Patients For Patient Safety US

Re: NIH RFI on the Development of a Strategic Plan for Artificial Intelligence (AI)
Submitted by: Patients for Patient Safety US (PFPS US)
Date: July 15, 2025

Dear Director Bhattacharya:

Patients for Patient Safety US (PFPS US) appreciates the opportunity to contribute to NIH's strategic planning for artificial intelligence (AI). PFPS US is a national network led by patients and families, working alongside healthcare providers, systems engineers, researchers, educators, and other key stakeholders to elevate patient safety as a national public health priority. Our network includes individuals, as well as nonprofit and for-profit organizations. Our shared goal is a healthcare system that centers the safety and well-being of patients in every setting, at every step of care. Our policy priorities align with the Administration's emphasis on stronger, technology-enabled transparency, measurement, and public reporting, fortified by patient/family engagement, as drivers of quality and safety improvement. Improving safety and accountability for safety of care will also reduce the conservatively estimated **$400 Billion +** in

Imagine This: A Safer, Smarter Health System in 2035

It's 2035, and a 72-year-old patient is admitted for a routine hip replacement. As she moves through the hospital, her care is quietly supported by an invisible network — an AI-powered safety nervous system.

Before surgery, the system checks her full medication history — not just from her EHR, but from her pharmacy records, her wearable health app, and even her self-reported symptoms — to flag a potentially dangerous interaction that no one had caught.

During recovery, **the system notices subtle changes in her mobility and breathing patterns** — the kind of early signals that, in 2024, would have been missed until they became a full-blown pulmonary embolism. A care team is notified. A test is ordered. Harm is prevented.

In the background, the AI system compares her experience with data from thousands of similar patients — surfacing real-world insights about which safety practices work best, for whom, and in what context. Those lessons are fed back into hospital protocols and national guidelines.

Critically, this patient knows it's happening. **She's informed**. She can see how the system is helping. And if something goes wrong, she has a clear, trusted way to report it — not to a black box, but to a system that learns from her voice.

direct costs for inappropriate or unnecessary care -- and associated taxpayer burden – attributable solely to patient harm events and their cascades.

Patients for Patient Safety US (PFPS US) urges NIH to lead a bold, forward-looking strategy that goes beyond accelerating discovery — one that ensures AI becomes a **nervous system for patient safety in health care**.

While AI is often promoted as a catalyst for precision medicine or biomedical innovation, its potential as an infrastructure for preventing error, identifying risk, and learning from harm remains underdeveloped. A strong, well-funded NIH strategy can change that — and we urge NIH to lead.

A North Star: Build the long overdue Nervous System for Patient Safety

Borrowing not from aviation but from biology, we envision AI as a nervous system for health care — one that continuously senses, signals, adapts, and learns. Just as the human body reflexively withdraws from danger, the AI-infused health system should:

- Detect when harm is imminent (a missed test, a drug error, a silent infection),
- Trigger protective actions (escalation, outreach, review),
- And remember near-misses and adverse events as fuel for learning — at both the local and system levels.

But just as a faulty nervous system can misfire, a poorly designed AI system can amplify errors, ignore early warnings, or worsen disparities. In short: AI will either help health care learn from harm and protect patients — or it will deepen silence and blind spots.

The NIH has a generational opportunity to establish a strategy that ensures AI strengthens the health system's capacity to sense, learn, respond, and protect. This will require NIH to invest not only in algorithm development but also in:

- The science of safety and usability,
- The architecture of transparency and trust, and
- The inclusive co-design of tools that reflect the diversity and complexity of patients' lives.

Patient Reporting of Harm = Sensory Nerve Endings

Just as the body relies on sensory receptors to detect when something may be wrong, **patient and family** experiences in real time are *frontline detectors* of harm or risk of harm, signals that internal systems may miss.

These are the “nerve endings” of the system:

- Distributed everywhere
- Sensitive to things automated monitors do not see
- Triggering protective reflexes when harm (or near-harm) occurs

Without sensory nerves, the body is numb. Without patient input, the AI nervous system is blind to our lived experience.

Patient-reported concerns — a missed test, a change in symptom, a medication error, or a serious but unreported adverse event — are vital inputs. They detect what automated systems often miss. These reports should not disappear into black boxes. They must be treated as integral and essential to the AI nervous system, and they must activate real responses, real learning, and real prevention.

NIH should fund research and model development that:

- Uses natural language processing and large language models to extract safety insights from:
    - Free-text survey responses (e.g., an updated HCAHPS)
    - Patient portal messages and complaints (e.g., “OurNotes”-style contributions)
    - Audio, video, or conversational data
- Incorporates Patient-Reported Incident Measures (PRIMs) into structured AI-ready datasets
- Tests patient-facing interfaces that invite safety alerts and prioritize follow-up

Patients are not just data sources — we are detectors. Any system we build must listen to what we and our families/caregivers know.

A Strategic Imperative: Make AI Safe, Accountable, and Equitable

To realize the vision of AI as a Nervous System for safety in health care, NIH must lead not only a whole-of-agency effort but catalyze an HHS-wide strategy to enhance safety, equity, and especially patient trust in AI. Patients must be integral—not incidental—to the AI learning loop.

While the National Academy of Sciences recently outlined a principled AI Code of Conduct for Health and Medicine, those commitments—though valuable—remain largely aspirational without implementation mechanisms. NIH has the opportunity to translate those principles into practice through investments that elevate patient voice as a primary input for detecting, preventing, and learning from harm. We propose the following four pillars:

1. Make AI Safe by Design

AI tools must be rigorously tested before they are deployed at scale in health care. NIH should fund research that:

- Develops human-centered safety evaluation frameworks for clinical AI.
- Supports independent validation of AI tools, ***particularly*** those used in diagnosis and treatment.
- Advances the science of AI usability, especially for high-risk settings and underserved populations.
- Builds capacity for prospective risk assessment, including simulation and scenario-based testing.
- Develops AI harm taxonomies and real-time reporting mechanisms
- Supports independent audits of safety and bias in both public and commercial models
- Creates meta-models that monitor AI performance drift over time

All can agree that a new source/category of potential harm deserves a intentional infrastructure for prevention and accountability.

2. Make AI Accountable in Practice

The deployment of AI must include mechanisms to detect and address harm — especially harms that may not be captured in claims data or traditional metrics. We recommend that NIH:

- Invest in developing, and validating more robust systems that enable real-time patient and clinician reporting of AI-related safety events.
- Support research on feedback loops that allow health systems to learn from and improve AI performance over time.
- Promote algorithmic transparency and explainability, including open science principles.
- Include patients, caregivers, and frontline clinicians as co-investigators in AI research and design

3. Make AI Equitable by Intention

AI should narrow, not widen, disparities. Yet many existing models encode systemic bias, worsen access, or fail to reflect the lived realities of marginalized populations. NIH should:

- Prioritize AI development and validation in diverse real-world settings, including safety net and rural health systems.
- Fund community-led research on equity-focused AI applications, including tools that address structural drivers of poor health.

- Ensure that funding decisions are informed by distributional impact, not just technical performance.
- Require disaggregated outcomes reporting to detect hidden harms to vulnerable groups.

Equity is not a byproduct of good AI — it is a precondition for trust and legitimacy.

4. Catalyze a Whole-of-HHS Approach to Safe AI Governance

Many of the greatest challenges — and greatest opportunities — lie not just in research, but in governance and deployment. NIH can lead by example and coordinate with CMS, FDA, ONC, and AHRQ to:

- Develop shared safeguards for AI validation and safety.
- Align incentives toward AI tools that improve measurable patient outcomes, not just throughput or billing.
- Establish national learning systems for AI-enabled harm detection and mitigation.
- Advance a common vision for safe, responsible, and patient-engaged AI across the entire health system.

Don't Plug AI Into a Broken System — Build a Safer One

If NIH wants to "go big" on AI, **there is no better place to start than preventable harm — the third leading cause of death in the U.S.** Every day, patients are hurt by systems that don't notice, do not speak up, and do not learn. AI can help change that.

We urge NIH to:

- Think like a system.
- Design for learning, not just prediction.
- Listen to patients — early, often, and structurally.

We believe AI can be transformative — if built on a foundation of safety science, human-centered design, and accountability. Patients must not be treated as passive recipients of algorithmic decisions, but as active partners in shaping the future of care.

Let's not waste this moment. Let's use AI to build a system that notices, adapts, and protects — like any good nervous system should.

We would welcome the opportunity to work with NIH and other federal partners to ensure that AI strengthens, rather than undermines, the promise of safe, evidence-based, patient-centered care.

Sincerely,
Patients for Patient Safety US (PFPS US)

References

1. Classen DC, et al. (2011). "'Global Trigger Tool' Shows That Adverse Events in Hospitals May Be Ten Times Greater Than Previously Measured." Health Affairs, 30(4):581–589.

2. OECD (2019). Patient-Reported Safety Indicators: Question Set and Data Collection Guidance. Link

3. Bell SK, et al. (2020). "Frequency and Types of Patient-Reported Errors in Electronic Health Record Ambulatory Notes." JAMA Netw Open, 3(6):e205867.

4. Giardina TD, et al. (2021). "Patients' Free-Text Comments on the CAHPS Hospital Survey Identify Hospital Priorities and Correlate With Hospital Performance." Health Serv Res, 56(1):156–165.

5. Makary MA, Daniel M. (2016). "Medical error—the third leading cause of death in the US." BMJ, 353:i2139.

6. National Academies of Medicine. (2015). Improving Diagnosis in Health Care. Washington, DC: National Academies Press.