIDENCE FILE

The research behind Hospitals Need Two AI Ledgers:
Value and Failure.

Six current sources. Clear evidence labels. The limitation that should travel with every claim.

VERSION 1.1

PUBLIC EDITION

EVIDENCE VERIFIED AUGUST 31, 2026

INSIDE

01

VALUE

Economics and outcomes

02

FAILURE

Errors and near misses

03

TRUST

Patient expectations



What survives the limitations.

The sources differ in method and authority, but three operating conclusions remain useful for hospital leaders.

01

AI VALUE IS LOCAL

Volume, workflow design, verification burden, implementation cost, and recurring operations can change the business case.

02

FAILURE NEEDS A PATH

Errors, near misses, drift, workarounds, and patient concerns cannot be managed if they remain scattered across informal channels.

03

THE LEDGERS MUST RECONCILE

Rework, remediation, abandonment, and trust failures are operating costs and belong in expansion and renewal decisions.

HOW TO READ THE EVIDENCE LABELS

Peer-reviewed systematic review

Synthesis of published research evaluated through scholarly peer review.

Preprint / modeled case

Public research not yet peer reviewed; modeled outputs are not realized results.

Authoritative industry initiative

First-party direction or intent; not a completed standard or proven outcome.

Government implementation

Official implementation account; results are limited to what the source reports.

Safety survey + reporting initiative

A patient-safety development bounded by its small, self-reported sample.

National public-opinion survey

Representative attitudes and expectations; not outcomes or legal requirements.

These are ALHQ operating interpretations of the source set. No single source establishes a universal hospital AI standard.

Value depends on operating conditions.

Two sources frame AI economics and the movement toward more disciplined outcome measurement.

1

FLARE: A Systematic, Uncertainty-Aware Framework for Evidence-Based Adoption of Artificial Intelligence in Healthcare

PREPRINT / MODELED CASE

PUBLISHED AUGUST 24, 2026

WHY IT MATTERS

- Prices human verification time, infrastructure, recurring operations, implementation, and workflow integration into the AI business case.
- In one illustrative stroke-imaging pathway, modeled break-even was about 3,992 patients per year and expected first-year ROI became positive near 5,000.

WHAT IT DOES NOT PROVE

This is a non-peer-reviewed, single modeled case. It does not show that a deployed hospital achieved the modeled savings.

ROLE: PRIMARY VALUE LEDGER EVIDENCE

[OPEN ARXIV SOURCE](#)

2

HIMSS AI Outcomes Framework founding agreements

INDUSTRY INITIATIVE

PUBLISHED AUGUST 24, 2026

WHY IT MATTERS

- HIMSS and five Asia-Pacific healthcare organizations signed founding agreements to develop a longitudinal, multi-institutional AI outcomes framework.
- Formal launch is planned for 2027, with measurement and data collection to begin then and peer-reviewed publication planned for 2028.

WHAT IT DOES NOT PROVE

This is a founding agreement, not a completed standard, validated benchmark, or evidence of financial results.

ROLE: SUPPORTING VALUE LEDGER EVIDENCE

[HIMSS ANNOUNCEMENT](#)[HIMSS-OWNED REPORT](#)

Scale requires stronger reporting paths.

A government implementation and a patient-safety initiative show why experimentation and incident visibility must travel together.

3

Singapore's AgentSea platform

GOVERNMENT IMPLEMENTATION

PUBLISHED AUGUST 24, 2026

WHY IT MATTERS

- Singapore reported that healthcare professionals had created more than 12,000 AI agents on Synapse's sector-wide AgentSea platform.
- The government account pairs experimentation with safety and performance evaluation, multidisciplinary governance, post-deployment monitoring, incident reporting, and human judgment.

WHAT IT DOES NOT PROVE

Agents created is an experimentation count, not validated production deployments, realized savings, or demonstrated safety outcomes.

ROLE: CONTEXT HELD OUTSIDE THE FLAGSHIP BODY

[OPEN SINGAPORE MOH SOURCE](#)

4

ECRI expands error reporting and calls for more data on AI errors in patient care

SAFETY SURVEY + REPORTING INITIATIVE

PUBLISHED AUGUST 25, 2026

WHY IT MATTERS

- ECRI expanded its Problem Reporting Network to capture AI errors, malfunctions, and near misses.
- Among 124 respondents, 31 percent suspected an incorrect or misleading AI output, 9 percent said an error reached a patient or affected care, and 35 percent were unsure whether they had encountered an AI error.

WHAT IT DOES NOT PROVE

The small, perception-based sample exposes a surveillance problem. It is not a national estimate of AI harm.

ROLE: PRIMARY FAILURE LEDGER EVIDENCE

[OPEN ECRI SOURCE](#)

Patient factors are still under-measured.

A systematic review and a national survey connect lifecycle oversight with disclosure and patient agency.

5

Patient factors in medical artificial intelligence: a systematic review

PEER-REVIEWED SYSTEMATIC REVIEW

PUBLISHED AUGUST 26, 2026

WHY IT MATTERS

- Across 330 medical AI papers, 16.7 percent examined trust, 10.9 percent examined safety, and only 3.9 percent considered patient factors during design and development.
- The review indicates that patient-centered questions are concentrated late in validation rather than built into earlier lifecycle decisions.

WHAT IT DOES NOT PROVE

These figures describe what published studies assessed. They do not measure the safety performance of deployed AI systems.

ROLE: SUPPORTING FAILURE LEDGER EVIDENCE

[OPEN NATURE HEALTH SOURCE](#)

6

Americans want transparency when AI is used in their healthcare

NATIONAL PUBLIC-OPINION SURVEY

PUBLISHED AUGUST 25, 2026

WHY IT MATTERS

- Pew surveyed 3,488 U.S. adults: 72 percent rated disclosure of healthcare AI use highly important, rising to 81 percent for diagnosis or medical scans and remaining 72 percent for AI note-taking.
- Fifty-three percent felt they had little or no say, and 46 percent were unsure whether AI had already been used in their care.

WHAT IT DOES NOT PROVE

The findings measure attitudes and expectations, not clinical outcomes, legal requirements, or the effectiveness of a disclosure model.

ROLE: PATIENT CONCERNS AND OVERSIGHT CONTEXT

[OPEN PEW SOURCE](#)

Use the evidence without outrunning it.

The credibility of the brief depends on keeping every claim inside the boundary established by its source.

CLAIM-BOUNDARY MATRIX

EVIDENCE TYPE	SAFE USE	DO NOT INFER
Modeled economics	Inputs, thresholds, scenario logic	Realized savings at a deployed hospital
Industry initiative	Direction, planned framework, timing	A validated standard or benchmark
Government account	Reported platform activity and controls	Validated deployments or outcomes
Small safety survey	A surveillance gap and reporting need	A national incidence rate of harm
Systematic review	What published studies examined	Safety performance of deployed systems
Public-opinion survey	Attitudes, disclosure expectations, agency	Clinical outcomes or legal requirements

SOURCE HANDLING AND FRESHNESS

- All six entries were verified against direct sources or the documented first-party verification path on August 31, 2026.
- No second freshness review is required for publication on September 2, 2026.
- If publication moves beyond September 14, 2026, every source marked for use must be rechecked before release.

ACCESS-LIMITED SOURCE

The HIMSS primary page was access-limited during verification. The announcement was confirmed through Healthcare IT News, which identifies HIMSS as its parent company. Unconfirmed framework-dimension claims were excluded.

USE NOTE

This Evidence File supports ALHQ Issue #2, Hospitals Need Two AI Ledgers: Value and Failure. It is an executive reference brief, not legal, clinical, privacy, security, or regulatory guidance. Local policy, applicable law, contracts, clinical risk, and regulatory requirements should determine final control design.