

## **From AI Assistance to Agentic Healthcare: A Framework for Autonomous Decision Support, Governance, and Human Oversight**

Mr.Mohammed Shahpurwala

DBA Scholar, mshahpurwala53@gmail.com

### **Abstract**

Artificial intelligence (AI) in healthcare is moving from systems that primarily provide predictions, recommendations, and analytical assistance toward more autonomous and agentic systems capable of interpreting information, reasoning across tasks, planning actions, and interacting with digital tools. This transition creates opportunities for faster, more personalized, and more coordinated healthcare decision support, but it also introduces concerns about accountability, transparency, patient safety, privacy, fairness, and human control. This conceptual paper develops an integrated framework explaining how agentic AI can support autonomous healthcare decision-making while remaining subject to effective AI governance and meaningful human oversight. The framework positions agentic AI capabilities as the technological driver, autonomous healthcare decision support as the central mechanism, and healthcare decision outcomes as the intended outcome. AI governance and human oversight are positioned as complementary control mechanisms. The framework draws on socio-technical systems theory, human–AI collaboration, and responsible AI governance perspectives and develops seven conceptual propositions. Three figures and three tables distinguish capability from authority, specify constructs and mechanisms, and illustrate risk-sensitive deployment. The contribution is a process-oriented account of controlled autonomy in which governance determines the permitted action space and oversight preserves effective intervention. The framework remains conceptual: it offers theoretically reasoned relationships and a future

validation agenda rather than evidence of clinical effectiveness. Its implications concern workflow design, professional accountability, patient participation, and evaluation across the AI lifecycle.

**Keywords:** Agentic AI; healthcare AI; autonomous decision support; AI governance; human oversight; human–AI collaboration; socio-technical systems; responsible AI.

## 1. Introduction

appropriate safeguards (He et al., 2019; Kelly

### 1.1 Background of the Study

et al., 2019). Greater autonomy introduces

Artificial intelligence (AI) has become an important component of modern healthcare, supporting diagnosis, clinical decision-making, patient monitoring, documentation, and healthcare management. Earlier applications performed specific analytical tasks, producing predictions, classifications, or recommendations for professionals to evaluate. Subsequent developments expanded AI involvement across clinical and organizational workflows (Jiang et al., 2017; Yu et al., 2018; Rajpurkar et al., 2022). Generative and agentic AI extend this trajectory toward connected, goal-oriented processes. Agentic systems can interpret information, reason across tasks, plan actions, use digital tools, and adapt to changing information. These capabilities may support clinical information synthesis, monitoring, care coordination, and documentation. Their practical value nevertheless depends on reliable implementation, workflow integration, and

additional concerns about accountability, transparency, privacy, fairness, and patient safety. Responsible implementation therefore requires attention to both technical performance and the healthcare context in which systems operate (Morley et al., 2020; Wiens et al., 2019). The World Health Organization (WHO, 2021) emphasizes human autonomy, safety, accountability, and equity, while the National Institute of Standards and Technology (NIST, 2023) provides guidance for identifying, assessing, and managing AI risks. Accordingly, this study conceptualizes agentic healthcare as controlled autonomy. AI governance establishes organizational boundaries, while meaningful human oversight preserves professional authority to review, intervene, and override system actions. Together, these mechanisms provide the foundation for integrating agentic capabilities into responsible autonomous healthcare decision support.

## 1.2 Problem Statement

The transition from AI assistance to agentic healthcare creates a challenge concerning technological capability and professional decision authority. Systems capable of interpreting information, planning tasks, and executing actions may influence connected healthcare workflows. However, these capabilities do not establish clinical appropriateness, authorization, or accountability. Errors may propagate across tasks, while technical performance alone cannot establish clinical value (He et al., 2019; Kelly et al., 2019). These concerns intensify when workflows involve sensitive patient information, uncertain evidence, or consequential decisions. Healthcare organizations must define permitted autonomy, allocate responsibility, and establish verification, monitoring, escalation, and intervention procedures. Ethical principles therefore require operational controls that protect patients and preserve professional accountability (Morley et al., 2020; WHO, 2021; NIST, 2023). The

literature examined addresses healthcare AI implementation, governance, and human–AI collaboration, but closer integration is needed to explain responsible autonomy across connected decision processes. Meaningful oversight requires sufficient information, authority, and opportunities to challenge system actions (Jarrahi, 2018). Accordingly, this study addresses the need for an integrated conceptual framework explaining how agentic AI capabilities may contribute to healthcare decision outcomes through autonomous decision support, and how governance and human oversight enable and condition these relationships.

## 1.3 Research Gap

### 1. AI Assistance–Agentic AI Gap:

Existing healthcare AI research has largely emphasized prediction, recommendation, and decision-support applications, while the conceptual transition toward multi-step, more autonomous agentic systems remains less developed.

**2. Autonomous Decision Support**

**Gap:** Limited conceptual attention has been given to how increasing AI autonomy may influence the accuracy, timeliness, personalization, efficiency, and consistency of healthcare decision support.

**3. AI Governance Gap:** Governance principles are widely discussed, but their relationship with increasing AI autonomy is not always integrated into a single decision-support framework covering transparency, accountability, privacy, fairness, and risk management.

**4. Human Oversight Gap:** As AI becomes more autonomous, healthcare professionals may increasingly act as reviewers, supervisors, and interveners. The requirements for meaningful oversight and override authority therefore require clearer conceptual treatment.

**5. Integrated Agentic Healthcare**

**Gap:** Agentic AI, autonomous decision support, governance, and human oversight are frequently discussed as separate topics. A unified framework is needed to explain how they can work together while maintaining human control and accountability.

**1.4 Research Aim and Objectives**

**Research Aim:** To develop a conceptual framework explaining how agentic AI can support autonomous healthcare decision-making through effective AI governance and human oversight.

**Research Objectives:**

**RO1:** To examine the theoretical foundations of agentic AI and autonomous decision support in healthcare.

**RO2:** To examine how agentic AI capabilities influence autonomous healthcare decision support.

**RO3:** To examine the role of AI governance in enabling responsible and accountable agentic healthcare.

**RO4:** To examine the role of human oversight in maintaining appropriate human control over autonomous healthcare decision support.

**RO5:** To develop an integrated framework linking agentic AI, autonomous healthcare decision support, AI governance, human oversight, and healthcare decision outcomes.

### 1.5 Research Questions

**RQ1:** What theoretical foundations explain the transition from AI assistance to agentic healthcare?

**RQ2:** How do agentic AI capabilities influence autonomous healthcare decision support?

**RQ3:** How does AI governance support responsible and accountable use of agentic AI in healthcare?

**RQ4:** What role does human oversight play in autonomous healthcare decision support?

**RQ5:** How can agentic AI, autonomous decision support, AI governance, and human oversight be integrated into a coherent healthcare framework?

### 1.6 Significance of the Study

The study is significant because it brings technological autonomy, decision support, governance, and professional control into one conceptual model. Theoretically, it connects agentic AI capabilities with autonomous decision support and explains why governance and human oversight should be treated as integral rather than secondary considerations. Practically, the framework can help healthcare organizations think about where greater automation may be appropriate, what controls should accompany it, and how professional accountability can be preserved. The framework also provides a structured basis for future empirical research.

### 1.7 Scope and meaning of autonomy

Autonomous healthcare decision support means bounded execution of information-processing and workflow tasks within an

approved objective. It can include assembling evidence, preparing a summary, flagging missing information, and routing a case for review. It does not imply independent professional licensure, unrestricted treatment decisions, or removal of patient participation. The unit of analysis is an AI-supported decision workflow embedded in a healthcare organization. Different applications can therefore have different permitted action spaces even when their underlying models have similar technical capabilities.

Table1 distinguishes forms of AI involvement by what the system contributes and who authorizes consequential action. The progression is analytical rather than a universal historical sequence: conventional, generative, and agentic systems may coexist. A system should not receive greater authority simply because it occupies a later category. The scope of delegation requires task-specific justification, including the ability to detect failure and restore human control.

**Table1: Forms of AI Involvement and the Separation of Capability from Authority**

| Form of AI involvement | System contribution                                                          | Authorization of consequential action                                                                    |
|------------------------|------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| <b>Conventional AI</b> | Predictions, classifications, and risk scores.                               | A healthcare professional evaluates the output and authorizes action.                                    |
| <b>Generative AI</b>   | Information synthesis, summaries, and draft recommendations.                 | A healthcare professional verifies the output and authorizes action.                                     |
| <b>Agentic AI</b>      | Planning, linked tasks, and permitted tool use within an approved objective. | The organization defines permissions; an authorized professional approves consequential clinical action. |

## 2. Literature Review

### 2.1 Agentic AI as an Emerging Healthcare Capability

Agentic AI can be understood as an emerging form of AI that performs connected, goal-

oriented tasks rather than only generating a single prediction or recommendation. Its defining capabilities include interpreting information, reasoning about objectives, planning sequences of tasks, taking permitted

actions through digital tools, and adapting to changing information. These characteristics extend the role of AI from isolated analytical support toward coordinated participation in workflows. In healthcare, the potential applications include synthesizing clinical information, supporting patient monitoring, coordinating routine processes, preparing decisions, and interacting with digital systems. However, the literature on healthcare AI consistently indicates that clinical impact depends on implementation context, workflow integration, data quality, and human interaction (He et al., 2019; Kelly et al., 2019). Agentic capability should therefore be treated as a potential organizational and technological capability rather than an automatic source of improved care. Its value depends on how safely and appropriately it is embedded in healthcare processes.

## **2.2 Autonomous Decision Support in Healthcare**

Clinical and healthcare decision support systems assist professionals in interpreting

patient information, identifying risks, and considering possible decisions. The development of more autonomous AI extends this function by enabling the system to connect information from multiple sources and perform sequential reasoning. Instead of providing a single recommendation, an agentic system may support several linked steps within a decision process. This form of decision support may improve timeliness, consistency, personalization, and workflow efficiency. However, decision quality is not determined by automation alone. Healthcare decisions often depend on patient context, uncertainty, clinical experience, communication, and ethical judgment. Explainability research also shows that users need appropriate information about AI-supported processes, especially when outputs influence consequential decisions (Amann et al., 2020; Loh et al., 2022). Autonomous decision support should therefore be conceptualized as controlled assistance in which the level of autonomy is matched to decision risk and professional review.

### 2.3 AI Governance in Agentic Healthcare

AI governance refers to the organizational structures, rules, processes, and controls used to guide the responsible development and use of AI. In healthcare, governance is especially important because AI systems may process sensitive information and influence decisions with direct consequences for patients. Core governance concerns include transparency, accountability, privacy, fairness, safety, and risk management (WHO, 2021; NIST, 2023). Agentic systems make governance more demanding because the system may perform several connected actions rather than one recommendation. Governance therefore needs to address the boundaries of permitted autonomy, monitoring, documentation, escalation, accountability, and intervention. Recent work on governance frameworks emphasizes the need for lifecycle controls and organizational responsibility as healthcare AI becomes more capable (Wang et al., 2026). Governance should not be viewed only as a restriction on innovation; it can provide the conditions under which

autonomous systems can be deployed responsibly.

### 2.4 Human Oversight in Agentic Healthcare

Human oversight becomes more important as AI moves from recommendation toward action. Effective oversight requires more than having a professional somewhere in the workflow. Professionals need the authority, information, and opportunity to review outputs, identify inappropriate behavior, intervene, and override AI-supported decisions when necessary. Oversight should also include monitoring for changes in system performance and emerging risks. The appropriate level of oversight should be proportional to autonomy and potential harm. Low-risk administrative tasks may permit greater automation, whereas high-risk clinical decisions may require stronger review and explicit human authorization. Recent research on human-in-the-loop clinical decision support highlights the importance of oversight across the AI

lifecycle rather than only at the final decision point (Bakker et al., 2026).

## **2.5 Human–AI Collaboration and Responsible Decision-Making**

Human–AI collaboration provides a useful perspective for understanding how autonomous systems can complement rather than replace professional expertise. AI can contribute rapid information processing, pattern recognition, and structured decision support, while healthcare professionals contribute contextual knowledge, clinical judgment, communication, ethical reasoning, and responsibility. Jarrahi (2018) describes this relationship as a form of human–AI symbiosis in organizational decision-making. As AI autonomy increases, the professional role may shift from completing every task to supervising, evaluating, and intervening in AI-supported processes. This shift requires clearly defined responsibilities, understandable communication of AI outputs, and meaningful authority to question or override system actions. Responsible collaboration therefore depends on

complementarity and accountability rather than unrestricted automation.

## **3. Theoretical Foundations**

### **3.1 Socio-Technical Systems Theory**

Socio-Technical Systems Theory views organizational performance as the result of interaction between technological and social elements. Healthcare is inherently socio-technical because clinical decisions depend on technologies, professionals, organizational processes, and the wider institutional environment. Baxter and Sommerville (2011) emphasize that successful systems engineering requires attention to these interacting components. The theory supports the proposed framework by explaining why agentic AI cannot be treated as an isolated technical system. Its effectiveness depends on workflow integration, role design, governance, and professional interaction. AI governance provides organizational boundaries, while human oversight connects technological

action with clinical responsibility. The theory therefore supports an integrated human–technology view of agentic healthcare.

### 3.2 Human–AI Collaboration Perspective

The Human–AI Collaboration Perspective emphasizes complementary strengths. AI can process large volumes of data and identify patterns, while humans contribute contextual understanding, professional judgment, and accountability (Jarrahi, 2018). This perspective becomes particularly relevant when AI performs increasingly complex and connected tasks. Within the proposed framework, human oversight is the mechanism through which collaboration remains meaningful. Professionals retain the ability to assess AI-supported decisions, intervene in uncertain cases, and override outputs when necessary. The perspective therefore supports the view that increased AI autonomy should change human roles, not eliminate human responsibility.

### 3.3 Responsible AI Governance Perspective

Responsible AI governance emphasizes safe, transparent, accountable, fair, and human-aligned AI. This perspective is highly relevant to healthcare because AI decisions can affect patient safety and involve sensitive personal data. WHO (2021) identifies human autonomy, well-being and safety, transparency, accountability, inclusiveness, and sustainability as important principles for AI in health, while NIST (2023) provides a structured approach to managing AI risks. For agentic healthcare, governance should define the boundaries of autonomous action, responsibilities for outcomes, monitoring requirements, and escalation procedures. This perspective therefore supports AI governance as a central control mechanism that conditions how agentic capability is translated into autonomous decision support.

### 3.4 Integrated theoretical argument

The three perspectives operate at different explanatory levels. Socio-technical systems theory establishes interdependence between

technology and organization. Human–AI collaboration explains how judgment and task responsibilities can be allocated. Responsible governance supplies normative criteria and organizational controls for that allocation. Treating governance guidance as a perspective rather than a standalone causal theory avoids overstating its explanatory status. The integrated argument is that agentic capabilities become useful through configured decision-support processes, while governance and oversight determine whether those processes remain appropriate and

controllable. Separating these elements can obscure the conditions under which greater autonomy is safe and useful. An integrated perspective is therefore appropriate for agentic healthcare. It recognizes that technical capability alone cannot guarantee better outcomes. Instead, outcomes depend on the interaction among capability, decision processes, governance, and professional oversight. This integrated logic provides the conceptual basis for the framework and propositions developed below.

#### 4. Proposed Conceptual Framework

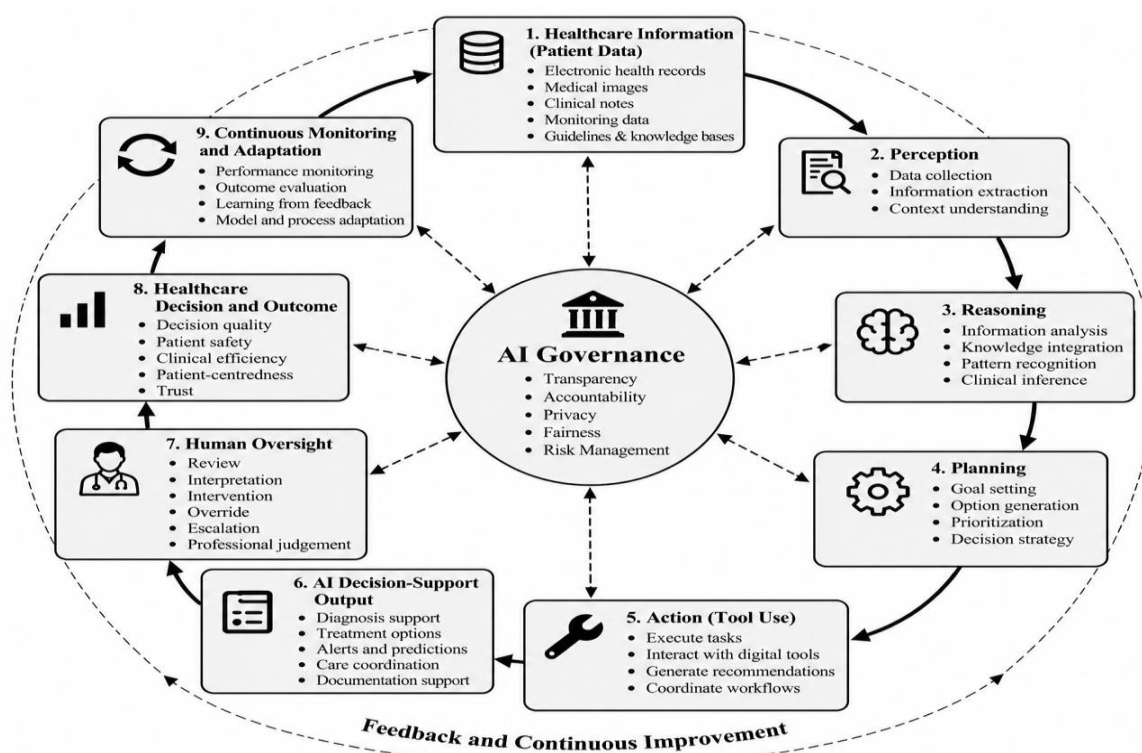


**Figure 1.** Controlled Agentic Healthcare Decision Cycle

#### 4.1 Agentic AI Capabilities

Agentic AI capabilities form the technological foundation of the framework. The construct is represented by perception, reasoning, planning, action, and adaptability. Perception refers to interpreting relevant healthcare information; reasoning refers to analyzing information and evaluating possible responses; planning refers to sequencing tasks toward a defined goal; action refers to executing permitted activities through digital tools; and adaptability refers to adjusting the process when information or circumstances change. These dimensions capture the move from single-output assistance toward coordinated, multi-step activity. They do not imply unrestricted clinical autonomy. Agentic capability must operate within defined objectives and organizational controls. The framework therefore treats capability as the driver of greater decision-support autonomy, not as a replacement for professional authority.

#### 4.2 Autonomous Healthcare Decision Support

Autonomous healthcare decision support refers to AI-enabled support for increasingly complex decision processes with limited but controlled human intervention. It is the central mechanism through which agentic capabilities are translated into practical healthcare decision support. The construct includes decision accuracy, timeliness, personalization, workflow efficiency, and decision consistency. The proposed framework does not define autonomy as complete replacement of healthcare professionals. Instead, autonomy is understood as a continuum in which the level of AI participation is matched to the complexity and potential consequences of the task. Appropriate review, escalation, and intervention remain available, particularly for high-risk decisions.

#### 4.3 AI Governance

AI governance represents the organizational mechanisms that guide the responsible development, deployment, and use of agentic

AI. Five dimensions are emphasized: transparency, accountability, privacy, fairness, and risk management. Transparency concerns the ability to document and communicate AI-supported processes; accountability identifies responsibility for decisions and actions; privacy protects sensitive patient information; fairness addresses potential bias and unequal effects; and risk management addresses monitoring, safeguards, and unintended consequences. Governance is positioned as a moderating mechanism because the same technical capability may produce different decision-support effects depending on the quality of the controls surrounding it. Stronger governance can define permitted actions, monitoring requirements, escalation pathways, and accountability, thereby creating conditions for responsible autonomy

#### 4.4 Human Oversight

Human oversight represents the mechanisms through which healthcare professionals maintain appropriate control over AI-supported decision processes. Five

dimensions are proposed: human review, intervention, override, escalation, and continuous monitoring. Review allows professionals to assess outputs; intervention allows modification or interruption; override provides authority to reject an AI recommendation; escalation transfers complex or high-risk cases to an appropriate decision-maker; and continuous monitoring identifies unexpected behavior or emerging risks. Oversight should be proportionate to autonomy and risk. It should not be designed to eliminate automation but to ensure that greater automation remains accountable and clinically appropriate. This position is consistent with human-in-the-loop approaches that treat oversight as a lifecycle governance requirement rather than a single final approval step (Bakker et al., 2026)

#### 4.5 Healthcare Decision Outcomes

Healthcare decision outcomes represent the expected consequences of effective AI-supported decision processes. The framework conceptualizes outcomes through decision quality, patient safety, efficiency, patient-

centeredness, and trust. Decision quality refers to the appropriateness and usefulness of supported decisions; patient safety concerns the reduction of avoidable risks; efficiency concerns time and resource use; patient-centeredness concerns alignment with individual needs; and trust concerns confidence in the decision process. These outcomes are not assumed to follow automatically from greater autonomy. Instead, the framework proposes that outcomes depend on how agentic capabilities are translated into decision support and how governance and human oversight shape that translation.

#### 4.6 Proposed Conceptual Framework

The framework integrates the five constructs into one model. Agentic AI capabilities provide the technological foundation; autonomous healthcare decision support is the central mechanism; healthcare

decision outcomes represent the intended consequences; and AI governance and human oversight provide complementary conditions that shape how autonomous decision support operates. The framework proposes a positive path from agentic capability to autonomous decision support and from decision support to outcomes, while positioning governance and oversight as important moderating mechanisms. This structure reflects a controlled-autonomy perspective. Stronger agentic capabilities may enable more coordinated decision support, but the value of that capability depends on governance and human control. Accordingly, the framework does not claim that agentic AI automatically improves healthcare outcomes. It instead identifies the organizational and professional conditions under which autonomy may contribute to responsible and effective decision-making.

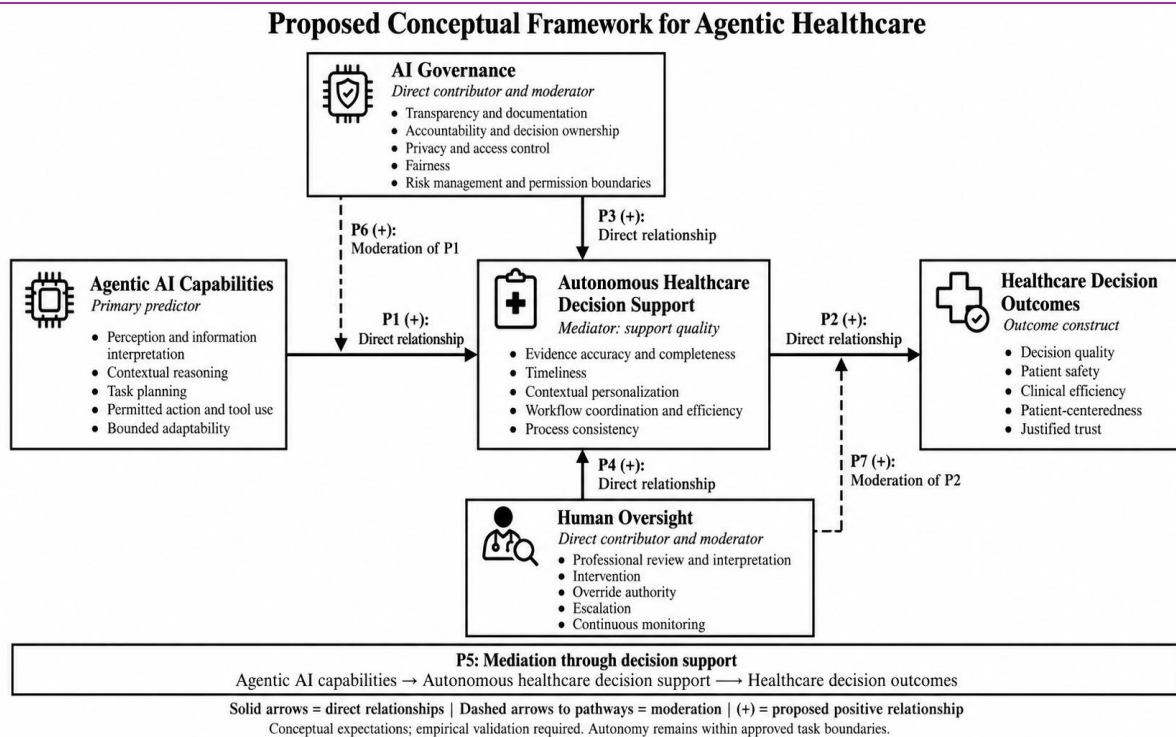


**Figure 2. Proposed Conceptual Framework for Agentic Healthcare**

## 5. Conceptual Propositions

The propositions are derived from the theoretical foundations and the relationships represented in Figure 2. They are conceptual expectations for future empirical testing and are not presented as empirical findings.

**P1:** Stronger agentic AI capabilities, reflected in perception, reasoning, planning, action, and adaptability, positively contribute to autonomous healthcare decision support by enabling more coordinated and multi-step decision-support activities.

**P2:** Stronger autonomous healthcare decision support, reflected in accuracy, timeliness, personalization, workflow efficiency, and consistency, positively contributes to healthcare decision outcomes by supporting more effective and timely decision-making.

**P3:** Effective AI governance, reflected in transparency, accountability, privacy, fairness, and risk management, positively contributes to responsible implementation of agentic AI by establishing appropriate boundaries and controls.

**P4:** Effective human oversight, reflected in review, intervention, override, escalation, and continuous monitoring, positively contributes to the quality and safety of autonomous healthcare decision support by maintaining professional control.

**P5:** Autonomous healthcare decision support mediates the relationship between agentic AI capabilities and healthcare decision outcomes by translating technological capabilities into practical and coordinated decision processes.

**P6:** AI governance moderates the relationship between agentic AI capabilities

and autonomous healthcare decision support, such that stronger governance strengthens the responsible and effective application of agentic capabilities (WHO, 2021; NIST, 2023).

**P7:** Human oversight moderates the relationship between autonomous healthcare decision support and healthcare decision outcomes, such that stronger oversight strengthens the positive contribution of autonomous decision support to outcomes. (Amann et al., 2020; Bakker et al., 2026).

**Table2: Five core constructs and uses a consistent P1–P7 mapping:**

| Proposition | Proposed relationship                                                                                     |
|-------------|-----------------------------------------------------------------------------------------------------------|
| <b>P1</b>   | Agentic AI capabilities → Autonomous healthcare decision support                                          |
| <b>P2</b>   | Autonomous healthcare decision support → Healthcare decision outcomes                                     |
| <b>P3</b>   | AI governance → Autonomous healthcare decision support                                                    |
| <b>P4</b>   | Human oversight → Autonomous healthcare decision support                                                  |
| <b>P5</b>   | Decision support mediates the relationship between agentic capabilities and healthcare decision outcomes. |
| <b>P6</b>   | AI governance moderates the P1 relationship.                                                              |
| <b>P7</b>   | Human oversight moderates the P2 relationship.                                                            |

## 6. Risk-sensitive delegation and action boundaries

The model requires an operational account of how permission changes with task risk. Risk reflects potential harm, uncertainty, reversibility, urgency, and available recovery options. This paper proposes an ordinal delegation logic, not validated numerical thresholds. Low-consequence, reversible support tasks may permit execution with monitoring. Tasks that could alter patient care require stronger verification and explicit professional authorization. Unexpected conflicts, missing evidence, or actions beyond the permitted scope should trigger escalation rather than improvised expansion

of authority. Figure 3 illustrates this logic as a decision gate. A request is assessed against its approved scope and evidence conditions before any consequential action proceeds. Escalation routes include requesting clarification, returning control to a clinician, or stopping the workflow. The framework does not prescribe a universal category for particular medical activities; local clinical evaluation must determine applicable boundaries. The analytical principle is that permissible autonomy decreases when consequence and uncertainty increase without adequate detection and recovery capacity. Table 3 translates this principle into illustrative controls and evaluation criteria.

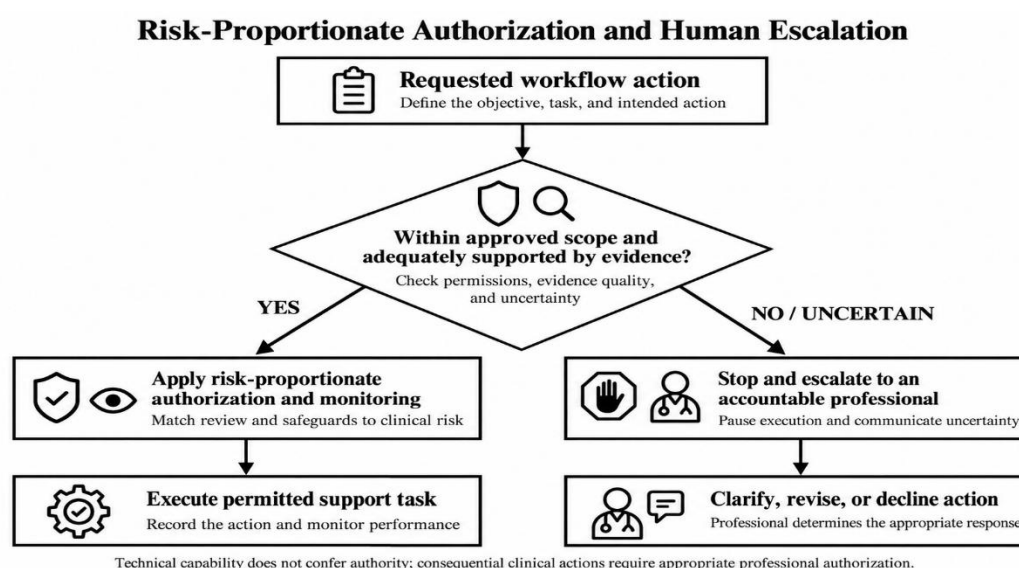


**Figure 3:** Risk-sensitive delegation and escalation gate

**Table 3 Illustrative delegation controls and evaluation criteria**

| Task condition                                  | Proposed delegation arrangement                                           | Evaluation focus                                                          |
|-------------------------------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------|
| Reversible information preparation              | Permit bounded execution; retain source records and monitoring.           | Completeness, correction burden, and unintended information exposure.     |
| Support influencing consequential care          | Require evidence verification and accountable professional authorization. | Decision appropriateness, review quality, and authorization compliance.   |
| Uncertain, conflicting, or out-of-scope request | Stop execution; return control through a defined escalation route.        | Escalation sensitivity, response time, and safe recovery.                 |
| Changed system or workflow conditions           | Reassess permissions and performance before continuing delegation.        | Drift detection, subgroup effects, and effectiveness of approved changes. |

## 7. Research Methodology

### 7.1 Research Design

This study adopts a conceptual research design based on an integrative review and synthesis of literature. The approach is appropriate because the purpose is to clarify an emerging phenomenon, identify relationships among constructs, and develop theoretically grounded propositions rather than estimate effects from primary data. The study synthesizes literature concerning healthcare AI, agentic AI, autonomous decision support, governance, human–AI collaboration, and human oversight. No primary data are collected and no statistical relationships are tested in the present paper.

P1–P7 are therefore conceptual propositions rather than empirical findings. This distinction is important because the framework is intended to guide future research, not to claim evidence that has not yet been generated.

### 7.2 Literature Search and Selection

The literature base should prioritize peer-reviewed research identified through major academic databases such as Scopus, Web of Science, PubMed, ScienceDirect, SpringerLink, IEEE Xplore, and Emerald. Search terms should combine the main constructs, including “agentic AI,” “AI agents in healthcare,” “autonomous healthcare decision support,” “clinical

decision support,” “AI governance,” and healthcare decision outcomes. Within “responsible AI,” “human oversight,” these constructs, recurring themes include “human–AI collaboration,” and “healthcare autonomy, reasoning, accountability, decision-making.” Boolean operators can be transparency, safety, privacy, intervention, used to combine related concepts. Recent decision quality, and human–AI studies should receive priority because collaboration. The analysis compares how agentic and generative AI are rapidly these themes are treated across the literature evolving, while foundational studies should and identifies relationships that support the be retained where they provide established proposed framework. Thematic synthesis is theoretical or ethical foundations. Sources used to move from individual findings and should be included when they directly theoretical perspectives to an integrated contribute to one or more core constructs or conceptual model. Socio-technical systems explain a relevant theoretical relationship. theory explains the interaction of technology and organizational context; the human–AI Records with limited healthcare relevance, collaboration perspective explains weak conceptual contribution, or duplication complementary human and AI roles; and should be excluded. The selected literature responsible AI governance explains the should then be organized around the controls required for safe and accountable framework constructs and synthesized thematically. deployment. Together, these perspectives support the logic underlying P1–P7.

### **7.3 Literature Coding and Thematic Synthesis**

The synthesis organizes the literature around five principal constructs: agentic AI capabilities, autonomous healthcare decision support, AI governance, human oversight,

### **8. Limitations and Future Research**

The framework is grounded in a supplied, bounded reference set. This limits comprehensiveness and particularly the depth of agent-specific technical evidence.

Established healthcare AI literature supports the implementation and governance foundations, but the proposed agentic extensions require direct validation. The paper makes no universal claim that the seven relationships are positive in every clinical setting. Data quality, infrastructure, professional capacity, and patient circumstances may alter their strength or direction. The framework also abstracts from differences between organizational settings and specialties. A hospital monitoring workflow and a community care-coordination workflow may have different urgency, integration needs, and oversight resources. Constructs such as governance and trust can be interpreted differently across institutions, requiring contextual refinement. Finally, the model cannot determine legal responsibility or clinical authorization for a particular deployment. Its purpose is conceptual explanation and research guidance, not a substitute for local clinical assessment or evidence of readiness for autonomous care.

Future research can empirically examine the framework using a quantitative survey of healthcare professionals, managers, or other knowledgeable respondents. Questionnaire items can be adapted from validated measures and refined through expert review and pilot testing. The proposed constructs can then be evaluated using PLS-SEM because the framework includes direct, mediation, and moderation relationships. Future testing should assess the effect of agentic AI capabilities on autonomous decision support, the effect of autonomous decision support on healthcare outcomes, the mediating role of decision support, and the moderating roles of governance and human oversight. The study should avoid reporting empirical coefficients, significance levels, or proposition decisions until actual data have been collected and analyzed.

## 9. Conclusion

This conceptual study develops an integrated framework for understanding the transition from AI assistance to agentic healthcare. The framework positions agentic AI

---

capabilities—perception, reasoning, continuous monitoring. These mechanisms planning, action, and adaptability—as the technological foundation for increasingly autonomous healthcare decision support. Autonomous decision support is positioned as the mechanism through which these capabilities may contribute to decision quality, patient safety, efficiency, patient-centeredness, and trust. The framework also emphasizes that greater AI autonomy should not be equated with unrestricted automation. AI governance provides organizational and ethical boundaries through transparency, accountability, privacy, fairness, and risk management. Human oversight provides professional control through review, intervention, override, escalation, and healthcare decision-making. These mechanisms are essential because healthcare decisions involve uncertainty, patient-specific context, and potentially significant consequences. Seven conceptual propositions capture the proposed direct, mediation, and moderation relationships. The contribution of the paper lies in integrating technological autonomy with governance and meaningful human oversight rather than treating agentic capability as an independent technological solution. Future empirical research using survey data and PLS-SEM can test the framework and determine whether, and under what conditions, agentic healthcare contributes to responsible and effective healthcare decision-making.

## References

1. Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1), 4–17. <https://doi.org/10.1016/j.intcom.2010.07.003>
2. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>
3. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25, 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
4. He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 25, 30–36. <https://doi.org/10.1038/s41591-018-0307-0>
5. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. <https://doi.org/10.1136/svn-2017-000101>
6. Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2, 719–731. <https://doi.org/10.1038/s41551-018-0305-z>
7. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17, 195. <https://doi.org/10.1186/s12916-019-1426-2>
8. Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., et al. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, 25, 1337–1340. <https://doi.org/10.1038/s41591-019-0548-6>

9. Morley, J., Machado, C. C. V., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine*, 260, 113172. <https://doi.org/10.1016/j.socscimed.2020.113172>
10. Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20, 310. <https://doi.org/10.1186/s12911-020-01332-6>
11. Loh, H. W., Ooi, C. P., Seoni, S., Barua, P. D., Molinari, F., & Acharya, U. R. (2022). Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022). *Computer Methods and Programs in Biomedicine*, 226, 107161. <https://doi.org/10.1016/j.cmpb.2022.107161>
12. Jung, J., Lee, H., Jung, H., & Kim, H. (2023). Essential properties and explanation effectiveness of explainable artificial intelligence in healthcare: A systematic review. *Heliyon*, 9(5), e16110. <https://doi.org/10.1016/j.heliyon.2023.e16110>
13. Ali, S., Akhlaq, F., Imran, A. S., Kastrati, Z., Daudpota, S. M., & Moosa, M. (2023). The enlightening role of explainable artificial intelligence in medical & healthcare domains: A systematic literature review. *Computers in Biology and Medicine*, 166, 107555. <https://doi.org/10.1016/j.compbimed.2023.107555>
14. Iniesta, R. (2023). The human role to guarantee an ethical AI in healthcare: A five-facts approach. *AI and Ethics*, 5, 385–397. <https://doi.org/10.1007/s43681-023-00353-x>
15. Sallam, M. (2023). The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. *Healthcare*, 11(6), 887. <https://doi.org/10.3390/healthcare11060887>

16. World Health Organization. (2021). Ethics and governance of artificial intelligence for health. World Health Organization.  
<https://www.who.int/publications/i/item/9789240029200>
17. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). U.S. Department of Commerce.  
<https://doi.org/10.6028/NIST.AI.100-1>
18. Wang, A., Freeman, S., & Magrabi, F. (2026). Governance for safe and responsible AI in healthcare organisations: A scoping review of frameworks. *npj Digital Medicine*, 9, 516. <https://doi.org/10.1038/s41746-026-02679-2>
19. Bakker, M., Van Garderen, A., Paget, T., Zhao, L., Naicker, L., Radhakrishnan, A., & Bidargaddi, N. (2026). Human in the loop in AI-enabled clinical decision support: A systematic scoping review and reporting checklist for lifecycle governance. *International Journal of Medical Informatics*, 222, 106701.  
<https://doi.org/10.1016/j.ijmedinf.2026.106701>
20. Meskó, B., & Topol, E. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digital Medicine*, 6, 120.  
<https://doi.org/10.1038/s41746-023-00873-0>
21. Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28, 31–38. <https://doi.org/10.1038/s41591-021-01614-0>