Review article

DIGITAL HEALTH

DIGITAL HEALTH
Volume 12: 1–13
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076261450385
journals.sagepub.com/home/dhj
Sage

Natural language processing-based chatbots for chronic disease self-management: A systematic review of implementation and health outcomes

Ga In Han, RN, MSN[1], Hi Jae Lee, RN, APN, MSN[2], and Youn-Jung Son, PhD, RN, CCAPN, FAAN[3]

Abstract

Background: Conversational agents (chatbots) are increasingly used as digital health interventions to support chronic disease self-management. Advances in natural language processing (NLP) have improved their capacity for interactive dialogue and personalization, yet evidence regarding their implementation and clinical impact remains limited. **Objectives:** This systematic review identifies and synthesizes studies implementing NLP-based chatbots for chronic disease self-management. **Methods:** We searched seven electronic databases (PubMed, Embase, CINAHL, Web of Science, Scopus, Cochrane Library, and IEEE Xplore) and Google Scholar for studies published between January 2010 and November 2025. Studies evaluating NLP-based chatbots designed to support chronic disease self-management were deemed eligible. Study quality and risk of bias were assessed using the Mixed Methods Appraisal Tool and the Quality Assessment with Diverse Studies instrument. **Results:** Six studies met the inclusion criteria; most were published in 2023 and targeted conditions such as cancer, diabetes, and hypertension. Chatbot functions primarily focused on symptom monitoring and disease-related education. Reported outcomes included improvements in disease-related knowledge, symptom burden, mental well-being, and self-care adherence. Usability and acceptability were generally favorable, with high satisfaction, perceived usefulness, and engagement. However, evidence of objective clinical benefits, including laboratory outcomes, was limited. Technical architectures varied widely, and advanced NLP capabilities—such as free-text natural language understanding—were rarely implemented. **Conclusions:** NLP-based chatbots show promise for supporting chronic disease self-management, particularly for psychosocial and behavioral outcomes. However, evidence of clinical efficacy remains limited. Future research should prioritize adaptive, context-aware designs and standardized outcome frameworks aligned with real-world self-management needs.

Keywords

chatbot, chronic disease, conversational agent, natural language processing, self-management

Received: 6 January 2026; Revised: 1 April 2026; Accepted: 27 April 2026

[1]Graduate School of Nursing, Chung-Ang University, Seoul, Republic of Korea
[2]Yonsei University Health System, Severance Hospital, Seoul, Republic of Korea
[3]Red Cross College of Nursing, Chung-Ang University, Seoul, Republic of Korea

Corresponding author:
Youn-Jung Son, Professor, PhD, RN, CCAPN, FAAN Red Cross College of Nursing, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea.
Email: yjson@cau.ac.kr

Introduction

A conversational agent, or chatbot, is a software application that communicates with users through text or voice and has become increasingly prominent in patient care and health management.[1] Chatbots can broaden access to health information and support individuals in regulating their health behaviors and decisions by simulating elements of human conversation.[2–4] Recent advances in system architecture and functionalities enable exchanges that closely resemble human-to-human communication, and their adoption across diverse healthcare settings has expanded rapidly.[5]

Most contemporary chatbots are underpinned by advances in natural language processing (NLP), which converts free-text inputs into machine-readable linguistic features, and natural language understanding (NLU), which infers users' intents and meanings in context, combined with machine learning techniques to interpret unstructured user inputs and deliver contextually appropriate responses.[5] Implementation strategies range from rule-based systems that follow predetermined scripts to more complex artificial intelligence (AI)-based retrieval and generative models, which occasionally integrate multimodal inputs such as text, speech, and images.[6] This technological foundation enables chatbots to extract clinically relevant information and provide context-aware and empathetic interactions that may foster user engagement and sustained use.[7,8]

In healthcare delivery, chatbots have been employed to improve access to timely information and reduce temporal and geographic barriers by providing personalized, on-demand support.[9,10] Such support can be directed toward patients, families, caregivers, and health professionals and may function as an adjunct tool for education, counseling, and communication.[10] Emerging studies have documented chatbot interventions in domains such as emergency response guidance, perinatal health management, and caregiving support.[11–13] However, the current evidence base remains fragmented, partly due to heterogeneity in study designs, relatively short intervention durations, and the absence of standardized outcome measures.[13,14]

Chronic diseases are a leading cause of morbidity and mortality worldwide, and impose long-term, complex demands on patients, families, and health systems.[15,16] As populations age and medical advances prolong survival, increasing numbers of individuals continue to live with chronic conditions for many years, thereby frequently experiencing recurrent exacerbations and clinical uncertainty.[15–17] These patterns underscore the necessity for tailored education, continuous monitoring, and timely feedback throughout the illness cycle.[17] Chronic disease self-management refers to the ongoing process through which individuals with chronic conditions actively manage the symptoms, treatments, lifestyle modifications, and cope with the psychosocial consequences of their illness.[16] Digital health interventions have emerged as a critical platform for delivering technology-enabled strategies for chronic disease self-management, providing dynamic responses to evolving needs, and offering personalized recommendations.[18]

Conventional rule-based chatbots rely solely on fixed scripts or button-based inputs, whereas NLP-based chatbots can process both predefined responses and free-text input, infer user intent, and, in some cases, detect emotional cues. This capacity enables more adaptive and personalized interactions.[19] Such systems are well-suited for chronic disease self-management because they can deliver tailored educational content, monitor symptoms and health behaviors, support goal-setting, and guide appropriate use of healthcare services.[10,20,21] Emerging research suggests that such interventions may enhance self-efficacy, promote health-supporting behaviors, and potentially reduce the long-term healthcare burden.[20,21] By incorporating emotional exchanges and leveraging longitudinal user data, NLP-driven chatbots can adapt to changing disease stages and needs, thereby fostering trust, continuity of care, and effective disease management over time.[18,19,22]

Previous reviews have primarily catalogued chatbot types, core functions, or usability features in the context of chronic disease management,[18,21,23] while often treating chatbots as a single, broad category rather than examining NLP-based systems separately. Consequently, evidence on how NLP-based chatbots are implemented for chronic disease self-management and the resultant health-related outcomes remains fragmented. Therefore, this systematic review aims to: (1) characterize the implementation of NLP-based chatbots designed to support chronic disease self-management, and (2) synthesize the health outcomes reported across these studies.

Methods

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Multimedia Appendix 1).[24] The review protocol was prospectively registered with PROSPERO (registration number: CRD420251058858).

Search strategy

A comprehensive search was conducted across seven electronic databases (PubMed, EMBASE, CINAHL, Web of Science, SCOPUS, the Cochrane Library, and IEEE Xplore) and one gray literature source (Google Scholar). The search included articles published between January 2010 and November 2025. Key terms targeted the population and intervention, including "patient," "chronic," and "chronic disease," combined with "chatbot," "natural language processing," "conversational agent," "conversational AI," "agent-based system," and "natural language understanding." These terms were searched in titles, abstracts, and full texts when available. We did not impose restrictions on comparison groups or predefined clinical endpoints to capture the full range of contexts in which NLP-based chatbots have been used for chronic disease self-management and identify exploratory evidence across diverse outcome domains. Furthermore, the reference lists of all eligible articles were manually screened to identify additional relevant studies. The complete search strategy is provided in Multimedia Appendix 2.

Inclusion and exclusion criteria

The inclusion criteria are summarized in Table 1. For this review, NLP-based chatbots were defined as applications or programs that enable bidirectional communication with users, rather than as one-way information-delivery tools. Considering the heterogeneous reporting of chatbot architectures and underlying AI techniques across studies, a consistent and reproducible classification approach is required. Therefore, this review primarily adopted the terminology reported in the original publications to minimize misclassification.

In the healthcare chatbot literature, NLP-enabled systems often include hybrid architectures that combine NLP modules with predefined dialogue flows or structured response options.[21] Accordingly, we operationally defined NLP-based chatbots as systems capable of processing unstructured or semi-structured user input using natural language techniques, including but not limited to intent classification, entity recognition, and semantic interpretation. This definition encompasses both hybrid rule-based–NLP systems and machine-learning–driven conversational agents, provided that the system includes explicit natural-language processing functions (e.g., intent recognition, entity extraction, or semantic parsing) applied to user input. Therefore, this review included studies only if they explicitly reported the use of NLP-related techniques in processing user input, regardless of the user interface modality (e.g., free-text or predefined input formats).

The World Health Organization defines chronic diseases as long-lasting conditions that typically progress slowly and cannot be fully cured with treatment alone.[25] Although chronic pain and mental health conditions may share psychosocial characteristics with other chronic diseases, chatbot interventions in these areas are frequently designed as dedicated mental health interventions with specific therapeutic aims and outcome measures.[26] These approaches often emphasize

Table 1. Eligibility criteria for the study according to the PICO[a] question.

Criteria	Inclusion	Exclusion
Population/Patient	• Adults aged 19 years and above with chronic conditions	• Participants with chronic pain or psychiatric disorders • Pediatric or adolescent populations (<19 years)
Intervention	• NLP[b]-based chatbot interventions	• Simple mobile applications or programs • Exclusively evaluating technical performance
Control/Comparison	• No restrictions based on comparison	
Outcome	• No restrictions based on outcomes	
Additional exclusion criteria	• Books, gray literature, conference proceedings, and abstracts • Reviews, protocols, editorials, letters, and non-human studies • Non-English publications and studies without accessible full texts (e.g., corporate property requiring purchase)	

[a]PICO: Population/patient, intervention, comparison, and outcome.
[b]NLP: Natural language processing.

psychological or behavioral therapy rather than disease-oriented self-management. Accordingly, this review focused on chatbot interventions that support disease-oriented self-management of clearly defined chronic diseases. Studies reporting only the usability, satisfaction, or technical performance of chatbot systems without assessing any health-related outcomes were excluded.

Study selection and data extraction

After completing the search, all identified records were imported into EndNote 21, a reference management software, where duplicate entries were removed using automated and manual procedures. The titles and abstracts of each study were exported to Microsoft Excel for further analysis. Two authors (GI and HJ) screened each record to assess potential eligibility and categorized the studies as "potentially relevant" or "not relevant." Full texts of studies deemed "potentially relevant" were subsequently reviewed to determine final eligibility. In cases of disagreement, the third author (YJ) was consulted to reach a consensus.

The study selection procedures were documented using a PRISMA 2020 flow diagram. Data from the included studies were extracted independently by two reviewers using a standardized piloted form that captured the study design, sample characteristics, chatbot characteristics, comparators (where applicable), and reported health and user experience outcomes.

Data synthesis and analysis

Data were extracted from abstracts, full texts, and supplementary materials in a standardized format and subsequently checked to ensure accuracy and consistency. Given the heterogeneity of NLP-based chatbots and study designs, this review was designed to integrate all reported systems and user-level and health-related outcomes, rather than evaluate a single predetermined clinical endpoint. The techniques and attributes of NLP-based chatbots for self-management were coded based on descriptions in the studies, which focused on implementation characteristics and targeted outcome domains.

The extracted data were comparatively analyzed and summarized in tables and narrative texts to identify core concepts across studies. A meta-analysis was not conducted due to substantial heterogeneity across intervention types, outcome measures, and study contexts. Consequently, narrative and descriptive syntheses were conducted.

Quality appraisals and risk of bias assessment

The methodological quality and risk of bias of the included studies were assessed using the Mixed Methods Appraisal Tool (MMAT),[27] supplemented by the Quality Assessment with Diverse Studies (QuADS)[28] for heterogeneous designs. Two authors (GI and HJ) independently conducted each appraisal, and any discrepancies were resolved through discussion with the third author (YJ), who also verified the consistency of the assessments (Multimedia Appendix 3).

Ethical approval statement

The study protocol was approved by the Institutional Research Board of Chung-Ang University (No. 1041078-20250522-HR-150).

Results

Literature search results

In the initial search across validated databases, 8,332 records were identified. After removing 3,348 duplicate records and 278 records that automation tools marked as ineligible, 4,706 records were screened. Based on the titles and abstracts, 4,546 records were excluded. The full texts of the remaining 160 articles were reviewed, and six studies that met the inclusion criteria were incorporated in the final analysis. Moreover, nine potentially eligible records were identified from citation lists; of these, two overlapped with previously extracted records, and the remaining seven were excluded due to mismatches in participant criteria or a primary focus on technical performance rather than intervention (Figure 1).

Quality assessment of the included studies

Overall, the six included studies reflected a heterogeneous but generally moderate methodological quality. According to the MMAT, randomized controlled trials demonstrated a low risk of bias in allocation and outcome assessment, although

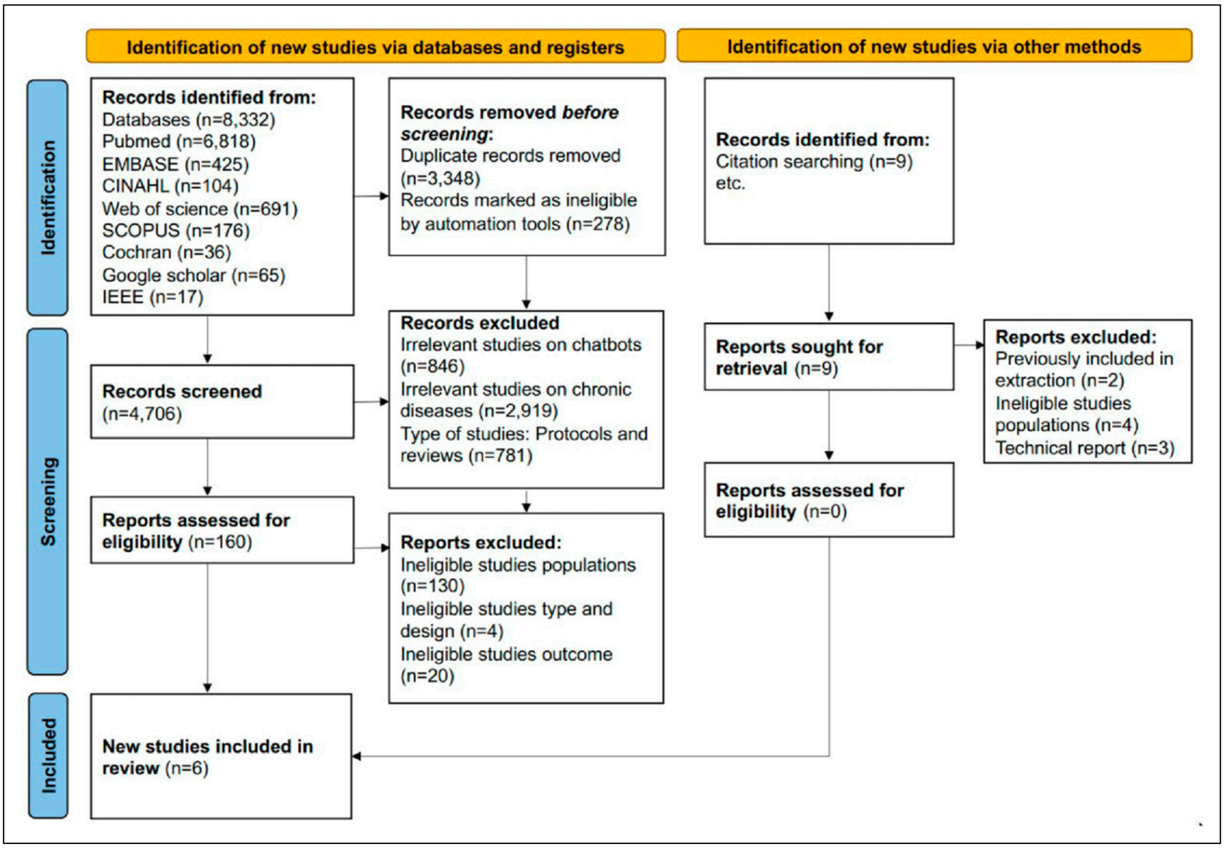

Figure 1. PRISMA flow diagram of study selection.

uncertainties persisted regarding blinding procedures and attrition.[29,30] Meanwhile, technical and feasibility studies were constrained by small sample sizes and limited transparency in recruitment processes.[31,32] According to QuADS, five studies scored between 27 and 37 out of a possible 42 points, and one study scored 31 out of 48 points, thereby collectively indicating moderate methodological quality.

Across study designs, common strengths included clearly defined aims, appropriate data collection procedures, strong alignment between research questions and analytic methods, and explicit reporting of study limitations.[29,33,34] Conversely, weaknesses included limited use of theoretical or conceptual frameworks, incomplete reporting of the reliability and validity of outcome measures, and minimal involvement of end users in the development and refinement of NLP-based chatbot interventions (Multimedia Appendix 3).

No study was excluded solely on the basis of quality appraisal; instead, risk-of-bias and quality assessments were used to inform the interpretation and weighting of the findings.

Characteristics of the included studies

Table 2 summarizes the characteristics of the six studies included in this review. Of these, five were published in 2023, including three development or feasibility studies[31–33] and two randomized controlled trials (RCTs),[29,30] and one mixed-methods study.[34] NLP-based chatbots were implemented across community and hospital settings, with cancer emerging as the most frequently targeted condition.[29,30,33] The sample sizes ranged from 10 to 156 participants, and only one study was designed to allow access to users other than patients.[29]

NLP-based chatbot attributes of the included studies

Table 3 summarizes the technical attributes of the NLP-based chatbots deployed for chronic disease self-management, as reported in the included studies. Web-based interfaces were the predominant delivery channel,[29,31,33,34] although several interventions were available as mobile or tablet applications.[32–34] The reported study periods ranged from three weeks to one

Table 2. Baseline characteristics of studies included for this review ($N = 6$).

First author, year [ref]	Country	Study design	Setting	Sample size	Mean age or range	Targeted user
Bibault et al., (2019)[29]	France	RCT[a]	Community	142	Mean: 42	Patients with breast cancer/and family
Görtz et al., (2023)[33]	Germany	Development and feasibility	Hospital	10	Mean: 68 Range: 49–81	Patients with prostate cancer
Tawfik et al., (2023)[30]	Egypt	RCT[a]	Hospital	150	Mean: 45.68 Range: 36–74	Patients with breast cancer
Bruijnes et al., (2023)[31]	Netherlands	Feasibility study	Community	156	Mean: 37 Range: 18–70	Patients with DM[b]
Au et al., (2023)[34]	Australia	Feasibility study	Hospital	20	Mean: 55.5	Patients with chronic liver disease
Griffin et al., (2023)[32]	United States	Mixed Methods	Community	10	Mean: 60	Patients with hypertension

[a]RCT: Randomized controlled trial.
[b]DM: Diabetes mellitus.

year; however, two studies did not specify the intended duration of chatbot use, thereby limiting the comparability of exposure across trials.[31–34] Engagement was typically limited to a single-use session, thus, leaving uncertainty regarding whether most systems were designed to support sustained or repeated use.[29,31,33]

Regarding user input modality, fixed text and free-text entries were used, with several studies combining these formats.[31–34] Text-based output was predominant for the agent output modality, whereas other studies incorporated images or videos.[32,34] In language processing, most chatbots incorporate NLP or NLU modules. Nevertheless, details of the underlying algorithms and the degree of automation or adaptive tailoring in responses were sparsely reported and varied across the six studies.

Table 3. Technical attributes of NLP-based chatbots in the included studies. ($N = 6$).

First author, year [ref]	Chatbot name	Channels of chatbot	Study period	Chatbot usage frequency	User input modality	Language processing technologies	Agent output modality
Bibault et al., (2019)[29]	Vik	Web app	1 month	Single-use session	Fixed texts	NLP[a]	Texts
Görtz et al., (2023)[33]	PROSCA	Web and mobile app	1 year	Single-use session	Free and fixed texts	NLP[a]	Texts
Tawfik et al., (2023)[30]	ChemoFreeBot	Mobile app	4 months	3 sessions only	Fixed texts	NLP[a]	Not reported
Bruijnes et al., (2023)[31]	Not reported	Web app	3 weeks	Single-use session	Free and fixed texts	NLU[b]	Texts
Au et al., (2023)[34]	Lucy Liver Bot	Tablet-based app	Not reported	Not reported	Free and fixed texts	NLP[a]	Texts, images, videos
Griffin et al., (2023)[32]	Medicagent	Web and mobile app	Not reported	Not reported	Free and fixed texts and speech	NLU[b]	Texts, images

[a]NLP: Natural language processing.
[b]NLU: Natural language understanding.
Notes: "Fixed texts" refer to predefined user interface input options (e.g., menus, keywords), which do not preclude the use of NLP techniques for processing user input.

Applications and outcomes of NLP-based chatbots in chronic disease self-management

Table 4 presents the patterns of NLP-based chatbot utilization by domain in chronic disease self-management. Overall, the most frequently reported function was providing information on disease and treatment. In particular, studies focusing on cancer classified information provision and clinical management support as primary or supportive functions.[29,30,33] Some studies on women's cancers included emotional support functions in a supportive role.[30] Studies designed for patients and their families (or caregivers) did not include psychosocial support.[29] In contrast, studies targeting chronic diseases, such as hypertension and diabetes, classified monitoring functions and behavioral change support as primary functions.[31,32] Some studies included all five domains, with chatbots providing clinical information, such as the management of chemotherapy-related side effects and medication guidance, while suggesting self-management strategies.[30] Conversely, a single-domain approach focused solely on information provision was observed in a study on breast cancer.[29] Overall, NLP-based chatbots were primarily used for information provision, self-management support, and clinical management support, with some studies also incorporating behavior change and emotional support functions (Multimedia Appendix 4).

As illustrated in Table 5, the outcomes of NLP-based chatbots were categorized into four domains—cognitive, clinical, psychosocial, and behavioral—as well as user experience. Across the included studies, chatbot use was generally associated with improvements in disease-related knowledge and health literacy, self-care behaviors such as medication adherence, and symptom management, as well as reductions in emotional distress. In terms of user experience, chatbots were generally reported to be acceptable and usable, with high levels of satisfaction, perceived usefulness, ease of use, and engagement. Figure 2 visually summarizes these results and illustrates the distribution of reported improvements across the result domains and the study as a whole.

Discussion

Principal findings

This systematic review analyzed six studies that utilized NLP-based chatbots to support self-management in patients with chronic diseases. Across these studies, NLP-based chatbots were primarily used to provide disease- and treatment-related information, monitor symptoms, alleviate emotional burden, and promote self-care behaviors. These findings are consistent with previous studies on NLP-based chatbots for lifestyle modification[35] and AI-based chatbots designed to enhance mental health and well-being.[36]

The most consistent finding across the included studies was that NLP-based chatbots improved cognitive outcomes, such as disease-specific knowledge and health literacy. Previous studies have emphasized the role of health literacy as a fundamental prerequisite for informed decision-making and subsequent behavioral changes in the management of chronic diseases.[37,38] This is particularly important for enabling patients to make informed decisions regarding self-management. Low health literacy may hinder sustained self-management, with its effects being less evident in short-term clinical outcomes but more likely to emerge as delayed behavioral or clinical improvements.[35,39] Despite these encouraging cognitive, emotional, and experiential outcomes, evidence for clinical outcomes is limited. Among the included studies, only Tawfik et al. reported patient-reported clinical outcomes, such as reduced severity of anticancer therapy–related symptoms and

Table 4. Domains of emphasis in NLP-based chatbot applications for chronic disease self-management.

First author, year [ref]	Information support	Clinical management support	Behavioral support	Monitoring support	Psychosocial support
Bibault et al., (2019)[29]	Primary	Supportive	–	–	–
Görtz et al., (2023)[33]	Primary	Primary	–	–	Limited
Tawfik et al., (2023)[30]	Supportive	Primary	Primary	Primary	Supportive
Bruijnes et al., (2023)[31]	–	–	Supportive	Supportive	Primary
Au et al., (2023)[34]	Primary	Supportive	Limited	Limited	Limited
Griffin et al., (2023)[32]	Supportive	Primary	Primary	Primary	–

Notes: Within each domain, functions were classified as primary, supportive, or limited based on how prominently they were described in the original studies (primary = core function; supportive = secondary role; limited = minimally described).

Table 5. Health outcomes and user experience with NLP-based chatbots.

Outcome domain	Measurements	Key findings	Studies (n)
Cognitive	· Knowledge · Health literacy	Chatbots consistently improved disease-related knowledge and understanding	5
Behavioral	· Self-care behaviors · Medication adherence	Chatbots supported adherence and self-management behaviors	4
Clinical	· Symptom management · Adverse event control	Chatbots facilitated symptom monitoring and management	1
Psychosocial	· Emotional distress · Psychosocial well-being	Chatbots reduced distress and improved emotional support	4
User experience	· Satisfaction · Usefulness · Ease of use · Engagement	Chatbots were generally well accepted, usable, and engaging across studies	6

improved symptom management.[30] However, none of the studies evaluated objective and clinically important outcomes, including blood pressure, hospitalization rates, disease progression, or mortality. This gap suggests current limitations in demonstrating clinically meaningful effects of NLP-based chatbots beyond intermediate or patient-reported outcomes.

An assessment of the reported studies reveals that, despite the relatively advanced technical capabilities of NLP-based chatbots, findings regarding their effectiveness remain fragmented and limited. User experience outcomes (satisfaction, perceived usefulness, ease of use, and engagement) were consistently positive across all studies. This supports previous findings that conversational agents are generally well accepted by patients.[40,41] Furthermore, NLP-based chatbots can interpret free-text input and tailor responses to the user's context,[19] which may explain the consistently high levels of perceived naturalness and responsiveness reported in this review. Such acceptability is a likely critical prerequisite for long-term engagement and treatment adherence in the self-management of chronic diseases.[20] These features should be understood as complementing, rather than replacing, nurse-led clinical judgment and assessment.[42,43]

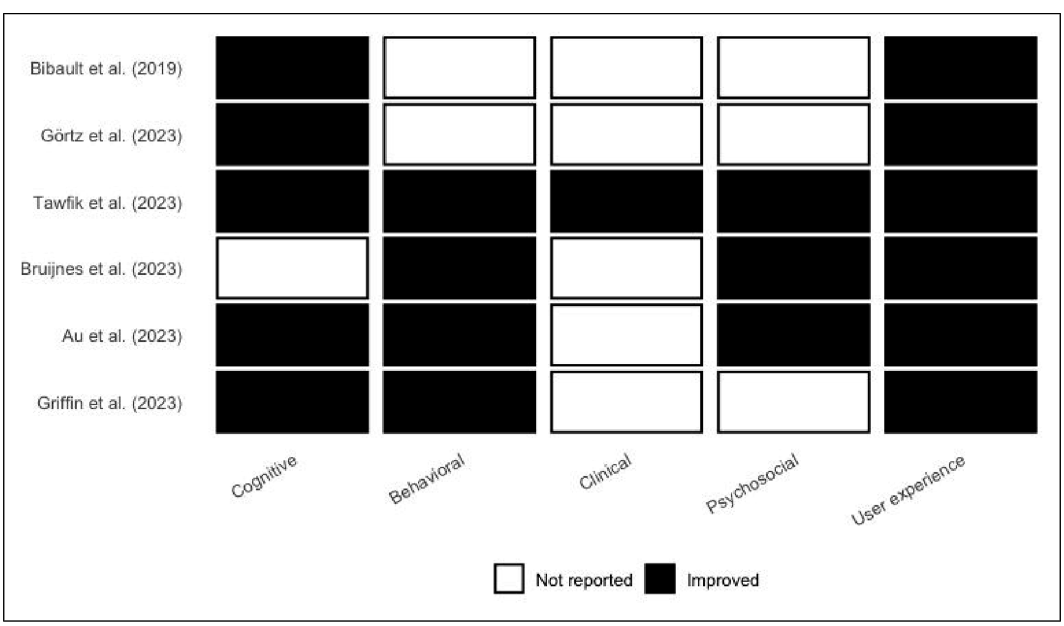

Figure 2. Heatmap of reported outcomes across NLP-based chatbot studies. Black boxes indicate that improvements were reported in the corresponding outcome domains, whereas white boxes indicate that the outcomes were not reported. Cognitive and user experience outcomes were most consistently observed, but evidence for clinical outcomes was limited.

NLP-based chatbots designed to support self-management in patients with chronic diseases are gradually expanding beyond their role as simple information-delivery tools to also coordinate overall self-management.[20,44] However, their ability to directly induce and sustain meaningful behavioral change in patients remains limited.[42,45] Given these limitations, current NLP-based chatbots must be understood within the continuum of functional maturity. Specifically, the roles expected of chatbots in nursing practice can be categorized into three levels: (1) information providers,[20] (2) interaction-based behavior coaches,[44] and (3) adaptive self-management partners.[46] Currently, most chatbots remain primarily at Level 1, with only a few implementing Level 2 functions to a limited extent. This imbalance suggests that clinical utilization lags behind the speed of technological advancements.

Consistent with this interpretation, NLP-based chatbots have predominantly been used as adjunctive tools rather than as standalone interventions. This aligns with prior findings indicating that most AI-based chatbots remain at the prototype stage and are primarily used for education, coaching, data collection, and support.[42] In our review, while information provision and clinical support functions were consistently implemented across studies, psychosocial support and monitoring functions were only partially represented in the literature. These findings are consistent with previous research demonstrating that empathetic and nonjudgmental interactions mediated by NLP-based chatbots can help alleviate psychosocial burdens associated with chronic disease self-management.[31] Such effects may be particularly pronounced in community settings where access to clinical care is limited,[36] thereby extending prior work on conversational agents by illustrating how context-aware dialogue can provide emotional support.[47,48]

Utilization of NLP-based chatbots in chronic disease self-management

Our findings further indicate that current NLP-based chatbots are distributed unevenly across the chronic disease spectrum, with a clear concentration in education-focused self-management support but limited use in conditions requiring complex, long-term care.

In this review, the target conditions included cancer, diabetes, chronic liver disease, and hypertension, which represent only a fraction of the overall burden of chronic diseases.

Therefore, NLP-based chatbots remain underutilized in high-need populations that require continuous management.[20] A crucial limitation that may reduce the system's impact within the broader healthcare network is its design for use by patients in isolation, without integration with family caregivers, community health workers, or clinical staff.[43] Even within these constraints, NLP-based chatbots have demonstrated the capacity to strengthen the core elements of self-management by improving patients' understanding of their condition, supporting the early recognition of emerging problems, and offering emotional reassurance. Furthermore, their capacity to process free-text input and adapt information to individual language and comprehension levels underscores their potential as a coordination tool that bridges home, community, and clinical settings.[49,50]

Technical potential and future research

Previous reviews have broadly examined AI and conversational agents, while describing their functions and potential. However, they rarely link specific NLP capabilities, such as free-text processing or emotion recognition, to patient-centered outcomes and user experiences.[18,21,23] Conversely, this review focused exclusively on NLP-based chatbots capable of processing unstructured text input, identifying user intent, and detecting affective cues, thereby allowing an in-depth examination of how these advanced features were implemented in chronic disease care and their association with self-management outcomes.

Although current implementations remain limited, capabilities such as NLU, intent recognition, context-sensitive response generation, and responsiveness to emotional cues can simultaneously enhance informational clarity, personalize feedback, and boost psychosocial support.[51,52] These functions align directly with the core mechanisms required for the self-management of chronic diseases, including education, monitoring, and emotional support.[44,53] However, as reported by Bruijnes et al., none of the studies included in this review successfully implemented real-time emotion detection or dynamically tailored interventions based on users' affective states.[31] This highlights the substantial gap between the theoretical capabilities of advanced NLP-based chatbots and their current implementation in clinical settings.[20,23] Addressing this gap requires translating technical capabilities into clinically meaningful and context-sensitive interventions. In this regard, nurses are well positioned to play a critical role in identifying real-world self-management needs in chronic disease care and operationalizing these needs into chatbot functions.[54] For example, these advanced NLP capabilities could be integrated into nurse-led care pathways to enable continuous updating of personalized goals,[55] escalation of unresolved or complex patient queries, and nurse-guided review of clinical data with timely, tailored feedback.[23,56] Such implementation may involve identifying challenges related to self-management during the development phase and evaluating whether the chatbot intervention—designed based on scenarios proposed by nurses—functions effectively in a real-world setting.[54,57]

Progress in this area requires longitudinal study designs that capture changes in clinically meaningful outcomes and document long-term engagement with chatbot systems.[42,58] This requires randomized controlled trials (RCTs) with sufficient sample sizes and low risk of bias. Such RCTs should compare NLP-based chatbots with other non-pharmacological interventions and evaluate the effectiveness of the overall solution. As most of the included interventions reflected only a partial implementation of NLP capabilities, future research should investigate whether more advanced features, such as contextual reasoning, emotion detection, and adaptive response generation, yield greater benefits for self-management.[51,52] Additionally, expanding the range of target conditions and integrating family caregivers or clinical personnel into chatbot-supported care pathways may provide a more realistic, real-world setting for chronic disease management.

Limitations

Despite its strengths, this study also had several limitations. First, most of the included studies were conducted at early developmental or feasibility stages, and NLP-based chatbot use relied on single- or short-term interactions with modest sample sizes, which reduced statistical power to detect clinically meaningful effects and precluded firm conclusions about sustained engagement or long-term clinical outcomes. Second, the reporting of technical implementation details was often insufficient, thereby constraining the interpretability and reproducibility of the findings. In several studies, NLP-based chatbots have been developed primarily from an engineering perspective, with minimal integration of clinical expertise. Future research should adopt co-design approaches involving nurses and patients to ensure that chatbot behaviors align with real-world self-management needs in home and community settings.

Third, the scope of this review was limited to studies published in English and excluded research on chronic pain and mental health conditions. These restrictions may have constrained the cultural and regional diversity of the evidence base, consequently limiting the generalizability of the findings to broader populations living with chronic diseases. Moreover, given the rapidly evolving nature of AI technologies, some studies included in this review may reflect the technical capabilities of earlier AI-based chatbot systems. Finally, few studies have quantitatively examined the relationship between primary outcome measures and satisfaction, which underscores the need for standardized outcome frameworks that address self-management behaviors and clinically meaningful endpoints in future NLP-based chatbot research.

Conclusions

NLP-based chatbots for chronic disease self-management remain at a relatively early stage of development but show potential as adjuncts to routine care. Across the six included studies, NLP-based chatbots were associated with improvements in disease-related knowledge, health and digital literacy, and health outcomes such as symptom distress, emotional well-being, and self-management behaviors. Additionally, they consistently received high ratings for satisfaction, perceived usefulness, usability, and engagement. However, evidence on the effects of clinically important endpoints (e.g., blood pressure control, hospitalization, or disease progression) remains sparse and methodologically limited. Future research should prioritize co-designed, robust longitudinal studies that integrate chatbots within nurse-led care pathways, employ advanced NLP capabilities, and use standardized self-management and clinical outcome measures to clarify the specific contributions of these systems to chronic disease care.

Acknowledgements

We would like to express our sincere gratitude to all those who contributed to this research.

ORCID iD

Youn-Jung Son https://orcid.org/0000-0002-0961-9606

Ethical considerations

The institutional research board of Chung-Ang University, Seoul, Korea (No. 1041078-20250522-HR-150) approved the study protocol.

Author contributions

Han GI. *Conceptualization, Data curation, Visualization, Validation*, Software. *Formal analysis, Writing – original draft, Writing – review & editing*. Lee HJ. *Formal analysis, Data curation, Validation, Writing –Writing – review & editing*. Son

Y-J. *Conceptualization, Methodology, Validation, Supervision, Project administration, Writing –Writing – original draft, Writing – review & editing*.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a grant from the National Research Foundation of Korea (NRF), funded by the Korean government (MSIT) (Grant No. RS-2025-24803234, RS-2025-00562535).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

Data supporting the findings of this study are available from the corresponding author on reasonable request.

Use of AI

Generative AI tools were not used in the preparation of this manuscript.

Supplemental material

Supplemental material for this article is available online.

References

1. Clark M and Bailey S. Chatbots in health care: connecting patients to information: emerging health technologies, Internet. Ottawa, ON: Canadian Agency for Drugs and Technologies in Health, 2024. Report No. EH0122. https://www.ncbi.nlm.nih.gov/books/NBK602381/
2. Aggarwal A, Tam CC, Wu D, et al. Artificial intelligence–based chatbots for promoting health behavioral changes: systematic review. *J Med Internet Res* 2023; 25: e40789. https://doi.org/10.2196/40789
3. Hindelang M, Sitaru S and Zink A. Transforming health care through chatbots for medical history-taking and future directions: comprehensive systematic review. *JMIR Med Inform* 2024; 12(1): e56628. https://doi.org/10.2196/56628
4. Haque MDR and Rubya S. An overview of chatbot-based mobile mental health apps: insights from app description and user reviews. *JMIR Mhealth Uhealth* 2023; 11: e44838. https://doi.org/10.2196/44838
5. Ayanouz S, Abdelhakim B and Benhmed M. A smart chatbot architecture based on NLP and machine learning for health care assistance. In: Proceedings of the 3rd International Conference on Networking; 2020 Mar 31–Apr 2; Marrakech, Morocco. Association for Computing Machinery, 2020, pp. 1–6. https://doi.org/10.1145/3386723.3387897
6. Leong LY, Hew TS, Ooi KB, et al. Generative AI: current status and future directions. *J Comput Inf Syst* 2025: 1–34. https://doi.org/10.1080/08874417.2025.2482571
7. Kocaballi AB, Berkovsky S, Quiroz JC, et al. The personalization of conversational agents in health care: systematic review. *J Med Internet Res* 2019; 21: e15360. https://doi.org/10.2196/15360
8. Schmager S, Pappas IO and Vassilakopoulou P. Understanding human-centred AI: a review of its defining elements and a research agenda. *Behav Inf Technol* 2025; 44: 3771–3810. https://doi.org/10.1080/0144929X.2024.2448719
9. Brown JEH and Halpern J. AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare. *SSM Ment Health* 2021; 1: 100017. https://doi.org/10.1016/j.ssmmh.2021.100017
10. Kowatsch T, Schachner T, Harperink S, et al. Conversational agents as mediating social actors in chronic disease management involving health care professionals, patients, and family members: multisite single-arm feasibility study. *J Med Internet Res* 2021; 23: e25060. https://doi.org/10.2196/25060
11. Zhu Z, Xu Q and Beheshti A. HEAL: healthcare emergency assistants leveraging large language models. In: Companion Proceedings of the ACM Web Conference 2025, 2025. https://doi.org/10.1145/3701716.3715181
12. Leitner K, Cutri-French C, Mandel A, et al. A conversational agent using natural language processing for postpartum care for new mothers: development and engagement analysis. *J Med Internet Res AI* 2025; 4: e58454. https://doi.org/10.2196/58454
13. Shankar R, Bundele A and Mukhopadhyay A. Effectiveness of chatbot interventions for reducing caregiver burden: protocol for a systematic review and meta-analysis. *MethodsX* 2025; 14: 103272. https://doi.org/10.1016/j.mex.2025.103272
14. He Y, Yang L, Qian C, et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2023; 25: e43862. https://doi.org/10.2196/43862
15. Barrett M, Boyne J, Brandts J, et al. Artificial intelligence supported patient self-care in chronic heart failure: a paradigm shift from reactive to predictive, preventive and personalised care. *EPMA J* 2019; 10: 445–464. https://doi.org/10.1007/s13167-019-00188-9

16. Hacker K. The burden of chronic disease. *Mayo Clin Proc Innov Qual Outcomes* 2024; 8: 112–119. https://doi.org/10.1016/j.mayocpiqo.2023.08.005
17. Celli BR and Barnes PJ. Exacerbations of chronic obstructive pulmonary disease. *Eur Respir J* 2007; 29: 1224–1238. https://doi.org/10.1183/09031936.00109906
18. Uetova E, Hederman L, Ross R, et al. Exploring the characteristics of conversational agents in chronic disease management interventions: a scoping review. *Digit Health* 2024; 10: 20552076241277693. https://doi.org/10.1177/20552076241277693
19. Gunawan TS, Falelmula Babiker ABF, Ismail N, et al. Development of intelligent Telegram chatbot using natural language processing. 7th International Conference on Wireless and Telematics (ICWT), 2021, pp. 1–5. https://doi.org/10.1109/ICWT52862.2021.9678471
20. Kurniawan MH, Handiyani H, Nuraini T, et al. A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. *Ann Med* 2024; 56: 2302980. https://doi.org/10.1080/07853890.2024.2302980
21. Schachner T and Keller Rv WF. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. *J Med Internet Res* 2020; 22: e20701. https://doi.org/10.2196/20701
22. Wang YF, Hsu MH and Wang MYF. HealthLINE: a messaging chatbot for supporting chronic disease self-management. *Digit Health* 2025; 11: 20552076251394891. https://doi.org/10.1177/20552076251394891
23. Peerbolte TF, van Diggelen RJ, van den Haak P, et al. Conversational agents supporting self-management in people with a chronic disease: systematic review. *J Med Internet Res* 2025; 27: e72309. https://doi.org/10.2196/72309
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372: n71. https://doi.org/10.1136/bmj.n71
25. World Health Organization. *Noncommunicable diseases [Internet]*. : WHO Regional Office for the Western Pacific, 2023. [cited 2025 Oct 3]. Available from: https://www.who.int/westernpacific/health-topics/noncommunicable-diseases
26. Korwisi B, Barke A, Rief W, et al. Chronic pain in the 11th Revision of the International Classification of Diseases: users' questions answered. *Pain* 2022; 163: 1675–1687. https://doi.org/10.1097/j.pain.0000000000002551
27. Hong QN, Gonzalez-Reyes A and Pluye P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *J Eval Clin Pract* 2018; 24: 459–467. https://doi.org/10.1111/jep.12884
28. Harrison R, Jones B, Gardner P, et al. Quality assessment with diverse studies (QuADS): an appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. *BMC Health Serv Res* 2021; 21: 144. https://doi.org/10.1186/s12913-021-06122-y
29. Bibault JE, Chaix B, Guillemassé A, et al. A chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial. *J Med Internet Res* 2019; 21: e15787. https://doi.org/10.2196/15787
30. Tawfik E, Ghallab E and Moustafa A. A nurse versus a chatbot—the effect of an empowerment program on chemotherapy-related side effects and the self-care behaviors of women living with breast cancer: a randomized controlled trial. *BMC Nurs* 2023; 22: 102. https://doi.org/10.1186/s12912-023-01243-7
31. Bruijnes M, Kesteloo M and Brinkman WP. Reducing social diabetes distress with a conversational agent support system: a three-week technology feasibility evaluation. *Front Digit Health* 2023; 5: 1149374. https://doi.org/10.3389/fdgth.2023.1149374
32. Griffin AC, Khairat S, Bailey SC, et al. A chatbot for hypertension self-management support: user-centered design, development, and usability testing. *JAMIA Open* 2023; 6: ooad073. https://doi.org/10.1093/jamiaopen/ooad073
33. Görtz M, Baumgärtner K, Schmid T, et al. An artificial intelligence-based chatbot for prostate cancer education: design and patient evaluation study. *Digit Health* 2023; 9: 20552076231173304. https://doi.org/10.1177/20552076231173304
34. Au J, Falloon C, Ravi A, et al. A beta-prototype chatbot for increasing health literacy of patients with decompensated cirrhosis: usability study. *JMIR Hum Factors* 2023; 10: e42506. https://doi.org/10.2196/42506
35. Chen S, Wang Y, Liu W, et al. Mechanism of health literacy impact on self-management behaviors in patients with chronic disease: a self-efficacy mediated model moderated by disease duration. *Front Public Health* 2025; 13: 1673723. https://doi.org/10.3389/fpubh.2025.1673723
36. Li H, Zhang R, Lee YC, et al. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digit Med* 2023; 6: 236. https://doi.org/10.1038/s41746-023-00979-5
37. Shao Y, Hu H, Liang Y, et al. Health literacy interventions among patients with chronic diseases: a meta-analysis of randomized controlled trials. *Patient Educ Couns* 2023; 114: 107829. https://doi.org/10.1016/j.pec.2023.107829
38. Liu L, Qian X, Chen Z, et al. Health literacy and its effect on chronic disease prevention: evidence from China's data. *BMC Public Health* 2020; 20: 690. https://doi.org/10.1186/s12889-020-08804-4
39. Muscat DM, Song W, Cvejic E, et al. The impact of the chronic disease self-management program on health literacy: a pre-post study using a multi-dimensional health literacy instrument. *Int J Environ Res Public Health* 2019; 17: 58. https://doi.org/10.3390/ijerph17010058

40. Wutz M, Hermes M, Winter V, et al. Factors influencing the acceptability, acceptance, and adoption of conversational agents in health care: integrative review. *J Med Internet Res* 2023; 25: e46548. https://doi.org/10.2196/46548
41. Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. *J Med Internet Res* 2020; 22: e20346. https://doi.org/10.2196/20346
42. Altom DS, Awad TAI, et al.; Mahmoud Hussein AAA Artificial intelligence-based chatbots in chronic disease management: a systematic review of applications and challenges. *Cureus* 2025; 17: e81001. https://doi.org/10.7759/cureus.81001
43. Laymouna M, Ma Y, Lessard D, et al. Roles, users, benefits, and limitations of chatbots in health care: rapid review. *J Med Internet Res* 2024; 26: e56930. https://doi.org/10.2196/56930
44. Du Y, Yang P, Liu Y, et al. Artificial intelligence in chronic disease self-management: current applications and future directions. *Front Public Health* 2025; 13: 1689911. https://doi.org/10.3389/fpubh.2025.1689911
45. Hou J, Lin S, Teng P, et al. Application of chatbots in chronic disease management: a scoping review. *Digit Health* 2026; 12: 20552076251411287. https://doi.org/10.1177/20552076251411287
46. Hwang M, Zheng Y, Cho Y, et al. AI applications for chronic condition self-management: scoping review. *J Med Internet Res* 2025; 27: e59632. https://doi.org/10.2196/59632
47. Kelly A, Noctor E, Ryan L, et al. The effectiveness of a custom AI chatbot for type 2 diabetes mellitus health literacy: development and evaluation study. *J Med Internet Res* 2025; 27: e70131. https://doi.org/10.2196/70131
48. Kim H and Wang Y. Unveiling the human touch: how AI chatbots' emotional support and human-like profiles reduce psychological reactance to promote user self-disclosure in mental health services. *Int J Advert* 2025: 1–25. https://doi.org/10.1080/02650487.2025.2558479
49. Serugunda HM, Jianquan O, Kasujja Namatovu H, et al. Using large language models for chronic disease management tasks: scoping review. *JMIR Med Inform* 2025; 13: e66905. https://doi.org/10.2196/66905
50. Zhang S and Song J. A chatbot based question and answer system for the auxiliary diagnosis of chronic diseases based on large language model. *Sci Rep* 2024; 14: 17118. https://doi.org/10.1038/s41598-024-67429-4
51. Al-Anezi FM. Exploring the use of ChatGPT as a virtual health coach for chronic disease management. *Learn Health Syst* 2024; 8: e10406. https://doi.org/10.1002/lrh2.10406
52. Soto-Chávez MJ, Bustos MM, Fernández-Ávila DG, et al. Evaluation of information provided to patients by ChatGPT about chronic diseases in Spanish language. *Digit Health* 2024; 10: 20552076231224603. https://doi.org/10.1177/20552076231224603
53. Vangibhurathachhi SK. Evaluating the effectiveness of AI chatbots in chronic disease self-management. *IJETCSIT* 2025: 38–45. https://doi.org/10.63282/WCAI25-110
54. Park G, Lee H and Khang AR. The development of automated personalized self-care (APSC) program for patients with type 2 diabetes mellitus. *J Korean Acad Nurs* 2022; 52(5): 535–549. https://doi.org/10.4040/jkan.22046
55. Han GI, Choi H and Son YJ. Advantages and challenges for utilization of generative artificial intelligence in clinical nursing practice: an integrative review. *BMC Nurs* 2026, Epub ahead of print. https://doi.org/10.1186/s12912-026-04518-x
56. McGrow K. Artificial intelligence in nursing: a journey from data to wisdom. *Nursing (Lond)* 2025; 55(4): 16–24. https://doi.org/10.1097/NSG.0000000000000165
57. Son YJ, Choi J, Kim HM, et al. Development and evaluation of a patient-family caregiver dyad mHealth intervention for heart failure self-care: quasi-experimental study. *J Med Internet Res* 2025; 27: e74922. https://doi.org/10.2196/74922
58. Huo B, Boyle A, Marfo N, et al. Large language models for chatbot health advice studies: a systematic review. *JAMA Netw Open* 2025; 8: e2457879. https://doi.org/10.1001/jamanetworkopen.2024.57879