Agentic AI for Modern Healthcare: A Comprehensive Review

Dr. Prerna Agrawal
GLS University, FCAIT-PG, Ahmedabad, India.

Abstract: *The demand for healthcare around the world is being amplified as a result of rising patient demand, an aging population, workforce shortages, and personalized healthcare. Traditional forms of AI have been successfully used in disciplines like medical imaging, diagnosis, and decision support but only complete discrete tasks, whereas there is no ability to understand the whole context in which they are being used or to integrate with regular healthcare operational tasks. Agentic AI enables smart systems to operate autonomously, establish objectives, and apply skills such as comprehension, reasoning, planning, memory, and task completion in complex healthcare environments. This paper provides an overview of the various applications of agentic AI in healthcare by reviewing the architectural frameworks, empirical validations, and specific clinical applications. Healthcare agents will be grouped by how smart they are and how they are used; recent research will be summarized, and important areas for more study—like testing in real healthcare settings, setting standards, getting regulatory approvals, and ensuring safety for agentic AI—will be pointed out. The results of this study show that agentic AI has a strong potential to help with clinical decision support, health care analytics, biomedical research, and providing quality long-term care. However, agentic AI has not yet achieved widespread use in routine practices because of insufficient evaluation of its effectiveness through the use of standardized methods. Lastly, this study offers avenues for future research to develop agentic AI systems for future healthcare delivery that are safe, scalable, interoperable, and ethically congruent.*

Key Words: *Agentic AI, healthcare AI, medical agents, clinical decision support, multi-agent systems, retrieval-augmented generation (RAG), healthcare workflow automation, artificial intelligence in medicine, and autonomous healthcare systems.*

I. INTRODUCTION

Health systems globally are facing significant pressures in terms of rising patient volumes, an aging population, a shortage of qualified staff, and demand for greater personalization of healthcare, which are all contributing to these pressures. Many areas of healthcare, such as medical imaging, disease diagnosis, risk prediction, and administrative decision support, are beginning to make use of artificial intelligence (AI), which has demonstrated the potential to complement many existing healthcare services. However, many of the traditional AI applications are reactive, task-based systems that lack an understanding of context, do not easily adapt to changing circumstances, and cannot integrate into clinical workflows or processes without complication. Large language models (LLMs), the use of multiple types of data, and the emergence of new tools have led to the development of a new category of AI in healthcare, known as agentic AI. This AI can act autonomously, relying on its own sensory information, to achieve specific goals. Agentic AI can perform different tasks on its own or work together like a virtual team of healthcare providers, using a patient's ongoing health record and other data from electronic health records (EHR), imaging studies, patient feedback, wearable devices, and clinical guidelines.

Agentic systems have recorded superior performance to traditional language models when completing complex clinical tasks. Some real-world examples of agentic AI have enhanced healthcare analysis by automatically accessing data and offering useful insights; they also aid advanced biomedical research by using retrieval-augmented generation (RAG) frameworks. In some cases, such as the use of agentic AI for geriatric patient care, researchers have expressed concern over ethical safeguards and emphasized there is a need to prioritize patients when using agentic systems. While some early chatbots hinted at the rise of agentic AI, they went through fewer safety checks and were not as well integrated into clinical practices as most AI applications today. There has been a lot of research on creating categories and rules for using agentic AI, but we still don't agree on the standards needed for using it in real life, measuring its performance, getting regulatory approvals, or ensuring its safety in the mentioned areas. The quick rise of agentic AI, along with the lack of research, indicates that we need to bring together what we already know and create a system that guarantees the safe and effective use of agentic AI in healthcare. Figure 1 shows country-wise usage of agentic AI in healthcare. The country with the largest amount of agentic AI research/application in healthcare is India at 33.3%, followed by the USA and China/Hong Kong (each at 22.2%), and the UK and Australia (each at 11.1%).

The main goals of this paper are: (1) To carefully examine and combine current research on agentic AI in healthcare, including its design frameworks, real-world tests, and specific uses. (2) To compare and group healthcare agentic systems based on how much freedom they give users, how they are built, how strong their validation is, and how they are governed. (3) To identify critical research gaps related to clinical validation, safety evaluation, regulatory frameworks, benchmarking standards, and workflow integration. (4) To propose future research directions for developing safe, scalable, interoperable, and ethically

aligned agentic AI systems in healthcare. (5) To provide a structured foundation that supports researchers, clinicians, and policymakers in advancing next-generation autonomous healthcare systems.
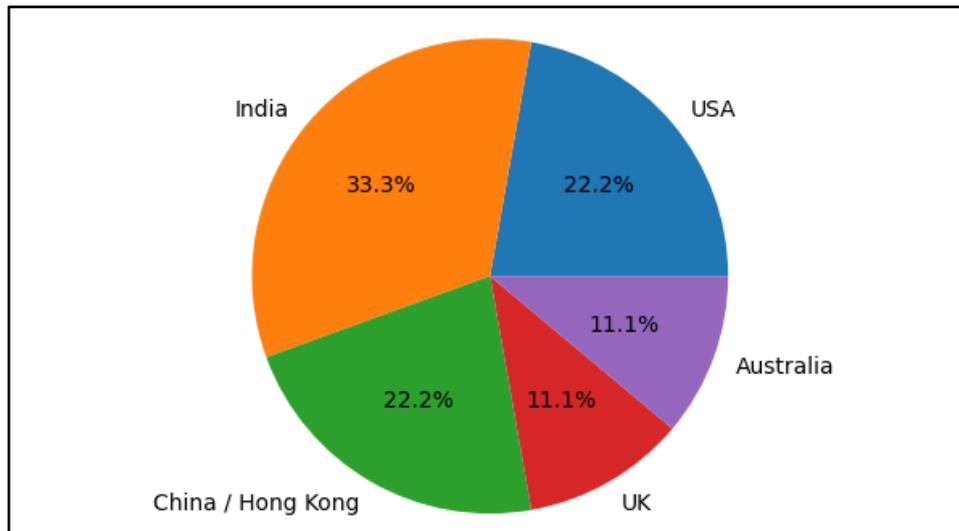

Figure 1: Country-wise use of Agentic AI in healthcare [1-9]

The paper is divided into the following sections. Section 2 represents the classification of the different healthcare agents. Section 3 shows the literature review, and section 4 illustrates the conclusion and the future scope.

II. CLASSIFICATION OF HEALTHCARE AGENTS

Figure 2 illustrates the types of AI agents within the healthcare setting. It also shows two main ways to classify health-related AI agents: (1) by how smart they are and (2) by what they do in healthcare, to create more advanced AI systems.

Classification by Intelligence Level

- Simple reflex agents follow a condition-action rule based on their current state. An example of a simple reflex agent would be an abnormal vital sign alerting system that will send alerts based on a defined threshold [8].
- Model-based agents use an internal representation of the patient to make contextual decisions in addition to immediate decision-making based on their state, especially in complex scenarios [1].
- Goal-based agents take action to achieve clearly defined clinical goals, including optimizing treatment or recovery [5].
- Utility-based agents examine the possible outcomes of their actions and select the one that produces the highest clinical outcome and lowest risk in the case of medication dosage [6].
- Learning agents improve their performance through the receipt of feedback from the environment and by using techniques such as reinforcement learning or adaptive training [4].

The characteristics of these agents demonstrate an evolutionary progression from reactive systems to adaptive intelligent agents capable of complex, long-term reasoning.
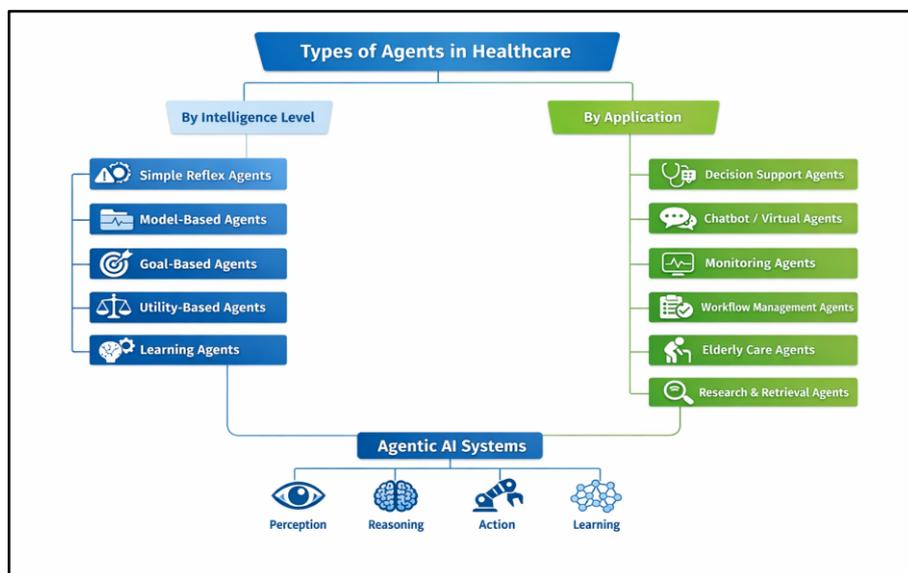

Figure 2: Types of Agents in Healthcare [1-9]

Classification by Application in Healthcare

- Decision support agents assist with diagnosis and treatment planning [1, 6].
- Chatbots, or virtual agents, interact with patients, providing triage and mental health support [8].
- Monitoring agents continuously monitor a patient's health status while in an intensive care unit or remotely until the patient is discharged from those settings [1].
- Workflow management agents automate the scheduling of patients within healthcare systems (e.g., hospitals), documentation of patient encounters, and allocation of resources [9].
- Elderly care agents help individuals remain in their home by assisting with medication compliance, helping them remain independent, and providing cognitive support [3].
- Research and retrieval agents assist clinicians in the research and retrieval of pertinent medical literature to provide evidence-based decision-making [7].

These classifications illustrate how agents operate in the clinical, administrative, and research areas of the healthcare continuum.

Agentic AI Systems

The bottom of the diagram also illustrates that agentic AI systems integrate the perception, reasoning, action, and learning capabilities of the agents discussed earlier into one system. Agentic AI systems utilize structured cycles of operation to process information, such as sense–think–act–learn and perceive–plan–act [1, 5, and 9]. Agentic systems are different from regular reactive agents because they can combine different types of information, use outside tools, have memory, and change based on the healthcare situation. The next section shows the literature review.

III. LITERATURE REVIEW

This section shows the literature review of different studies done by the authors. Paper [1] reviews the concept of agentic AI in healthcare and presents a taxonomy to categorize the different types of agentic AI. The authors provide an architectural framework to develop effective agentic AI systems in the area of healthcare using a sense/thinking/acting/learning framework. Additionally, the authors show how AI is evolving from merely making predictions to becoming independent agents that can process and understand various types of data from different sources, analyze the context of that data, utilize tools, and retain the information. The authors present a review of different systems, grouping them into three types: single-agent, multi-agent, and human/AI hybrid architectures used in clinical settings. However, the review does not provide empirical validation through real-world clinical experience for the proposed agentic AI systems. Paper [2] moves beyond theoretical evaluation and describes a healthcare application utilizing agentic AI to create a data product to predict patient no-show appointments. The author utilizes a Random Forest predictive model and enhances it with a layer for autonomous insight generation—resulting in a prediction accuracy of 77.33%. This project demonstrates the capacity of agentic AI to enhance the ability to interpret data and discharge automated decision-support insights. However, the way the system is set up means it can only provide automated responses after making predictions, and it doesn't have features like working with others. Paper [3] reviews how agentic AI systems will affect the care of older adults. The authors note how LLMs can be used to build agentic AI systems providing proactive monitoring, personalized assistance, and cognitive support to older adults. The authors discuss significant ethical and privacy issues, as well as human-centered design considerations, when deploying agentic AI systems with at-risk older populations. Although the authors present an optimistic vision of using agentic AI to change the way we provide care to older adults, the review does not provide empirical data to support the developed concept as a prototype within the domain of older adult care. Paper [4] gives a broader overview of agentic AI and explains how it differs from traditional AI systems, highlighting features of agentic systems like independence, goal-oriented actions, and flexible thinking. Although the discussion surrounding agentic AI in healthcare is an important application domain, the analysis is more theoretical than model-based or experimental, as it does not include modeling healthcare workflows or standardized evaluations. This analysis helps to clearly explain the differences between AI agents and agentic AI systems. Paper [5] focuses on how agentic AI can be used in biomedical fields and proposes a system that includes a process of observing, planning, and taking action. The authors suggest that future development of agentic AI systems in biomedical applications will be based on foundational models. The model includes memory modules, retrieval-augmented generation (RAG), tools, and governance mechanisms. The authors also present the need for uncertainty, human oversight, and provenance when developing agentic AI systems in high-stakes medical environments. The agents in the capability-centered framework are well-made, but there isn't a clear way to evaluate them consistently, and there's no research showing that agentic AI systems perform better in large clinical settings than claimed by the capability-centered framework. Paper [6] presents research demonstrating the effectiveness of using agentic AI agents in providing clinical care. The authors reviewed 20 studies on how well agentic AI works in healthcare, and the results showed that these systems did better than a standard LLM-based approach by an average of 53 percentage points. Additionally, the results of the systematic review suggest that maximum performance is achieved when there are matching levels of complexity for the agent and the clinical task. However, the authors noted that the nonsystematic design and outcome measures of the studies prevent a quantitative meta-analysis from being completed. They also pointed out that there aren't any common metrics for assessing how well agentic AI systems perform. The paper [7] shows that using publicly available datasets with agentic retrieval-augmented generation (RAG) models greatly improves medical research by making literature searches more accurate and thorough. The findings attest to the feasibility of RAG systems acting to enhance clinical research workflows and decision-making processes. Unfortunately, the full incorporation of RAG systems into the real-time clinical environment and bedside decision support systems has yet to be investigated. The review in paper [8] found 17 studies on chatbots that assist with self-care and patient involvement, including randomized controlled trials. Almost all systems were finite-state or frame-based, limiting their autonomy vs. current agentic architectures. Lastly, safety evaluations were performed infrequently;

therefore, early systems were installed without appropriate governance frameworks. Finally, paper [9] presents a three step plan that describes how to help knowledge-based agents, combine workflows with decision-making, and carry out semi-autonomous workflows. Medical agents are positioned to provide the infrastructure for a learning healthcare system with an emphasis placed on redesigning the workflow rather than just focusing on performance via benchmarking. Although the roadmap provides direction, the establishment of standardized deployment protocols or evaluation benchmarks remains outstanding. Table 1 shows a comparative analysis of agentic AI studies in healthcare. The next section represents the conclusion and the future directions of the study.

Paper No	Paper Type	Core Technical Focus	Level of Agent Autonomy	Governance/ Ethics Coverage	Key Research Contribution	Identified Research Gap
[1]	Survey (Healthcare specific)	Taxonomy, architectures, multimodal reasoning, sense–think–act–learn framework	High (Single, Multi agent, Hybrid systems)	Strong (bias, safety, interoperability, accountability)	Multi dimensional taxonomy and architectural mapping	Lack of real-world deployment validation
[2]	Experimental Study	Random Forest + autonomous insight generation layer	Low Moderate (Post prediction autonomous insights)	Limited	Practical demonstration of agentic insight automation	No multi-agent architecture or longitudinal evaluation
[3]	Domain Specific Survey	LLM-based agentic AI for elderly monitoring & care	Moderate (Human centered agentic support)	Very Strong (privacy, autonomy, ethics)	First focused survey on elderly-care agentic AI	Lack of empirical implementation studies
[4]	Broad Cross-Domain Survey	Definition of agentic AI vs traditional AI, autonomy principles	Conceptual High-Level	Moderate Strong	Clear theoretical distinction between AI agents & agentic AI	Healthcare specific implementation depth limited
[5]	Biomedical Technical Survey	Perceive–Plan–Act loop, RAG, foundation models, tool use	High (Governable biomedical agents)	Strong (uncertainty, provenance, oversight)	Capability centered biomedical agent design pathway	Clinical validation still emerging
[6]	Systematic Review (PRISMA based)	Agent vs Base-LLM performance comparison	Moderate High (Single and Multi-agent systems)	Moderate	Empirical evidence that agents outperform LLMs	Heterogeneity prevents meta-analysis
[7]	Experimental + Review	Retrieval Augmented Generation (RAG) for medical research	Moderate (Research support agent)	Moderate (bias & privacy concerns)	Shows RAG improves medical literature retrieval	Limited real-time clinical deployment

[8]	Systematic Review	NLP-based conversational agents	Low (mostly finite-state/ frame-based)	Weak (Safety rarely evaluated)	Early evidence base for healthcare chatbots	Limited autonomy & safety evaluation
[9]	Comprehensive Survey	Developmental roadmap (Knowledge → Workflow → Semi autonomous execution)	High (Workflow integrated and semi autonomous agents)	Strong (regulation, workflow governance)	Structured 3-level roadmap for medical agents	Lack of standardized evaluation benchmarks

Table 1: Comparative Analysis of Agentic AI Studies in Healthcare

IV. CONCLUSION AND FUTURE SCOPE

The current study looked at how healthcare-focused agentic AI has changed from simply reacting to tasks to becoming independent systems that can see, think, plan, act, and learn on their own. The study shows that agentic systems vary based on how they are used and proves that today's smart agent systems are more effective than older decision-support systems, working well with people in clinical, administrative, and research tasks. Research provides evidence that agent-based systems provide higher-quality solutions for complex clinical tasks when compared to conventional large language model-based systems. Evidence of the potential for meaningful change can be seen in the practical applications of agentic AI in healthcare analytics and biomedical research, elder care, and related health domains.

While agentic AI has significant potential, there are still numerous barriers that limit widespread adoption. Research should concentrate on creating and confirming standard measures for agentic AI systems, agreeing on how to measure their performance, ensuring safety checks for these systems, including regulations, finding ways for agentic AI systems to work with electronic health records, reducing bias in agentic AI, setting up ethical oversight groups, and creating uniform methods for evaluating these systems so that different studies can be compared.

Future studies should incorporate multiple clinical settings to test and evaluate the reliability and safety of agentic AI over time. We need to create reliable data sets that allow us to keep track of how well agentic AI is performing, and we must follow rules that match the regulations for checking how effective agentic AI is. In addition, the study of enhancing teamwork among multiple agents, reinforcement learning, adaptive memory, and explainability will contribute to improving personalized and long-lasting patient care. Finally, understanding how to integrate agentic AI effectively into the processes of healthcare providers while addressing transparency, accountability, and patient trust is essential to ensuring the successful development of agentic AI within the healthcare industry. We should also conduct economic evaluations and cost-effectiveness studies to assess the effectiveness of agentic AI in various healthcare systems. In general, agentic AI can greatly change how healthcare services are provided, but to use it safely and ethically on a large scale, we need thorough testing, effective management rules, and teamwork among different parts of the healthcare system.

References

1. S. Banerjie, Y. Zhu, I. Freeman, J. Villa Machado, A. Ahmed, A. Sarker, and M. Al-Garadi, "Agentic AI in healthcare: A comprehensive survey of foundations, taxonomy, and applications," *TechRxiv*, preprint, 2025, doi: 10.36227/techrxiv.176238073.31262603.
2. S. Murumkar, "Agentic AI-driven data product for automated healthcare insights generation," *Int. J. Artif. Intell., Data Sci., Mach. Learn.*, vol. 6, no. 1, pp. 226–230, 2025, doi: 10.63282/3050-9262.IJAIDSML-V6I1P125.
3. R. A. Khalil, K. Ahmad, and H. Ali, "Redefining elderly care with agentic AI: Challenges and opportunities," *IEEE Open J. Comput. Soc.*, vol. 7, pp. 326–? 2026, doi: 10.1109/OJCS.2026.3650842.
4. A. K. Pati, "Agentic AI: A comprehensive survey of technologies, applications, and societal implications," *IEEE Access*, vol. 13, pp. 151824–? 2025, doi: 10.1109/ACCESS.2025.3585609.
5. A. Abdollahi, M. A. Rezaei, X. Wang, S. He, S. M. Ayyoubzadeh, Y. Nie, and H. Chen, "The next paradigm in medical AI: A survey of agentic AI in biomedicine," *TechRxiv*, preprint, 2025, doi: 10.36227/techrxiv.176472621.18843904.
6. A. Gorenshtein, M. Omar, B. S. Glicksberg, G. N. Nadkarni, and E. Klang, "AI agents in clinical medicine: A systematic review," *medRxiv*, preprint, 2025, doi: 10.1101/2025.08.22.25334232.
7. A. Singh, "Future prospects of agentic RAG in medical research," *SSRN Electron. J.*, working paper, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5212343
8. L. Laranjo, A. G. Dunn, H. L. Tong, A. B. Kocaballi, J. Chen, R. Bashir, D. Surian, B. Gallego, F. Magrabi, A. Y. S. Lau, and E. Coiera, "Conversational agents in healthcare: A systematic review," *J. Amer. Med. Inform. Assoc.*, vol. 25, no. 9, pp. 1248–1258, 2018, doi: 10.1093/jamia/ocy072.