



Patients For Patient Safety US

Submitted by: Patients for Patient Safety US (PFPS US)

Date: February 23, 2026

Re: Request for Information: *Accelerating the Adoption and Use of Artificial Intelligence as part of Clinical Care*, Department of Health and Human Services, Office of the Secretary, RIN 0955-AA13

Patients for Patient Safety US (PFPS US) appreciates the opportunity to comment on the Department of Health and Human Services' Request for Information on accelerating the adoption of artificial intelligence (AI) into clinical care.

PFPS US is led by patients and family caregivers who have experienced preventable harm in healthcare and are motivated to contribute our lived experience to systemic learning. We work in partnership with clinicians, health systems, and policymakers to make care safer and more trustworthy for all Americans, and to reduce outcomes disparities experienced by vulnerable populations.

PFPS US has a special interest in improving diagnostic timeliness and accuracy, as it is the largest opportunity at hand to prevent harm and strengthen the well-being of patients and families. A growing body of research indicates that missed, delayed or inaccurate diagnoses occur more often than all other preventable harms in healthcare combined. More than 900,000 Americans die or are seriously harmed annually by such diagnostic failures. (Newman-Toker DE. Just how many diagnostic errors and harms are out there, really? It depends on how you count. *BMJ Qual Saf.* 2025 May 19;34(6):355-360). The potential cost savings to patients, families, Medicare and Medicaid, and ultimately taxpayers are enormous.

Just as healthcare providers are increasingly integrating AI into their operations, so are patients using consumer-facing AI to gather information not provided by our clinicians, identify risks, reconcile gaps or inconsistencies in our medical records, and inform decision-making. In many ways, AI is democratizing patient access to actionable information that is making us better patients, mitigating our risks of harm, and empowering us as individuals to make the decisions that work best for us. From this vantage point, we enthusiastically support responsible innovation in AI. At the same time, we urge HHS to anchor its AI strategy in a clear and shared purpose: **reducing preventable patient harm and strengthening the health system's capacity for continuous learning and improvement.**

From the patient perspective, the central question is not how quickly AI is adopted, but for what purpose, whose benefit, and with what accountability once it is in use. In health care, adoption follows incentives. **Aligning payment and accountability structures with patient safety goals**

would help ensure that AI is optimized not only for efficiency and throughput, but also for preventing harm and strengthening trust in the technology.

We encourage HHS to view the question of accelerating AI adoption through a simple, patient-centered theory of change:

First, meaningful adoption in health care follows meaningful incentives. Health systems, clinicians, and vendors respond to the business case embedded in payment policy, accreditation, and regulatory expectations. Absent clear incentives, patient safety is not prioritized in the value proposition, as manifested in the rollout of electronic health records which introduced significant new risks of patient harm. Even promising technologies with significant potential to prevent harm remain peripheral if the incentives for technology developers are focused on other goals, e.g. cost saving, maximizing revenue, or shifting liability.

Second, preventable patient harm remains one of the most costly and under-addressed problems in the health care system—measured not only in hundreds of billions of dollars in wasteful and unnecessary spending, but in avoidable suffering, loss of trust, long-term disability, and healthcare workforce distress. Yet patient safety is still treated as a regulatory obligation rather than a core value proposition of care delivery.

Third, AI is uniquely suited to function as a “central nervous system” for patient safety—continuously sensing risk across fragmented data sources, detecting emerging harm and near-misses, and supporting timely corrective action across clinicians and care settings. Few other technologies offer comparable potential to help health systems see and respond to safety failures as they unfold.

Fourth, if CMS and HHS meaningfully prioritize patient safety within payment policy and program design, health systems will seek out solutions that help them meet those expectations. In practical terms, this means that patient safety can—and should—become the primary, high-value use case for AI in clinical care. Aligning incentives around harm prevention and detection would immediately change what kinds of AI tools are built, and the speed at which they are purchased, and implemented.

Finally, this approach will only succeed if providers can adopt AI with confidence that these tools are themselves safe, accountable, and supported by robust post-deployment learning systems. A laissez-faire approach to oversight and monitoring is unlikely to accelerate adoption. To the contrary, when technologies introduce unmanaged risk or opaque liability, organizations rationally avoid them. Clear expectations for safety, transparency, and continuous learning are therefore not barriers to adoption; they are preconditions for trust.

Taken together, these principles point toward a coherent national strategy: make patient safety the central use case for AI in health care, align incentives accordingly, and ensure that AI-enabled safety infrastructure is supported by strong governance and learning systems. This approach reflects how patients experience harm—and a radical new opportunity to prevent it.

In practical terms, this means that accelerating AI adoption will depend less on exhortations to innovate and more on whether payment policy, accountability structures, and safety expectations make patient harm prevention a central business priority for health care organizations.

AI as a Central Nervous System for Patient Safety

PFPS US encourages HHS to view AI not primarily as a collection of point tools, but as a potential **central nervous system (CNS) for patient safety**—one that can continuously sense risk, detect emerging harm, learn from near-misses and failures, and support timely corrective action across care settings.

Patients and caregivers experience harm longitudinally—across encounters, clinicians, settings, and payers. Unlike claims systems, EHRs, or specialty silos, **patients cross boundaries**, and are often the first to recognize breakdowns long before those breakdowns appear in claims data or quality reports.

AI has unprecedented potential to:

- Detect missed diagnoses, delayed follow-up, and care coordination failures
- Identify early deterioration and system-level breakdowns
- Aggregate patient-reported health care safety concerns/signals at scale

Yet today, most AI tools are designed as **isolated point solutions**, optimized for efficiency or prediction rather than **harm prevention and detection**.

Absent this safety-oriented framing, accelerated AI adoption risks reinforcing fragmentation, obscuring accountability, and scaling existing inequities rather than addressing them.

Responses to specific Questions:

In the responses that follow, PFPS US addresses those questions most directly related to patient safety and the conditions needed for AI to meaningfully reduce preventable harm in clinical care.

1. Barriers to Innovation and Adoption

From a patient safety perspective, the greatest barriers to effective AI innovation and adoption are at least threefold:

- 1. The absence of meaningful incentives to invest in safety infrastructure. Health systems have little reason to absorb the organizational and legal risk associated with early adoption of AI tools that represent potential breakthroughs in patient safety.** Where patient safety is treated as a compliance obligation rather than a strategic priority, technologies oriented toward harm prevention and detection are understandably deprioritized.
- 2. Opacity of performance combined with hospital risk aversion. Today, there is no** consistent national approach to AI performance and safety monitoring, or identifying, aggregating, or learning from AI-related harm or near-misses. Healthcare organizations are on their own in an environment characterized by diffused and unclear accountability across the developer, vendor, and organizational ecosystem. In this environment, neither providers nor developers are well positioned to innovate safely and with confidence. PFPS US applauds the Joint Commission/CHAI Responsible Use of AI framework and the Patient AI Collaborative's recently released AI Care Standard for Patient Communication to try to fill the gap; however, these are strictly voluntary guidances. Given the stakes, and given that HHS has stepped back from AI oversight, PFPS US advises that CMS should consider developing an aligned Condition of Participation (CoP) around AI governance that gives providers leverage to compel performance information from developers.
- 3. Archaic, hospital-centric thinking about how health care should be organized.** A distributed care model supported by agentic workflow and real teamwork holds enormous promise for better serving rural and other under-resourced areas.

2. Payment and Programmatic Policy Changes as a Catalyst for Safety-Oriented AI

Payment policy is not merely a financing mechanism; it is one of the strongest signals of what the system values. If patient safety is not meaningfully valued in payment, the tools designed to improve it—AI included—will remain marginal.

Today, most patient safety infrastructure—such as tools and processes designed to detect missed diagnoses, delayed follow-up, care coordination failures, or early clinical deterioration—are treated as unreimbursed overhead rather than value-generating assets.

PFPS US encourages HHS and CMS to explore **creative payment models that explicitly prioritize patient safety**, including approaches that allow organizations to “earn back” value through demonstrated improvements in patient-reported safety experience. Such approaches could meaningfully shift the internal business case for adopting AI tools that serve safety goals.

As one illustrative example, PFPS US’s *Project PIVOT* has explored the introduction of patient safety, diagnostic accuracy/timeliness, and bias-related questions into patient experience surveys, notably the Consumer Assessment of Health Provider Services (CAHPS) surveys used across the U.S. healthcare system. When paired with an earn-back or shared-savings framework, such measures could create a powerful incentive for healthcare organizations to adopt AI-enabled tools that continuously detect safety risks, surface patient-reported harm signals, and support timely corrective action.

Importantly, such approaches would not require CMS to mandate specific technologies. Instead, they would allow organizations flexibility to invest in AI and other innovations that best meet patient safety goals—while aligning payment with outcomes that matter most to patients and caregivers.

Programmatically, HHS needs a place to aggregate, analyze and disseminate post-deployment learnings. While FDA is the appropriate lifecycle overseer for AI-as-medical-device, AHRQ, as the nation’s healthcare quality and safety agency, should be resourced to steward the implementation and evaluation science.

3. Legal, Governance, and Accountability Considerations

A prime objective of PFPS US and our network of patient safety advocates is that patient experience of harm, or bias that contributes to harm, is learned from to fuel future prevention. PFPS US therefore urges HHS to consider policy approaches that normalize disclosure and learning, including safe harbors for good-faith reporting of AI failures and near-misses, expectations for auditability and traceability of AI-supported decisions, and prohibition of confidentiality agreements that hide AI safety concerns from being seen and understood. **These mechanisms reduce unmanaged risk for providers while making real-world failures visible and**

correctable. However, in the absence of clear expectations for AI safety, monitoring, and accountability, health care organizations face rational disincentives to adopt tools that could introduce unmanaged liability risk.

A laissez-faire approach to oversight is therefore unlikely to accelerate adoption. To the contrary, it risks creating precisely the conditions under which providers avoid new technologies out of fear that failures will be opaque, reputationally damaging, or legally ambiguous.

Diffuse liability, opaque decision provenance, and silent failure modes pose real risks as AI use expands, particularly for tools that influence clinical workflows without being classified as medical devices. Without safe pathways for reporting problems, failures are likely to remain hidden.

4. Evaluation Beyond Deployment

Evaluation that ends at deployment will not build provider confidence in AI-enabled safety tools. **Health systems will only invest in AI as safety infrastructure if they can trust that performance will be monitored, failures surfaced, and learning shared over time.**

Traditional approaches to AI evaluation in health care—such as pre-deployment validation and periodic audits—are not sufficient for technologies that continuously interact with clinical workflows and evolve over time. From a patient safety perspective, AI tools should be embedded within integrated safety management systems that monitor performance, detect drift, surface near-misses, and enforce design constraints in real-world use. Systems safety frameworks developed in other high-risk industries, such as the System-Theoretic Accident Model and Processes (STAMP), emphasize continuous, system-level monitoring and feedback rather than episodic spot checks. HHS can play an important role in advancing post-deployment evaluation approaches that treat AI safety as an ongoing operational responsibility, supported by routine monitoring, reporting infrastructure, and learning loops embedded in clinical care.

6. AI tools with the greatest potential to improve healthcare outcomes

From a patient safety perspective, the greatest unrealized potential lies in AI that:

- **Aggregates and analyzes patient-reported harm signals**
- **Detects process breakdowns (e.g. during transitions), not just discrete clinical findings**
- **Recognizes patterns across encounters**
- **Does automated surveillance for missed follow-up and diagnostic delays**

- **Uses predictive analytics to avert harm events**
- **Extends provider capacity/availability through agentic workflow – i.e., identifying remote patient care needs in real time and triggering outreach to the appropriate member of the care team to respond.**

8. Interoperability for Patient Safety and Learning

Interoperability will only accelerate meaningful AI development if it supports the flow of safety-relevant information across fragmented care settings. From the patient perspective, the most consequential gaps are not simply in data exchange, but in the ability to surface and learn from safety signals that cross organizational boundaries—missed or delayed diagnoses, breakdowns in follow-up, adverse events and near-misses, and patient-reported experiences of harm and bias. AI functions as a “central nervous system” for patient safety only when these signals can move across teams, tools, and settings in near real time. Without interoperability that supports safety signal flow and learning, AI risks reinforcing fragmentation rather than helping to correct it.

As an example, a consequential interoperability gap is one that could well frustrate the use of AI for a critical use case: improving timeliness and accuracy of diagnosis for patients who move between settings. Right now, researchers cannot reliably track presenting symptoms across care settings. Patients who present to the emergency department with symptoms such as dizziness may be discharged and later admitted with a diagnosis such as stroke, yet the system often captures “what the patient presented with” differently— or not at all —across ED and inpatient workflows. This semantic and longitudinal fragmentation prevents AI tools from linking symptom presentation to subsequent diagnosis, blinding them to patterns of missed or delayed diagnosis across care transitions. Without consistent, computable representation of presenting symptoms and chief complaints across settings, AI cannot function as a safety “central nervous system” for diagnostic error detection at precisely the points of highest risk. To begin to address this problem, we urge ONC to adopt into the USCDI the proposed new Symptom data class and standardized symptom data elements that HL-7 currently is balloting.

9. Patient and Caregiver Priorities and Concerns

Patients and caregivers consistently express that they want AI to help ensure that things do not fall through the cracks. We want tools that notice when the system is failing—when follow-up does not happen, when conditions worsen, when information is lost between clinicians, or when bias shapes care. This absolutely includes the ability for patients to contribute our own

safety experiences as signals to such an AI-powered “central nervous system”. Patients also want safe and reliable AI tools that can empower us and our caregivers with knowledge, insight, and strategies as we engage with payers, providers, and others.

Patients worry about opaque decisions that cannot be questioned, AI being used to deny or delay care, and data being extracted without tangible benefit to those whose experiences generated it. We fear that technology may further distance us from accountability rather than bringing problems to light.

Patients also are keenly aware that what health systems prioritize follows what is paid for. When patient safety is not treated as a core value in payment and accountability structures, we reasonably worry that AI will be optimized for efficiency, throughput, utilization control, or short-term revenue margin rather than for harm prevention.

10 Research Priorities to Support Safe Adoption

PFPS US encourages HHS to prioritize research that supports the safe and trustworthy integration of AI into real-world care, not only the development of new models. From the patient safety perspective, high-impact research priorities include implementation science for safety-oriented AI, approaches to human–AI teaming that reduce cognitive burden rather than shifting risk, and the design of learning systems that incorporate patient-reported safety signals. Investment in these areas would help ensure that AI adoption strengthens, rather than undermines, the health system’s capacity to prevent harm and continuously improve.

In closing

PFPS US urges HHS to ensure that efforts to incentivize AI adoption accelerate rather than outpace the nation’s ability to detect harm and protect patients. We are confident that AI can earn trust — and ultimately achieve scale — if it is embedded in a patient-centered safety infrastructure that treats harm prevention, transparency, and continuous learning as a core value proposition. Aligning incentives around patient safety, supporting AI as safety infrastructure, and ensuring robust oversight and learning systems are not competing goals — they are mutually reinforcing prerequisites for adoption that earns patient trust.

We appreciate the opportunity to comment and welcome continued engagement as HHS advances this important work.

Respectfully submitted,

Patients for Patient Safety US