AI ADOPTION TRENDS · FIELD REPORT 02

Healthcare AI 2026

From pilots to an operating system

A decision-grade field study for provider executives, clinical leaders, IT, finance, and compliance.

23 pages · evidence, economics, controls, worksheets

Research cutoff: 18 September 2026

aiadoptiontrends.com

THE DECISION MEMO

Healthcare has crossed from experimentation into operating exposure. The near-term winners will not be the organizations with the most pilots; they will be the ones that can retire weak tools, measure workflow outcomes, and preserve clinical accountability.

81%

of surveyed physicians reported awareness or professional use of at least one AI use case in 2026. [1]

71%

of U.S. non-federal acute-care hospitals reported predictive AI integrated with the EHR in 2024. [2]

1,614

entries in FDA's AI-enabled device table downloaded for this report; 1,230 were radiology-led. [3]

- **Start where the work is already reviewed.**
  Ambient documentation, chart summarization, scheduling, billing support, and inbox triage have owners, artifacts, and measurable baselines. They are easier to constrain than autonomous diagnosis.
- **Buy an operating change, not a model.**
  The contract, EHR integration, exception queue, reviewer behavior, monitoring, and decommission plan determine most realized value.
- **Local evidence beats a famous logo.**
  Randomized studies show meaningful but product- and setting-specific differences. Measure the exact workflow, users, specialties, and failure modes you will run.

BOARD-LEVEL RECOMMENDATION

Fund a 90-day portfolio: one documentation workflow, one administrative workflow, and one clinical decision-support use case in shadow mode. Give each a baseline, accountable owner, safety case, economic threshold, and stop rule. Do not grant a system write authority simply because it can draft.

WHAT NOT TO DO

Do not create a central "AI innovation" queue that accumulates pilots without line ownership. Do not average adoption across hospitals, specialties, or products. Do not call model output accuracy a business case.

[1] American Medical Association, 2026 Physician Survey on Augmented Intelligence.

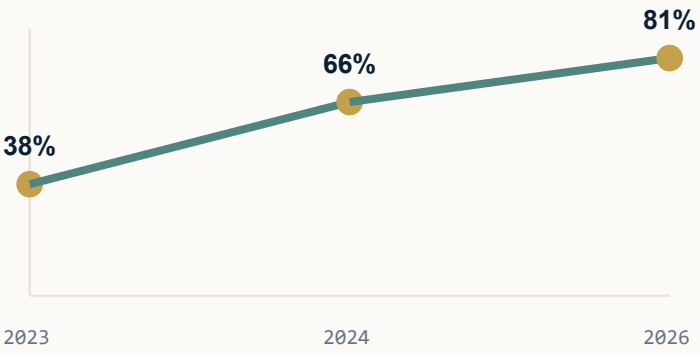
[2] ASTP/ONC, Hospital Trends in the Use, Evaluation, and Governance of Predictive AI, 2023–2024.

[3] FDA AI-Enabled Medical Device List; table downloaded and counted 18 Sep 2026; latest listed decision 29 Jun 2026.

ADOPTION ACCELERATED FASTER THAN OPERATING MATURITY

Three different measures—physician self-report, hospital EHR deployment, and regulatory authorizations—show breadth. None alone proves safety, ROI, or routine use.

PHYSICIAN AI AWARENESS OR USE · AMA SELF-REPORT

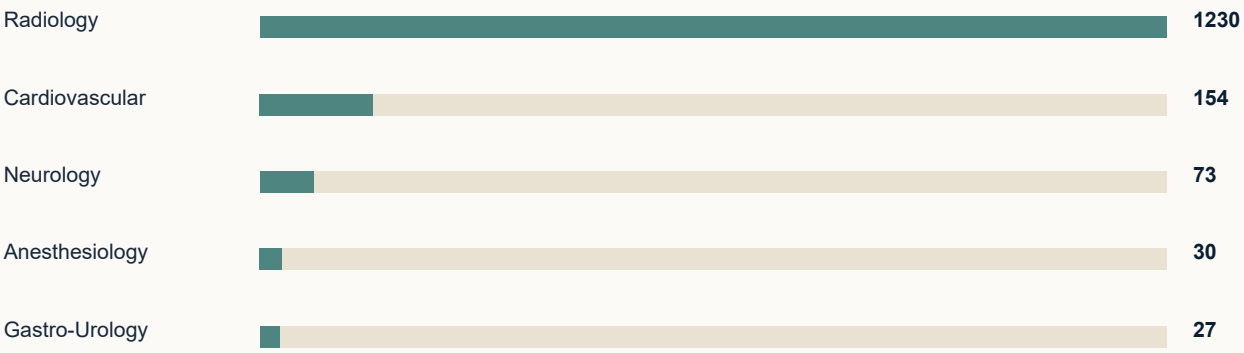


Instrument and response options changed across waves; read directionally, not as a controlled longitudinal cohort. [1]

HOSPITAL PREDICTIVE AI IN THE EHR

ASTP's measure covers non-federal acute-care hospitals and predictive AI integrated with the EHR—not every generative AI pilot. [2]

FDA AUTHORIZATION LANDSCAPE · COUNTED FROM CURRENT TABLE



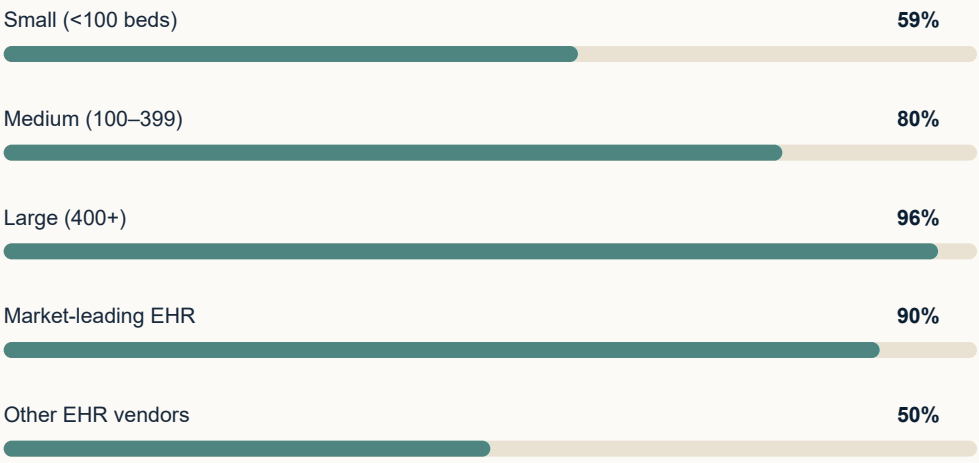
INTERPRETATION DISCIPLINE

Authorization means the product met the applicable premarket requirements for its intended use. It does not mean a hospital's implementation is effective, unbiased locally, integrated well, or economically justified. The FDA also says its list is not comprehensive. [3]

"HEALTHCARE" IS NOT ONE BUYER

Capability concentrates in large, system-affiliated institutions and in organizations using the market-leading EHR. A national average hides the implementation gap.

HOSPITALS USING PREDICTIVE AI · 2024 [2]



THE OPERATIONAL READ

Large hospitals can spread governance, interface, validation, and training cost across more use cases. Smaller and independent providers need narrower products, clearer implementation support, and shared evidence—not a smaller version of the enterprise portfolio.

WHAT THE AVERAGE OBSCURES

CAPACITY

Clinical informatics, integration, security, analytics, legal, training.

DATA

Local labels, workflow metadata, outcome capture, stable interfaces.

LEVERAGE

Volume to negotiate terms, demand evidence, and fund monitoring.

RISK

Small organizations can have less redundancy when a workflow fails.

Decision implication: segment the evidence by setting, specialty, user maturity, and EHR. A result from a large academic ambulatory network is a hypothesis for a rural hospital—not a transferable ROI guarantee.

EIGHT MARKETS HIDING INSIDE “HEALTHCARE AI”

Each lane has a different buyer, evidence standard, system boundary, and downside. Score and govern them separately.

01 Ambient documentation

VALUE Clinician time
ARTIFACT Draft note
FAILURE Inaccuracy; omission

02 Patient messaging

VALUE Access; inbox
ARTIFACT Draft response
FAILURE Unsafe advice; tone

03 Revenue cycle

VALUE Cash; denials
ARTIFACT Code / packet
FAILURE Upcoding; denial drift

04 Scheduling & access

VALUE Capacity
ARTIFACT Appointment / queue
FAILURE Equity; no-show shift

05 Predictive operations

VALUE Flow; staffing
ARTIFACT Risk score
FAILURE Feedback loops

06 Clinical decision support

VALUE Detection
ARTIFACT Recommendation
FAILURE Automation bias

07 Imaging & devices

VALUE Diagnosis
ARTIFACT Device output
FAILURE Distribution shift

08 Research & knowledge

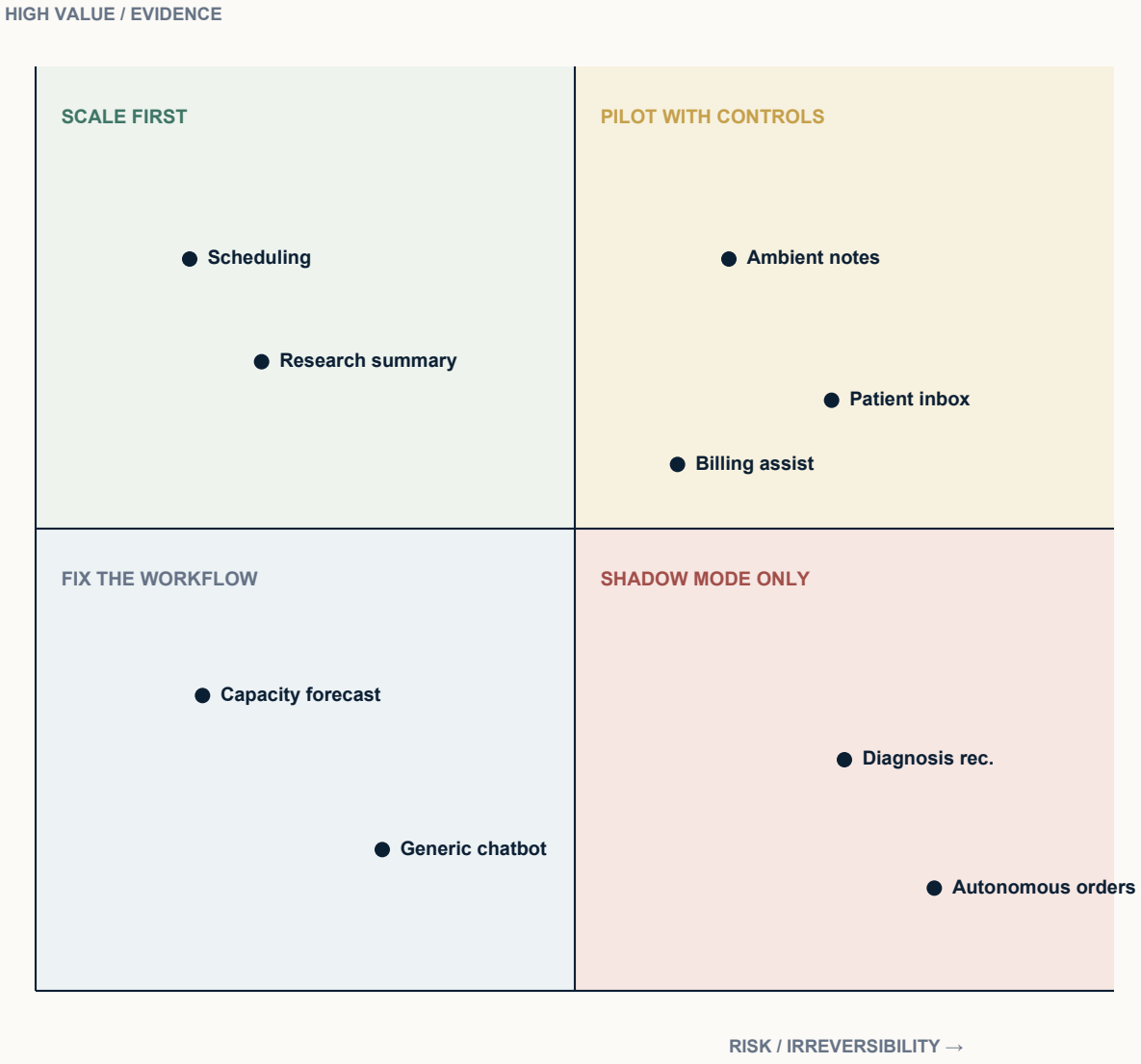
VALUE Evidence speed
ARTIFACT Summary
FAILURE Citation / recency

PORTFOLIO RULE

Put no more than one-third of the 90-day portfolio into clinical recommendation or autonomous action. Balance it with workflows where a human already reviews the artifact and where value appears in time, throughput, or cash—not only model metrics.

SEQUENCE BY VALUE, REVERSIBILITY, AND EVIDENCE

The first use case should not be the most impressive. It should teach the organization how to contract, integrate, validate, monitor, and stop.



Placement is a starting hypothesis, not a universal ranking. A well-controlled diagnostic device may carry a stronger evidence base than a loosely deployed patient-message generator. Move every candidate after reviewing intended use, local workflow, evidence, and failure recovery.

AMBIENT DOCUMENTATION: PROMISE, VARIANCE, VIGILANCE

This is the best-documented generative workflow in care delivery. The evidence supports testing—not blanket rollout.

California academic RCT 238 physicians; 2 products; 2 months	**RESULT** Nabla: −9.5% time-in-note vs control; DAX: no significant change	**READ WITH** Any scribe improved survey measures; occasional clinically significant inaccuracies. [7]
Wisconsin pragmatic RCT 66 practitioners; 71,487 notes	**RESULT** −0.36 hours/day on notes; lower work exhaustion	**READ WITH** Professional fulfillment did not significantly increase under the prespecified threshold. [8]
Six-system QI study 263 paired survey participants; 30 days	**RESULT** Burnout 51.9% → 38.8%; −0.90 hours after-hours documentation	**READ WITH** Observational pre/post design; volunteer and response effects remain. [9]
Mayo ED pilot 5 early adopters; AI vs human scribe	**RESULT** AI required more EHR note time per patient	**READ WITH** Small pilot, but a useful warning that setting and comparator matter. [10]

WHAT THIS MEANS

Do not contract on “hours saved.” Run the same product across specialties and the same specialty across products where feasible. Measure adoption, editing, note time, after-hours work, note quality, coding changes, and safety events together.

THE AMBIENT-SCRIBE TRIAL YOU CAN ACTUALLY RUN

A credible local trial separates eligibility, exposure, editing behavior, outcomes, and safety. It also preserves a control or stepped rollout long enough to learn.

1 · BASELINE	Four weeks	EHR time-in-note, after-hours work, note closure, visit volume, coding, patient experience, burnout pulse.
2 · STRATIFY	Before access	Specialty, visit type, language, clinician baseline burden, note template, patient consent requirements.
3 · RANDOMIZE	If feasible	Product/control or stepped-wedge assignment; avoid giving only enthusiasts the tool.
4 · OBSERVE	Six–eight weeks	Usage per eligible visit, edit distance, unsupported details, omissions, downstream corrections, opt-outs.
5 · ADJUDICATE	Weekly	Clinical reviewer samples notes; safety lead triages incidents; operations reviews adoption barriers.
6 · DECIDE	Pre-set gate	Scale, modify, or stop by specialty. Do not average a strong clinic with a harmful workflow.

MINIMUM STOP RULES

Pause if clinically consequential inaccuracies exceed the agreed threshold, protected information leaves the approved boundary, note closure worsens materially, or users cannot reliably review the output.

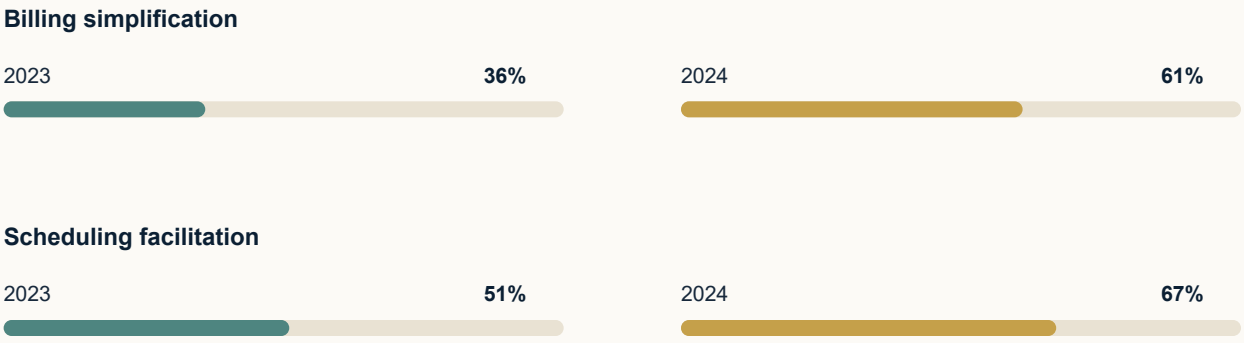
MINIMUM SCALE RULES

Require sustained use on eligible encounters, non-inferior note quality, measurable burden reduction, manageable edit load, and no adverse trend in coding or patient complaints.

ADMINISTRATIVE AI IS BECOMING THE ADOPTION WEDGE

Among hospitals already using predictive AI, billing and scheduling were the fastest-growing use cases from 2023 to 2024. The value is attractive because outcomes are observable; the risks are simply different.

SHARE OF AI-USING HOSPITALS APPLYING PREDICTIVE AI [2]



SCHEDULING

Track filled slots, time-to-appointment, no-show displacement, language access, and queue fairness. A lower call volume can hide worse access.

PRIOR AUTHORIZATION

Track complete packets, first-pass approvals, turnaround, appeal rate, and clinician touches. Generated evidence must trace to the chart.

CODING AND BILLING

Track coder agreement, denial rate, net yield, rebills, and post-payment audit. More codes are not automatically more value.

THE CONTROL MOST TEAMS MISS

Measure displacement. If automation moves work from call center to clinic, coder to physician, or patient to appeals, the local metric improves while total cost and experience worsen.

THE INBOX IS A WORKFLOW, NOT A TEXT-GENERATION TASK

Patient messaging combines urgency detection, clinical context, empathy, language, routing, and liability. A fluent draft can still be operationally unsafe.

1 **INTAKE**

Identity, channel, consent, language

2 **TRIAGE**

Intent, urgency, red flags, queue

3 **GROUND**

Chart + approved policy + current care plan

4 **DRAFT**

Response with uncertainty visible

5 **REVIEW**

Named role; escalation rules

6 **CLOSE**

Send, document, monitor reply

- **Faceplant: urgency disappears.**
  A long portal message is summarized as an administrative request and routed to the wrong queue.
- **Faceplant: the chart wins over the patient.**
  A stale medication or diagnosis in the record is repeated without reconciling the new message.
- **Faceplant: readability masquerades as safety.**
  The response sounds empathetic but gives advice beyond the approved scope.
- **Faceplant: language is treated as translation.**
  Literal translation misses cultural, literacy, and emergency-context cues.

RELEASE THRESHOLD

Run a red-flag test set by language and intent. Require sensitivity targets for urgent categories, zero autonomous closure on excluded intents, and routine review of false negatives—not only average accuracy.

PREDICTIVE AI NEEDS A SERVICE LINE, NOT A DASHBOARD

The model is one component in a detection-to-action pathway. Performance can be excellent while the program fails through alert routing, capacity, adoption, or feedback loops.

SILENT COHORT GAP

Missingness, payer mix, site, language, or device access changes who receives a score.

THRESHOLD THEATER

AUC is reported; alert volume, positive predictive value, and capacity are not.

ACTION AMBIGUITY

The alert has no named responder, deadline, or supported intervention.

FEEDBACK CONTAMINATION

Interventions change the observed label; retraining treats treated outcomes as ground truth.

LOCAL DRIFT

Documentation, population, workflow, or upstream software changes after validation.

EQUITY BY AVERAGE

Overall performance is acceptable while a subgroup carries the false-negative burden.

ASTP reports that 82% of hospitals using predictive AI evaluated at least some models for accuracy, 74% for bias, and 79% conducted post-implementation evaluation or monitoring in 2024. “At least some” is not portfolio coverage. [2]

THE FDA LANDSCAPE IS LARGE—AND SHARPLY CONCENTRATED

Regulatory status, intended use, local implementation, and post-deployment performance are separate questions. Procurement should preserve all four.

76.2%

of the 1,614 entries counted in FDA's current table were radiology-led (1,230 entries). [3]

335

listed final decisions in 2025, versus 235 in 2024; 2026 is partial in the downloaded table.

4 files

to keep per device: authorization, intended use, local validation, live monitoring record.

DEVICE DUE-DILIGENCE CARD

Regulatory identity	Submission number, product code, panel, decision date, software version.
Intended use	Population, modality, setting, user, inputs, outputs, exclusions, workflow claim.
Evidence	Comparator, sites, sample, prevalence, subgroups, reader design, confidence intervals.
Integration	Where the result appears, latency, failure display, override, downtime process.
Local validation	Representative cases, readers, equipment, protocol, subgroup and site analysis.
Monitoring	Volume, missing output, sensitivity/PPV proxies, overrides, downstream action, incidents.
Change control	Version notice, revalidation triggers, model/data changes, rollback, decommission.

FDA explicitly notes that its list is not comprehensive and is derived in part from AI-related terms in public summaries. Treat the table as an authoritative landscape resource, not a denominator for every AI function in U.S. care. [3]

GOVERNANCE BELONGS IN THE DELIVERY PATH

A committee can set policy. It cannot replace the service-line owner, the workflow designer, the security review, or the person monitoring a live system.

ENTERPRISE AI COUNCIL

Portfolio rules, risk tiers, exceptions, inventory, public commitments.

CLINICAL SERVICE OWNER

Intended workflow, staffing, threshold, adoption, outcomes, stop decision.

MODEL / PRODUCT OWNER

Version, validation, integration, defects, vendor relationship, lifecycle.

SAFETY & QUALITY

Hazard analysis, incident review, sampling, patient impact, corrective action.

PRIVACY & SECURITY

Data map, BAA, access, retention, subprocessors, logging, breach response.

FINANCE & OPERATIONS

Baseline, total cost, displacement, capacity benefit, contract realization.

MINIMUM ARTIFACTS AT EACH GATE

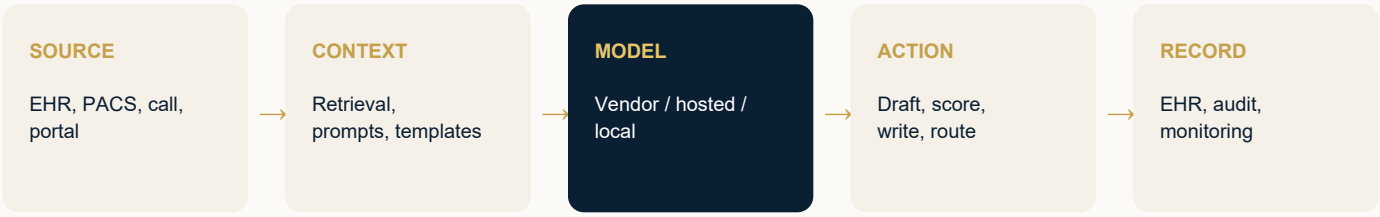
- Use-case charter
- Data-flow map
- Evidence appraisal
- Hazard log
- Local test protocol
- Human factors plan
- Monitoring spec
- Incident runbook
- Economic baseline
- Decommission plan

REGULATORY ALIGNMENT

Use HTI-1 transparency for predictive decision support in certified health IT as a floor, not a complete governance system. Pair it with NIST’s Govern–Map–Measure–Manage cycle and workflow-specific clinical safety practice. [5][6]

DRAW THE DATA PATH BEFORE DISCUSSING INTELLIGENCE

A BAA is necessary when a vendor handles ePHI on behalf of a covered entity; it is not a product certification, security architecture, or permission design. [4]



Identity

Whose authority does the system use? Does retrieval enforce the caller's permissions?

Data minimization

Which fields are required? Are recordings, prompts, and outputs retained?

Subprocessors

Where does data travel, in which region, and under which contractual terms?

Learning

Can customer data train shared models? What opt-out and proof exist?

Logging

Can logs contain PHI? Who can inspect them, and for how long?

Write authority

Can the system draft, queue, sign, send, order, code, or post? Separate every verb.

Downtime

What remains possible if the vendor, interface, or model fails?

Deletion

Can the organization delete source, prompt, derived output, embedding, and backup copies?

CONTRACT TRAP

"HIPAA compliant" is marketing language unless the exact service, configuration, data flow, BAA, retention, and permitted use match the deployment. HHS specifically lists AI chatbot vendors handling PHI as potential business associates. [4]

MAKE EVERY FINALIST RUN THE SAME EVIDENCE ROOM

The procurement document should be a test protocol with commercial terms attached—not a feature checklist assembled from vendor websites.

01	**Intended use**	Exact workflow, user, setting, output, exclusions, prohibited uses.
02	**Evidence**	Peer-reviewed or customer evidence; comparator; sample; confidence; limitations.
03	**Local test**	Sandbox, replay data, acceptance thresholds, subgroup and workflow cases.
04	**Human factors**	Review burden, override, uncertainty, alert design, training, patient notice.
05	**Data & security**	BAA, flow, retention, training use, encryption, access, subprocessors, incidents.
06	**Integration**	EHR write paths, FHIR/API, latency, downtime, identity, version compatibility.
07	**Monitoring**	Logs, version notice, drift, defects, adverse events, exports, audit access.
08	**Economics**	Implementation, interface, training, review, support, minimums, overages, exit cost.
09	**Liability**	Indemnity, warranty, reliance language, cyber coverage, IP, regulatory change.
10	**Exit**	Data export, deletion, workflow fallback, knowledge transfer, termination assistance.

COMMERCIAL RULE

Tie expansion pricing and renewal to measured eligible use and agreed operating outcomes. Secure a termination path before interfaces and training create switching cost. Ask for version-change notice and revalidation support in writing.

BUILD THE ECONOMICS FROM MINUTES, QUEUES, AND CAPACITY

A model-accuracy improvement has no financial value until it changes a staffed workflow, an avoided cost, a collected dollar, or a clinical outcome the organization can responsibly attribute.

ANNUAL VALUE MODEL

Eligible volume	visits / cases / messages
× Adoption	share actually using
× Net minutes	baseline minus review + rework
× Loaded rate	role cost per minute
+ Capacity value	only if schedulable / collectible
− Total cost	license + interface + ops + risk

WORKED EXAMPLE · ILLUSTRATION ONLY

100 clinicians × 18 eligible visits/day × 220 days × 60% sustained adoption × 6 net minutes saved = 142,560 clinician minutes, or 2,376 hours. At a $150 loaded hourly rate, gross time value is $356,400. Subtract licenses, interfaces, implementation, training, monitoring, and added review/rework. Do not count capacity revenue unless the freed time becomes filled appointments or measurable retention. Replace every input with local evidence.

FOUR NUMBERS FINANCE SHOULD REFUSE TO ACCEPT

“HOURS SAVED”

without measured review and rework

“ROI PER USER”

without eligible-use denominator

“REVENUE ENABLED”

without capacity conversion

“DENIALS PREVENTED”

without net yield and audit risk

A 90-DAY PORTFOLIO, WITH STOP RULES

Run three different risk shapes at once so the organization learns reusable controls—not three versions of the same chatbot.

DAYS 0–15
Frame
Select three workflows; baseline outcomes; tier risk; name owners; map data; define prohibited actions.

DAYS 16–30
Prove
Contract sandbox terms; build test sets; replay historical cases; validate integration and permissions.

DAYS 31–60
Pilot
Limited users; human review; weekly safety and operations meeting; track adoption and displacement.

DAYS 61–75
Stress
Subgroups, downtime, language, adversarial inputs, stale data, unusual cases, version-change drill.

DAYS 76–90
Decide
Scale, modify, stop, or extend with a named uncertainty. Publish the evidence packet and owner.

PORTFOLIO A · LOW

Scheduling or research summarization. Optimize the full queue and citation chain.

PORTFOLIO B · MEDIUM

Ambient note or patient-message drafting with named review and sampling.

PORTFOLIO C · HIGH

Clinical prediction in shadow mode with service-line action design.

ONE SCORECARD, FIVE LENSES

Report a balanced operating view. A tool that wins time and loses safety—or wins adoption and loses equity—has not passed.

OUTCOME	What changed for the patient, clinician, queue, cash, or capacity?	**EXAMPLES** Clinical outcome; access; net yield; resolved demand
WORKFLOW	Did the intended user adopt it and did work move elsewhere?	**EXAMPLES** Eligible use; edit/rework; handoffs; closure time
SAFETY	What can go wrong, how often, and how quickly is it caught?	**EXAMPLES** Severity-weighted events; false negatives; unsupported content
EQUITY	Who receives the benefit and who carries the error?	**EXAMPLES** Performance, access, and overrides by meaningful subgroup/site
ECONOMICS	Did realized value exceed the full operating cost?	**EXAMPLES** Net minutes; capacity conversion; interface + monitoring + exit cost

DASHBOARD RULE

Show numerator, denominator, cohort, period, product version, and owner. Separate eligible use from licensed users. Separate suggested, reviewed, accepted, sent, and acted-on outputs. Keep safety events visible next to productivity.

WHERE HEALTHCARE AI FACEPLANTS

The expensive failures are rarely “the model gave a weird answer.” They are system failures where a plausible output meets unclear authority, weak review, or an overloaded queue.

Plausible fabrication

A note, message, or evidence summary includes a detail not present in source material.

CONTROL Grounding + explicit uncertainty + sampled source comparison

Critical omission

Social context, negation, red flag, or exception is filtered as noise.

CONTROL High-risk test set + severity-weighted false-negative review

Automation bias

A user reviews faster and stops independently checking.

CONTROL Blind samples + override analysis + periodic unaided cases

Alert displacement

The model finds more risk than the service can act on.

CONTROL Capacity gate + queue aging + threshold tied to action

Coding inflation

Suggested codes improve apparent yield and create denial/audit exposure.

CONTROL Coder agreement + denial + audit + net collections

Data leakage

PHI reaches an unapproved service, log, subprocessor, or training path.

CONTROL Data map + BAA + technical restriction + log review

Version drift

A vendor changes behavior without local revalidation.

CONTROL Change notice + canary set + rollback right

Workflow decay

Users route around friction and the measured pilot no longer resembles practice.

CONTROL Usage telemetry + observation + frontline owner

The safest response is often not a better prompt. It is a narrower intended use, lower authority, clearer source, better queue, or explicit refusal.

THE COPIED ALLERGY

Run this 45-minute exercise before a system can write to the chart, send to a patient, or influence an order.

SCENARIO

An ambient system drafts a note that carries forward an outdated allergy from the chart while omitting the patient's verbal correction. The clinician signs after a rapid review. A patient-message assistant later uses the signed note when answering a medication question. No harm has yet been reported.

0–5 min	**Detect**	Which logs, correction signals, patient replies, or downstream checks reveal the issue?
5–15 min	**Contain**	Can you stop sending, stop write-back, identify affected notes, and preserve evidence?
15–25 min	**Assess**	Who determines clinical significance, scope, notification, and whether this is a reportable event?
25–35 min	**Correct**	How are records, messages, prompts, templates, and vendor behavior repaired without obscuring history?
35–45 min	**Learn**	What changes to review design, sampling, training, thresholds, contract, or intended use?

EVIDENCE TO LEAVE THE ROOM WITH

Named incident commander · affected-system inventory · decision log · search/replay query · patient-safety owner · vendor escalation path · correction workflow · regulatory/privacy assessment · restart criteria · learning owner

PASS CONDITION

The team can contain the system without the vendor, identify affected records by version and time, preserve the original output, correct the record visibly, and name who decides restart.

TWELVE QUESTIONS BEFORE THE NEXT APPROVAL

A board or executive committee does not need to pick models. It does need to know what the system may do, what evidence exists, and how the organization will discover that it is wrong.

01 What exact decision, artifact, or action is the system changing?

02 Who owns the clinical or operational outcome—not merely the contract?

03 What is the baseline, denominator, comparator, and time horizon?

04 What evidence transfers to our population, workflow, and setting—and what does not?

05 Which actions are draft, recommend, queue, send, sign, order, code, or post?

06 Where does PHI travel, persist, train, log, and delete?

07 How are subgroup performance and access differences measured?

08 What human review is required, and how is review quality observed?

09 What new queue, alert, correction, or downstream work does the system create?

10 How will version changes trigger testing, communication, or rollback?

11 What event pauses the system, and who has the kill switch?

12 How do we exit without losing records, workflow continuity, or bargaining leverage?

DECISION RECORD

Approve / approve with conditions / extend evidence / stop. Record the dissent, unresolved uncertainty, owner, next review date, and evidence required to change the decision.

FOUR COMPANIES, FOUR HEALTH-SYSTEM REALITIES

Named deployments make scale visible; published studies make effects interpretable. Keep the two forms of evidence separate.

KAISER PERMANENTE × ABRIDGE

40 hospitals · 600+ medical offices

Kaiser announced systemwide availability of assisted documentation after a year of work with Abridge. The scale is real; the announcement is not a controlled outcome study. [13]

Lesson: deployment scale proves integration and change capacity, not a universal effect size.

PROVIDENCE × MICROSOFT DAX

Randomized step-wedge study

Providence researchers reported 2.5 fewer hours per week of off-hours documentation plus improvements in frustration and burnout. Vendor licenses were supplied, a disclosed conflict that belongs beside the result. [14]

Lesson: objective EHR metadata plus randomized rollout is stronger than testimonial evidence.

UPMC × ABRIDGE

Investor, early user, enterprise adopter

UPMC first described a small clinician cohort, later reported more than 2,000 recordings in three weeks in Central Pennsylvania, and in 2026 described Abridge as its primary ambient AI tool. UPMC also discloses a financial interest. [15]

Lesson: strategic alignment can speed integration; financial relationships must stay visible.

STANFORD HEALTH CARE

Emergency department adoption

Across 8,740 eligible ED encounters, ambient AI was used in 11.2%. Thirty-five of 92 attendings used it, with use concentrated among a small group and in lower-acuity, noninterpreted encounters. When used, median on-shift documentation time was 28% shorter. [16]

Lesson: optional use creates selection. Adoption pattern is part of the result, not a footnote.

HOW TO USE THESE STORIES

Borrow the operating pattern—phased access, embedded workflow, disclosed conflicts, objective metadata, specialty segmentation—not the vendor conclusion. Reproduce the evidence locally before expanding authority.

WHAT THIS REPORT IS—AND IS NOT

A practical synthesis of public primary sources, federal data, peer-reviewed research, and clearly labeled secondary evidence. It is not medical, legal, regulatory, or investment advice.

METHOD

Research cutoff: 18 September 2026. FDA counts were computed from the agency's live HTML table downloaded on that date; the latest listed decision was 29 June 2026. The table contained 1,614 rows, of which 1,230 were radiology-led. AMA trend figures are self-reported and the instrument evolved across waves. ASTP hospital figures use the American Hospital Association IT Supplement and defined predictive AI as described in the source. Study designs and limitations are shown beside reported findings. Recommendations are AI Adoption Trends synthesis.

PRIMARY SOURCES

[1] AMA (2026), Physician Survey on Augmented Intelligence. https://www.ama-assn.org/sites/ama-assn.org/files/2026-03/physician-ai-sentiment-report_0.pdf

[2] ASTP/ONC (2025), Hospital Trends in the Use, Evaluation, and Governance of Predictive AI, 2023–2024. https://healthit.gov/data/data-briefs/hospital-trends-use-evaluation-and-governance-predictive-ai-2023-2024/

[3] FDA, Artificial Intelligence-Enabled Medical Devices. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices

[4] HHS OCR, Business Associates; and Guidance on HIPAA & Cloud Computing. https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/business-associates/

[5] ASTP/ONC, HTI-1 Final Rule: Algorithm Transparency. https://healthit.gov/regulations/hti-rules/hti-1-final-rule/

[6] NIST, AI Risk Management Framework 1.0 and Generative AI Profile. https://www.nist.gov/itl/ai-risk-management-framework

[7] Lukac et al. (2025), randomized trial of two ambient AI scribes. PubMed 40672471. https://pubmed.ncbi.nlm.nih.gov/40672471/

[8] Pragmatic randomized trial of ambient AI and practitioner well-being (2025). PubMed 41625485. https://pubmed.ncbi.nlm.nih.gov/41625485/

[9] Olson et al. (2025), ambient AI scribes across six health systems. PubMed 41037268. https://pubmed.ncbi.nlm.nih.gov/41037268/

[10] Morey et al. (2025), ambient AI versus human scribes in the emergency department. PubMed 41251650. https://pubmed.ncbi.nlm.nih.gov/41251650/

[11] Schiff (2025), AI-Driven Clinical Documentation—Driving Out the Chitchat? NEJM. https://www.nejm.org/doi/10.1056/NEJMp2416064

[12] Deloitte / NEJM Catalyst (2025), Generative AI in health care: from proof of concept to transformational impact. Secondary survey and examples; used with attribution.

[13] Kaiser Permanente (2024), assisted clinical documentation deployment with Abridge. https://about.kaiserpermanente.org/news/press-release-archive/kaiser-permanente-improves-member-experience-with-ai-enabled-clinical-technology

[14] Providence (2025), randomized study of DAX Copilot and documentation burden. https://digitalcommons.providence.org/publications/10587/

[15] UPMC (2023–2026), Abridge deployment and training materials; financial interest disclosed. https://inside.upmc.com/upmc-in-central-pa-uses-ai-based-solution-to-reduce-physician-burnout-improve-patient-care/

[16] Stanford Health Care / Preiksaitis et al. (2026), ambient scribe adoption in the ED. https://stanfordhealthcare.org/publications/972/972065.html