Medicine Plus 2 (2025) 100084

Medicine Plus

Perspective

Harnessing large language model agents for healthy aging

Pengfei Li[a,b,c,*,1], Jingyi Wu[c,1], Shaomei Shang[b,*], Qimin Zhan[a,*]

[a] National Institute of Health Data Science, Peking University, Beijing 100191, China

[b] School of Nursing, Peking University, Beijing 100191, China

[c] Advanced Institute of Information Technology, Peking University, Hangzhou 311215, China

The global population is aging rapidly, with those aged 65 years and older projected to reach 2.2 billion by the 2070s, surpassing the number of children.[1] This demographic shift presents profound healthcare and socioeconomic challenges, driven by the growing burden of chronic diseases such as diabetes, cardiovascular conditions, and neurodegenerative disorders among the elderly. Managing these conditions requires continuous monitoring and proactive intervention; however, traditional healthcare systems relying on in-clinic visits and periodic testing are inadequate to meet the growing demands of an aging population.[2] This widening gap highlights the urgent need for innovative solutions that enhance healthcare efficiency and effectiveness, to ensure healthy aging and improved quality of life.

Emerging large language model (LLM) technology offers a promising path for transforming geriatric care.[3] LLM-driven agents, equipped with perception, memory, reasoning, and interaction modules, have demonstrated exceptional performance in solving complex tasks within specific environments. These agents can integrate comprehensive medical knowledge, vast real-world healthcare data, and advanced reasoning and natural language processing capabilities to improve healthcare accessibility by automating various clinical tasks.[4] For example, they can provide real-time and personalized health guidance, monitor chronic health conditions, and support medication adherence and daily self-care.[5] Additionally, LLM-driven systems can alleviate the burden on healthcare professionals by automating routine tasks, improving decision-making, and enhancing the overall efficiency of geriatric care delivery.

Here, we provide an overview of the current advancements, ongoing challenges, and future prospects in leveraging LLM agents to enhance geriatric care and promote healthy aging.

LLM agents for healthy aging. In the context of healthy aging, the primary goal of LLM-driven agents is to enable older adults to live safely and comfortably in their homes and communities while benefiting from continuous medical supervision through virtual services. Numerous studies have explored the

* Corresponding authors.
E-mail addresses: pfli@bjmu.edu.cn (P. Li), mei916@263.net (S. Shang), zhanqimin@bjmu.edu.cn (Q. Zhan).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.medp.2025.100084
Received 27 February 2025; Received in revised form 18 March 2025; Accepted 7 April 2025
Available online 14 April 2025

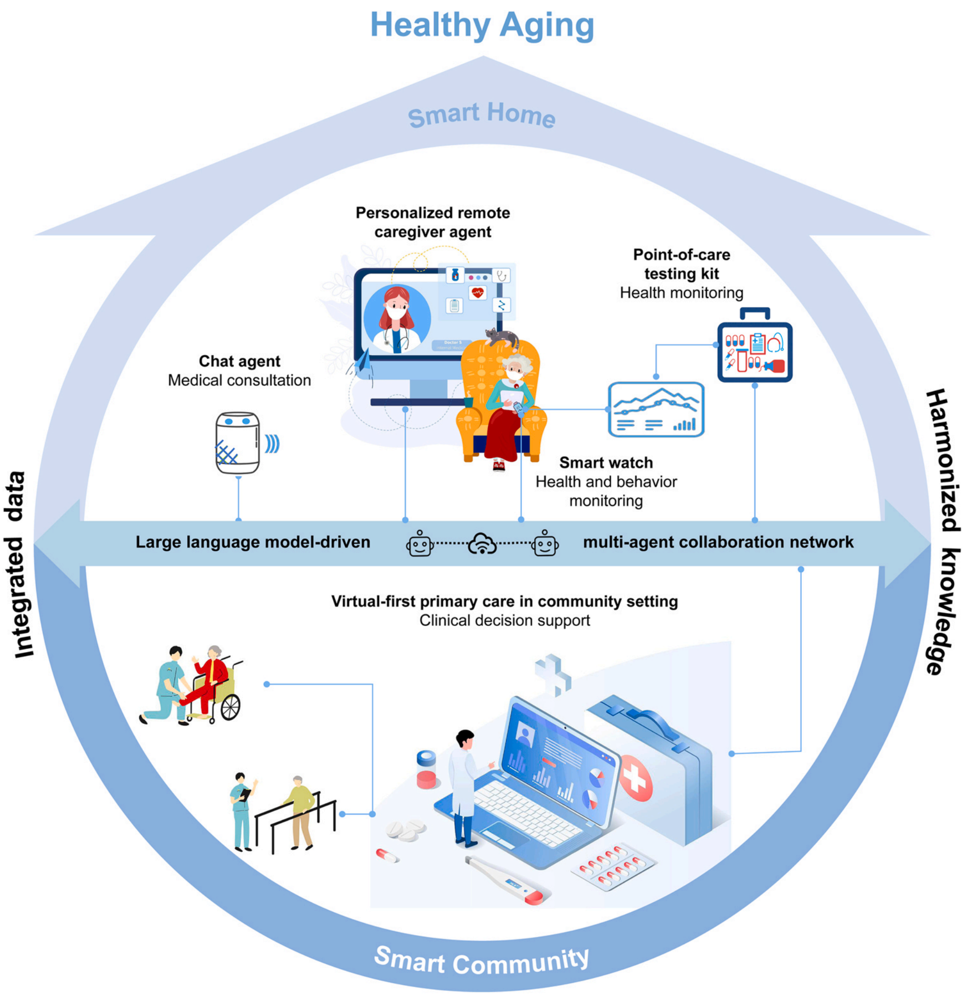

Fig. 1. LLM-driven healthy aging in home and community settings. LLM, large language model.

adoption of LLM agents to optimize the full spectrum of geriatric care, from preventive measures to ongoing management and support. These agents can form a collaboration network that integrates health data and decision-making knowledge to deliver personalized geriatric care, contributing to the development of smart homes and communities for healthy aging (Fig. 1).

LLM agents have significantly transformed health education and medical consultations for older adults by leveraging advanced information retrieval and natural language generation capabilities. These agents can distill complex medical knowledge into clear, accessible language, reducing both manpower and economic costs while broadening the reach and effectiveness of health education and medical consultation. This helps patients gain a deeper understanding of disease prevention, healthy lifestyles, and strategies for managing common health issues. Many studies have demonstrated that LLM-based agents used for remote health education and online consultations can substantially reduce the workload on medical staff, enhance the consistency of health education, and improve both the clarity and educational value of communication, ultimately delivering higher-quality medical support.[6] For example, Parmanto et al.[7] proposed a caregiving language model (CaLM), which can be deployed as front-line help tools for geriatric caregivers.

Combining body-worn sensors and electronic surveillance devices, LLM agents can continuously monitor chronic health conditions for the aged population in home settings and deliver personalized health management advice based on each patient's unique data and behavioral patterns.[8] Such capabilities foster a proactive approach to personal health management and improve adherence to prescribed interventions, ultimately enhancing outcomes and reducing the overall healthcare burden. For instance, in the management of hypertension, LLM agents can prompt regular monitoring of key physiological indicators, perform trend analyses, and generate intuitive feedback reports. This data-driven process allows for tailored dietary and physical activity recommendations, as well as scientifically grounded guidance for clinical interventions. When abnormal fluctuations occur, the agents can proactively conduct an initial risk evaluation and synthesize the patient's medical history, recent symptoms, and physiological data into detailed reports for physician review, thereby facilitating timely adjustments to treatment plans. An illustrative case is the Personal Health Insights Agent (PHIA) proposed by Google.[9] PHIA seamlessly integrates multidimensional, continuous personal health data from wearable devices and leverages advanced code generation and information retrieval techniques for precise data analysis. This integration enables it to deliver personalized health recommendations and support patients in managing their health from home.

In addition to advancements in promoting self-care at home, LLM agents can also help streamline the workflow of primary care in community settings, which is crucial for the elderly. By automating routine clinical tasks and providing personalized intervention guidance, LLM agents improve access to and the quality of community-based healthcare for older adults, particularly in rural or underserved areas.[10,11] Recent studies have actively explored the "virtual-first" primary care model driven by LLM agents to enhance healthcare efficiency and improve patient experiences. This approach leverages LLM agents to assist healthcare professionals in conducting initial consultations and triage, enabling rapid processing and analysis of patients' symptom descriptions and medical histories. An example is the voice-interactive agent Talk2Care, which efficiently collects health information through patient-side voice assistants and generates professional, structured medical summaries for healthcare providers.[10] This simplifies communication, improves consultation efficiency, and reduces barriers for elderly patients. Moreover, LLM agents have shown significant potential in clinical decision support by assisting physicians in disease diagnosis, risk prediction, and treatment planning. For instance, Google's Articulate Medical Intelligence Explorer (AMIE) excels in medical diagnostic reasoning, enhancing the quality of diagnostic and management decisions, particularly for complex cases in underserved communities.[12] ZODIAC, an LLM-powered multi-agent collaboration framework with cardiologist-level professionalism, has been successfully integrated into electrocardiography devices to assist cardiological diagnostics.[13] Du et al.[14] evaluated the effectiveness of LLM agents in enhancing early detection of cognitive decline in the elderly using clinical notes. Another noteworthy multimodal agent, the Health Large Language Model for Multimodal Understanding (HeLM), can integrate diverse patient data, including demographic, clinical, and high-dimensional temporal features, to assess and predict disease progression risks.[15] These advancements highlight the transformative role of LLM agents in improving primary care delivery for older adults.

In precision medicine, LLM agents are increasingly used to analyze omics data to accelerate the identification of targets and drug discovery for age-related diseases.[3] For example, Precious1GPT, a multimodal LLM agent, leverages methylation and transcriptomic data for interpretable age prediction and target discovery.[16] Similarly, AlphaFold 3, an advanced LLM-driven structural biology tool developed by DeepMind, has been applied to model the structures of proteins linked to Alzheimer's disease, Parkinson's disease, and stroke, aiding in the discovery of new disease mechanisms and potential drug targets.[17]

Evaluating LLM agents using clinical simulations. The rapid advancements in LLM technologies have presented significant challenges in establishing a comprehensive evaluation framework for their applications in healthcare. Currently, there is no unified framework to rigorously assess the effectiveness, potential risks, feedback mechanisms, and health-economic value of LLM agents. While some researchers have proposed quantitative evaluation frameworks using methods like the Delphi technique and literature reviews, these traditional static approaches are ill-suited for the dynamic and evolving nature of LLMs. Furthermore, existing evaluation frameworks for general LLMs fail to address the patient-centered and safety-critical requirements of geriatric care. To address these gaps, emerging dynamic simulation-based evaluation methods have been proposed.[4] One such approach is Agent-Based Modeling (ABM), which models LLM agents as autonomous entities to analyze their decision-making processes, adaptability, and collaboration within diverse clinical scenarios. ABM provides a dynamic and adaptive framework, enabling a holistic assessment of how LLM agents function in complex clinical environments. Another promising framework is Artificial Intelligence Structured Clinical Examinations (AI-SCEs), inspired by the Objective Structured Clinical Examination (OSCE) used in medical education. AI-SCEs involve multidisciplinary teams designing scenario-based benchmarks to evaluate not only the final outputs of LLM agents but also their reasoning processes, data integration capabilities, and interactions with healthcare providers and patients. By focusing on critical metrics like reliability, interpretability, feasibility, and effectiveness, AI-SCEs provide a comprehensive and practical evaluation system for LLM-based agents in real-world clinical applications. These innovative frameworks pave the way for developing standardized and robust evaluation systems, ensuring that LLM agents meet the complex demands of geriatric care while maintaining safety and efficacy.

What is next? Despite their potential, deploying LLM agents in aging-related healthcare raises several critical concerns. General issues include hallucinations, interpretability, and ethical considerations.[3,18] For instance, hallucinations—where LLM agents generate plausible but incorrect information—pose a significant risk for older adults who may struggle to differentiate accurate health information from misinformation.[6] Ensuring the reliability and stability of these systems is essential to safeguard their well-being and foster trust in digital healthcare solutions. Moreover, equity in healthcare delivery remains a major concern, given the socioeconomic disparities within the elderly population. Prioritizing accessibility and affordability is key to ensuring a fair distribution of healthcare resources.[3]

There are also issues specific to the geriatric population. Older adults often have limited familiarity with digital technologies and unique health needs that demand tailored solutions. Enhancing these agents with the ability to integrate multimodal health data—such as speech, text, and sensor-based inputs—can significantly boost their utility. In addition, designing user-friendly devices with intuitive interfaces (e.g., larger text, fewer buttons, improved color contrast) and developing passive agents that require minimal user interaction can simplify use and improve adherence.[2] Addressing the complexities of geriatric healthcare and the diverse needs of the aged across their full lifecycle necessitates multidisciplinary collaboration and multidimensional decision-making. A range of software and hardware agents, such as wearable patient monitoring devices, machine learning-based decision-support systems, and robotic therapy tools, can work alongside LLM-driven agents to enhance healthcare delivery. Integrating these various functional agents into a multi-agent healthcare collaboration network can enable efficient, cross-disciplinary, and cross-system collaboration in complex geriatric care scenarios. For example, integrating LLM-based virtual medical agents with physical systems such as nursing robots could facilitate end-to-end intelligent healthcare services—from information processing to hands-on patient care. Finally, widespread adoption of LLM agents in geriatric care will require rigorous validation through large-scale clinical trials focusing on cost-effectiveness and safety. It is also essential to seamlessly integrate these agents into existing geriatric healthcare workflows and assess how their systematic addition or removal affects overall outcomes and quality of life for older adults.[4]

In conclusion, the rapid growth of the global geriatric population is outpacing healthcare systems' ability to address the complex needs of older adults. LLM agents have significant potential to revolutionize geriatric care by enhancing access, personalizing interventions, and improving interactions between older adults and healthcare providers. However, their current value remains largely theoretical, with clinical efficacy yet to be validated through prospective research and real-world applications. Misuse could lead to unintended health consequences, emphasizing the need for a robust, multidimensional evaluation framework that accommodates their dynamic nature. It is crucial to define LLM agents as supportive tools for healthcare professionals—not as replacements for human judgment—and to provide appropriate training for both users and caregivers. Additionally, addressing the financial implications of integrating digital care is essential to ensure these technologies remain accessible and affordable for diverse populations. With continued advancements in digital innovation, LLM agents are poised to enable timely, effective, and scalable interventions, ultimately reshaping the landscape of geriatric care.

CRediT authorship contribution statement

Pengfei Li: Writing – review & editing, Funding acquisition, Conceptualization, Writing – original draft. **Jingyi Wu:** Writing – review & editing, Writing – original draft. **Shaomei Shang:** Writing – review & editing, Supervision. **Qimin Zhan:** Writing – review & editing, Funding acquisition, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (72474010) and Strategic Research and Consulting Project of Chinese Academy of Engineering (2022-XBZD-30).

References

1. United Nations. World population prospects 2024: summary of results. 2024. https://desapublications.un.org/publications/world-population-prospects-2024-summary-results. Accessed March 10, 2025.
2. Chen C, Ding S, Wang J. Digital health for aging populations. *Nat Med.* 2023;29(7):1623–1630.
3. Gao S, Fang A, Huang Y, et al. Empowering biomedical discovery with AI agents. *Cell.* 2024;187(22):6125–6151.
4. Mehandru N, Miao BY, Almaraz ER, Sushil M, Butte AJ, Alaa A. Evaluating large language models as agents in the clinic. *NPJ Digit Med.* 2024;7(1):84.
5. Abadir PM, Battle A, Walston JD, Chellappa R. Enhancing care for older adults and dementia patients with large language models. In: Proceedings of the National Institute on Aging—Artificial Intelligence Technology Collaboratory for Aging Research Symposium. 2024.
6. Chen S, Kann BH, Foote MB, et al. Use of artificial intelligence chatbots for cancer treatment information. *JAMA Oncol.* 2023;9(10):1459–1462.
7. Parmanto B, Aryoyudanta B, Soekinto TW, et al. A reliable and accessible Caregiving Language Model (CaLM) to support tools for caregivers: development and evaluation study. *JMIR Form Res.* 2024;8:e54633.
8. Crista V, Martinho D, Marreiros G. A multi-agent system approach with generative AI for improved elderly daily living. In: EPIA Conference on Artificial Intelligence. 2024. pp. 128–140.
9. Merrill MA, Paruchuri A, Rezaei N, et al. Transforming wearable data into health insights using large language model agents. arXiv:240606464. 2024.
10. Yang Z, Xu X, Yao B, et al. Talk2Care: An LLM-based voice assistant for communication between healthcare providers and older adults. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2024. pp. 1–35.
11. Wan P, Huang Z, Tang W, et al. Outpatient reception via collaboration between nurses and a large language model: a randomized controlled trial. *Nat Med.* 2024;30(10):2878–2885.
12. Tu T, Palepu A, Schaekermann M, et al. Towards conversational diagnostic AI. arXiv:240105654. 2024.
13. Zhou Y, Zhang P, Song M, et al. Zodiac: a cardiologist-level LLM framework for multi-agent diagnostics. arXiv:241002026. 2024.
14. Du X, Novoa-Laurentiev J, Plasek JM, et al. Enhancing early detection of cognitive decline in the elderly: a comparative study utilizing large language models in clinical notes. *eBioMedicine.* 2024;109:105401.
15. Belyaeva A, Cosentino J, Hormozdiari F, et al. *Multimodal LLMs for health grounded in individual-specific data.* Switzerland: Springer Nature; 2023:86–102.
16. Urban A, Sidorenko D, Zagirova D, et al. Precious1GPT: multimodal transformer-based transfer learning for aging clock development and feature importance analysis for aging and age-related disease target discovery. *Aging.* 2023;15(11):4649–4666.
17. Gadde N, Dodamani S, Altaf R, Kumar S. Leveraging AlphaFold 3 for structural modeling of neurological disorder-associated proteins: a pathway to precision medicine. *bioRxiv.* 2024 20241118624211.
18. Qiu J, Lam K, Li G, et al. LLM-based agentic systems in medicine and healthcare. *Nat Mach Intell.* 2024;6(12):1418–1420.