The Physician-as-Pilot Framework: A Governance Architecture for AI-Mediated Home Care

Andy Squire

Harvard Medical School, Executive Education (AI in Healthcare); Founder, PatientCentricCare.AI

Abstract

The deployment of artificial intelligence (AI) in unsupervised home healthcare environments introduces a governance challenge that existing performance-focused and human-in-the-loop models do not adequately resolve. This paper introduces the Physician-as-Pilot (AI-in-the-Loop; AIITL) framework, an accountability-explicit governance architecture that formalizes clinical authority, escalation pathways, and responsibility boundaries for AI-mediated care. This manuscript presents a conceptual governance and systems architecture and does not report clinical outcomes, patient-level data, or validated medical device performance.

This work contributes a governance architecture situated within health informatics and health systems research, addressing the safe operationalization of AI-enabled home care technologies.

Throughout all phases of deployment, protection of patient safety and preservation of patient autonomy are treated as paramount design constraints. The framework is operationalized through a Safety OS - a governance layer that integrates consent state, oversight, escalation logic, auditability, and real-world evidence governance across the AI lifecycle.

A non-medical, caregiver-in-the-loop Phase I implementation illustrates how safety-governed engagement and observational engagement indicators may support quality

of life and earlier clinical conversations - without asserting diagnostic or therapeutic claims, and within a regulator-legible pathway.

Introduction

The proliferation of Artificial Intelligence (AI) in healthcare promises to revolutionize diagnostics, treatment, and patient monitoring. However, this promise is tempered by a significant governance challenge, particularly in the context of unsupervised home care. As AI models become more autonomous, the lines of clinical responsibility blur, creating a potential safety and accountability vacuum. Current “human-in-the-loop” (HITL) models, while a step in the right direction, often fail to adequately define the locus of control, leaving clinicians in a reactive, rather than proactive, role. This paper argues for a paradigm shift, from the passive HITL model to an active, clinician-led governance framework.

The Problem: Autonomous AI and the Accountability Gap

The fundamental challenge of AI in healthcare is not one of technical performance, but of governance. An autonomous AI, by its nature, operates without direct human supervision. In a clinical setting, this raises critical questions:

- **Who is responsible** when an autonomous AI makes an error?
- **How is patient consent** managed for dynamic, AI-driven interventions?
- **What are the mechanisms** for oversight and escalation when a patient’s condition changes?

Existing regulatory frameworks, such as the FDA’s guidelines for Software as a Medical Device (SaMD), are still evolving to address the unique challenges of autonomous AI. The risk is a future where AI-driven healthcare is a black box, with neither clinicians nor patients having clear visibility or control. This is not a tenable or ethical path forward.

The Physician-as-Pilot (AIITL) Framework

To address this accountability gap, we propose the **Physician-as-Pilot (AIITL)** framework. This model draws a direct analogy from the aviation industry, where a human pilot retains ultimate command and control authority over a highly automated aircraft. The autopilot can handle routine operations, but the pilot is always present, able to intervene, and ultimately responsible for the safety of the flight.

In the AIITL model, the clinician is the "pilot." The AI system functions as the "autopilot," handling routine monitoring and tasks, but always under the direct oversight and authority of the clinician. This is a critical distinction from the traditional HITL model. The physician is not merely "in the loop"; they are **in command**.

This framework is built on four key principles:

1. **Clinician Authority:** The clinician has the ultimate authority to override, modify, or disengage the AI system at any time.
2. **Clear Accountability:** The clinician, as the "pilot," is the single point of accountability for the patient's care.
3. **Continuous Oversight:** The AI system provides continuous data and alerts to the clinician, enabling proactive, rather than reactive, intervention.
4. **Designed Responsibility:** The system is designed from the ground up to facilitate, not replace, clinical judgment.

Failure Modes and Responsibility Boundaries

A credible governance framework must explicitly acknowledge its limits.

Failure Mode	System Response	Responsibility Locus
AI misclassification	Escalation to clinician	Clinician (operational authority)
Clinician unavailable	Predefined backup escalation	Institution + system design
Consent ambiguity	Conservative default action	System governance
AI system failure	Graceful degradation	System designers
Patient refusal	Documented consent override	Patient autonomy

What This Framework Does Not Prevent:

- Clinician error
- Patient non-compliance
- Infrastructure outages
- Malicious attacks (security remains a distinct domain)

This explicit boundary-setting shifts the framework from aspirational ethics to operational safety.

The Safety OS: A Governance Layer for AI-Mediated Care

The Physician-as-Pilot framework is operationalized through a foundational **Safety OS**. This is not a traditional operating system, but a governance layer that serves as the system of record for all AI-mediated actions. The Safety OS has four core components:

- **Consent:** A dynamic, auditable record of patient consent for all AI-driven interventions.
- **Oversight:** Real-time monitoring and data streams provided to the clinician "pilot."
- **Escalation:** Pre-defined pathways for escalating alerts and anomalies to the clinician.
- **Auditability:** A complete, immutable log of all AI actions and clinician decisions.

Governance Pathway Across Deployment Phases

Figure 1. Phased Governance Pathway for AI-Mediated Home Care

This figure illustrates a phased governance model for the deployment of artificial intelligence in home care environments. Phase I represents non-medical, non-SaMD functionality focused on caregiver support and observational engagement, governed by explicit consent, transparency, and safety constraints. Phase II introduces clinician-supervised AI operating within a SaMD-ready framework, with defined escalation pathways, auditability, and physician-retained authority. Phase III represents

regulated clinical AI subject to jurisdictional approval and formal medical device oversight.

Across all phases, the Safety OS functions as a continuous governance layer enforcing consent state management, escalation logic, oversight boundaries, and real-world evidence capture throughout the AI lifecycle. Device illustrations are shown for conceptual context only and do not represent specific commercial products or clinical claims.

The phased structure shown in Figure 1 is intentionally designed to separate governance maturity from technical capability. Rather than assuming a binary distinction between non-clinical and clinical AI, the framework enables progressive deployment aligned with regulatory readiness, clinical risk, and accountability requirements. Early-phase implementations prioritize patient autonomy and safety without asserting diagnostic or therapeutic intent, while later phases formalize clinician authority and regulatory compliance. This approach allows real-world learning and governance validation to occur prior to escalation into regulated clinical use.

A Phased Pathway to Implementation

The Physician-as-Pilot framework is designed for pragmatic, real-world implementation. We propose a three-phase pathway:

- **Phase I: The Non-Medical Home Companion.** A non-SaMD, caregiver-in-the-loop system that provides non-clinical support and monitoring. This phase allows for the refinement of the Safety OS and user interface in a low-risk environment.
- **Phase II: The Clinician-Supervised, SaMD-Ready System.** The introduction of clinical monitoring and decision support, with the clinician acting as the “pilot.” This phase requires adherence to SaMD development guidelines and prepares the system for regulatory submission.
- **Phase III: The Regulated Clinical AI.** A fully regulated SaMD, with the Physician-as-Pilot framework serving as the core of its safety and governance architecture.

Regulatory Alignment

The Physician-as-Pilot framework aligns with emerging regulatory expectations, including FDA SaMD guidance, EU MDR and AI Act provisions on human oversight, and UK MHRA governance principles. By embedding accountability and escalation into system design, it provides a regulator-legible pathway for AI deployment beyond controlled clinical environments.

Illustrative Implementation: Phase I Home Companion

The Home Companion, a non-medical, caregiver-in-the-loop device, serves as an illustrative implementation of Phase I of the Physician-as-Pilot framework. In this implementation:

- **Signals Monitored:** Ambient sensors, voice interactions, and wearable data (e.g., fall detection).
- **Alerts:** Non-clinical alerts (e.g., missed medication reminders, potential fall) are sent to a designated caregiver via a mobile application.
- **Consent:** Patient consent for data collection and caregiver notifications is logged via a verbal and digital consent process, recorded in the Safety OS.
- **Escalation:** If a caregiver is unresponsive or the situation requires clinical attention, the system escalates to a pre-designated on-call clinician or emergency services, as defined in the patient's care plan.

Implications for Policy and Industry

For policymakers, the framework offers a concrete model for operationalizing human oversight. For industry, it provides a blueprint for building trust, enabling adoption, and reducing governance ambiguity without stalling innovation.

Conclusion

AI will increasingly mediate healthcare delivery beyond traditional clinical settings. The Physician-as-Pilot framework reframes this evolution not as a question of

autonomy, but of governance. By explicitly assigning authority, responsibility, and escalation across the AI lifecycle, it offers a pragmatic, ethically grounded pathway for safe AI deployment in home healthcare.

References

[1] Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), 491-497. https://academic.oup.com/jamia/article-abstract/27/3/491/5612169

[2] Kim, J. Y., Hasan, A., Kueper, J., Tang, T., Hayes, C., & Lau, F. (2025). Establishing organizational AI governance in healthcare: a case study in Canada. *npj Digital Medicine*, *8*(1), 1-9. https://www.nature.com/articles/s41746-025-01909-3

[3] Bakken, S. (2023). AI in health: keeping the human in the loop. *Journal of the American Medical Informatics Association*, *30*(7), 1225-1226. https://academic.oup.com/jamia/article-abstract/30/7/1225/7200020

[4] Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., & Moret-Bonillo, V. (2023). Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review*, *56*(1), 1-57. https://link.springer.com/article/10.1007/s10462-022-10246-w

Authorship and IP Statement

Andy Squire is the sole author of the Physician-as-Pilot (AIITL) governance framework, including the Safety OS architecture, regulatory pathway design, and Phase II/III clinical AI governance model.

The **Phase I Home Companion** illustrative implementation was developed in collaboration with:

- **Dr. Cristina Crisan Tran, MD** - Clinical adoption strategy and commercialization pathways
- **Carla Maldonado, PhD** - Compliance design and operating governance

The author used AI-based tools for editorial assistance; all ideas, framing, and conclusions are the author's own.