Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ31038

The Impact of Multi Modal AI Consultation Systems on Middle-Aged and Older Adults Experience in Accessing Health Information

Yue Cao

University of Sydney, Sydney, Australia
770491141@qq.com

Abstract. With the evolution of digital health technologies, artificial intelligence (AI) consultation systems are increasingly applied to enhance the accessibility and efficiency of medical communication. This study explores how different interaction modes, specifically multimodal AI consultation systems (integrating text, audio, and pictures) versus traditional text-based self-service consultations, influence middle-aged and older adults' experiences in understanding cardiovascular health information. This study collected data from 120 participants aged 45 and above via a structured questionnaire. Four scales measured health information comprehension, trust, usability, and engagement. Compared to text-based interactions, multimodal AI consultations significantly enhanced engagement and trust, while differences in comprehension and usability were not statistically significant. Findings indicate that multimodal interaction enhances emotional engagement and perceived credibility, delivering a more humanized and trustworthy user experience. However, cognitive benefits may be limited when multiple information channels increase cognitive load. In conclusion, carefully designed multimodal AI systems hold significant potential for improving middle-aged and older adults' access to and acceptance of health information, particularly in managing chronic conditions such as cardiovascular disease.

Keywords: AI-assisted diagnosis, multimodal interaction, middle-aged and older adult, health information comprehension, trustworthiness

1. Introduction

According to the Report on Nutrition and Chronic Disease Status of Chinese Residents [1], chronic diseases account for 88.5% of all deaths in China, with cardiovascular diseases (CVDs) being the leading cause, responsible for over 40% of mortality. More than 300 million people in China are affected by cardiovascular conditions, which pose a particularly severe health threat to the middle-aged and elderly population. Effective management of these conditions remains challenging, as many older adults confront difficulties in understanding health information due to low health literacy, limited cognitive abilities, and reduced access to healthcare resources [2].

Artificial intelligence (AI) presents promising solutions to these challenges. AI-driven multimodal systems that integrate text, audio, and visual information have emerged as effective tools

DOI: 10.54254/2755-2721/2026.TJ31038

for health communication [3]. By employing multiple formats, these systems can improve comprehension and engagement among middle-aged and older adults, provide personalized health advice, and make complex medical information more accessible. For cardiovascular disease management specifically, AI-powered chatbots can offer tailored guidance, explain diagnostic results, and suggest lifestyle modifications through user-friendly interfaces.

Despite these advantages, empirical evidence on the effectiveness of multimodal AI systems in improving middle-aged and older adults' comprehension of health information remains limited. Most existing research focuses on single-modality approaches, such as text-based interfaces or static website searches, which may not fully address the diverse needs of middle-aged and older users. Furthermore, the interactions between different information delivery modes and their combined effects on comprehension require further investigation.

This study aims to evaluate the effectiveness of a multimodal AI chatbot system in enhancing health information access for middle-aged and older adults with cardiovascular diseases. Using simulated online consultations and structured questionnaires, the study assess the system's usability and its impact on users' understanding of key health concepts. The findings will inform the development of more inclusive and effective digital health tools tailored to the needs of the middle-aged and older adult population.

2. Literature review

Middle-aged and older adults confront multiple challenges in using AI platforms for health information, including difficulties in information acquisition, high cognitive load, limited interactivity, and trust deficits. Addressing these issues requires an integrated strategy that prioritizes emotional and psychological needs, emphasizes clarity and simplicity in design, and fosters interactive, personalized experiences. While AI counseling systems hold significant potential to improve health information access and comprehension for older adults, their effectiveness depends on overcoming cognitive, technical, emotional, and cultural barriers. By incorporating multimodal communication, establishing trust, and balancing technical capabilities with human-centric design, AI health counseling systems can better support the middle-aged and older adult population and enhance their ability to manage health autonomously.

Older adults encounter significant barriers in accessing and comprehending health information. Physical and cognitive challenges—such as visual or hearing impairment and slower information processing [2]—hinder their engagement with health content. Additionally, complex terminology and information overload can overwhelm older users, impairing knowledge retention and practical application [4]. Skepticism toward emerging technologies, including AI-driven health systems, further complicates adoption. As Berkowsky et al. [5] noted, such distrust is often compounded by concerns over data privacy and a perceived lack of human empathy in digital interfaces.

Traditional health information platforms, such as static websites or printed brochures, also suffer from limited interactivity, offering little opportunity for clarification or personalized guidance. Furthermore, as Sheng [6] suggests, systems that fail to establish credibility or align with older adults' cultural and experiential backgrounds are unlikely to gain their trust or acceptance.

3. The usage of AI in counseling

AI counseling systems hold significant potential to improve middle-aged and older adults' access to and comprehension of health information, though maximizing their efficacy requires addressing cognitive, technical, emotional, and cultural barriers.

DOI: 10.54254/2755-2721/2026.TJ31038

Artificial Intelligence (AI) refers to the ability of computer systems to simulate human intelligence, including learning, reasoning, problem solving, and language comprehension [7]. AI Chatbot (Chatbot) is an application of Artificial Intelligence that is capable of engaging in a conversation with a user through Natural Language Processing (NLP) technology [8]. They can be used in diverse scenarios, with increasing adoption in healthcare, especially in disease diagnosis, treatment advice, health management and drug development. Key applications of AI Chatbot in healthcare include: providing round-the-clock health consultation services, addressing questions about symptoms, medication, and lifestyle; conducting preliminary screenings for potential illnesses by analyzing symptoms to guide patients toward medical care; offering personalized dietary and exercise recommendations for chronic disease patients; featuring dedicated mental health support to help alleviate anxiety and depression; and providing medication reminders and tracking drug efficacy and side effects.

3.1. Text-only self-service search communication model

Text-based models are widely used in healthcare counseling systems due to their simplicity and scalability. Studies indicate that text-based interfaces can effective deliver detailed and organized health information, making them particularly suitable for explaining complex medical conditions such as cardiovascular diseases [6]. However, for middle-aged and older adults, text-only models pose significant challenges:

Text-only information increases cognitive load. Middle-aged and older users often have difficulty reading and comprehending lengthy or technical text, which may prevent them from understanding important health information [9].

Text-based communication is unengaging. Text-based communication lacks the emotional and contextual cues found in face-to-face communication, which reduces user engagement and trust [10].

Medical websites rely heavily on static content such as articles and infographics. These resources, while informative, require users to actively search for and interpret information, which may be overwhelming for middle-aged and older adults. Accessibility is also an important influencing factor. Visual impairments and lower digital health literacy among middle-aged and older adults further limit the dissemination of eHealth information. Many websites are not optimized for middle-aged and older users, with small fonts, dense text and complex navigation structures [11].

3.2. Multimodal AI consultation models integrating video and images

Multimodal AI consultation systems demonstrate substantial benefits in improving the comprehension of health information for middle-aged and older adults with cardiovascular disease. These systems can provide diverse interaction modes [12]:

Users can describe symptoms in speech, which the AI transcribes into text for understanding in real time, or they can choose to hear a voice response from the AI. For visual supports, users upload relevant images (e.g., dermatological conditions, CT films, ECGs, etc.) alongside text descriptions; AI can analyze the symptoms together with image recognition and natural language processing. During interactions, users combine voice narration, photo uploads, and AI's intelligent follow-up questions to provide multi-channel information. The AI then make comprehensive judgments based on Multimodal Deep Learning.

Multimodal systems can enhance comprehension of complex health information. Multimodal systems disseminate information through multiple sensory channels, making it easier for older adults

DOI: 10.54254/2755-2721/2026.TJ31038

to process and retain important health information through integrating video, imagery, and AI technologies, with visual aids such as heart or vascular mapping combined with video interpretation to simplify complex medical concepts [13].

Artificial intelligence anthropomorphization can enhance the trust of the other party and increase the motivation to participate. Artificial intelligence avatars and voice can simulate human interactions, promoting trust and engagement by providing personalized and empathetic communication [14].

Based on above literature, this study proposes the following hypotheses:

H1: Participants who consult with the multimodal AI chatbot possess a higher level of engagement compared to participants who consult with a website.

H2: Participants who consult with the multimodal AI chatbot had higher comprehension compared to those who communicated through text-only communication

H3: Participants who consult with the multimodal AI chatbot possess a higher level of trust to the information compared to participants who consult with a website.

H4: Participants who consult with the multimodal AI chatbot perceive the information more accessible compared to participants who consult with a website.

Different modes of AI communication exhibit distinct strengths and limitations in delivering cardiovascular health information to middle-aged and older adults. While text-based systems offer fundamental support, multimodal AI counseling demonstrates significant advantages over traditional text-only and website search models. These integrated systems enhance the clarity, engagement, and accessibility of health information through combined sensory channels. Future research and development should prioritize user-centered design and seamless integration of multimodal features to better serve the middle-aged and older adult population.

4. Method

This study aims to evaluate the effectiveness of AI consultation system in promoting middle-aged and older adults' understanding of cardiovascular health. It compares two interaction modes: a multimodal (text, audio, images) system and a text-only self-service website. An experimental design was employed, combined with structured questionnaires to assess both interaction modes.

4.1. Participants and procedure

Participants were recruited according to the following criteria: (1) aged 45 years or older; (2) diagnosed with cardiovascular disease or associated risk factors (hypertension, diabetes, or hypercholesterolemia); (3) prior experience with AI-based health counseling services (text-based or multimodal); (4) with basic familiarity with digital technology, though advanced proficiency was not required. Individuals with significant cognitive or physical impairments that could hinder the use of digital platforms were excluded.

A total of 60 participants were enrolled and randomly assigned to two groups, with 30 participants in each group.

4.2. Experimental design

Participants were randomly assigned to two experimental conditions: one group utilized a multimodal AI consultation system (integrating text, audio, and visual elements), while the other group interacted with a text-only AI consultation system.

DOI: 10.54254/2755-2721/2026.TJ31038

The multimodal group engaged with a comprehensive AI system capable of processing and responding through multiple communication channels (text, speech, images) to deliver cardiovascular health information. In contrast, the text-based group accessed information exclusively through textual messages and instructions.

4.3. Measures and procedure

Evaluation of system effectiveness focused on four key dimensions: engagement, usability, comprehension, and trust. These measures were assessed through a structured questionnaire comprising four validated scales.

Prior to completing the questionnaire, participants received a brief study description and visual to help recall their AI consultation experiences.

Health Information Comprehension Scale: A customized Likert scale for perceptual comprehension.

System Satisfaction and Usability Scale: Ease of use and satisfaction assessed using the System Usability Scale (SUS).

User Engagement Scale: Simplified User Engagement Scale (UES) and Likert Scale [15]:

To assess attention, interest and willingness to continue use.

A questionnaire were administered to a sample of participants and questions will address these areas

Usability: Is the system easy to navigate and access?

Comprehension: Does the system help you understand cardiovascular health information in a more straightforward manner?

Engagement: Does the system keep your attention and interest? And are you willing to engage in dialog and communication?

Trust: Do you trust the health advice provided by the system?

4.4. Comprehension

The Health Information Comprehension Scale was used to measure participants' comprehension of cardiovascular health knowledge after using the AI Health Questioning System. The scale adopted an adapted S-TOFHLA (Short Test of Functional Health Literacy in Adults) structure that focuses on assessing information clarity, logical organization, the effectiveness of pictorial aids, and the comprehensibility of action recommendations [16]. The scale included 7 items, rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Example items: “The system helped me understand cardiovascular health information in a more straightforward way.”

4.5. Engagement

The User Engagement Scale (simplified version of the UES) is used to assess the user's level of concentration, interface attractiveness, willingness to interact, and pleasurable experience during the AI questioning process [15]. The simplified version of the UES contained 12 items covering the four core dimensions. All items were rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Example items: “The system maintained my attention and interest”; “I was willing to engage in dialogue with the system.”

DOI: 10.54254/2755-2721/2026.TJ31038

4.6. System Usability Scale (SUS)

User satisfaction was measured using the System Usability Scale (SUS), a tool for assessing the overall usability and user satisfaction feelings of an AI system [17]. The SUS consists of 10 items scored on a 5-point likert scale (1=strongly disagree, 5=strongly agree). Example item: “The system was easy to navigate and access.”

4.7. Trust

The System Trust Scale was used to assess the level of user trust in the AI health inquiry system, including perceptions of information accuracy, motivational credibility, and willingness to adopt recommendations [18]. The scale included 5 items covering the reliability of the system's recommendations, the credibility of the explanations, and the perceived safety of the AI system. All items were rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Example item: “I trust the health advice provided by the system.”

5. Data analysis

Descriptive statistics (means, standard deviations, and frequencies) were first calculated to summarize participants' demographic characteristics and responses to each measurement scale, including comprehension, trust, satisfaction, and engagement.

Descriptive statistics were calculated for demographic information (e.g., age, education, and level of use of the multimodal system) and scale scores for each of the four measurement constructs:

·Health information comprehension

·Trust in the AI system

·System usability and satisfaction

·User engagement

Means, standard deviations, and score distributions were calculated for each subscale to determine overall multimodal user perceptions of system effectiveness.

Table 1. Group statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
engage	1.00	58	4.3448	.23018	.03022
	2.00	57	4.1520	.43817	.05804
compre	1.00	58	4.3128	.27739	.03642
	2.00	57	4.2155	.45346	.06006
trust	1.00	58	4.1839	.33434	.04390
	2.00	57	3.9620	.67555	.08948
use	1.00	58	4.0486	.27098	.03558
	2.00	57	4.0319	.32392	.04290

Table 1 summarizes descriptive statistics for the two participant groups across four measurement dimensions: engagement, comprehension, trust, and usability. Participants using the multimodal AI consultation system (Group 1) generally scored higher on most dimensions than those using the self-service text-based medical search (Group 2). Specifically, the most significant difference was observed in engagement ($M_1 = 4.34$, $SD = 0.23$; $M_2 = 4.15$, $SD = 0.44$). Regarding comprehension,

DOI: 10.54254/2755-2721/2026.TJ31038

the multimodal group's average score (M_1 = 4.31, SD = 0.28) was also marginally higher than the text group's (M_2 = 4.22, SD = 0.45). Regarding trustworthiness, the multimodal group scored significantly higher than the text group (M_1 = 4.18, SD = 0.33 vs. M_2 = 3.96, SD = 0.68).

Finally, usability ratings were largely comparable across groups (M_1 = 4.05, SD = 0.27; M_2 = 4.03, SD = 0.32).

Reliability analysis is a statistical method used in quantitative research to assess whether a questionnaire scale is consistent, stable, and reliable. It ensures whether the scale measures accurately and produces relatively consistent results across different contexts.

The t-test compares whether the mean values of two groups exhibit statistical differences. The significance level is determined by the p-value [19].

If $p < .05$, the results are statistically significant, meaning the difference is unlikely to be caused by random factors;

If $p \geq .05$, the difference is considered non-significant.

To test the effect of interaction modes on users' experiences, a list of independent samples t-test were conducted. H1 states that users experience a higher level of engagement when using multimodal xx compared to text condition. Results showed the multimodal group had significantly higher engagement scores (M = 4.34, SD = 0.23) than the text-based group (M = 4.15 , SD = 0.44), $t(113) = 2.96$, $p < .01$. H1 is supported.

H2 states that users experience a higher level of comprehension when interacting with Multimodal AI Consultation System compared to text condition. However, Results showed no statistically significant difference in comprehension scores between the multimodal group (M = 4.31, SD = 0.28) and the text-based group(M = 4.22 , SD = 0.45), $t(113) = 1.39$, $p = .08>.005$. H2 is not supported.

H3 states that users experience a higher level of trust when interacting with the Multimodal AI Consultation System compared to text conditions. Results showed the multimodal group had significantly higher trust scores (M = 4.18, SD = 0.33) than the text-based group (M = 3.96 , SD = 0.67), $t(113) = 2.23$, $p < .005$. H3 is supported.

H4 states that users experience a higher level of useability when interacting with Multimodal AI Consultation System compared to text condition. Results showed no statistically significant difference in usability scores between the multimodal group (M = 4.04, SD = 0.27) and the text-based group (M = 4.03 , SD = 0.32), $t(113) = 0.3$, $p = .038>.005$. H4 is not supported.

6. Discussion

This study examines how multimodal AI consultation systems and text-based self-care searches influence user experience in terms of engagement, comprehension, trust, and usability. Four hypotheses (H1-H4) were validated using two independent samples. Results indicate that middle-aged and elderly groups demonstrate higher engagement and trust levels, suggesting multimodal presentation formats (text + audio + pictures) may play a positive role in supporting their interaction with digital health technologies.

6.1. Engagement

Findings support Hypothesis H1, indicating participants using the multimodal AI consultation system demonstrated significantly higher engagement than those relying on text-based medical search (M=4.34 vs. M=4.15, p<.01). This aligns with prior research showing multimodal interactions, integrating voice, visuals, and interactive feedback, enhance users' emotional

DOI: 10.54254/2755-2721/2026.TJ31038

involvement and attentional focus [15]. The richer sensory experiences and conversational dynamics of multimodal systems may strengthen users' sense of presence and active participation, fostering stronger connection and motivation during health consultations.

6.2. Comprehension

Contrary to expectations, H2 was not supported. Although users of the multimodal system reported higher comprehension scores (M = 4.31) than text-based users (M = 4.22), this difference was not statistically significant within the sample. One possible explanation is that while multimodal content may boost engagement, it could also increase cognitive load—particularly for middle-aged and older adults who must process audiovisual information simultaneously. According to Meyer's cognitive load theory [20], excessive information channels can exceed working memory capacity, diminishing deep comprehension even during immersive experiences. Future designs should optimize the pacing and clarity of multimodal interfaces to balance stimulation with comprehension.

Trust, H3 was validated: Users interacting with multimodal AI systems demonstrated significantly higher trust levels (M = 4.18) than those using text-only interactions (mean = 3.96, $p < 0.005$). This suggests multimodal systems enhance users' social presence and perceived credibility through natural speech and human-like responses [21]. Incorporating empathetic vocal intonations or visual cues can make systems appear more reliable and authoritative, thereby increasing user trust in their recommendations.

6.3. Usability

H4 was not supported, with only a marginal difference between the multimodal condition (M=4.04) and the text-only condition (M=4.03). This suggests both formats were perceived as relatively easy to use, potentially due to participants' prior experience with digital health tools. The lack of significant usability improvement implies that while multimodal systems enhance engagement and trust, they do not necessarily simplify the interface experience compared to well-designed text-based platforms.

Overall, the findings indicate that multimodal AI consultations enhance engagement and trust, but their impact on comprehension and usability remains unclear. These mixed results suggest multimodal technology benefits users at the emotional and relational levels but does not guarantee advantages at the cognitive or operational levels. Future research should explore how specific design features influence comprehension and perceived usability.

6.4. Limitations and future research

While this study offers meaningful insights into multimodal AI consultation systems' impact on user experience, several limitations require acknowledgment. First, reliance on self-reported data introduces potential biases, including recall inaccuracies and social desirability effects. Second, the online experimental setting may not fully capture the complexities of real-world clinical consultations, limiting ecological validity. Third, the sample composition presents another constraint, as participants demonstrated relative homogeneity in digital literacy, which may restrict generalizability to populations with lower technological proficiency or older age groups.

Future investigations should implement mixed-methods designs that integrate quantitative measures with qualitative interviews and behavioral observations to comprehensively examine user interactions with multimodal AI systems. Additional research should also examine how individual

DOI: 10.54254/2755-2721/2026.TJ31038

characteristics, including age, educational background, and cognitive capacity, moderate the effectiveness of such systems, enabling more tailored implementation strategies.

7. Conclusion

This study demonstrates that multimodal AI consultation systems (integrating text, audio, and visual elements) significantly enhance user engagement and trust compared to traditional text-based interfaces, though their advantages in improving comprehension and usability remain less pronounced. The findings suggest that while enriched sensory channels and human-like interaction foster more positive user experiences, simply increasing modality complexity does not automatically translate to better understanding or operational ease. A key implication is that multimodal systems achieve their primary benefits through strengthening emotional connection and perceived credibility, yet this must be balanced against potential cognitive overload. For middle-aged and older adults processing complex health information, optimal system design must navigate the tension between interface simplicity and interactive richness. Adaptive multimodal systems tailored to users' cognitive characteristics and sensory preferences are recommended. Future implementations should prioritize clear visual explanations, natural voice interactions, and streamlined information architecture to simultaneously enhance trustworthiness and reduce usage barriers. Multimodal AI consultation represents a promising direction for digital health innovation, particularly for supporting middle-aged and older adult patients with chronic cardiovascular conditions. When appropriately designed, such systems can make critical medical information more accessible, trustworthy, and emotionally resonant for this vulnerable population.

References

[1] National Health Commission of the People's Republic of China. (2020). Report on Nutrition and Chronic Disease Status of Chinese Residents (2020). The State Council Information Office of the People's Republic of China. Retrieved from http: //www.scio.gov.cn/xwfb/gwyxwbgsxwfbh/wqfbh_2284/2020n_4408/2020n12y23rsw/twzb_5755/202207/t20220716_230693.html

[2] AlSaad, R., Abd-alrazaq, A., Boughorbel, S., Ahmed, A., Renault, M.-A., Damseh, R., & Sheikh, J. (2024). Multimodal large language models in health care: Applications, challenges, and future outlook. Journal of Medical Internet Research, 26, e59505. https: //doi.org/10.2196/59505

[3] Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2023). Normal cognitive aging. Clinics in Geriatric Medicine, 29(4), 737–752.

[4] Federman, A. D., et al. (2009). Health literacy and cognitive performance in older adults. Journal of the American Geriatrics Society, 57(8), 1475–1480.

[5] Berkowsky, R. W., et al. (2013). Attitudes towards and limitations to ICT use in assisted and independent living communities: Findings from a specially-designed technological intervention. Educational Gerontology, 39(11), 797–811.

[6] Sheng, X. (2015). Trust and credibility in web-based health information: A review of the literature. Proceedings of the 2015 International Conference on Social Science and Technology Education.

[7] Xu, L., et al. (2021). Artificial intelligence: A survey on evolution, models, and applications. IEEE Transactions on Neural Networks and Learning Systems, 32(9), 3904–3921.

[8] Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In Artificial Intelligence Applications and Innovations (pp. 373–383). Springer.

[9] Kobayashi, L. C., et al. (2008). The role of cognitive function in the relationship between age and health literacy: A cross-sectional analysis of older adults. Journal of Health Communication, 13(8), 741–756.

[10] Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications. John Wiley & Sons.

[11] Tennant, B., et al. (2015). eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. Journal of Medical Internet Research, 17(3), e70.

DOI: 10.54254/2755-2721/2026.TJ31038

[12] Chin, J., et al. (2017). The role of multimedia in improving health literacy. Journal of Health Communication, 22(10), 827–835.
[13] Marey, S., et al. (2023). Visual aids and multimedia in patient education: A systematic review. Patient Education and Counseling, 106, 107–115.
[14] Paiva, A., et al. (2017). Empathy in virtual agents and robots: A survey. ACM Transactions on Interactive Intelligent Systems, 7(3), 1–40.
[15] O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. Journal of the American Society for Information Science and Technology, 61(1), 50–69.
[16] Parker, R. M., et al. (1995). The test of functional health literacy in adults: A new instrument for measuring patients' literacy skills. Journal of General Internal Medicine, 10(10), 537–541.
[17] Brooke, J. (1996). SUS: A quick and dirty usability scale. Usability Evaluation in Industry, 189(194), 4–7.
[18] Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. International Journal of Cognitive Ergonomics, 4(1), 53–71.
[19] Field, A. (2013). Discovering statistics using IBM SPSS Statistics (4th ed.). Sage Publications.
[20] Meyer, R. E. (2005). The Cambridge handbook of multimedia learning. Cambridge University Press.
[21] Charness, N., & Boot, W. R. (2016). Aging and information technology use: Potential and barriers. Current Directions in Psychological Science, 25(2), 96–101.