

Review

Usability and User Experience Assessment of Health Care Conversational Agents Using Validated Subjective Instruments: Systematic Review and Comparative Analysis

João Pavão¹, Prof Dr; Rute Bastardo², Prof Dr; Anabela Gonçalves Silva³, Prof Dr; Nelson Pacheco Rocha⁴, Prof Dr

¹Institute for Systems and Computer Engineering, Technology and Science—INESC-TEC, Science and Technology School, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal

²UNIDCOM, Science and Technology School, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal

³Center for Health Technology and Services Research, School of Health Sciences, University of Aveiro, Aveiro, Portugal

⁴Institute of Electronics and Informatics Engineering of Aveiro, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal

Corresponding Author:

Nelson Pacheco Rocha, Prof Dr
Institute of Electronics and Informatics Engineering of Aveiro
Department of Medical Sciences
University of Aveiro
Campus Universitário de Santiago
Aveiro, 3810-193
Portugal
Phone: 351 234 370 200
Email: npr@ua.pt

Abstract

Background: Digital applications based on health care conversational agents (HCAs) are increasingly being developed to support health care provision. The usability and user experience of these solutions are critical determinants of their acceptability and, consequently, their impact on health-related outcomes.

Objective: This systematic review aims to synthesize current evidence on the use of valid and reliable subjective instruments for assessing the usability and user experience of HCAs and to examine whether assessment outcomes vary according to their technical characteristics.

Methods: A systematic search was conducted in PubMed, Web of Science, and Scopus from inception to February 2026. Studies were included if they used subjective instruments to evaluate the usability or user experience of HCAs.

Results: A total of 127 studies met the inclusion criteria. The studies examined 3 categories of HCAs—text-based, voice-based, and embodied—applied to patient care, health education and prevention, health data collection, and support for daily activities among older adults. The System Usability Scale (SUS) was the most frequently used assessment instrument. Comparative analysis of SUS scores indicated higher usability ratings for text-based HCAs relative to voice-based and embodied systems. However, SUS and other subjective instruments used in the included studies may not fully capture key dimensions of usability and user experience of HCAs. Additionally, substantial heterogeneity was observed in assessment methodologies across studies.

Conclusions: Comparative analysis suggested that text-based HCAs were associated with significantly higher SUS scores than voice-based and embodied HCAs. However, this finding should be interpreted with caution given the substantial heterogeneity across the included studies in health care application domains, study designs, evaluation contexts, participant populations, and HCAs' implementation and use characteristics, as well as the limitations of the SUS in evaluating the usability of modern HCAs. The variability in assessment approaches underscores the need for standardized protocols and the development of more context-specific evaluation frameworks to enhance methodological consistency and comparability across studies.

(JMIR Hum Factors 2026;13:e102911) doi: [10.2196/102911](https://doi.org/10.2196/102911)

KEYWORDS

digital health; health care provision; usability; user experience; usability assessment; user experience assessment; conversational agents; older adults

Introduction

Since the development of ELIZA in 1966 [1], a rule-based system developed to emulate text-based interactions with a psychotherapist, advancements in information technology have driven the widespread adoption of health care conversational agents (HCAs) [2,3], which are applications designed to support interactive, multiturn communication through written or spoken language [4]. Consumer solutions such as Siri, Google Now, Cortana, and Alexa are well-known examples of agents designed to help users find information and accomplish tasks [5].

Modern HCAs extend well beyond simple rule-based input-output interactions. Advances in large language models have facilitated their evolution from narrowly defined, task-oriented tools into open-domain systems capable of retaining conversational context, following topics across multiple turns, and sustaining coherent, continuous dialogue for more natural interactions [6,7]. In addition, AI techniques have enabled more advanced dialogue management, facilitating the exhibition of human-like behaviors, the recognition of users' emotions and intentions, and the use of various information sources for reasoning and decision-making [8-10].

Building on these advancements, researchers have explored anthropomorphized interfaces, leading to the development of embodied agents, or virtual humans [11,12]. Unlike traditional agents, which rely solely on text or voice, embodied agents combine verbal communication with nonverbal cues (eg, posture, gestures, eye movements, or facial expressions) to foster more lifelike interactions [13-15]. This embodiment enhances users' perception of social presence and engagement, making interactions feel more intuitive and emotionally rich [16,17]. Moreover, the introduction of virtual reality technologies has paved the way for increasingly sophisticated agents designed to promote user engagement through immersive experiences [18].

HCAs are widely used in health care [19], reflecting the widespread adoption of similar agents across other human-centered sectors, including customer service [20], entertainment [21], and education [22].

From a health care research perspective, the assessment of HCAs should consider health-related outcomes, such as symptom reduction, disease knowledge, disease self-management, and quality of life [8]. This assessment should be conducted in the context of continuous use of the applications and preferably supported by randomized controlled trials. However, before conducting randomized controlled trials, it is important to safeguard aspects that might affect health-related outcomes. This means that preliminary assessments (ie, feasibility studies) should be conducted [23] to evaluate the technical characteristics of HCAs or their quality in terms of usability and user experience and, consequently, their potential to promote user engagement among the target users [8,24].

System usability comprises several components (ie, learnability, efficiency, memorability, errors, and satisfaction) and relates

to the ease of using the system interface [25]. The rise of new technologies, devices, and interaction methods has expanded the concept of usability into the broader notion of user experience, which encompasses the cognitive, affective, social, and physical aspects of interaction [26-28].

There are 3 main methods for evaluating usability and user experience, namely, inspection, testing, and inquiry [29]. The inspection method consists of a systematic analysis of interaction mechanisms carried out by experts according to specific guidelines, while the testing and inquiry methods are based on the collection of usage data [29]. The testing method involves observing users while they perform tasks with a given system to collect primarily quantitative data (eg, the number of completed tasks) that can provide empirical evidence on how to improve the interaction mechanisms [29]. By contrast, the inquiry method involves collecting qualitative data from users, namely through interviews or subjective measurement instruments such as questionnaires or scales [29]. Examples of commonly used instruments include the Post-Study System Usability Questionnaire (PSSUQ) [30] and the System Usability Scale (SUS) [31]. These instruments have had their measurement properties assessed, particularly validity (ie, how effectively usability or user experience is measured) and reliability (ie, the consistency of measurements when repeated under comparable conditions) [30-33].

In this context, this review aims to synthesize the current evidence on usability and user experience assessments of HCAs using valid and reliable subjective instruments, to identify the purposes and categories of HCAs being assessed, and to determine, through a comparative analysis, how this typification affects their usability and user experience.

Several reviews [34-41] have already focused on different aspects of HCAs (eg, the use of AI or large language models), including usability and user experience assessments (eg, [42-45]); however, these reviews did not perform comparative analyses of usability and user experience across different categories of HCAs. Such analyses are important because different characteristics of HCAs affect how users interact with them (eg, by shaping communication style, pacing, and information processing), as well as cognitive load and perceptions of trust and privacy. Therefore, this review may provide insights to inform better design decisions and the development of standardized evaluation frameworks, enabling researchers and practitioners to benchmark usability and user experience across different HCAs.

Methods

Study Protocol Registration and Scope

The research protocol for this systematic review was registered in the Open Science Framework and followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; see [Multimedia Appendix 1](#)) guidelines [46]. The research scope ([Textbox 1](#)) was structured using the Population, Concept, and Context framework [47].

Textbox 1. Research scope.

1. Population

Adults (≥ 18 years) managing their health conditions using digital health applications.

2. Concept

Usability and user experience assessments of health care conversational agents designed for interactive, multiturn communication through written or spoken language to support the management of users' health conditions, supported by validated subjective instruments (ie, questionnaires or scales used to systematically collect users' self-reported evaluations of usability or user experience).

3. Context

Community or institutional care settings.

Literature Search

Scopus, Web of Science, and PubMed were selected as the online databases for retrieving studies included in this review. The search query included keywords representing the following 3 main components:

- HCAs: conversational agent, relational agent, virtual human, virtual agent, virtual assistant, virtual companion, virtual coach, chatbot, enhanced communication agent, embodied communication agent, ECA, and avatar.

- Health care: health care, health care, care, patient, digital health, electronic health, e-health, eHealth, older adults, and elderly.
- Usability or user experience assessment: usability, user experience, acceptance, evaluation, assessment, and test.

The keywords within each component were linked using the Boolean operator OR, whereas the expressions derived from different components were linked using the Boolean operator AND. As an example, [Figure 1](#) presents the search query prepared for the Scopus database.

Figure 1. Search query for the Scopus database.

TITLE-ABS-KEY(("conversational agent" OR "relational agent" OR "virtual human" OR "virtual agent" OR "virtual assistant" OR "virtual companion" OR "virtual coach" OR chatbot OR "enhanced communication agent" OR "embodied communication agent" OR ECA OR avatar) AND (patient OR "digital health" OR "electronic health" OR "e-health" OR ehealth OR healthcare OR "health care" OR care OR "older adults" OR elderly) AND (usability OR "user experience" OR acceptance OR evaluation OR assessment OR test)).

Inclusion and Exclusion Criteria

To be included in the review, studies had to meet the eligibility criteria presented in [Textbox 2](#).

Textbox 2. Inclusion and exclusion criteria.

1. Inclusion criteria

- Empirical studies published in peer-reviewed scientific journals or conference proceedings.
- Adults (≥ 18 years).
- Studies evaluating health care conversational agents (HCAs) to support the management of users' health conditions.
- Studies performing usability or user experience assessments of HCAs based on validated subjective instruments.

2. Exclusion criteria

- Review articles, opinion pieces, letters, and editorials.
- Children or adolescents.
- Studies evaluating the integration of HCAs into clinician-oriented digital health applications or applications focused on domains other than health care.
- Studies assessing usability or user experience without using validated subjective instruments or not focusing on this type of assessment.
- Preliminary versions of studies for which a more mature version was included in the review (eg, when a study was published as both a conference paper and later a journal article, the conference version was excluded).
- Studies published in languages other than English.
- Studies for which the full text was not available.

Identification Process

All references retrieved from the databases were imported into a Microsoft Excel spreadsheet. The selection process then followed the general guidelines of the PRISMA flowchart [46]: (1) exclusion of duplicates, (2) title and abstract screening, and (3) full-text screening. Additionally, a backward search was performed based on the reference lists of the articles included for full-text screening.

All stages were performed manually by 2 reviewers (JP and NPR) without the use of automation tools. Disagreements between the reviewers were discussed and resolved collaboratively.

Quality Assessment

The methodological quality assessment of the reviewed studies was conducted using a subset of the Critical Assessment of Usability Studies Scale (CAUSS) [48]. As this scale evaluates both testing and inquiry methods, the reviewers (AGS and NPR) selected 5 items related to assessments using inquiry methods: (1) representativeness of the participants, (2) proximity to the real-world context of use, (3) adequacy of the number of participants, (4) representativeness of the functionalities assessed, and (5) continuous and prolonged use of the app.

The excluded items comprised (1) items related to the validation of measurement instruments, as the eligibility criteria for this review explicitly excluded studies that did not use validated instruments; (2) items related to testing methods involving the direct observation of participants while they performed tasks to collect objective measures, such as error rates or mean task completion time; and (3) the item related to the triangulation of testing and inquiry methods, as testing methods were not considered in this review.

As only a subset of the CAUSS was applied, the resulting scores should not be interpreted as reflecting the overall quality of the usability studies, but rather the quality of the use of subjective measurement instruments.

Data Extraction

The following data were extracted:

- Demographic characteristics, namely, the author(s), nationality of the authors' affiliations, type of publication, and year of publication.
- Purposes of the assessed HCAs.
- Technical characteristics of the HCAs and supporting technologies (eg, AI or virtual reality).
- Types of experimental study designs (eg, cross-sectional or longitudinal studies).
- Duration of the longitudinal studies.
- Number and characteristics of the study participants.
- Recruitment and distribution of the participants.
- Validated subjective instruments assessing usability or user experience.
- Results of the use of the validated subjective assessment instruments (ie, scores and SDs).

Data were independently extracted from all included studies by 2 reviewers (JP and RB) using a customized Microsoft Excel

form. Discrepancies were resolved through discussion and consensus.

Data Synthesis

Results were synthesized into summary tables, figures, and descriptive text according to the following domains: (1) demographic characteristics (ie, publication types and publication years); (2) purposes of HCAs; (3) technical characteristics of HCAs; (4) study experimental characteristics (ie, experimental design, number and characteristics of the participants, and measurement instruments); and (5) results of usability and user experience assessments.

A quantitative analysis was performed only for studies using the SUS to assess usability. The number of studies using other instruments was insufficient for quantitative analysis. Studies using the SUS were subdivided based on the categories of HCAs being assessed. To compare SUS scores and the number of participants across the categories of HCAs, the Kruskal-Wallis test was used because the data did not follow a normal distribution (Kolmogorov-Smirnov test, $P < .05$). Medians and IQRs were provided for each group. Additionally, a subgroup analysis comparing SUS scores between studies involving older adults and those involving adults was performed for all studies and within each category of HCAs using the Mann-Whitney U test. The Bonferroni correction for multiple tests was applied. The significance level was set at $P < .05$. Analyses were performed using SPSS.

In addition, the effect size r was manually calculated as the standardized test statistic from the Mann-Whitney U test or the Kruskal-Wallis test (z) divided by the square root of the total sample size (N) ⁽²⁶⁾ and interpreted as small (≥ 0.1), medium (≥ 0.3), or large (≥ 0.5) [49]. For the Kruskal-Wallis test, the ordinal eta-squared (η^2H) was calculated as $\eta^2H = (H - K + 1)/(N - K)$, where H is the test statistic, N is the sample size, and K is the number of groups. It was interpreted as small ($\eta^2H \geq 0.01$), medium ($\eta^2H \geq 0.06$), or large ($\eta^2H \geq 0.14$) [49].

Given the substantial heterogeneity among the included studies—including differences in health care application domains, study designs, evaluation contexts, participant populations, and HCAs' implementation and use characteristics—the quantitative comparison of SUS scores was intended as an exploratory synthesis of patterns across studies rather than as evidence of causal differences in usability among HCA categories. Moreover, the SUS may not adequately capture all relevant aspects of the usability of the evaluated HCAs, such as conversational quality. Consequently, the comparative analyses should be interpreted as identifying associations at the study level rather than establishing the intrinsic usability of specific HCA categories.

Results

Identification of Studies

The online search was performed in February 2026, and 8639 references were retrieved. As presented in Figure 2, after the exclusion of 3391 duplicates, the titles and abstracts of 5248 references were analyzed, and 5094 were excluded according

to the eligibility criteria. Therefore, 154 references were included for full-text analysis. During the full-text analysis, 30 references [50-79] were excluded because they did not meet the eligibility criteria (eg, studies that did not apply valid and reliable instruments to measure usability or user experience, or

studies focused on the assessment of clinician-oriented applications). Additionally, the backward search identified 3 references, resulting in 127 references [80-206] being included in the review.

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.

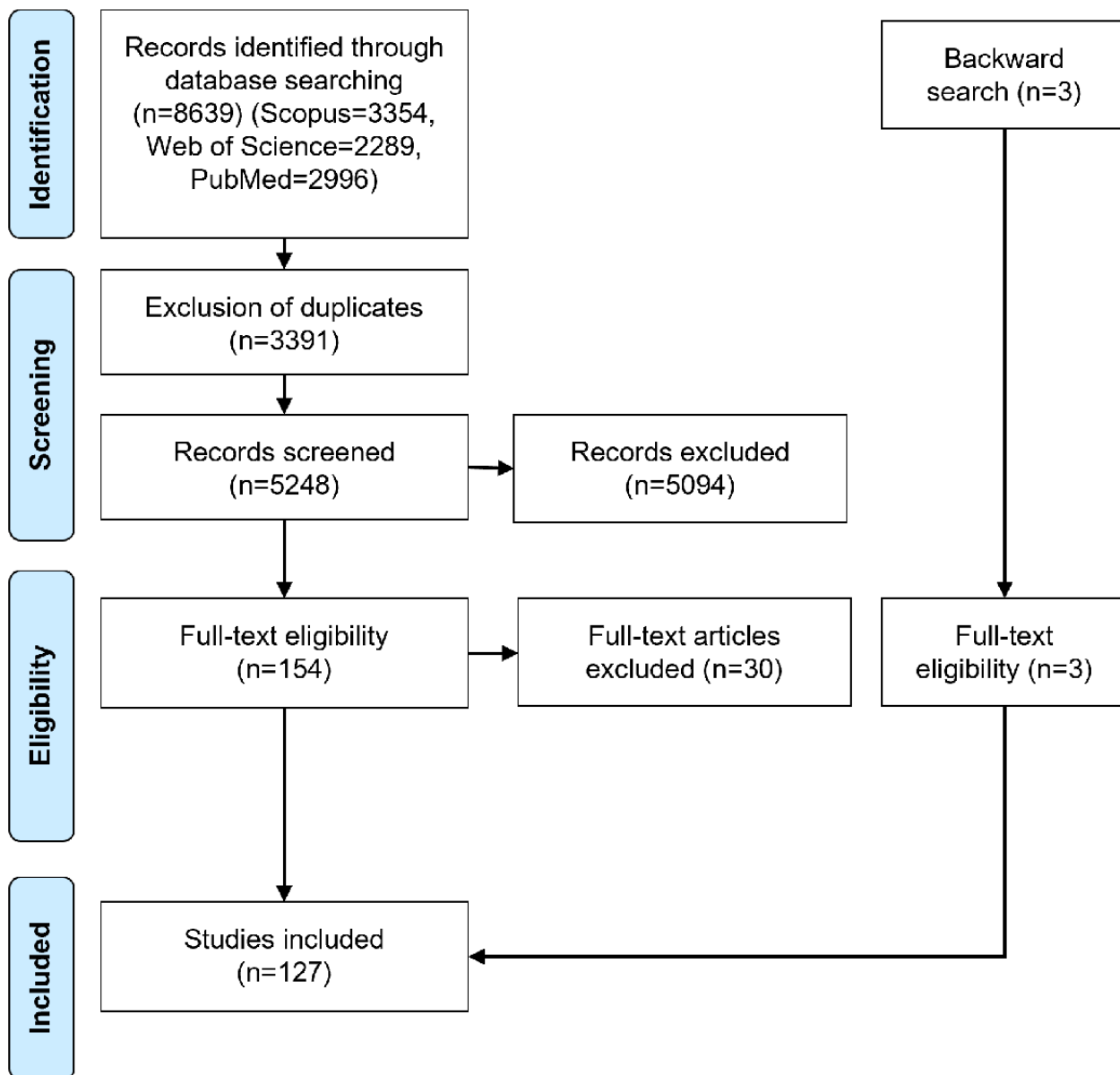
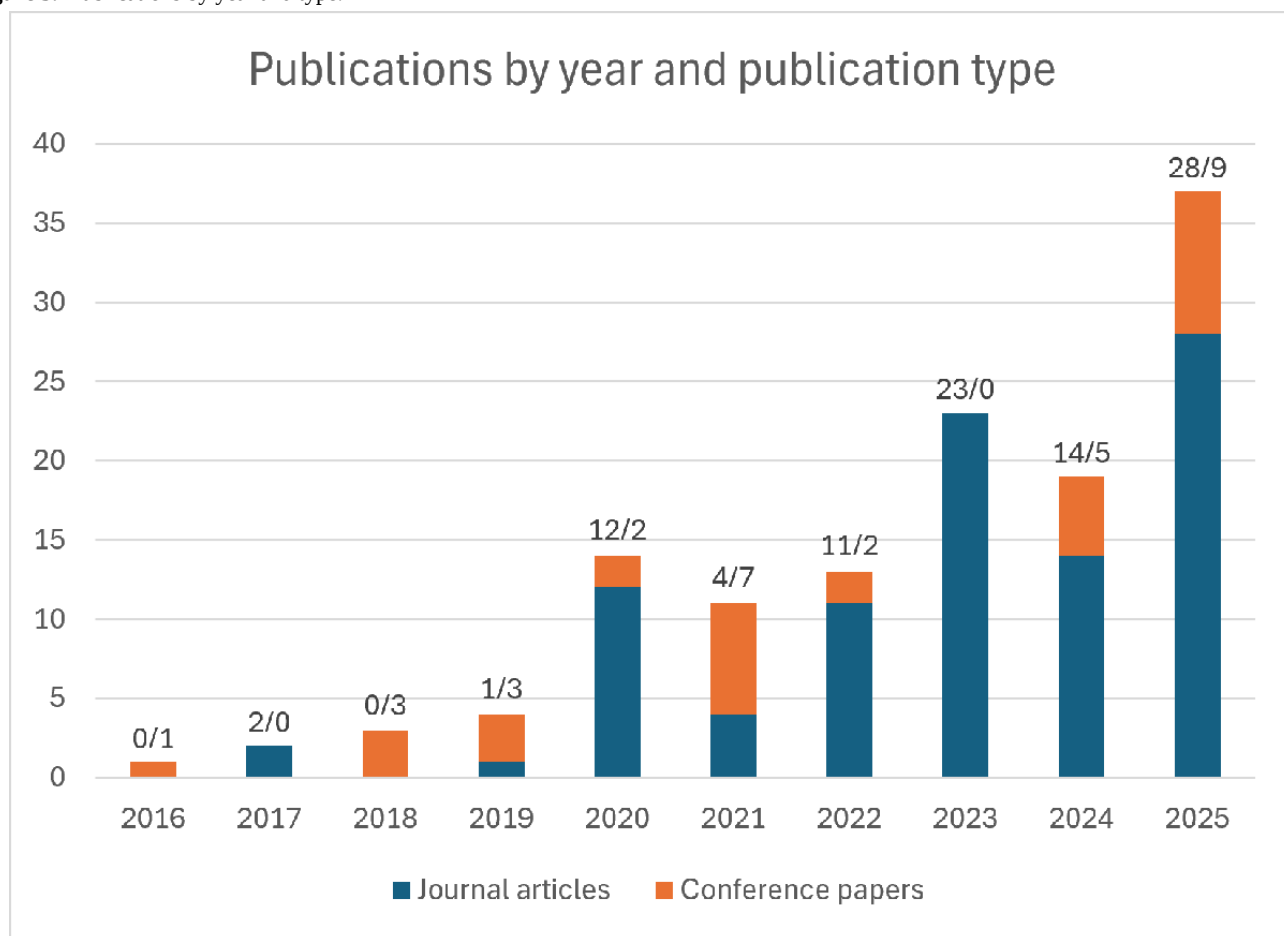


Figure 3 presents the distribution of publication years by publication type (ie, journal articles and conference proceedings).

Figure 3. Publications by year and type.

Quality Assessment Results

Most CAUSS criteria considered for the methodological quality assessment were met by a majority of studies: (1) representativeness of the participants (113/127, 89%); (2) proximity to the real-world context of use (105/127, 82.7%); (3) adequacy of the number of participants (90/127, 70.9%); and (4) representativeness of the functionalities assessed (96/127, 75.6%). The exception was the criterion related to the continuous use of the application, which was met by only 44/127 (34.6%) studies. In turn, 28/127 (22%) studies met all 5 criteria.

Purposes of the HCAs

In terms of the purposes of the HCAs described in the included articles, 4 major categories were identified:

- Patient care: 73 studies [81,82,85,87,89,92,95-97,104,106,107,111,113-115,119,121,123,124,127,128,130,132,133,135,136,138,140,141,144-151,153-157,159,162,165-170,172-175,177,179-184,187-189,191,194,197,198,203-206].
- Health education and prevention: 33 studies [86,90,93,103,109,110,112,116-118,122,125,129,131,134,137,142,143,158,160,163,164,171,176,185,192,193,195,196,199-202].
- Health data collection: 15 studies [88,91,98,99,101,102,105,108,120,126,139,152,161,178,186].
- Daily activities of older adults: 6 studies [80,83,84,94,100,190].

Concerning patient care, almost half of the respective articles (ie, 34 [81,82,85,92,95,97,103,104,119,121,123,124,132,147,

149,150,153,155,165-170,172,173,179,181,184,188,191,204-206]) targeted mental health disorders (eg, depression, anxiety, stress, posttraumatic stress disorder, eating disorders, and suicidal risk), while the remaining 39 [87,89,96,106,107,111,114,115,127,128,130,133,135,136,138,140,141,144-146,148,151,154,156,157,159,162,174,175,177,180,182,183,187,189,194,197,198,203] articles targeted a range of other conditions, including cancer [138,145,182,183,197], chronic kidney disease [144], chronic pain [115,162,198], COVID-19 [135], delirium [146], diabetes [89,96,111,115,140,194], hemophilia [136], heart-related diseases [128,130,133,154,159,174,187,203], hydrocephalus [177], intestinal ostomy [148], neurodegenerative diseases [107,141,157,175], psoriasis [127], sickle cell disease [114], traumatic injuries [106,151,180], and vestibular disorders [189].

Looking at health education and prevention, beginning with health education, some articles [117,131,137,158,160,171,176,185,193,195,196,202] focused on general health education topics, while others focused on specific topics, such as (1) combating misinformation about COVID-19 [109,113,134], (2) endoscopy explanation [164], (3) hematologic malignancies [110], (4) hereditary breast and ovarian cancer [201], (5) perinatal information for parents [143], and (6) vaccine pharmacovigilance [142]. In the studies that focused on health prevention, the assessed HCAs were designed to promote behavior change (eg, physical activity and weight maintenance) [86,93,112,116,118,125,129,199,200], support vaccination (ie, COVID-19 [122] and human papillomavirus [192]), and provide balance training [90,163].

In terms of health data collection, 11 [88,98,101,105,108,120,126,139,152,178,186] studies assessed HCAs for collecting patient-reported outcome measures, 3 [91,99,161] focused on collecting family health history, and 1 [102] described an application designed to collect data for the evaluation of health care services and quality of care.

Finally, 6 [80,83,84,94,100,190] studies assessed HCAs designed to promote the independence and autonomy of older adults by assisting them in managing activities of daily living, including problem-solving [100] and cooking [190].

Technical Characteristics of HCAs

Three categories of HCAs were identified: (1) text-based HCAs, (2) voice-based HCAs, and (3) embodied HCAs.

Five studies [87,98,123,153,169] assessed 2 categories of HCAs: (1) voice-based and embodied [87], (2) text- and voice-based [98,123], and (3) text-based and embodied [153,169]. The remaining studies focused on a single category, according to the following distribution:

- Text-based HCAs: 69 studies [85,86,91,99,100,102,103,105,106,109,110,112-117,119,120,122,125,126,129-134,

137-145,147-150,154,157-159,161,164,166,168,170,177,184-187,191,194-206].

- Voice-based HCAs: 18 studies [93,97,101,104,111,118,124,127,135,136,146,155,165,167,172,183,190,192].
- Embodied HCAs: 35 studies [80-84,88-90,92,94-96,107,108,121,128,151,152,156,160,162,163,171,173-176,178-182,188,189,193].

Some studies [91,99,100,102,105,108,109,112,124,126,137,138,147,152,155,161,163,173,178], although maintaining the HCA categories, established different comparisons, namely, (1) HCAs versus regular computer-based questionnaires [91,99,102,105,126,161]; (2) similar applications with different functionalities [108,109,112,137,152,155,173,178] or different intervention strategies [100,124,147,163]; and (3) HCAs versus usual care [138].

Figure 4 illustrates the distribution of studies across publication years by HCA category. According to the figure, the number of studies assessing voice-based HCAs has remained approximately constant, whereas the number of studies assessing text-based and embodied HCAs has increased over the past 3 years. Table 1 presents the relationship between HCA categories and their purposes.

Figure 4. Distribution of studies across publication years, categorized by types of health care conversational agents (HCAs).

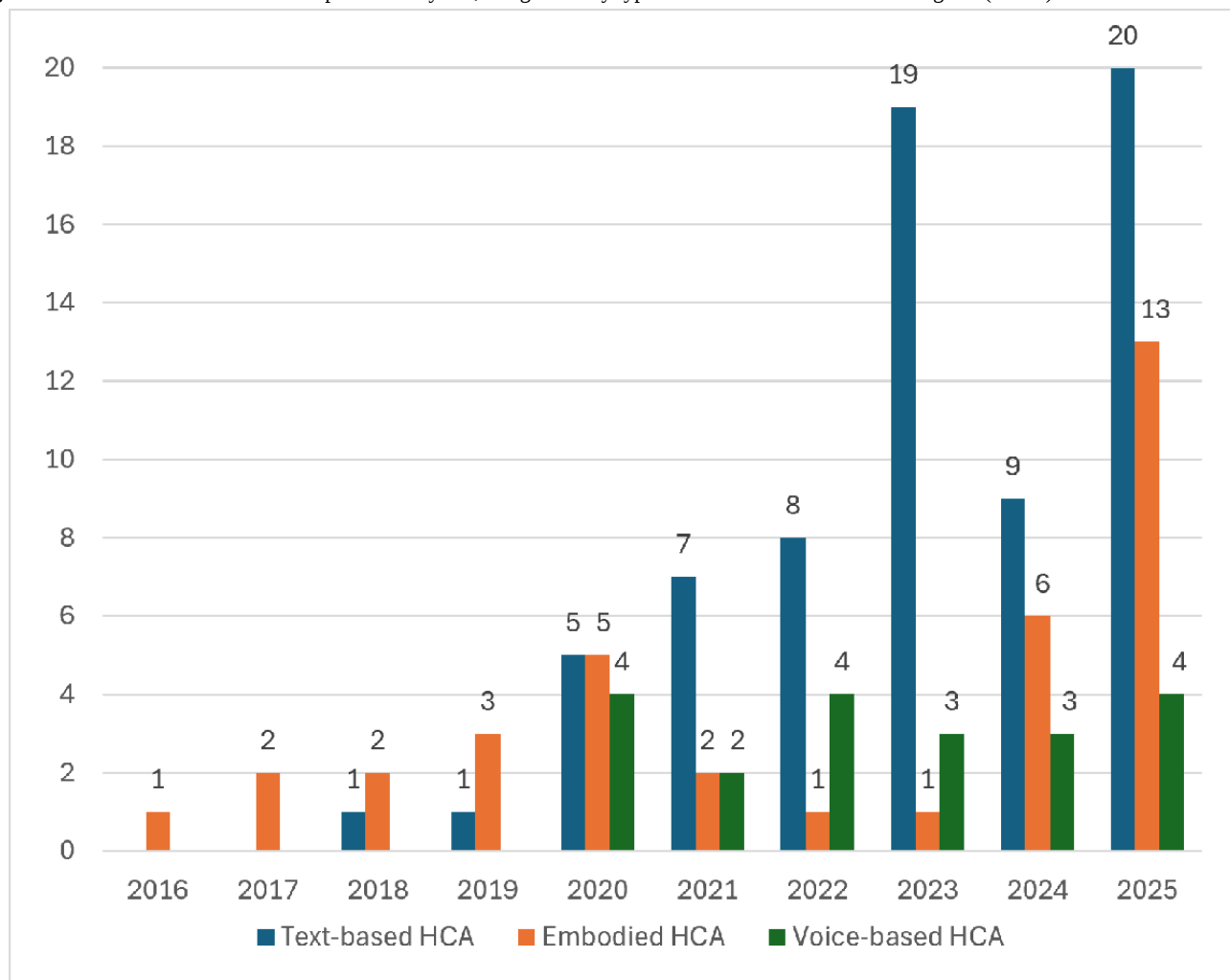


Table 1. Relation of the categories of HCAs and the purposes of the applications.

Purposes	Text-based HCAs	Voice-based HCAs	Embodied HCAs
Patient care—mental health	[85,103,119,123,132,147,149,150,153,166,168-170,184,191,204-206]	[97,104,123,124,155,165,167,172]	[81,82,92,95,121,153,169,173,179,181,188]
Patient care—other	[87,106,114,115,130,133,138,140,141,144,145,148,154,157,159,177,187,194,197,198,203]	[111,127,135,136,146,183]	[87,89,96,107,128,151,156,162,175,180,189]
Health care education and prevention	[86,112,113,116,122,125,129,164,185,195,196,199-202]	[93,118,192]	[90,163,171,193]
Health data collection	[91,98,99,102,105,120,126,139,161,186]	[98,101]	[88,108,152,178]
Older adult daily activities	[100]	[190]	[80,83,84,94]

^aHCA: health care conversational agent.

Eleven of the embodied HCAs [87,90,92,108,163,171,174,181,182,188,189] were implemented using virtual reality technologies. Of these, 7 [92,108,163,174,181,182,189] were implemented with fully immersive virtual reality technologies, while the remaining 4 used augmented reality [90,171] and mixed reality [87,188] technologies.

Moreover, some studies [80,83,84,89,93,94,96,103,109,110,129,134,155,156,172,179-181] implemented algorithms to recognize users' emotional states. The main reasons for emotion recognition were to enable the HCAs to provide appropriate empathetic responses [80,83,84,89,93,94,96,109,110,129,134,155,156,179] or to suggest specific activities for emotion regulation [155,172,180,181].

Natural language processing based on AI frameworks (eg, Google Dialogflow, Microsoft Language Understanding Intelligent Service, Rasa, or Oskova) was used in several studies to improve dialogue flow [85,86,101,110,112,117,120,124,136,145,167,168,178,179,185,191,195-197,200,202]. In addition, large language models were used in some of the more recent studies [179,185,191,195,196,202].

AI techniques were implemented not only to improve dialogue flow, but also for other purposes, namely, (1) text-based emotion recognition [103,155]; (2) facial emotion recognition [172]; (3) definition of contextualized instructions [171]; (4) avatar generation [176]; (5) prediction of diabetes self-management [194]; and (6) improvement of the consistency and reliability of the information provided [204]. However, most studies did not report evidence of the efficacy of these techniques.

Studies' Experimental Characteristics

Experimental Design

A total of 43 studies [83,84,96,100,104,107,110,112,115,116,118,121,122,124,127-130,132,138,140,141,144,145,149,151,154,160,162,165,170,173,175,177,183,184,197,199-201,203-205] used a longitudinal design, among which 15 studies [100,116,121,124,138,140,141,145,162,173,175,177,184,204,205] randomized participants. The remaining 84 studies used a cross-sectional design.

Cross-sectional studies were conducted in controlled settings, including research or academic settings [81,86,87,89-91,93,97-99,101,103,105,106,108,109,111,113,114,117,120,123,125,126,131,133-137,139,142,143,148,152,153,155-159,161,163,166-169,171,172,174,176,178-181,185-196,198,202], outpatient clinical

settings [82,85,88,92,95,102,119,164], hospital clinical settings [146,147,150,182,206], and residential care facilities for older adults [80,94]. In the longitudinal studies, participants used the applications as part of their daily routines, although in some cases data collection was performed in controlled settings. The duration of the longitudinal studies varied from 1 week to 24 weeks: (1) 1 week, 6 [104,110,122,124,129,201] studies; (2) 2 weeks, 8 [100,121,132,151,173,183,197,204] studies; (3) 3 weeks, 1 [140] study; (4) 4 weeks, 7 [96,107,115,128,154,203,205] studies; (5) 5 weeks, 1 [199] study; (6) 6 weeks, 4 [112,160,162,165] studies; (7) 8 weeks, 2 [116,145] studies; (8) 9 weeks, 1 [170] study; (9) 12 weeks, 12 [83,84,118,127,133,138,141,144,149,175,184,200] studies; and (10) 24 weeks, 1 [177] study.

Number and Characteristics of the Participants

Studies included participants from specific age groups (ie, adults or older adults) and health status categories (ie, healthy participants or patients): (1) healthy adults (ie, participants between 18 and 65 years of age) [85-87,92,98,99,101,102,104,105,108-113,117,120-122,124-126,131,132,134,137,142,143,149,152,153,155,158,160,161,163,164,166-168,171-173,177-181,185,186,189,191,192,195,196,199,200,202,204,205]; (2) healthy older adults (ie, participants older than 65 years) [83,84,88,90,93,94,100,115,116,118,123,129,139,141,146,190]; (3) adult patients [81,82,89,95,97,114,119,127,128,130,133,135,136,138,140,144,145,147,150,151,154,156,169,170,176,183,184,187,194,197,198,201,203]; (4) older adult patients [162,165,175,182,206]; (5) healthy adults and adult patients [103,106,148,157,159,174,193]; (6) healthy adults and older adults [91,188]; and (7) healthy adults and older adult patients [80,96,107].

In terms of patients, the following conditions were identified: mental health disorders [81,82,85,92,95,97,103,104,119,121,123,124,132,147,149,150,153,155,165-170,172,173,179,181,184,188,191,204-206], cancer [138,145,182,183,197], heart disease [154], chronic kidney disease [144], chronic pain [105,162,198], COVID-19 [95], delirium [146], dementia [107], diabetes [89,96,111,115,140,194], heart failure [130,187], hydrocephalus [177], hypertension [133,203], intestinal ostomy [148], mild cognitive impairment [141], Parkinson disease [157,175], psoriasis [127], sickle cell disease [114], stroke [128,159,174], traumatic brain injury [151], traumatic injuries [106,180], vestibular disorders [189], and other unspecified chronic diseases [87,156].

Concerning the number of participants, the studies included a total of 7040 participants, of whom 875 were older adults. However, the sample sizes varied considerably across studies. Specifically, 56 studies [82,83,87,91,92,94,95,98-102,107,109,112-116,121,122,124,126,130,134-142,144,145,149,152,153,158,160,161,167,169,170,173,177,180,181,191,194,195,197,200,203,204,206] included more than 30 participants, and 18 of these studies [82,95,100,112,126,139,142,149,158,160,161,170,173,180,181,194,200,204] enrolled more than 100 participants. The remaining 71 studies included 30 or fewer participants.

Validated Subjective Assessment Instruments

Three categories of validated subjective instruments were identified:

- Usability assessment instruments: (1) PSSUQ [30], (2) SUS [31], (3) Chatbot Usability Questionnaire (CUQ) [207], (4)

Computer System Usability Questionnaire (CSUQ) [208], (5) mHealth App Usability Questionnaire (MAUQ) [209], (6) Usability Metric for User Experience—Lite Version (UMUX-Lite) [210], and (7) Usefulness, Satisfaction, and Ease of Use Questionnaire (USE) [211].

- User experience assessment instruments: (1) User Experience Questionnaire (UEQ) [212] and (2) User Experience Questionnaire—Short Version (UEQ-S) [213].
- Acceptability assessment instrument: Acceptability E-Scale (AES) [214], which includes a subscale for assessing usability.

Nine studies [86,105,117,119,145,146,174,175,191] applied more than 1 assessment instrument: (1) UEQ and SUS [86,105,117,145,174,175,191]; (2) SUS and USE [119]; and (3) CUQ and SUS [146]. See Table 2 for the validated subjective assessment instruments.

Table 2. Validated subjective assessment instruments.

Instruments	Studies	Total, n
System Usability Scale	[80,81,83-90,92,96-98,100,105-109,111-123,126-130,132,133,135,136,140-142,144-148,150,151,153-157,160-168,170-172,174-177,179,180,182,184,186,190-202]	91
User Experience Questionnaire	[86,101-103,105,110,117,143,145,174,175,178,181,189,191]	15
Chatbot Usability Questionnaire	[125,134,138,146,203,204]	6
Computer System Usability Questionnaire	[91,99,131,139,205]	5
Acceptability E-Scale	[82,95,149,169,206]	5
mHealth App Usability Questionnaire	[159,173,185,187]	4
Usefulness, Satisfaction, and Ease of Use Questionnaire	[94,104,119,183]	4
User Experience Questionnaire—Short Version	[93,152,158]	3
Usability Metric for User Experience—Lite Version	[124,188]	2
Post-Study System Usability Questionnaire	[137]	1

Results of Usability and User Experience Assessments

The usability and user experience scores were generally high. The PSSUQ was applied in only 1 [137] study, which reported

a total score of 6.0 (above average, ≥ 6). Additionally, Table 3 summarizes the ranges of usability and user experience scores for the remaining instruments, based on the studies that reported these scores.

Table 3. Range of the reported usability and user experience scores according to subjective assessment instruments and the standardized scores for each instrument indicating good usability.

Instruments	Score range	Standardized scores indicating good usability or user experience
System Usability Scale	51.4-93.8	≥ 70.0
Chatbot Usability Questionnaire	40.6-85.9	≥ 70.0
Computer System Usability Questionnaire	4.6-5.8	≥ 4.5
Acceptability E-Scale	22.0-25.4	≥ 24.0
mHealth App Usability Questionnaire	4.8-6.7	≥ 5.0
Usefulness, Satisfaction, and Ease of Use Questionnaire	4.5-5.7	≥ 5.0
Usability Metric for User Experience—Lite Version	82.4-86.1	≥ 70.0
User Experience Questionnaire and User Experience Questionnaire—Short Version	0.4-1.7	≥ 1.0

Comparative Analysis of the SUS Results

The following comparative analysis should be interpreted as exploratory because it synthesizes SUS scores reported by independent studies rather than direct comparisons conducted under common experimental conditions. Consequently, the observed differences among HCA categories may be influenced not only by the interaction modality itself but also by variations in health care application domains, study designs, evaluation contexts, participant populations, and HCA implementation and use characteristics. In addition, although the SUS is a well-established measure of perceived usability, it may not adequately capture important aspects of HCAs, such as conversational quality, which is influenced by factors including

turn-taking, the ability to sustain coherent and continuous dialogue, and the naturalness of interactions.

Considering the studies that applied the SUS and reported usability scores, a significant difference among groups was found when comparing usability scores across the 3 HCA categories ($P=.001$; $\eta^2H=0.11$; Figure 5, Table 4). Pairwise comparisons showed significant differences between text-based and embodied HCAs ($P=.02$) and between text- and voice-based HCAs ($P=.006$). No significant difference was found between voice-based and embodied HCAs ($P>.99$). When comparing the number of participants across the 3 HCA categories, no significant difference was found ($P=.18$; $\eta^2H=0.016$).

Figure 5. Median and IQR for System Usability Scale (SUS) scores (group 1, text-based HCAs; group 2, voice-based HCAs; and group 3, embodied HCAs). HCA: health care conversational agent.

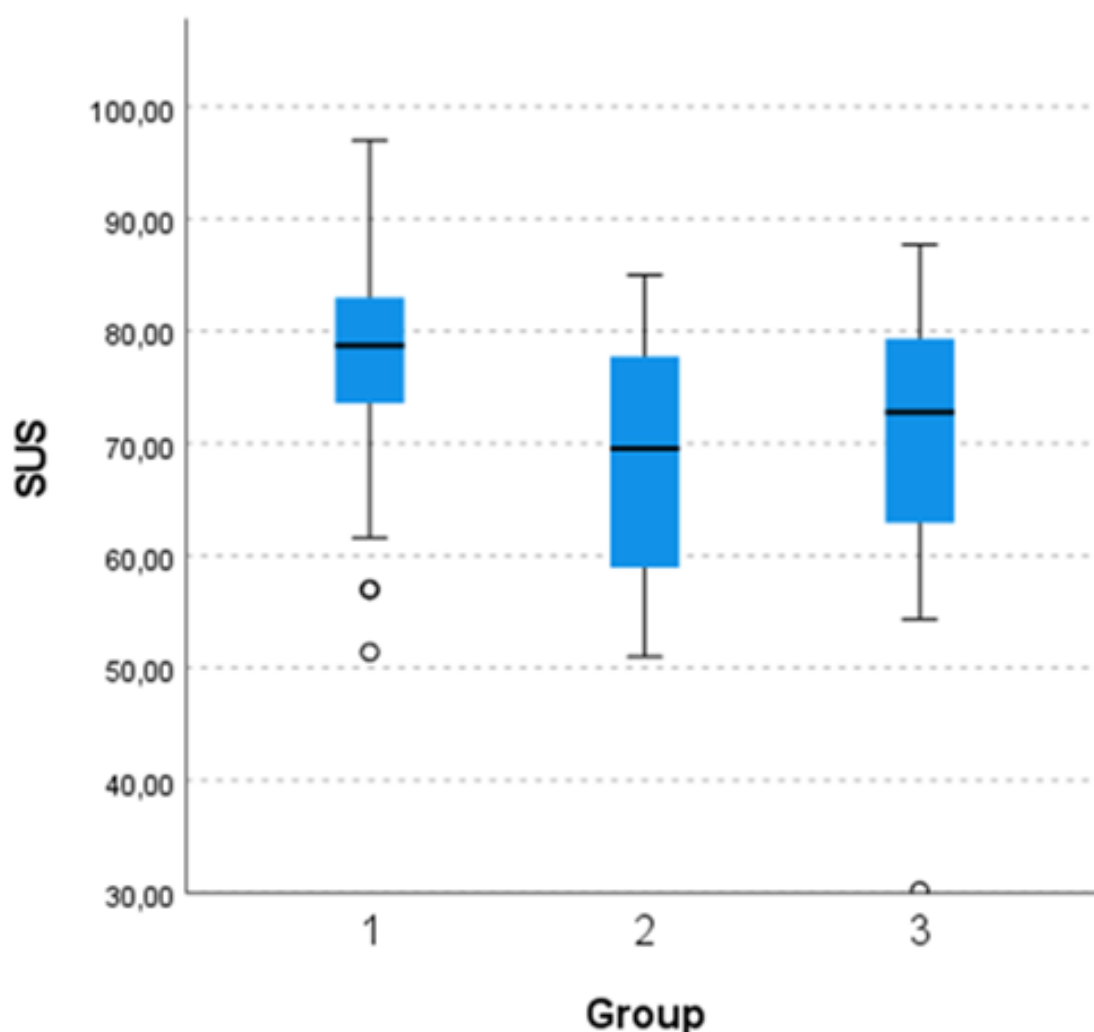


Table 4. Comparative analysis of SUS^a scores across the 3 categories of HCAs^b.

Category	Number of studies, n	Number of participants, median (IQR ^c)	SUS scores, median (IQR ^c)	SUS range
Text-based HCAs	58	30.00 (11.50-52.75)	78.70 (73.43-83.06)	51.40-97.00
Voice-based HCAs	18	22.00 (13.0-34.25)	69.55 (58.90-78.47)	51.00-85.00
Embodied HCAs	26	18.00 (11.00-26.25)	72.79 (62.29-79.38)	30.13-87.70

^aSUS: System Usability Scale.^bHCA: health care conversational agent.^c25th-75th percentiles.

A significant difference in SUS scores was found when comparing studies involving adults with those involving older adults in the overall sample of studies ($P<.001$; $r=0.39$; Figure 6, Table 5). Similarly, a significant difference in SUS scores between studies involving older adults and those involving

adults was found for text-based HCAs ($P=.009$; $r=0.33$), indicating lower SUS scores in studies involving older adults. No significant between-group difference was found for voice-based HCAs ($P=.10$; $r=0.38$) or embodied HCAs ($P=.20$; $r=0.25$).

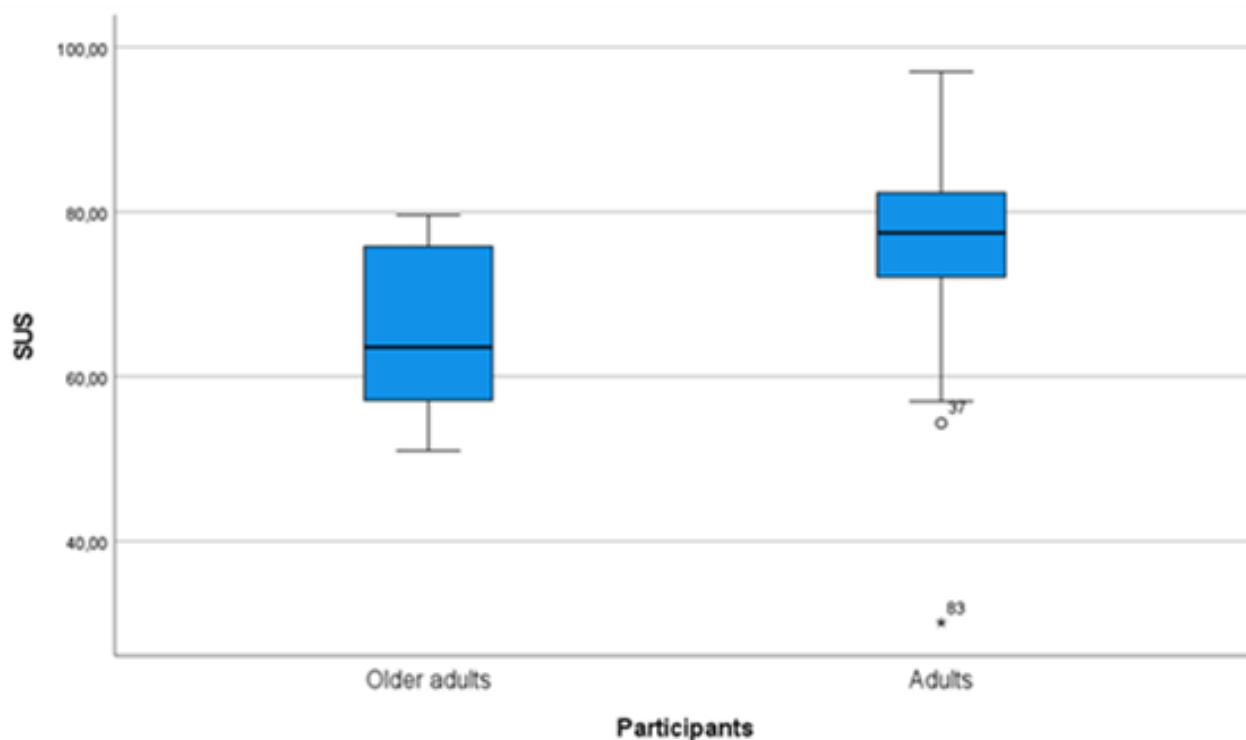
Figure 6. Median and IQR for System Usability Scale (SUS) scores in studies using adults and older adults.

Table 5. Comparative analysis of SUS^a scores across age range (adults vs older adults).

Category and age group	Number of studies, n	SUS, median (IQR ^b)
All HCAs^c		
Adults	82	77.45 (72.06-82.43)
Older adults	19	63.56 (57.07-76.19)
Text-based HCAs		
Adults	51	79.50 (75.00-83.80)
Older adults	6	64.05 (55.58-78.28)
Voice-based HCAs		
Adults	13	72.10 (59.97-81.17)
Older adults	5	59.69 (52.17-72.00)
Embodied HCAs		
Adults	18	74.41 (67.36-81.08)
Older adults	8	63.75 (57.65-77.78)

^aSUS: System Usability Scale.

^b25th-75th percentiles.

^cHCA: health care conversational agent.

Discussion

This review provides an overview of the current evidence on the usability and user experience of HCAs assessed using validated subjective instruments. As HCAs become increasingly widespread, a growing number of individuals interact with them daily, and it is anticipated that these agents will play an important role in future interactive systems. Therefore, it is crucial to investigate the assessment of their usability and user experience, particularly in the context of emerging human-machine interaction, because well-designed assessments contribute to the overall improvement of interactions. Although well-designed interactions do not necessarily mean that digital health applications are effective, their efficacy and effectiveness may be limited by usability and user experience constraints [215].

Patient care emerged as the predominant application domain for HCAs, particularly in mental health. Other reviews [8,216] have already identified the importance of patient care, while the efficacy of HCAs in supporting mental health disorders has been the subject of several systematic reviews (eg, [217,218]) and a meta-analysis (eg, [219]). Among the other application domains addressed, HCAs may support health education and prevention by optimizing the delivery of understandable, personalized health education at scale and by providing interactive and engaging information across platforms (eg, mobile, web, or smart devices) [220,221]. Similarly, health data collection has already been identified as a core use case of HCAs [216,222], namely, in terms of automating diagnostic interviews and longitudinal monitoring of patient-reported outcomes (eg, anxiety or depression scales) to track symptoms and adherence to treatments over time. Furthermore, the use of HCAs designed to promote the independence and autonomy of older adults by assisting them in managing activities of daily living and self-management of health conditions is in line with current

interest in developing digital applications to support the care and well-being of older adults [223,224]. Another relevant point is that a high proportion of the studies included older adults, who accounted for 875 of the 7040 (12.43%) participants.

Although text-based HCAs were the most frequently evaluated, all 3 identified categories were applied across different health care domains. This suggests that developers select interaction modalities according to technological feasibility and user requirements rather than the clinical purpose itself.

The increasing integration of virtual reality with embodied HCAs reflects growing interest in immersive human-computer interaction. However, the relatively limited number of studies indicates that these technologies remain at an early stage of adoption in health care, likely due to higher development costs, hardware requirements, and implementation complexity.

Although AI was increasingly incorporated into HCAs, most of these implementations appear to remain at a relatively immature stage, which is consistent with other studies demonstrating the difficulties of translating AI implementations into clinical workflows [225-227].

SUS clearly dominated the assessment of HCAs, despite having been developed for generic interactive systems. SUS is widely used for usability evaluation of different types of digital health applications [215,228-230] because it is an easy and quick tool with a scoring system that has been translated into multiple languages [228]. However, it was not originally developed to assess complex multimodal interaction mechanisms and does not capture the holistic user experience because it focuses only on usability aspects such as ease of use, satisfaction, and learnability.

Although several alternative instruments were identified, CUQ was the only instrument specifically appropriate for assessing HCAs, even though it does not fully capture the multimodal

and AI-driven characteristics of modern HCAs. The absence of recently developed instruments [46], such as the Conversational Agents Scale [231], the Conversational Agent's Usage Scale [232], the User Experience Evaluation of Conversational Agents [233], or the Conversational Agent Scale for User Experience [234], further highlights the need for greater methodological standardization in future evaluations. Moreover, a relevant number of studies were excluded from this review because usability and user experience assessments were based on questionnaires consisting of Likert scale questions developed by the researchers. However, these questionnaires were not validated instruments because they had not undergone psychometric analysis, which is frequently a lengthy process that can be particularly challenging for researchers who lack experience in developing and validating subjective instruments.

Usability and user experience assessments were conducted in both controlled (ie, research or academic settings, outpatient clinical settings, hospital clinical settings, and institutional residences for older adults) and uncontrolled (eg, participants' homes) environments. Controlled experiments generally focus on intensive sessions to capture immediate usability or user experience feedback. These experiments enable consistent data collection, but they may constrain the observation of more natural interaction patterns. By contrast, longitudinal studies conducted in uncontrolled settings allow participants to engage with HCAs more organically, integrating their use into daily routines. Therefore, uncontrolled environments provide authentic user experiences and valuable insights, as well as functional challenges, including connectivity issues, reflecting authentic usage scenarios. Moreover, many longitudinal studies strategically combined uncontrolled environments with data collection in controlled environments.

The participants comprised both patients and healthy adults or older adults. However, the reviewed studies often did not adequately justify their choice of sample sizes. Small samples can undermine the reliability of results by increasing the risk of sampling bias and reducing statistical power. Consequently, the findings may not capture the full variability of the population, thereby restricting their generalizability.

The quality assessment of the components related to the application of subjective measurement instruments highlighted methodological limitations, including insufficient evaluation of prolonged real-world use and inadequate justification of sample sizes, the 2 CAUSS items with lower percentages. These shortcomings reduce confidence in the ecological validity and generalizability of the reported usability findings.

Looking at the assessment results, the usability and user experience of HCAs tend to be above average; however, the results indicate differences among the various HCA categories. The comparison of usability scores across the 3 HCA modalities revealed statistically significant differences, with pairwise comparisons indicating that text-based HCAs achieved significantly higher usability scores than both voice-based and embodied HCAs, whereas no significant differences were observed between voice-based and embodied HCAs. The estimated eta-squared ($\eta^2 H=0.11$) indicates a moderate effect size, suggesting a moderate association between interaction

modality and users' perceptions of usability. No significant differences were found in the number of participants across the 3 HCA categories, indicating that the studies included broadly comparable sample sizes. Consequently, systematic differences in study size are unlikely to fully explain the observed differences in usability.

Several factors may contribute to the higher usability observed for text-based HCAs. Although the reviewed studies did not directly investigate the underlying mechanisms, one possible explanation is users' greater familiarity with text-based communication, particularly through messaging applications. Moreover, text interfaces provide a persistent conversation history that enables users to review previous exchanges, potentially increasing their sense of control and reducing memory demands during interaction. In addition, text-based interactions allow users to engage with HCAs at their own pace, giving them more time to read, reflect on, and formulate responses. By contrast, voice-based HCAs may be affected by speech recognition inaccuracies, environmental noise, and privacy concerns, whereas embodied HCAs introduce additional visual and social interaction elements that may increase interaction complexity without necessarily improving perceived usability.

The results also suggest that older adults tend to perceive HCAs as less usable than adults. This may be associated with older adults generally having lower digital literacy, as well as age-related accessibility needs. For example, age-related sensory, cognitive, and motor changes, together with the higher prevalence of multimorbidity among older adults, may affect their ability to interact effectively with HCAs. Nevertheless, the available evidence from the included studies does not allow this explanation to be confirmed.

Interestingly, the difference between adults and older adults was statistically significant only for text-based HCAs, despite this modality achieving the highest usability scores overall. A potential explanation is that text-based HCAs are currently more widespread than voice-based or embodied HCAs, which may lead to greater differences in familiarity and prior experience between age groups.

Although no statistically significant age-related differences were observed for voice-based or embodied HCAs, the estimated effect sizes were of comparable magnitude across modalities. Therefore, the absence of statistical significance may reflect the smaller number of available studies and limited statistical power rather than provide evidence that age has no influence on perceived usability in these interaction modalities.

These results should be interpreted with caution. The included studies exhibited substantial heterogeneity in terms of application domains (patient care, health education and prevention, health data collection, and support for daily activities), study designs (eg, cross-sectional or longitudinal), evaluation contexts (controlled versus real-world settings), target populations, underlying technologies (eg, AI, large language models, or virtual reality), task complexity, and duration of HCA use (single-session vs prolonged use). Consequently, the observed differences in SUS scores may have been influenced by factors beyond the interaction modality itself. Furthermore,

the SUS assesses perceived usability rather than objective performance or long-term acceptance, and therefore high SUS scores do not necessarily translate into sustained engagement or continued use. Future studies should investigate which specific design characteristics contribute most to usability across different HCA modalities while accounting for user characteristics and contextual factors.

Despite the use of extensive search strategies to locate relevant studies, some articles may have been overlooked due to variations in terminology. Furthermore, the exclusion of non-English publications and gray literature may have limited the overall representativeness of the findings and introduced language and publication bias, as studies published in other languages or unpublished studies with null or negative findings may have been underrepresented.

The wide variety of assessment instruments used meant that it was generally not possible to identify enough studies to enable a comparative analysis of usability or user experience outcomes by instrument across different HCA categories. The only exception was SUS; however, as it was developed in the 1990s, it may not adequately capture relevant aspects of the usability of modern HCAs.

Finally, this review focused solely on the usability and user experience of HCAs and did not address their efficacy or effectiveness.

In conclusion, this review included 127 studies applying validated subjective instruments to assess the usability and user experience of 3 HCA categories: text-based, voice-based, and embodied HCAs. In terms of health care applications, HCAs aimed to support patient care, health education and prevention, health data collection, and daily activities of older adults.

The usability and user experience of HCAs were assessed in cross-sectional and longitudinal studies using the following

instruments: SUS, UEQ, UEQ-S, CUQ, AES, MAUQ, UMUX-Lite, USE, CSUQ, and PSSUQ.

This review highlights that, except for CUQ, the identified subjective assessment instruments were not specifically developed to evaluate the usability and user experience of HCAs. Therefore, the findings suggest that such evaluations may be limited by the inadequacy of these instruments.

Considering the comparative analysis of SUS scores, text-based HCAs tend to show higher usability scores than voice-based and embodied HCAs, and older adults tend to perceive HCAs as less usable than adults. However, these findings represent associations observed across the included studies rather than direct evidence that one HCA category is inherently more usable than another, because the studies differ substantially in their health care applications, design, evaluation contexts, participant populations, and implementation and use characteristics. Moreover, a relevant number of studies used small sample sizes, which reduce statistical power, may fail to capture population variability, limit the generalizability of the results, and increase the risk of misleading conclusions. Most HCAs were also assessed in controlled environments at a single point in time, without considering the impact of continuous and long-term use in real-world settings.

Considering the substantial differences in how evaluations were designed and conducted across studies, especially regarding data collection methods, participant selection, testing conditions, and reporting of results, there is a need for standardized protocols and established guidelines to measure the usability and user experience of HCAs. Such standardization would improve methodological consistency and enhance the comparability of findings across studies, enabling researchers and practitioners to benchmark usability and user experience.

Funding

This research received no external funding.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.

[\[PDF File \(Adobe PDF File\), 502 KB-Multimedia Appendix 1\]](#)

References

1. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun ACM*. Jan 1966;9(1):36-45. [doi: [10.1145/365153.365168](https://doi.org/10.1145/365153.365168)]
2. Caldarini G, Jaf S, McGarry K. A literature survey of recent advances in chatbots. *Information*. Jan 15, 2022;13(1):41. [doi: [10.3390/info13010041](https://doi.org/10.3390/info13010041)]
3. Leschanowsky A, Rech S, Popp B, Bäckström T. Evaluating privacy, security, and trust perceptions in conversational AI: a systematic review. *Computers in Human Behavior*. Oct 2024;159:108344. [doi: [10.1016/j.chb.2024.108344](https://doi.org/10.1016/j.chb.2024.108344)]
4. McTear M. The rise of the conversational interface: a new kid on the block? Cham, Switzerland. Springer; 2017. Presented at: International Workshop on the Future and Emerging Trends in Language Technology; 30 November to 2 December 2016:38-49; Seville, Spain. [doi: [10.1007/978-3-319-69365-1_3](https://doi.org/10.1007/978-3-319-69365-1_3)]

5. McTear M, Callejas Z, Griol D. Future directions. In: McTear M, Callejas Z, Griol D, editors. *The Conversational Interface*. Cham, Switzerland. Springer; 2016:403-418.
6. Abbasian M, Azimi I, Rahmani A, Jain R. Conversational health agents: a personalized large language model-powered agent framework. *JAMIA Open*. Aug 2025;8(4):ooaf067. [FREE Full text] [doi: [10.1093/jamiaopen/ooaf067](https://doi.org/10.1093/jamiaopen/ooaf067)] [Medline: [40620478](https://pubmed.ncbi.nlm.nih.gov/40620478/)]
7. Wang J, Ma B, Nan Y, He Y, Sun H, Liu C, et al. Improving matching models with contextual attention for multi-turn response selection in retrieval-based chatbots. *IEEE Trans Netw Sci Eng*. May 2025;12(3):1497-1509. [doi: [10.1109/tnse.2025.3532663](https://doi.org/10.1109/tnse.2025.3532663)]
8. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc*. Sep 01, 2018;25(9):1248-1258. [FREE Full text] [doi: [10.1093/jamia/ocy072](https://doi.org/10.1093/jamia/ocy072)] [Medline: [30010941](https://pubmed.ncbi.nlm.nih.gov/30010941/)]
9. Liu B. *Sentiment Analysis: Mining opinions, Sentiments, and Emotions*. Cambridge, UK. Cambridge University Press; 2020.
10. Shah M, Kavathiya H. Exploring conversational AI. In: Al-Marzouqi A, Salloum S, Al-Saidat M, Aburayya A, Gupta B, editors. *Artificial Intelligence in Education: The Power and Dangers of ChatGPT in the Classroom*. Cham, Switzerland. Springer; 2024:511-526.
11. Cassell J, Bickmore T, Billinghamurst M, Campbell L, Chang K, Vilhjálmsdóttir H, et al. Embodiment in conversational interfaces: Rea. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY. Association for Computing Machinery; 1999. Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 15-20, 1999:520-527; Pittsburgh, PA. [doi: [10.1145/302979.303150](https://doi.org/10.1145/302979.303150)]
12. Swartout W, Gratch J, Hill JR, Hovy E, Marsella S, Rickel J, et al. Toward virtual humans. *The USC Institute for Creative Technologies*. 2006. URL: <https://people.ict.usc.edu/~gratch/AI-mag06.pdf> [accessed 2026-07-27]
13. Wolfert P, Robinson N, Belpaeme T. A review of evaluation practices of gesture generation in embodied conversational agents. *IEEE Trans Human-Mach Syst*. Jun 2022;52(3):379-389. [doi: [10.1109/THMS.2022.3149173](https://doi.org/10.1109/THMS.2022.3149173)]
14. Dai X, Liu Z, Liu T, Zuo G, Xu J, Shi C, et al. Modelling conversational agent with empathy mechanism. *Cognitive Systems Research*. Mar 2024;84:101206. [doi: [10.1016/j.cogsys.2023.101206](https://doi.org/10.1016/j.cogsys.2023.101206)]
15. Nguyen TTN, Dam QT, Tran DT, Lee J. A survey on generative nonverbal facial behavior for highly realistic embodied agents. *Intel Serv Robotics*. Jan 28, 2026;19(2):27. [doi: [10.1007/s11370-025-00674-2](https://doi.org/10.1007/s11370-025-00674-2)]
16. Sajjadi P, Hoffmann L, Cimiano P, Kopp S. A personality-based emotional model for embodied conversational agents: effects on perceived social presence and game experience of users. *Entertainment Computing*. Dec 2019;32:100313. [doi: [10.1016/j.entcom.2019.100313](https://doi.org/10.1016/j.entcom.2019.100313)]
17. Kroczeck LO, May A, Hettenkofer S, Ruider A, Ludwig B, Mühlberger A. The influence of persona and conversational task on social interactions with a LLM-controlled embodied conversational agent. *Computers in Human Behavior*. Nov 2025;172:108759. [doi: [10.1016/j.chb.2025.108759](https://doi.org/10.1016/j.chb.2025.108759)]
18. Yang F, Acevedo P, Guo S, Choi M, Mousas C. Embodied conversational agents in extended reality: a systematic review. *IEEE Access*. 2025;13:79805-79824. [doi: [10.1109/access.2025.3566698](https://doi.org/10.1109/access.2025.3566698)]
19. Huynh AL, Roy TJ, Jackson KN, Lee AG, Liaw W, Hossain MM. Applications of artificial intelligence-based conversational agents in healthcare: a systematic umbrella review. *Int J Med Inform*. Mar 01, 2026;207:106204. [FREE Full text] [doi: [10.1016/j.ijmedinf.2025.106204](https://doi.org/10.1016/j.ijmedinf.2025.106204)] [Medline: [41337874](https://pubmed.ncbi.nlm.nih.gov/41337874/)]
20. Bălan C. Chatbots and voice assistants: digital transformers of the company–customer interface—a systematic review of the business research literature. *JTAER*. May 18, 2023;18(2):995-1019. [doi: [10.3390/jtaer18020051](https://doi.org/10.3390/jtaer18020051)]
21. Xu Z, Liu F, Xia G, Wang S, Duan Y, Yu L, et al. Immersive HCI for intangible cultural heritage in tourism contexts: a narrative review of design and evaluation. *Sustainability*. Dec 23, 2025;18(1):153. [doi: [10.3390/su18010153](https://doi.org/10.3390/su18010153)]
22. Yusuf H, Money A, Daylamani-Zad D. Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art. *Education Tech Research Dev*. Jan 15, 2025;73(2):815-874. [doi: [10.1007/s11423-025-10447-4](https://doi.org/10.1007/s11423-025-10447-4)]
23. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. Sep 29, 2008;337(sep29 1):a1655-a1655. [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
24. Fu L, Burns R, Xie Y, Shen J, Zhe S, Estabrooks P, et al. The development and use of AI chatbots for health behavior change: scoping review. *J Med Internet Res*. Jan 28, 2026;28:e79677. [FREE Full text] [doi: [10.2196/79677](https://doi.org/10.2196/79677)] [Medline: [41604667](https://pubmed.ncbi.nlm.nih.gov/41604667/)]
25. Nielsen J. The usability engineering life cycle. *Computer*. Mar 1992;25(3):12-22. [doi: [10.1109/2.121503](https://doi.org/10.1109/2.121503)]
26. Hassan H, Galal-Edeen G. From usability to user experience. In: *Proceedings of the International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*. New York, NY. IEEE; 2017. Presented at: International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS); November 24-26, 2017:216-222; Okinawa, Japan.
27. International Organization for Standardization (ISO). ISO 9241-210-2019: Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems. ISO. Geneva, Switzerland. International Organization for Standardization; 2019. URL: <https://www.iso.org/standard/77520.html> [accessed 2026-07-22]

28. Ntoa S. Usability and user experience evaluation in intelligent environments: a review and reappraisal. *International Journal of Human-Computer Interaction*. Sep 12, 2024;41(5):1-30. [doi: [10.1080/10447318.2024.2394724](https://doi.org/10.1080/10447318.2024.2394724)]
29. Martins A, Queirós A, Silva A, Rocha N. Usability evaluation methods: a systematic review. In: Saeed S, Bajwa IS, Mahmood Z, editors. *Human Factors in Software Development and Design*. New York, NY: IGI Global; 2015:250-273.
30. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. *Proceedings of the Human Factors Society Annual Meeting*. 1992;36(16):1259-1260. [doi: [10.1177/154193129203601617](https://doi.org/10.1177/154193129203601617)]
31. Brooke J. SUS-A quick and dirty usability scale. In: Jordan P, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation in Industry*. Boca Raton, FL: CRC Press; 1996:4-7.
32. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*. Jul 30, 2008;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
33. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *Journal of Usability Studies*. 2009;4(3):114-123. [FREE Full text]
34. Branda F, Stella M, Ceccarelli C, Cabitza F, Ceccarelli G, Maruotti A, et al. The role of AI-based chatbots in public health emergencies: a narrative review. *Future Internet*. Mar 26, 2025;17(4):145. [doi: [10.3390/fi17040145](https://doi.org/10.3390/fi17040145)]
35. Orrù L, Mannarini S. The role of artificial intelligence in clinical psychology: how AI and NLP systems are reshaping psychological interventions. A systematic review. *Clin Psychol Psychother*. Feb 25, 2026;33(2):e70242. [doi: [10.1002/cpp.70242](https://doi.org/10.1002/cpp.70242)] [Medline: [41738110](https://pubmed.ncbi.nlm.nih.gov/41738110/)]
36. Mahajan P, Kadam S, Kachwala H, Tapikar N, Namde V, Poonawala E. Reimagining mental health support: the role of chatbots in bridging gaps and raising ethical questions. *Couns and Psychother Res*. Feb 15, 2026;26(1):e70095. [doi: [10.1002/capr.70095](https://doi.org/10.1002/capr.70095)]
37. Cesar Abrantes P, Netto AV, Kazuo Takahata A. Healthbots for conducting clinical screening and remote monitoring with patient mood assessment: a scoping review. *Int J Med Inform*. Mar 01, 2026;207:106186. [doi: [10.1016/j.ijmedinf.2025.106186](https://doi.org/10.1016/j.ijmedinf.2025.106186)] [Medline: [41308276](https://pubmed.ncbi.nlm.nih.gov/41308276/)]
38. Fiske S, Cody JL, Shen M, Choi J. AI conversational agents in older adults with chronic disease: a scoping review. *Geriatr Nurs*. Mar 2026;68:103856. [doi: [10.1016/j.gerinurse.2026.103856](https://doi.org/10.1016/j.gerinurse.2026.103856)] [Medline: [41576576](https://pubmed.ncbi.nlm.nih.gov/41576576/)]
39. Dong P, Zhang S, Zheng X, Yuan W, Wu L, Chen Y. Artificial intelligence robots for mental health applications: a scoping review. *Psychiatry Res*. Jun 2026;360:117059. [doi: [10.1016/j.psychres.2026.117059](https://doi.org/10.1016/j.psychres.2026.117059)] [Medline: [41819775](https://pubmed.ncbi.nlm.nih.gov/41819775/)]
40. Jalali S, You Q, Xu V, Chen L, Zini J, Zamorano T, et al. The use of artificial intelligence for personalized treatment in psychiatry. *Curr Psychiatry Rep*. Dec 29, 2025;28(1):7. [doi: [10.1007/s11920-025-01656-y](https://doi.org/10.1007/s11920-025-01656-y)] [Medline: [41457119](https://pubmed.ncbi.nlm.nih.gov/41457119/)]
41. Tahir H, Özdemir MK, Alhajj R. The role of LLM-powered chatbots in assisting elderly people: systematic review. *Netw Model Anal Health Inform Bioinforma*. Dec 27, 2025;15(1):26. [doi: [10.1007/s13721-025-00698-9](https://doi.org/10.1007/s13721-025-00698-9)]
42. Curtis RG, Bartel B, Ferguson T, Blake HT, Northcott C, Virgara R, et al. Improving user experience of virtual health assistants: scoping review. *J Med Internet Res*. Dec 21, 2021;23(12):e31737. [FREE Full text] [doi: [10.2196/31737](https://doi.org/10.2196/31737)] [Medline: [34931997](https://pubmed.ncbi.nlm.nih.gov/34931997/)]
43. Bastardo R, Pavão J, Rocha N. User-centred usability evaluation of embodied communication agents to support older adults: a scoping review. Cham, Switzerland: Springer; 2022. Presented at: International Conference on Information Technology Systems; February 9-11, 2022:509-518; San Carlos, Costa Rica. [doi: [10.1007/978-3-030-96293-7_42](https://doi.org/10.1007/978-3-030-96293-7_42)]
44. Li K, Wu S, Zhang Y, Zhu B, Qi Z, Hou S, et al. The usability and experience of artificial intelligence-based conversational agents in health education for cancer patients: a scoping review. *J Clin Nurs*. Jul 14, 2025;1-11. (forthcoming). [doi: [10.1111/jocn.70020](https://doi.org/10.1111/jocn.70020)] [Medline: [40654175](https://pubmed.ncbi.nlm.nih.gov/40654175/)]
45. Faruk LID, Babakerkhell MD, Mongkolnam P, Chongsuphajaisiddhi V, Funilkul S, Pal D. A review of subjective scales measuring the user experience of voice assistants. *IEEE Access*. 2024;12:14893-14917. [doi: [10.1109/access.2024.3358423](https://doi.org/10.1109/access.2024.3358423)]
46. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg*. 2010;8(5):336-341. [FREE Full text] [doi: [10.1016/j.ijsu.2010.02.007](https://doi.org/10.1016/j.ijsu.2010.02.007)] [Medline: [20171303](https://pubmed.ncbi.nlm.nih.gov/20171303/)]
47. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. *JBIM Evid Synth*. Oct 2020;18(10):2119-2126. [doi: [10.1112/JBIES-20-00167](https://doi.org/10.1112/JBIES-20-00167)] [Medline: [33038124](https://pubmed.ncbi.nlm.nih.gov/33038124/)]
48. Silva AG, Simões P, Santos R, Queirós A, Rocha NP, Rodrigues M. A scale to assess the methodological quality of studies assessing usability of electronic health products and services: Delphi study followed by validity and reliability testing. *J Med Internet Res*. Nov 15, 2019;21(11):e14829. [FREE Full text] [doi: [10.2196/14829](https://doi.org/10.2196/14829)] [Medline: [31730036](https://pubmed.ncbi.nlm.nih.gov/31730036/)]
49. Fiel PF. Effect sizes for nonparametric tests. *Biochemia Medica*. 2026;36(1):5-16. [doi: [10.11613/bm.2026.010101](https://doi.org/10.11613/bm.2026.010101)]
50. Mendu S, Boukhechba M, Gordon J, Datta D, Molina E, Arroyo G. Design of a culturally-informed virtual human for educating hispanic women about cervical cancer. In: *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*. New York, NY: Association for Computing Machinery; 2018. Presented at: 12th EAI International Conference on Pervasive Computing Technologies for Healthcare; 21-24 May 2018:330-336; New York, NY. [doi: [10.1145/3240925.3240968](https://doi.org/10.1145/3240925.3240968)]
51. Sun O, Chen J, Magrabi F. Using voice-activated conversational interfaces for reporting patient safety incidents: a technical feasibility and pilot usability study. *Stud Health Technol Inform*. 2018;252:139-144. [doi: [10.3233/978-1-61499-890-7-139](https://doi.org/10.3233/978-1-61499-890-7-139)]

52. Booth A, van der Krogt M, Buizer A, Steenbrink F, Harlaar J. The validity and usability of an eight marker model for avatar-based biofeedback gait training. *Clin Biomech (Bristol)*. Dec 2019;70:146-152. [doi: [10.1016/j.clinbiomech.2019.08.013](https://doi.org/10.1016/j.clinbiomech.2019.08.013)] [Medline: [31499394](https://pubmed.ncbi.nlm.nih.gov/31499394/)]
53. Huff JE, Mack N, Cummings R, Womack K, Gosha K, Gilbert J. Evaluating the usability of pervasive conversational user interfaces for virtual mentoring. Cham, Switzerland. Springer; 2019. Presented at: International Conference on Human-Computer Interaction; July 26-31, 2019; Orlando, FL. [doi: [10.1007/978-3-030-21817-1_7](https://doi.org/10.1007/978-3-030-21817-1_7)]
54. Ter Stal S, Broekhuis M, van Velsen L, Hermens H, Tabak M. Embodied conversational agent appearance for health assessment of older adults: explorative study. *JMIR Hum Factors*. Sep 04, 2020;7(3):e19987. [FREE Full text] [doi: [10.2196/19987](https://doi.org/10.2196/19987)] [Medline: [32886068](https://pubmed.ncbi.nlm.nih.gov/32886068/)]
55. Roman MK, Bellei EA, Biduski D, Pasqualotti A, De Araujo CDSR, De Marchi ACB. "Hey assistant, how can I become a donor?" The case of a conversational agent designed to engage people in blood donation. *J Biomed Inform*. Jul 2020;107:103461. [FREE Full text] [doi: [10.1016/j.jbi.2020.103461](https://doi.org/10.1016/j.jbi.2020.103461)] [Medline: [32504669](https://pubmed.ncbi.nlm.nih.gov/32504669/)]
56. Luengo-Polo J, Conde-Caballero D, Rivero-Jiménez B, Ballesteros-Yañez I, Castillo-Sarmiento CA, Mariano-Juárez L. Rationale and methods of evaluation for ACHO, a new virtual assistant to improve therapeutic adherence in rural elderly populations: a user-driven living lab. *Int J Environ Res Public Health*. Jul 26, 2021;18(15):7904. [FREE Full text] [doi: [10.3390/ijerph18157904](https://doi.org/10.3390/ijerph18157904)] [Medline: [34360194](https://pubmed.ncbi.nlm.nih.gov/34360194/)]
57. de Pennington N, Mole G, Lim E, Milne-Ives M, Normando E, Xue K, et al. Safety and acceptability of a natural language artificial intelligence assistant to deliver clinical follow-up to cataract surgery patients: proposal. *JMIR Res Protoc*. Jul 28, 2021;10(7):e27227. [FREE Full text] [doi: [10.2196/27227](https://doi.org/10.2196/27227)] [Medline: [34319248](https://pubmed.ncbi.nlm.nih.gov/34319248/)]
58. Bassi G, Donadello I, Gabrielli S, Salcuni S, Giuliano C, Forti S. Early development of a virtual coach for healthy coping interventions in type 2 diabetes mellitus: validation study. *JMIR Form Res*. Feb 11, 2022;6(2):e27500. [FREE Full text] [doi: [10.2196/27500](https://doi.org/10.2196/27500)] [Medline: [35147505](https://pubmed.ncbi.nlm.nih.gov/35147505/)]
59. Islam A, Chaudhry B. Early usability evaluation of a relational agent for the COVID-19 pandemic. In: Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology. New York, NY. Association for Computing Machinery; 2022. Presented at: 35th Annual ACM Symposium on User Interface Software and Technology; 29 October to 2 November 2022:1-3; Bend, OR. [doi: [10.1145/3526114.3558702](https://doi.org/10.1145/3526114.3558702)]
60. Nikou S, Chang M. Learning by building chatbot: a system usability study and teachers' views about the educational uses of chatbots. Cham, Switzerland. Springer; 2023. Presented at: International Conference on Intelligent Tutoring Systems; June 2-5, 2023:342-351; Corfu, Greece. [doi: [10.1007/978-3-031-32883-1_31](https://doi.org/10.1007/978-3-031-32883-1_31)]
61. Li J, Cesar P. Social virtual reality (VR) applications and user experiences. In: Valenzise G, Alain M, Zerman E, Ozcinar C, editors. *Immersive Video Technologies*. Cambridge, MA. Academic Press; 2023:609-648.
62. García AS, Fernández-Sotos P, Vicente-Querol MA, Sánchez-Reolid R, Rodríguez-Jimenez R, Fernández-Caballero A. Co-design of avatars to embody auditory hallucinations of patients with schizophrenia. *Virtual Reality*. Jul 08, 2021;27(1):217-232. [doi: [10.1007/s10055-021-00558-7](https://doi.org/10.1007/s10055-021-00558-7)]
63. Blasco J, Díaz-Díaz B, Igual-Camacho C, Pérez-Maletzki J, Hernández-Guilén D, Roig-Casasús S. Effectiveness of using a chatbot to promote adherence to home physiotherapy after total knee replacement, rationale and design of a randomized clinical trial. *BMC Musculoskelet Disord*. Jun 15, 2023;24(1):491. [doi: [10.1186/s12891-023-06607-3](https://doi.org/10.1186/s12891-023-06607-3)]
64. Lent HC, Ortner VK, Karmisholt KE, Wiegell SR, Nissen CV, Omland SH, et al. A chat about actinic keratosis: examining capabilities and user experience of ChatGPT as a digital health technology in dermatology. *JEADV Clinical Practice*. Oct 27, 2023;3(1):258-265. [doi: [10.1002/jvc.2.263](https://doi.org/10.1002/jvc.2.263)]
65. Knobel SEJ, Oberson R, Räber J, Schütz N, Egloff N, Botros A, et al. Evaluation of a new mobile virtual reality setup to alter pain perception: pilot development and usability study in healthy participants. *JMIR Serious Games*. Dec 11, 2024;12:e52340-e52340. [FREE Full text] [doi: [10.2196/52340](https://doi.org/10.2196/52340)] [Medline: [39661868](https://pubmed.ncbi.nlm.nih.gov/39661868/)]
66. Murawski A, Ramirez-Zohfeld V, Mell J, Tschoe M, Schierer A, Olvera C, et al. NegotiAge: development and pilot testing of an artificial intelligence-based family caregiver negotiation program. *J Am Geriatr Soc*. Apr 13, 2024;72(4):1112-1121. [doi: [10.1111/jgs.18775](https://doi.org/10.1111/jgs.18775)] [Medline: [38217356](https://pubmed.ncbi.nlm.nih.gov/38217356/)]
67. Macedo P, Madeira RN, Santos PA, Mota P, Alves B, Pereira CM. A conversational agent for empowering people with Parkinson's disease in exercising through motivation and support. *Applied Sciences*. Dec 30, 2024;15(1):223. [doi: [10.3390/app15010223](https://doi.org/10.3390/app15010223)]
68. Holderried F, Stegemann-Philipps C, Herschbach L, Moldt J, Nevins A, Griewatz J, et al. A generative pretrained transformer (GPT)-powered chatbot as a simulated patient to practice history taking: prospective, mixed methods study. *JMIR Med Educ*. Jan 16, 2024;10:e53961. [FREE Full text] [doi: [10.2196/53961](https://doi.org/10.2196/53961)] [Medline: [38227363](https://pubmed.ncbi.nlm.nih.gov/38227363/)]
69. Ahmmed A, Butts E, Naeji K, Thiamwong L, Daher S. System usability and technology acceptance of a geriatric embodied virtual human simulation in augmented reality. In: Proceedings of the 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). New York, USA. IEEE; 2024. Presented at: 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); October 21-25, 2024:594-603; Bellevue, WA. [doi: [10.1109/ISMAR62088.2024.00074](https://doi.org/10.1109/ISMAR62088.2024.00074)]
70. Arreola W, Rivas JJ, Castrejón L, Sucar LE, Rivas JJ, Rivas JJ, et al. Affective embodied agent for patient assistance in virtual rehabilitation. *IEEE Trans Affective Comput*. Oct 2025;16(4):3110-3121. [doi: [10.1109/taffc.2025.3569676](https://doi.org/10.1109/taffc.2025.3569676)]

71. Ajana K, Everard G, Lejeune T, Edwards MG. An immersive virtual reality serious game set for the clinical assessment of spatial attention impairments: effects of avatars on perspective? *Technol Health Care*. Jul 17, 2025;33(4):1626-1644. [doi: [10.1177/09287329241296380](https://doi.org/10.1177/09287329241296380)] [Medline: [39973878](https://pubmed.ncbi.nlm.nih.gov/39973878/)]
72. Tasleem U, Qamar T. Enhancing interview skills through virtual reality—a usability study. In: *Proceedings 2025 International Conference on Emerging Technologies in Electronics, Computing, and Communication (ICETECC)*. New York, NY. IEEE; 2025. Presented at: 2025 International Conference on Emerging Technologies in Electronics, Computing, and Communication (ICETECC); June 23-26, 2025:1-5; Brest, France. [doi: [10.1109/icetecc65365.2025.11070266](https://doi.org/10.1109/icetecc65365.2025.11070266)]
73. Laverde N, Grévisse C, Jaramillo S, Manrique R. Integrating large language model-based agents into a virtual patient chatbot for clinical anamnesis training. *Comput Struct Biotechnol J*. 2025;27:2481-2491. [doi: [10.1016/j.csbj.2025.05.025](https://doi.org/10.1016/j.csbj.2025.05.025)] [Medline: [40547455](https://pubmed.ncbi.nlm.nih.gov/40547455/)]
74. Liu M, Su Y. Mediguide: reducing cognitive load and improving elderly patients' healthcare experience. In: *Proceedings of the Extended Abstracts of the 2025 CHI Conference on Human Factors in Computing Systems*. New York, NY. Association for Computing Machinery; 2025. Presented at: 2025 CHI Conference on Human Factors in Computing Systems; April 26 to May 1, 2025:1-7; Yokohama, Japan. [doi: [10.1145/3706599.3706701](https://doi.org/10.1145/3706599.3706701)]
75. Gomes de Siqueira A, Gehling GM, Subramanian A, Le A, Garg R, Islam AR, et al. Methods to modernize a multimedia, web-based reproductive health education intervention for individuals with sickle cell disease or trait using virtual human narration and user-centered design. *Health Informatics J*. Nov 21, 2025;31(4):14604582251397311. [FREE Full text] [doi: [10.1177/14604582251397311](https://doi.org/10.1177/14604582251397311)] [Medline: [41270235](https://pubmed.ncbi.nlm.nih.gov/41270235/)]
76. Suri A, Sindwani G. Mothers' Assistant for Labor analgesia (MALA): a novel artificial intelligence interactive avatar for patient education in obstetric anesthesia. *Int J Obstet Anesth*. Nov 2025;64:104762. [doi: [10.1016/j.ijoa.2025.104762](https://doi.org/10.1016/j.ijoa.2025.104762)] [Medline: [40915118](https://pubmed.ncbi.nlm.nih.gov/40915118/)]
77. Valério MDP, Boschi SRMDS, Ribeiro DL, Fernandes da Silva AR, Moura LDA, Alexandre GD, et al. Proposal of a serious game for dynamic balance training using a force platform: a pilot study. *Games Health J*. Aug 01, 2025;14(4):312-320. [doi: [10.1089/g4h.2023.0194](https://doi.org/10.1089/g4h.2023.0194)] [Medline: [40126159](https://pubmed.ncbi.nlm.nih.gov/40126159/)]
78. Lehocki F, Dudasko S, Vrins A, Tirpakova V, Discantini I, Putekova S, et al. To physically embody or not? A comparison of virtual vs. physical robots as exercise coaches for older adults. In: *Proceedings of the 34th International Conference on Robot and Human Interactive Communication (RO-MAN)*. New York, NY. IEEE; 2025. Presented at: 34th International Conference on Robot and Human Interactive Communication (RO-MAN); August 25-29, 2025:2340-2345; Eindhoven, The Netherlands. [doi: [10.1109/RO-MAN63969.2025.11217653](https://doi.org/10.1109/RO-MAN63969.2025.11217653)]
79. Swallow V, Horsman J, Mazlan E, Campbell F, Zaidi R, Julian M, et al. DigiBete, a novel chatbot to support transition to adult care of young people/young adults with type 1 diabetes mellitus: outcomes from a prospective, multimethod, nonrandomized feasibility and acceptability study. *JMIR Diabetes*. Jul 23, 2025;10:e74032. [FREE Full text] [doi: [10.2196/74032](https://doi.org/10.2196/74032)] [Medline: [40699892](https://pubmed.ncbi.nlm.nih.gov/40699892/)]
80. Hanke S, Sandner E, Kadyrov S, Stainer-Hochgatterer A. Daily life support at home through a virtual support partner. In: *Proceedings of the 2nd IET International Conference on Technologies for Active and Assisted Living (TechAAL 2016)*. London, UK. IET; 2016. Presented at: 2nd IET International Conference on Technologies for Active and Assisted Living (TechAAL 2016); October 24-25, 2016:1-7; London, UK. [doi: [10.1049/ic.2016.0058](https://doi.org/10.1049/ic.2016.0058)]
81. Tielman ML, Neerincx MA, Bidarra R, Kybartas B, Brinkman W. A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories. *J Med Syst*. Aug 11, 2017;41(8):125. [FREE Full text] [doi: [10.1007/s10916-017-0771-y](https://doi.org/10.1007/s10916-017-0771-y)] [Medline: [28699083](https://pubmed.ncbi.nlm.nih.gov/28699083/)]
82. Philip P, Micoulaud-Franchi J, Sagaspe P, Sevin ED, Olive J, Bioulac S, et al. Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders. *Sci Rep*. Feb 16, 2017;7(1):42656. [FREE Full text] [doi: [10.1038/srep42656](https://doi.org/10.1038/srep42656)] [Medline: [28205601](https://pubmed.ncbi.nlm.nih.gov/28205601/)]
83. Tsiourti C, Quintas J, Ben-Moussa M, Hanke S, Nijdam N, Konstantas D. The CaMeLi framework—a multimodal virtual companion for older adults. Cham, Switzerland. Springer; 2018. Presented at: *IntelliSys 2016: SAI Intelligent Systems Conference*; 21-22 September 2016:196-217; London, UK. [doi: [10.1007/978-3-319-69266-1_10](https://doi.org/10.1007/978-3-319-69266-1_10)]
84. Tsiourti C, Ben MM, Quintas J, Loke B, Jochem I, Lopes J, et al. A virtual assistive companion for older adults: design implications for a real-world application. Cham, Switzerland. Springer; 2018. Presented at: *IntelliSys 2016: SAI Intelligent Systems Conference*; September 21-22, 2016:1014-1033; London, UK. [doi: [10.1007/978-3-319-56994-9_69](https://doi.org/10.1007/978-3-319-56994-9_69)]
85. Cameron G, Cameron D, Megaw G, Bond R, Mulvenna M, O'Neill S. Assessing the usability of a chatbot for mental health care. Cham, Switzerland. Springer; 2019. Presented at: *International Conference on Internet Science*; October 24-26, 2018:121-132; St. Petersburg, Russia. [doi: [10.1007/978-3-030-17705-8_11](https://doi.org/10.1007/978-3-030-17705-8_11)]
86. Holmes S, Moorhead A, Bond R, Zheng H, Coates V, McTear M. Usability testing of a healthcare chatbot: can we use conventional methods to assess conversational user interfaces? In: *Proceedings of the 31st European Conference on Cognitive Ergonomics*. New York, NY. Association for Computing Machinery; 2019. Presented at: 31st European Conference on Cognitive Ergonomics; September 10-13, 2019:207-214; Belfast, Northern Ireland, UK. [doi: [10.1145/3335082.3335094](https://doi.org/10.1145/3335082.3335094)]
87. Kim K, Norouzi N, Losekamp T, Bruder G, Anderson M, Welch G. Effects of patient care assistant embodiment and computer mediation on user experience. In: *Proceedings of the IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. New York, NY. IEEE; 2019. Presented at: *IEEE International Conference on Artificial*

- Intelligence and Virtual Reality (AIVR); December 9-11, 2019:17-177; San Diego, CA. [doi: [10.1109/AIVR46125.2019.00013](https://doi.org/10.1109/AIVR46125.2019.00013)]
88. Sun R, Aldunate RG, Paramathayalan VR, Ratnam R, Jain S, Morrow DG, et al. Preliminary evaluation of a self-guided fall risk assessment tool for older adults. *Arch Gerontol Geriatr*. May 2019;82:94-99. [doi: [10.1016/j.archger.2019.01.022](https://doi.org/10.1016/j.archger.2019.01.022)] [Medline: [30735851](https://pubmed.ncbi.nlm.nih.gov/30735851/)]
 89. Faddoul G, Chatterjee S. The virtual diabetician: a prototype for a virtual avatar for diabetes treatment using persuasion through storytelling. In: *Proceedings of the 25th Americas Conference on Information Systems, AMCIS 2019*. Atlanta, GA. Association for Information Systems; 2019. Presented at: 25th Americas Conference on Information Systems, AMCIS 2019; August 15-17, 2019:1-10; Cancún, Mexico.
 90. Mostajeran F, Steinicke F, Ariza NO, Gatsios D, Fotiadis D. Augmented reality for older adults exploring acceptability of virtual coaches for home-based balance training in an aging population. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, USA. ACM; 2020. Presented at: 2020 CHI Conference on Human Factors in Computing Systems; April 25-30, 2020:1-12; Honolulu, HI. [doi: [10.1145/3313831.3376565](https://doi.org/10.1145/3313831.3376565)]
 91. Ponathil A, Ozkan F, Bertrand J, Agnisarman S, Narasimha S, Welch B, et al. An empirical study investigating the user acceptance of a virtual conversational agent interface for family health history collection among the geriatric population. *Health Informatics J*. Dec 16, 2020;26(4):2946-2966. [FREE Full text] [doi: [10.1177/1460458220955104](https://doi.org/10.1177/1460458220955104)] [Medline: [32938275](https://pubmed.ncbi.nlm.nih.gov/32938275/)]
 92. Miloff A, Carlbring P, Hamilton W, Andersson G, Reuterskiöld L, Lindner P. Measuring alliance toward embodied virtual therapists in the era of automated treatments with the Virtual Therapist Alliance Scale (VTAS): development and psychometric evaluation. *J Med Internet Res*. Mar 24, 2020;22(3):e16660. [FREE Full text] [doi: [10.2196/16660](https://doi.org/10.2196/16660)] [Medline: [32207690](https://pubmed.ncbi.nlm.nih.gov/32207690/)]
 93. El KM, Angelini L, Lalanne D, Abou KO, Mugellini E. Multimodal conversational agent for older adults' behavioral change. In: *Proceedings of the 2020 International Conference on Multimodal Interaction*. New York, NY. Association for Computing Machinery; 2020. Presented at: 2020 International Conference on Multimodal Interaction; October 25-29, 2020:270-274; Utrecht, The Netherlands. [doi: [10.1145/3395035.3425315](https://doi.org/10.1145/3395035.3425315)]
 94. Jegundo AL, Dantas C, Quintas J, Dutra J, Almeida AL, Caravau H, et al. Perceived usefulness, satisfaction, ease of use and potential of a virtual companion to support the care provision for older adults. *Technologies*. Jul 25, 2020;8(3):42. [doi: [10.3390/technologies8030042](https://doi.org/10.3390/technologies8030042)]
 95. Philip P, Dupuy L, Auriacombe M, Serre F, de Sevin E, Sauteraud A, et al. Trust and acceptance of a virtual psychiatric interview between embodied conversational agents and outpatients. *NPJ Digit Med*. Jan 07, 2020;3(1):2. [FREE Full text] [doi: [10.1038/s41746-019-0213-y](https://doi.org/10.1038/s41746-019-0213-y)] [Medline: [31934646](https://pubmed.ncbi.nlm.nih.gov/31934646/)]
 96. Balsa J, Félix I, Cláudio AP, Carmo MB, Silva ICE, Guerreiro A, et al. Usability of an intelligent virtual assistant for promoting behavior change and self-care in older people with type 2 diabetes. *J Med Syst*. Jun 13, 2020;44(7):130. [doi: [10.1007/s10916-020-01583-w](https://doi.org/10.1007/s10916-020-01583-w)] [Medline: [32533367](https://pubmed.ncbi.nlm.nih.gov/32533367/)]
 97. Arem H, Scott R, Greenberg D, Kaltman R, Lieberman D, Lewin D. Assessing breast cancer survivors' perceptions of using voice-activated technology to address insomnia: feasibility study featuring focus groups and in-depth interviews. *JMIR Cancer*. May 26, 2020;6(1):e15859. [FREE Full text] [doi: [10.2196/15859](https://doi.org/10.2196/15859)] [Medline: [32348274](https://pubmed.ncbi.nlm.nih.gov/32348274/)]
 98. Chen J, Lyell D, Laranjo L, Magrabi F. Effect of speech recognition on problem solving and recall in consumer digital health tasks: controlled laboratory experiment. *J Med Internet Res*. Jun 01, 2020;22(6):e14827. [FREE Full text] [doi: [10.2196/14827](https://doi.org/10.2196/14827)] [Medline: [32442129](https://pubmed.ncbi.nlm.nih.gov/32442129/)]
 99. Ponathil A, Ozkan F, Welch B, Bertrand J, Chalil Madathil K. Family health history collected by virtual conversational agents: an empirical study to investigate the efficacy of this approach. *J Genet Couns*. Dec 03, 2020;29(6):1081-1092. [doi: [10.1002/jgc4.1239](https://doi.org/10.1002/jgc4.1239)] [Medline: [32125052](https://pubmed.ncbi.nlm.nih.gov/32125052/)]
 100. Bennion MR, Hardy GE, Moore RK, Kellett S, Millings A. Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: controlled study. *J Med Internet Res*. May 27, 2020;22(5):e16794. [FREE Full text] [doi: [10.2196/16794](https://doi.org/10.2196/16794)] [Medline: [32384055](https://pubmed.ncbi.nlm.nih.gov/32384055/)]
 101. Rehman UU, Chang DJ, Jung Y, Akhtar U, Razzaq MA, Lee S. Medical instructed real-time assistant for patient with glaucoma and diabetic conditions. *Applied Sciences*. Mar 25, 2020;10(7):2216. [doi: [10.3390/app10072216](https://doi.org/10.3390/app10072216)]
 102. Te Pas ME, Rutten WGMM, Bouwman RA, Buise MP. User experience of a chatbot questionnaire versus a regular computer questionnaire: prospective comparative study. *JMIR Med Inform*. Dec 07, 2020;8(12):e21982. [FREE Full text] [doi: [10.2196/21982](https://doi.org/10.2196/21982)] [Medline: [33284125](https://pubmed.ncbi.nlm.nih.gov/33284125/)]
 103. Denecke K, Vaaheesan S, Arulnathan A. A mental health chatbot for regulating emotions (SERMO) - concept and usability test. *IEEE Trans Emerg Topics Comput*. Jul 1, 2021;9(3):1170-1182. [doi: [10.1109/tetc.2020.2974478](https://doi.org/10.1109/tetc.2020.2974478)]
 104. Chung K, Cho HY, Park JY. A chatbot for perinatal women's and partners' obstetric and mental health care: development and usability evaluation study. *JMIR Med Inform*. Mar 03, 2021;9(3):e18607. [FREE Full text] [doi: [10.2196/18607](https://doi.org/10.2196/18607)] [Medline: [33656442](https://pubmed.ncbi.nlm.nih.gov/33656442/)]
 105. Iftikhar A, Bond RR, McGilligan V, Leslie SJ, Rjoob K, Knoery C, et al. Comparing single-page, multipage, and conversational digital forms in health care: usability study. *JMIR Hum Factors*. May 26, 2021;8(2):e25787. [FREE Full text] [doi: [10.2196/25787](https://doi.org/10.2196/25787)] [Medline: [34037531](https://pubmed.ncbi.nlm.nih.gov/34037531/)]

106. da Silva Lima Roque G, Roque de Souza R, Araújo do Nascimento JW, de Campos Filho AS, de Melo Queiroz SR, Ramos Vieira Santos IC. Content validation and usability of a chatbot of guidelines for wound dressing. *Int J Med Inform.* Jul 2021;151:104473. [doi: [10.1016/j.ijmedinf.2021.104473](https://doi.org/10.1016/j.ijmedinf.2021.104473)] [Medline: [33964703](https://pubmed.ncbi.nlm.nih.gov/33964703/)]
107. Stara V, Vera B, Bolliger D, Rossi L, Felici E, Di Rosa M, et al. Usability and acceptance of the embodied conversational agent Anne by people with dementia and their caregivers: exploratory study in home environment settings. *JMIR Mhealth Uhealth.* Jun 25, 2021;9(6):e25891. [FREE Full text] [doi: [10.2196/25891](https://doi.org/10.2196/25891)] [Medline: [34170256](https://pubmed.ncbi.nlm.nih.gov/34170256/)]
108. Yu K, Gorbachev G, Eck U, Pankratz F, Navab N, Roth D. Avatars for teleconsultation: effects of avatar embodiment techniques on user perception in 3D asymmetric telepresence. *IEEE Trans Vis Comput Graph.* Nov 2021;27(11):4129-4139. [doi: [10.1109/TVCG.2021.3106480](https://doi.org/10.1109/TVCG.2021.3106480)] [Medline: [34449373](https://pubmed.ncbi.nlm.nih.gov/34449373/)]
109. El HW, El BA, Herbert C, Abdennadher S. Chase away the virus: a character-based chatbot for COVID-19. In: *Proceedings of the 9th International Conference on Serious Games and Applications for Health (SeGAH)*. New York, NY. IEEE; 2021. Presented at: 9th International Conference on Serious Games and Applications for Health (SeGAH); August 4-6, 2021:1-8; Online. [doi: [10.1109/SEGAH52098.2021.9551895](https://doi.org/10.1109/SEGAH52098.2021.9551895)]
110. Chatzimina M, Papadaki H, Pontikoglou C, Koumakis L, Marias K, Tsiknakis M. Designing a conversational agent for patients with hematologic malignancies: usability and usefulness study. In: *Proceedings of the 2021 EMBS International Conference on Biomedical and Health Informatics (BHI)*. New York, NY. IEEE; 2021. Presented at: 2021 EMBS International Conference on Biomedical and Health Informatics (BHI); July 27-30, 2021:1-4; Athens, Greece. [doi: [10.1109/bhi50953.2021.9508587](https://doi.org/10.1109/bhi50953.2021.9508587)]
111. Nguyen TT, Sim K, Kuen AT, O'Donnell RR, Lim ST, Wang W, et al. Designing AI-based conversational agent for diabetes care in a multilingual context. In: *Proceedings of the Twenty-Fifth Pacific Asia Conference on Information Systems*. 2021. Presented at: Twenty-Fifth Pacific Asia Conference on Information Systems; July 12-14, 2021:1-14; Dubai, UAE.
112. To QG, Green C, Vandelanotte C. Feasibility, usability, and effectiveness of a machine learning-based physical activity chatbot: quasi-experimental study. *JMIR Mhealth Uhealth.* Nov 26, 2021;9(11):e28577. [FREE Full text] [doi: [10.2196/28577](https://doi.org/10.2196/28577)] [Medline: [34842552](https://pubmed.ncbi.nlm.nih.gov/34842552/)]
113. Roque G, Cavalcanti A, Nascimento J, Souza R, Queiroz S. BotCovid: development and evaluation of a chatbot to combat misinformation about COVID-19 in Brazil. In: *Proceedings of the 2021 International Conference on Systems, Man, and Cybernetics (SMC)*. New York, NY. IEEE; 2021. Presented at: 2021 International Conference on Systems, Man, and Cybernetics (SMC); October 17-20, 2021:2506-2511; Online. [doi: [10.1109/smc52423.2021.9658693](https://doi.org/10.1109/smc52423.2021.9658693)]
114. Issom D, Hardy-Dessources M, Romana M, Hartvigsen G, Lovis C. Toward a conversational agent to support the self-management of adults and young adults with sickle cell disease: usability and usefulness study. *Front Digit Health.* Jan 29, 2021;3:600333. [doi: [10.3389/FDGTH.2021.600333](https://doi.org/10.3389/FDGTH.2021.600333)]
115. Hurmuz MZ, Jansen-Kosterink SM, Beinema T, Fischer K, Op den Akker H, Hermens HJ. Evaluation of a virtual coaching system eHealth intervention: a mixed methods observational cohort study in the Netherlands. *Internet Interv.* Mar 2022;27:100501. [FREE Full text] [doi: [10.1016/j.invent.2022.100501](https://doi.org/10.1016/j.invent.2022.100501)] [Medline: [35198411](https://pubmed.ncbi.nlm.nih.gov/35198411/)]
116. Kramer LL, van Velsen L, Clark JL, Mulder BC, de Vet E. Use and effect of embodied conversational agents for improving eating behavior and decreasing loneliness among community-dwelling older adults: randomized controlled trial. *JMIR Form Res.* Apr 11, 2022;6(4):e33974. [FREE Full text] [doi: [10.2196/33974](https://doi.org/10.2196/33974)] [Medline: [35404255](https://pubmed.ncbi.nlm.nih.gov/35404255/)]
117. Minutolo A, Damiano E, De Pietro G, Fujita H, Esposito M. A conversational agent for querying Italian Patient Information Leaflets and improving health literacy. *Comput Biol Med.* Feb 2022;141:105004. [doi: [10.1016/j.combiomed.2021.105004](https://doi.org/10.1016/j.combiomed.2021.105004)] [Medline: [34774337](https://pubmed.ncbi.nlm.nih.gov/34774337/)]
118. Jansons P, Dalla Via J, Daly R, Fyfe J, Gvozdenko E, Scott D. Delivery of home-based exercise interventions in older adults facilitated by Amazon Alexa: a 12-week feasibility trial. *J Nutr Health Aging.* Jan 2022;26(1):96-102. [FREE Full text] [doi: [10.1007/s12603-021-1717-0](https://doi.org/10.1007/s12603-021-1717-0)] [Medline: [35067710](https://pubmed.ncbi.nlm.nih.gov/35067710/)]
119. Shah J, DePietro B, D'Adamo L, Firebaugh M, Laing O, Fowler LA, et al. Development and usability testing of a chatbot to promote mental health services use among individuals with eating disorders following screening. *Int J Eat Disord.* Sep 18, 2022;55(9):1229-1244. [FREE Full text] [doi: [10.1002/eat.23798](https://doi.org/10.1002/eat.23798)] [Medline: [36056648](https://pubmed.ncbi.nlm.nih.gov/36056648/)]
120. Afrizal S, Hakiem N, Permasari A, Albab H, Sanjaya G, Lazuardi L. A user-centered design of natural language processing for maternal monitoring chatbot system. In: *Proceedings of the 2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*. New York, NY. IEEE; 2022. Presented at: 2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS); November 16-17, 2022:244-248; Jakarta, Indonesia. [doi: [10.1109/ICIMCIS56303.2022.10017517](https://doi.org/10.1109/ICIMCIS56303.2022.10017517)]
121. Six SG, Byrne KA, Aly H, Harris MW. The effect of mental health app customization on depressive symptoms in college students: randomized controlled trial. *JMIR Ment Health.* Aug 09, 2022;9(8):e39516. [FREE Full text] [doi: [10.2196/39516](https://doi.org/10.2196/39516)] [Medline: [35943788](https://pubmed.ncbi.nlm.nih.gov/35943788/)]
122. Luk TT, Lui JHT, Wang MP. Efficacy, usability, and acceptability of a chatbot for promoting COVID-19 vaccination in unvaccinated or booster-hesitant young adults: pre-post pilot study. *J Med Internet Res.* Oct 04, 2022;24(10):e39063. [FREE Full text] [doi: [10.2196/39063](https://doi.org/10.2196/39063)] [Medline: [36179132](https://pubmed.ncbi.nlm.nih.gov/36179132/)]
123. Striegl J, Gotthardt M, Loitsch C, Weber G. Investigating the usability of voice assistant-based CBT for age-related depression. In: *Proceedings of the International Conference on Computers Helping People With Special Needs*. Cham,

- Switzerland. Springer; 2022. Presented at: International Conference on Computers Helping People With Special Needs; July 11-15, 2022:432-441; Lecco, Italy. [doi: [10.1007/978-3-031-08648-9_50](https://doi.org/10.1007/978-3-031-08648-9_50)]
124. He Y, Yang L, Zhu X, Wu B, Zhang S, Qian C, et al. Mental health chatbot for young adults with depressive symptoms during the COVID-19 pandemic: single-blind, three-arm randomized controlled trial. *J Med Internet Res*. Nov 21, 2022;24(11):e40719. [FREE Full text] [doi: [10.2196/40719](https://doi.org/10.2196/40719)] [Medline: [36355633](https://pubmed.ncbi.nlm.nih.gov/36355633/)]
 125. Larbi D, Denecke K, Gabarron E. Usability testing of a social media chatbot for increasing physical activity behavior. *J Pers Med*. May 20, 2022;12(5):828. [FREE Full text] [doi: [10.3390/jpm12050828](https://doi.org/10.3390/jpm12050828)] [Medline: [35629252](https://pubmed.ncbi.nlm.nih.gov/35629252/)]
 126. Soni H, Ivanova J, Wilczewski H, Bailey A, Ong T, Narma A, et al. Virtual conversational agents versus online forms: patient experience and preferences for health data collection. *Front Digit Health*. Oct 13, 2022;4:954069. [FREE Full text] [doi: [10.3389/fdgth.2022.954069](https://doi.org/10.3389/fdgth.2022.954069)] [Medline: [36310920](https://pubmed.ncbi.nlm.nih.gov/36310920/)]
 127. Roca S, Almenara M, Gilaberte Y, Gracia-Cazaña T, Morales Callaghan AM, Murciano D, et al. When virtual assistants meet tele dermatology: validation of a virtual assistant to improve the quality of life of psoriatic patients. *Int J Environ Res Public Health*. Nov 05, 2022;19(21):14527. [FREE Full text] [doi: [10.3390/ijerph192114527](https://doi.org/10.3390/ijerph192114527)] [Medline: [36361408](https://pubmed.ncbi.nlm.nih.gov/36361408/)]
 128. Richards D, Miranda Maciel PS, Janssen H. The co-design of an embodied conversational agent to help stroke survivors manage their recovery. *Robotics*. Aug 22, 2023;12(5):120. [doi: [10.3390/robotics12050120](https://doi.org/10.3390/robotics12050120)]
 129. Martinho D, Crista V, Carneiro J, Matsui K, Corchado JM, Marreiros G. Effects of a gamified agent-based system for personalized elderly care: pilot usability study. *JMIR Serious Games*. Nov 23, 2023;11:e48063. [FREE Full text] [doi: [10.2196/48063](https://doi.org/10.2196/48063)] [Medline: [37995116](https://pubmed.ncbi.nlm.nih.gov/37995116/)]
 130. Gingele AJ, Amin H, Vaassen A, Schnur I, Pearl C, Brunner-La Rocca H, et al. Integrating avatar technology into a telemedicine application in heart failure patients: a pilot study. *Wien Klin Wochenschr*. Dec 02, 2023;135(23-24):680-684. [FREE Full text] [doi: [10.1007/s00508-022-02150-8](https://doi.org/10.1007/s00508-022-02150-8)] [Medline: [36732377](https://pubmed.ncbi.nlm.nih.gov/36732377/)]
 131. Wanberg LJ, Kim A, Vogel RI, Sadak KT, Teoh D. Usability and satisfaction testing of game-based learning avatar-navigated mobile (GLAm), an app for cervical cancer screening: mixed methods study. *JMIR Form Res*. Aug 08, 2023;7:e45541. [FREE Full text] [doi: [10.2196/45541](https://doi.org/10.2196/45541)] [Medline: [37552527](https://pubmed.ncbi.nlm.nih.gov/37552527/)]
 132. Wrightson-Hester A, Anderson G, Dunstan J, McEvoy PM, Sutton CJ, Myers B, et al. An artificial therapist (Manage Your Life Online) to support the mental health of youth: co-design and case series. *JMIR Hum Factors*. Jul 21, 2023;10:e46849. [FREE Full text] [doi: [10.2196/46849](https://doi.org/10.2196/46849)] [Medline: [37477969](https://pubmed.ncbi.nlm.nih.gov/37477969/)]
 133. Griffin AC, Khairat S, Bailey SC, Chung AE. A chatbot for hypertension self-management support: user-centered design, development, and usability testing. *JAMIA Open*. Oct 2023;6(3):ooad073. [FREE Full text] [doi: [10.1093/jamiaopen/ooad073](https://doi.org/10.1093/jamiaopen/ooad073)] [Medline: [37693367](https://pubmed.ncbi.nlm.nih.gov/37693367/)]
 134. Powell L, Nour R, Sleibi R, Al Suwaidi H, Zary N. Democratizing the development of chatbots to improve public health: feasibility study of COVID-19 misinformation. *JMIR Hum Factors*. Dec 28, 2023;10:e43120. [FREE Full text] [doi: [10.2196/43120](https://doi.org/10.2196/43120)] [Medline: [37290040](https://pubmed.ncbi.nlm.nih.gov/37290040/)]
 135. Islam A, Chaudhry BM. Design validation of a relational agent by COVID-19 patients: mixed methods study. *JMIR Hum Factors*. Jun 08, 2023;10:e42740. [FREE Full text] [doi: [10.2196/42740](https://doi.org/10.2196/42740)] [Medline: [36350760](https://pubmed.ncbi.nlm.nih.gov/36350760/)]
 136. Babington-Ashaye A, de Moerloose P, Diop S, Geissbuhler A. Design, development and usability of an educational AI chatbot for people with haemophilia in Senegal. *Haemophilia*. Jul 22, 2023;29(4):1063-1073. [doi: [10.1111/hae.14815](https://doi.org/10.1111/hae.14815)] [Medline: [37347648](https://pubmed.ncbi.nlm.nih.gov/37347648/)]
 137. Biro J, Linder C, Neyens D. The effects of a health care chatbot's complexity and persona on user trust, perceived usability, and effectiveness: mixed methods study. *JMIR Hum Factors*. Feb 01, 2023;10:e41017. [FREE Full text] [doi: [10.2196/41017](https://doi.org/10.2196/41017)] [Medline: [36724004](https://pubmed.ncbi.nlm.nih.gov/36724004/)]
 138. Tawfik E, Ghallab E, 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*. Apr 06, 2023;22(1):102. [FREE Full text] [doi: [10.1186/s12912-023-01243-7](https://doi.org/10.1186/s12912-023-01243-7)] [Medline: [37024875](https://pubmed.ncbi.nlm.nih.gov/37024875/)]
 139. Wilczewski H, Soni H, Ivanova J, Ong T, Barrera JF, Bunnell BE, et al. Older adults' experience with virtual conversational agents for health data collection. *Front Digit Health*. 2023;5:1125926. [FREE Full text] [doi: [10.3389/fdgth.2023.1125926](https://doi.org/10.3389/fdgth.2023.1125926)] [Medline: [37006821](https://pubmed.ncbi.nlm.nih.gov/37006821/)]
 140. Bruijnes M, Kesteloo M, Brinkman W. Reducing social diabetes distress with a conversational agent support system: a three-week technology feasibility evaluation. *Front Digit Health*. Jun 13, 2023;5:1149374. [FREE Full text] [doi: [10.3389/fdgth.2023.1149374](https://doi.org/10.3389/fdgth.2023.1149374)] [Medline: [37383944](https://pubmed.ncbi.nlm.nih.gov/37383944/)]
 141. Bevilacqua R, Felici E, Cucchieri G, Amabili G, Margaritini A, Franceschetti C, et al. Results of the Italian RESILIEN-T pilot study: a mobile health tool to support older people with mild cognitive impairment. *J Clin Med*. Sep 22, 2023;12(19):6129. [FREE Full text] [doi: [10.3390/jcm12196129](https://doi.org/10.3390/jcm12196129)] [Medline: [37834775](https://pubmed.ncbi.nlm.nih.gov/37834775/)]
 142. de SCA, do NJ, Roque G, de SR, de MQS, Correa J. A virtual assistant in vaccine pharmacovigilance: content and usability validation. *CIN: Computers, Informatics, Nursing*. 2023;41(7):482-490. [doi: [10.1097/cin.0000000000000978](https://doi.org/10.1097/cin.0000000000000978)]
 143. Chua JYX, Choolani M, Chee CYI, Yi H, Chan YH, Lalor JG, et al. 'Parentbot - a digital healthcare assistant (PDA)': a mobile application-based perinatal intervention for parents: development study. *Patient Educ Couns*. Sep 2023;114:107805. [doi: [10.1016/j.pec.2023.107805](https://doi.org/10.1016/j.pec.2023.107805)] [Medline: [37245443](https://pubmed.ncbi.nlm.nih.gov/37245443/)]

144. Chen N, Huang C, Fan C, Lu L, Lin F, Liao J, et al. User evaluation of a chat-based instant messaging support health education program for patients with chronic kidney disease: preliminary findings of a formative study. *JMIR Form Res*. Sep 19, 2023;7:e45484. [[FREE Full text](#)] [doi: [10.2196/45484](https://doi.org/10.2196/45484)] [Medline: [37725429](https://pubmed.ncbi.nlm.nih.gov/37725429/)]
145. Albino de Queiroz D, Silva Passarello R, Veloso de Moura Fé V, Rossini A, Folchini da Silveira E, Aparecida Isquierdo Fonseca de Queiroz E, et al. A wearable chatbot-based model for monitoring colorectal cancer patients in the active phase of treatment. *Healthcare Analytics*. Dec 2023;4:100257. [doi: [10.1016/j.health.2023.100257](https://doi.org/10.1016/j.health.2023.100257)]
146. Alonso-Mencía J, Castro-Rodríguez M, Herrero-Pinilla B, Alonso-Weber JM, Rodríguez-Mañas L, Pérez-Rodríguez R. ADELA: a conversational virtual assistant to prevent delirium in hospitalized older persons. *J Supercomput*. May 09, 2023;79(15):17670-17690. [doi: [10.1007/s11227-023-05352-7](https://doi.org/10.1007/s11227-023-05352-7)]
147. Larkin C, Tulu B, Djasasbi S, Garner R, Varzgani F, Siddique M, et al. Comparing the acceptability and quality of intervention modalities for suicidality in the emergency department: randomized feasibility trial. *JMIR Ment Health*. Oct 24, 2023;10:e49783. [[FREE Full text](#)] [doi: [10.2196/49783](https://doi.org/10.2196/49783)] [Medline: [37874619](https://pubmed.ncbi.nlm.nih.gov/37874619/)]
148. Morato J, do Nascimento JWA, Roque G, de Souza RR, Santos ICRV. Development, validation, and usability of the chatbot ESTOMABOT to promote self-care of people with intestinal ostomy. *Comput Inform Nurs*. Dec 01, 2023;41(12):1037-1045. [doi: [10.1097/CIN.0000000000001075](https://doi.org/10.1097/CIN.0000000000001075)] [Medline: [37725781](https://pubmed.ncbi.nlm.nih.gov/37725781/)]
149. Sanchez-Ortuno MM, Pecune F, Coelho J, Micoulaud-Franchi JA, Salles N, Auriacombe M, et al. Predictors of users' adherence to a fully automated digital intervention to manage insomnia complaints. *J Am Med Inform Assoc*. Nov 17, 2023;30(12):1934-1942. [doi: [10.1093/jamia/ocad163](https://doi.org/10.1093/jamia/ocad163)] [Medline: [37672004](https://pubmed.ncbi.nlm.nih.gov/37672004/)]
150. Larkin C, Djasasbi S, Boudreaux ED, Varzgani F, Garner R, Siddique M, et al. ReachCare mobile apps for patients experiencing suicidality in the emergency department: development and usability testing using mixed methods. *JMIR Form Res*. Jan 27, 2023;7:e41422. [[FREE Full text](#)] [doi: [10.2196/41422](https://doi.org/10.2196/41422)] [Medline: [36705961](https://pubmed.ncbi.nlm.nih.gov/36705961/)]
151. Hocking J, Maeder A, Powers D, Perimal-Lewis L, Dodd B, Lange B. Mixed methods, single case design, feasibility trial of a motivational conversational agent for rehabilitation for adults with traumatic brain injury. *Clin Rehabil*. Mar 06, 2024;38(3):322-336. [[FREE Full text](#)] [doi: [10.1177/02692155231216615](https://doi.org/10.1177/02692155231216615)] [Medline: [38058144](https://pubmed.ncbi.nlm.nih.gov/38058144/)]
152. Ashrafi N, Vona F, Hinzmann S, Graf P, Harnisch P, Voigt-Antons J. Single vs dual: influence of the number of displays on user experience within virtually embodied conversational systems. In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*. New York, NY. Association for Computing Machinery; 2024. Presented at: 30th ACM Symposium on Virtual Reality Software and Technology; October 9-11, 2024:1-2; Trier, Germany. [doi: [10.1145/3641825.3689700](https://doi.org/10.1145/3641825.3689700)]
153. Thunström AO, Carlsen HK, Ali L, Larson T, Hellström A, Steingrímsson S. Usability comparison among healthy participants of an anthropomorphic digital human and a text-based chatbot as a responder to questions on mental health: randomized controlled trial. *JMIR Hum Factors*. Apr 29, 2024;11:e54581. [[FREE Full text](#)] [doi: [10.2196/54581](https://doi.org/10.2196/54581)] [Medline: [38683664](https://pubmed.ncbi.nlm.nih.gov/38683664/)]
154. Martins A, Velez Lapão L, Nunes IL, Paula Giordano A, Semedo H, Vital C, et al. A conversational agent for enhanced self-management after cardiothoracic surgery. *Int J Med Inform*. Dec 2024;192:105640. [[FREE Full text](#)] [doi: [10.1016/j.ijmedinf.2024.105640](https://doi.org/10.1016/j.ijmedinf.2024.105640)] [Medline: [39321492](https://pubmed.ncbi.nlm.nih.gov/39321492/)]
155. Striegl J, Richter JW, Grossmann L, Bråstad B, Gotthardt M, Rück C, et al. Deep learning-based dimensional emotion recognition for conversational agent-based cognitive behavioral therapy. *PeerJ Comput Sci*. 2024;10:e2104. [doi: [10.7717/peerj-cs.2104](https://doi.org/10.7717/peerj-cs.2104)] [Medline: [38983201](https://pubmed.ncbi.nlm.nih.gov/38983201/)]
156. Schlieter H, Gand K, Weimann TG, Sandner E, Kreiner K, Thoma S, et al. Designing virtual coaching solutions. *Bus Inf Syst Eng*. May 29, 2024;66(3):377-400. [doi: [10.1007/s12599-024-00871-w](https://doi.org/10.1007/s12599-024-00871-w)]
157. Alves B, Mota PR, Sineiro D, Carmo R, Santos P, Macedo P, et al. MoveONParkinson: developing a personalized motivational solution for Parkinson's disease management. *Front Public Health*. Aug 19, 2024;12:1420171. [[FREE Full text](#)] [doi: [10.3389/fpubh.2024.1420171](https://doi.org/10.3389/fpubh.2024.1420171)] [Medline: [39224558](https://pubmed.ncbi.nlm.nih.gov/39224558/)]
158. Ali SH, Rahman F, Kuwar A, Khanna T, Nayak A, Sharma P, et al. Rapid, tailored dietary and health education through a social media chatbot microintervention: development and usability study with practical recommendations. *JMIR Form Res*. Dec 09, 2024;8:e52032. [[FREE Full text](#)] [doi: [10.2196/52032](https://doi.org/10.2196/52032)] [Medline: [39652870](https://pubmed.ncbi.nlm.nih.gov/39652870/)]
159. Devittori G, Akeddar M, Retevoi A, Schneider F, Cvetkova V, Dinacci D. Towards RehabCoach: design and preliminary evaluation of a conversational agent sup-porting unsupervised therapy after stroke. In: *Proceedings of the 10th RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*. New York, NY. IEEE; 2024. Presented at: 10th RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob); September 1-4, 2024:569-574; Heidelberg, Germany. [doi: [10.1109/BioRob60516.2024.10719881](https://doi.org/10.1109/BioRob60516.2024.10719881)]
160. Nkabane-Nkholongo E, Mpata-Mokgatle M, Jack BW, Julce C, Bickmore T. Usability and acceptability of a conversational agent health education app (Nthabi) for young women in Lesotho: quantitative study. *JMIR Hum Factors*. Mar 12, 2024;11:e52048. [[FREE Full text](#)] [doi: [10.2196/52048](https://doi.org/10.2196/52048)] [Medline: [38470460](https://pubmed.ncbi.nlm.nih.gov/38470460/)]
161. Nguyen MH, Sedoc J, Taylor CO. Usability, engagement, and report usefulness of chatbot-based family health history data collection: mixed methods analysis. *J Med Internet Res*. Sep 30, 2024;26:e55164. [[FREE Full text](#)] [doi: [10.2196/55164](https://doi.org/10.2196/55164)] [Medline: [39348188](https://pubmed.ncbi.nlm.nih.gov/39348188/)]

162. Reilly ED, Kelly MM, Grigorian HL, Waring ME, Quigley KS, Hogan TP, et al. Virtual coach-guided online acceptance and commitment therapy for chronic pain: pilot feasibility randomized controlled trial. *JMIR Form Res*. Nov 08, 2024;8:e56437. [FREE Full text] [doi: [10.2196/56437](https://doi.org/10.2196/56437)] [Medline: [39514264](https://pubmed.ncbi.nlm.nih.gov/39514264/)]
163. Segear S, Chheang V, Baron L, Li J, Kim K, Barmaki RL. Visual feedback and guided balance training in an immersive virtual reality environment for lower extremity rehabilitation. *Comput Graph*. Apr 2024;119:103880. [doi: [10.1016/j.cag.2024.01.007](https://doi.org/10.1016/j.cag.2024.01.007)] [Medline: [38645661](https://pubmed.ncbi.nlm.nih.gov/38645661/)]
164. Morita Y, Wong Z. Designing and adopting a video-based LINE chatbot system for endoscopy explanation in a real-world hospital: a mixed method approach. In: *Proceedings of the 2024 Australasian Computer Science Week*. New York, NY. Association for Computing Machinery; 2024. Presented at: 2024 Australasian Computer Science Week; January 29 to February 2, 2024:87-91; Sydney, Australia. [doi: [10.1145/3641142.3641161](https://doi.org/10.1145/3641142.3641161)]
165. Song S, Seo Y, Hwang S, Kim H, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. *JMIR Mhealth Uhealth*. Jun 17, 2024;12:e55842. [FREE Full text] [doi: [10.2196/55842](https://doi.org/10.2196/55842)] [Medline: [38885033](https://pubmed.ncbi.nlm.nih.gov/38885033/)]
166. Wan Ab.Rahman WN, Abdul Hamid NM. Rule-based chatbot for early self-depression indication: a promising approach. *Int J Inform Visualization*. Nov 24, 2024;8(3-2):1625. [doi: [10.62527/joiv.8.3-2.1628](https://doi.org/10.62527/joiv.8.3-2.1628)]
167. De la Puente G, Silva A, Felix R. Development of a chatbot powered by artificial intelligence to diagnose and improve stress and anxiety levels in university students. In: *Proceedings of the XXXI International Conference on Electronics, Electrical Engineering and Computing (IN-TERCON)*. New York, NY. IEEE; 2024. Presented at: XXXI International Conference on Electronics, Electrical Engineering and Computing (IN-TERCON); November 6-8, 2024:1-8; Lima, Peru. [doi: [10.1109/intercon63140.2024.10833503](https://doi.org/10.1109/intercon63140.2024.10833503)]
168. Abilkaiyrkyzy A, Laamarti F, Hamdi M, Saddik AE. Dialogue system for early mental illness detection: toward a digital twin solution. *IEEE Access*. 2024;12:2007-2024. [doi: [10.1109/access.2023.3348783](https://doi.org/10.1109/access.2023.3348783)]
169. Menard C, Coelho J, De SE, Levavasseur Y, Philip P, Pecune F. Quantitative and qualitative evaluation of the acceptability of an embodied conversational agent for insomnia screening. In: *Proceedings of the 24th ACM International Conference on Intelligent Virtual Agents*. New York, NY. Association for Computing Machinery; 2024. Presented at: 24th ACM International Conference on Intelligent Virtual Agents; September 16-19, 2024:1-4; Glasgow, Scotland, UK. [doi: [10.1145/3652988.3673940](https://doi.org/10.1145/3652988.3673940)]
170. Papiernik P, Dzula S, Zimanyi M, Millgate E, Bouazzaoui M, Buttmer J, et al. Acceptability of a conversational agent-led digital program for anxiety: mixed methods study of user perspectives. *JMIR Hum Factors*. Nov 04, 2025;12:e76377. [FREE Full text] [doi: [10.2196/76377](https://doi.org/10.2196/76377)] [Medline: [40824528](https://pubmed.ncbi.nlm.nih.gov/40824528/)]
171. Shi J, Jain R, Chi S, Doh H, Chi H, Quinn A, et al. Caring-AI: towards authoring context-aware augmented reality instruction through generative artificial intelligence. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. New York, NY. Association for Computing Machinery; 2025. Presented at: 2025 CHI Conference on Human Factors in Computing Systems; April 26 to May 1, 2025:1-23; Yokohama, Japan. [doi: [10.1145/3706598.3713348](https://doi.org/10.1145/3706598.3713348)]
172. Mendes C, Pereira R, Frazao L, Ribeiro J, Rodrigues N, Costa N. Emotionally intelligent customizable conversational agent for elderly care: development and impact of Chatto. In: *Proceedings of the 11th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*. New York, NY. Association for Computing Machinery; 2025. Presented at: 11th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion; November 13-15, 2024:208-214; Abu Dhabi, UAE. [doi: [10.1145/3696593.3696619](https://doi.org/10.1145/3696593.3696619)]
173. Six S, Schlesener E, Hill V, Babu SV, Byrne K. Impact of conversational and animation features of a mental health app virtual agent on depressive symptoms and user experience among college students: randomized controlled trial. *JMIR Ment Health*. Apr 11, 2025;12:e67381-e67381. [FREE Full text] [doi: [10.2196/67381](https://doi.org/10.2196/67381)] [Medline: [40215483](https://pubmed.ncbi.nlm.nih.gov/40215483/)]
174. Lazem H, Harris D, Hall A, Mansoubi M, Pontes RG, de Mello Monteiro CB, et al. Validity, safety, usability, and user experience of virtual reality gamified home-based exercises in stroke. *Clin Rehabil*. Nov 02, 2025;39(11):1527-1540. [FREE Full text] [doi: [10.1177/02692155251371435](https://doi.org/10.1177/02692155251371435)] [Medline: [40891724](https://pubmed.ncbi.nlm.nih.gov/40891724/)]
175. Del Pino R, de Echevarría AO, Díez-Cirarda M, Ustarroz-Aguirre I, Caprino M, Liu J, et al. Virtual coach and telerehabilitation for Parkinson's disease patients: vCare system. *J Public Health (Berl)*. Nov 13, 2023;33(7):1583-1596. [doi: [10.1007/s10389-023-02082-1](https://doi.org/10.1007/s10389-023-02082-1)]
176. Haider SA, Prabha S, Gomez-Cabello CA, Genovese A, Collaco B, Wood N, et al. Artificial intelligence physician avatars for patient education: a pilot study. *J Clin Med*. Dec 04, 2025;14(23):8595. [FREE Full text] [doi: [10.3390/jcm14238595](https://doi.org/10.3390/jcm14238595)] [Medline: [41375898](https://pubmed.ncbi.nlm.nih.gov/41375898/)]
177. Düzgün MV, İşler A, Kazan MS. Effect of avatar-based education program in hydrocephalus on ventriculoperitoneal shunt complications and parents' knowledge and care skills: multicenter randomized controlled trials. *Pediatr Neurol*. Aug 2025;169:131-139. [doi: [10.1016/j.pediatrneurol.2025.05.019](https://doi.org/10.1016/j.pediatrneurol.2025.05.019)] [Medline: [40505427](https://pubmed.ncbi.nlm.nih.gov/40505427/)]
178. Ashrafi N, Vona F, Hinzmann S, Henning J, Vergari M, Warsinke M. Size matters: the impact of avatar size on user experience in healthcare applications. In: *Proceedings of the 17th International Conference on Quality of Multimedia*

- Experience (QoMEX). New York, NY. IEEE; 2025. Presented at: 17th International Conference on Quality of Multimedia Experience (QoMEX); September 29 to October 3, 2025:1-7; Madrid, Spain. [doi: [10.1109/QoMEX65720.2025.11219980](https://doi.org/10.1109/QoMEX65720.2025.11219980)]
179. Pinto G, De SJ, Rosa R, Rodriguez D. Embodied multimodal chatbot for mental health support in web-based 3D environments. In: Proceedings of the 2025 International Conference on Software, Telecommunications and Computer Networks (SoftCOM). New York, NY. IEEE; 2025. Presented at: 2025 International Conference on Software, Telecommunications and Computer Networks (SoftCOM); September 18-20, 2025:1-6; Split, Croatia. [doi: [10.23919/softcom66362.2025.11197343](https://doi.org/10.23919/softcom66362.2025.11197343)]
 180. Hopman K, Richards D, Norberg MM. A staged approach to the development of an embodied conversational agent to support wellbeing after injury. *Behaviour & Information Technology*. Mar 20, 2025;45(8):1490-1514. [doi: [10.1080/0144929x.2025.2477757](https://doi.org/10.1080/0144929x.2025.2477757)]
 181. Elliott TC, Yang Y, Knibbe J, Henry JD, Baghaei N. Avatar customization and embodiment in virtual reality self-compassion therapy for depressive symptoms: three-part mixed methods study. *JMIR Form Res*. Oct 02, 2025;9:e71004-e71004. [FREE Full text] [doi: [10.2196/71004](https://doi.org/10.2196/71004)] [Medline: [41037663](https://pubmed.ncbi.nlm.nih.gov/41037663/)]
 182. Hernandez R, Nisar H, Kesavadas, McGee MC, Gerstner GJ, Martinez A, et al. Assessing safety and feasibility of virtual reality intervention in patients with lung cancer: a pilot study. *Support Care Cancer*. Mar 25, 2025;33(4):318. [doi: [10.1007/s00520-025-09338-4](https://doi.org/10.1007/s00520-025-09338-4)] [Medline: [40131484](https://pubmed.ncbi.nlm.nih.gov/40131484/)]
 183. Groninger H, Arem H, Ayangma L, Gong L, Zhou E, Greenberg D. Development of a voice-activated virtual assistant to improve insomnia among young adult cancer survivors: mixed methods feasibility and acceptability study. *JMIR Form Res*. Mar 10, 2025;9:e64869. [FREE Full text] [doi: [10.2196/64869](https://doi.org/10.2196/64869)] [Medline: [40063947](https://pubmed.ncbi.nlm.nih.gov/40063947/)]
 184. Sharp G, Dwyer B, Randhawa A, McGrath I, Hu H. The effectiveness of a chatbot single-session intervention for people on waitlists for eating disorder treatment: randomized controlled trial. *J Med Internet Res*. May 21, 2025;27:e70874. [FREE Full text] [doi: [10.2196/70874](https://doi.org/10.2196/70874)] [Medline: [40397933](https://pubmed.ncbi.nlm.nih.gov/40397933/)]
 185. Petrati H, Chan G, Orji R. Leveraging generative and rule-based models for persuasive STI education: a multi-chatbot mobile application. In: Proceedings of the 7th ACM Conference on Conversational User Interfaces. New York, NY. Association for Computing Machinery; 2025. Presented at: 7th ACM Conference on Conversational User Interfaces; July 8-10, 2025:1-9; Waterloo, ON, Canada. [doi: [10.1145/3719160.3737619](https://doi.org/10.1145/3719160.3737619)]
 186. Isaacs K, Shifflett A, Patel K, Karpisek L, Cui Y, Lawental M, et al. Women empowered to connect with addiction resources and engage in evidence-based treatment (WE-CARE)-an mHealth application for the universal screening of alcohol, substance use, depression, and anxiety: usability and feasibility study. *JMIR Form Res*. Feb 07, 2025;9:e62915. [FREE Full text] [doi: [10.2196/62915](https://doi.org/10.2196/62915)] [Medline: [39918861](https://pubmed.ncbi.nlm.nih.gov/39918861/)]
 187. Sivakumar B, Ricupero M, Mahajan A, Jefferson K, Wenger J, Code J, et al. A mobile app intervention to support nutrition education for heart failure management: co-design, development and user-testing. *BMC Nutr*. Jul 12, 2025;11(1):139. [doi: [10.1186/s40795-025-01117-y](https://doi.org/10.1186/s40795-025-01117-y)] [Medline: [40652288](https://pubmed.ncbi.nlm.nih.gov/40652288/)]
 188. Zhou D, Globa A. Social connectedness in older adults: the role of MR and haptic interaction in remote digital storytelling. In: Proceedings of the 37th Australian Conference on Human-Computer Interaction. New York, NY. Association for Computing Machinery; 2025. Presented at: 37th Australian Conference on Human-Computer Interaction; November 29 to December 3, 2025:807-818; Sydney, Australia. [doi: [10.1145/3764687.3769932](https://doi.org/10.1145/3764687.3769932)]
 189. Kuswik H, Poletaykina A, Rittmann A, Kruse L, Steinicke F. Instruct me! Comparing virtual agents and static picture instructions in therapeutic virtual reality exercises. In: Proceedings of the 2025 ACM Symposium on Spatial User Interaction. New York, NY. Association for Computing Machinery; 2025. Presented at: 2025 ACM Symposium on Spatial User Interaction; November 10-11, 2025:1-10; Montreal, QC, Canada. [doi: [10.1145/3694907.3765934](https://doi.org/10.1145/3694907.3765934)]
 190. Carneiro P, Águia I, Palricas D, Rocha A, Teixeira A, Almeida N. Kitchen assistant for active ageing at home. In: Proceedings of the 11th International Conference on Software Development Technologies for Enhancing Accessibility Fighting Info-exclusion. New York, NY. Ass; 2025. Presented at: 11th International Conference on Software Development Technologies for Enhancing Accessibility Fighting Info-exclusion; November 13-15, 2024:193-202; Abu Dhabi, UAE. [doi: [10.1145/3696593.3696624](https://doi.org/10.1145/3696593.3696624)]
 191. Rosteck N, Striegl J, Loitsch C. Bridging the treatment gap: a novel LLM-driven system for scalable initial patient assessments in mental healthcare. In: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. New York, NY. Association for Computing Machinery; 2025. Presented at: CHI Conference on Human Factors in Computing Systems; April 26 to May 1, 2025:1-8; Yokohama, Japan. [doi: [10.1145/3706599.3720043](https://doi.org/10.1145/3706599.3720043)]
 192. Tang L, Weng C, Hou S, Savas L, Neumann A, Moore N. Constructing a conversation-first dialogue model for comprehensive, evidence-based counseling for HPV vaccination for young adults. In: Proceedings of the International Conference on Human-Computer Interaction. Cham, Switzerland. Springer; 2025. Presented at: International Conference on Human-Computer Interaction; June 22-27, 2025:72-84; Gothenburg, Sweden. [doi: [10.1007/978-3-031-93502-2_5](https://doi.org/10.1007/978-3-031-93502-2_5)]
 193. Ejem D, Bakitas M, Durant RW, Parker TN, Oppong KD, Esterson J, et al. Exploring the acceptability and feasibility of a self-directed approach to identifying health priorities in a sample of southern older African American Adults with multiple chronic conditions. *J Racial Ethn Health Disparities*. Aug 23, 2026;13(4):2940-2957. [doi: [10.1007/s40615-025-02469-8](https://doi.org/10.1007/s40615-025-02469-8)] [Medline: [40410637](https://pubmed.ncbi.nlm.nih.gov/40410637/)]

194. Begum TUS, Devi RR, Haridas D, Bacanin N, Jovicic MD, Nikolic B. My diabetes care: an AI-based mobile app with conversational agent for type 2 diabetes self-management. *Sci Rep*. Nov 17, 2025;15(1):40228. [[FREE Full text](#)] [doi: [10.1038/s41598-025-24172-8](https://doi.org/10.1038/s41598-025-24172-8)] [Medline: [41249394](#)]
195. Suffoletto B, Clark DB, Lee C, Mason M, Schultz J, Szeto I, et al. Development and preliminary testing of a secure large language model-based chatbot for brief alcohol counseling in young adults. *Drug Alcohol Depend*. Jul 01, 2025;272:112697. [doi: [10.1016/j.drugalcdep.2025.112697](https://doi.org/10.1016/j.drugalcdep.2025.112697)] [Medline: [40334327](#)]
196. Khairat S, Geracitano J, Mendis K. Acceptability of academic large language model for patients seeking health information. *Studies in Health Technology and Informatics*. 2025;328:31. [doi: [10.3233/shti250667](https://doi.org/10.3233/shti250667)]
197. Leung YW, So J, Sidhu A, Asokan V, Gancarz M, Gajjar VB, et al. The extent to which artificial intelligence can help fulfill metastatic breast cancer patient healthcare needs: a mixed-methods study. *Curr Oncol*. Mar 02, 2025;32(3):145. [[FREE Full text](#)] [doi: [10.3390/curroncol32030145](https://doi.org/10.3390/curroncol32030145)] [Medline: [40136349](#)]
198. Parry M, Huang T, Clarke H, Bjørnnes AK, Harvey P, Parente L, et al. Development and systematic evaluation of a progressive web application for women with cardiac pain: usability study. *JMIR Hum Factors*. Apr 17, 2025;12:e57583. [[FREE Full text](#)] [doi: [10.2196/57583](https://doi.org/10.2196/57583)] [Medline: [40245401](#)]
199. Daniels K, Vonck S, Robijns J, Quadflieg K, Bergs J, Spooren A, et al. Exploring the feasibility of a 5-week mHealth intervention to enhance physical activity and an active, healthy lifestyle in community-dwelling older adults: mixed methods study. *JMIR Aging*. Jan 27, 2025;8:e63348. [[FREE Full text](#)] [doi: [10.2196/63348](https://doi.org/10.2196/63348)] [Medline: [39869906](#)]
200. Vandelanotte C, Maher C, Hodgetts D, Imam T, Rashid M, To Q, et al. Impact of iterative development and beta-testing on the usability and acceptability of a novel just-in-time adaptive digital physical activity intervention. *J Phys Act Health*. Oct 01, 2025;22(10):1315-1321. [[FREE Full text](#)] [doi: [10.1123/jpah.2025-0107](https://doi.org/10.1123/jpah.2025-0107)] [Medline: [40467054](#)]
201. Wolff D, Kupka T, Reichert C, Ammon N, Oeltze-Jafra S, Vajen B. Personalized support in hereditary breast and ovarian cancer after genetic counseling by the chatbot-based GENIE mobile app: proof-of-concept wizard of oz study. *JMIR Form Res*. Jun 05, 2025;9:e69115-e69115. [[FREE Full text](#)] [doi: [10.2196/69115](https://doi.org/10.2196/69115)] [Medline: [40471748](#)]
202. Chisty S, Miami A, Noor J. Smart Sheba: enhancing elderly user experience with LLM-enabled chatbots and user-centered design. In: *Proceedings of the 13th International Conference on Information & Communication Technologies and Development*. New York, NY. Association for Computing Machinery; 2025. Presented at: 13th International Conference on Information & Communication Technologies and Development; December 9-11, 2024:69-93; Nairobi, Kenya. [doi: [10.1145/3700794.3700803](https://doi.org/10.1145/3700794.3700803)]
203. Antia SE, Ugwu CN, Ghodka V, Chori BS, Nazir MS, Odili CA, et al. Healthy Heart Assistant, a WhatsApp-based generative pretrained transformer technology, for self-care in hypertensive patients. *Mayo Clinic Proceedings: Digital Health*. Sep 2025;3(3):100243. [doi: [10.1016/j.mcpgdig.2025.100243](https://doi.org/10.1016/j.mcpgdig.2025.100243)]
204. Vowels LM, Vowels MJ, Sweeney SK, Hatch SG, Darwiche J. The efficacy, feasibility, and technical outcomes of a GPT-4o-based chatbot Amanda for relationship support: a randomized controlled trial. *PLOS Ment Health*. Sep 24, 2025;2(9):e0000411. [[FREE Full text](#)] [doi: [10.1371/journal.pmen.0000411](https://doi.org/10.1371/journal.pmen.0000411)] [Medline: [41662052](#)]
205. Yslado-Méndez R, Escobar-Agreda S, Villarreal-Zegarra D, Trejo Flores WM, Sánchez-Broncano JD, Vilela-Estrada AL, et al. Effectiveness, usability, and satisfaction of a self-administered digital intervention for reducing depression, anxiety, and stress in a university community in the Andean region of Peru: randomized controlled trial. *JMIR Form Res*. Oct 15, 2025;9:e71465-e71465. [[FREE Full text](#)] [doi: [10.2196/71465](https://doi.org/10.2196/71465)] [Medline: [41092084](#)]
206. Chanteclair A, Lartigau M, Salles N, Coelho J, Pécune F, Philip P, et al. Assessing the acceptability of a sleep-targeted digital intervention among geriatric inpatients: a preliminary study. *Digital Health*. Jan 29, 2025;11:1-10. [doi: [10.1177/20552076241293935](https://doi.org/10.1177/20552076241293935)]
207. Borsci S, Schmettow M, Malizia A, Chamberlain A, van der Velde F. A confirmatory factorial analysis of the Chatbot Usability Scale: a multilanguage validation. *Pers Ubiquit Comput*. Aug 04, 2022;27(2):317-330. [doi: [10.1007/s00779-022-01690-0](https://doi.org/10.1007/s00779-022-01690-0)]
208. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*. Jan 1995;7(1):57-78. [doi: [10.1080/10447319509526110](https://doi.org/10.1080/10447319509526110)]
209. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. *JMIR Mhealth Uhealth*. Apr 11, 2019;7(4):e11500. [[FREE Full text](#)] [doi: [10.2196/11500](https://doi.org/10.2196/11500)] [Medline: [30973342](#)]
210. Lewis J, Utesch B, Maher D. UMUX-LITE: when there's no time for the SUS. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY. Association for Computing Machinery; 2013. Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27 to May 2, 2013:2099-2102; Paris, France. [doi: [10.1145/2470654.2481287](https://doi.org/10.1145/2470654.2481287)]
211. Lund A. Measuring usability with the USE questionnaire. *Usability Interface*. 2001;8(2):3-6. [[FREE Full text](#)]
212. Rauschenberger M, Schrepp M, Perez-Cota M, Olschner S, Thomaschewski J. Efficient measurement of the user experience of interactive products. how to use the User Experience Questionnaire (UEQ). Example: Spanish language version. *IJIMAI*. 2013;2(1):39-45. [doi: [10.9781/ijimai.2013.215](https://doi.org/10.9781/ijimai.2013.215)]
213. Mitchell-Box K, Braun KL. Fathers' thoughts on breastfeeding and implications for a theory-based intervention. *J Obstet Gynecol Neonatal Nurs*. 2012;41(6):E41-E50. [doi: [10.1111/j.1552-6909.2012.01399.x](https://doi.org/10.1111/j.1552-6909.2012.01399.x)] [Medline: [22861175](#)]

214. Tariman JD, Berry DL, Halpenny B, Wolpin S, Schepp K. Validation and testing of the Acceptability E-scale for web-based patient-reported outcomes in cancer care. *Appl Nurs Res*. Feb 2011;24(1):53-58. [FREE Full text] [doi: [10.1016/j.apnr.2009.04.003](https://doi.org/10.1016/j.apnr.2009.04.003)] [Medline: [20974066](https://pubmed.ncbi.nlm.nih.gov/20974066/)]
215. Galavi Z, Montazeri M, Khajouei R. Which criteria are important in usability evaluation of mHealth applications: an umbrella review. *BMC Med Inform Decis Mak*. Nov 29, 2024;24(1):365. [FREE Full text] [doi: [10.1186/s12911-024-02738-2](https://doi.org/10.1186/s12911-024-02738-2)] [Medline: [39614283](https://pubmed.ncbi.nlm.nih.gov/39614283/)]
216. Griffin A, Xing Z, Khairat S, Wang Y, Bailey S, Arguello J, et al. Conversational agents for chronic disease self-management: a systematic review. In: *Proceedings of the AMIA Annual Symposium*. Washington, DC. AMIA; 2021. Presented at: AMIA Annual Symposium; October 30 to November 3, 2021:504-513; San Diego, CA.
217. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *Can J Psychiatry*. Jul 2019;64(7):456-464. [FREE Full text] [doi: [10.1177/0706743719828977](https://doi.org/10.1177/0706743719828977)] [Medline: [30897957](https://pubmed.ncbi.nlm.nih.gov/30897957/)]
218. Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. *JMIR Ment Health*. Oct 18, 2019;6(10):e14166. [FREE Full text] [doi: [10.2196/14166](https://doi.org/10.2196/14166)] [Medline: [31628789](https://pubmed.ncbi.nlm.nih.gov/31628789/)]
219. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. *J Med Internet Res*. Jul 13, 2020;22(7):e16021. [FREE Full text] [doi: [10.2196/16021](https://doi.org/10.2196/16021)] [Medline: [32673216](https://pubmed.ncbi.nlm.nih.gov/32673216/)]
220. Tudor Car L, Dhinakaran DA, Kyaw BM, Kowatsch T, Joty S, Theng Y, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res*. Aug 07, 2020;22(8):e17158. [FREE Full text] [doi: [10.2196/17158](https://doi.org/10.2196/17158)] [Medline: [32763886](https://pubmed.ncbi.nlm.nih.gov/32763886/)]
221. Singh B, Olds T, Brinsley J, Dumuid D, Virgara R, Matricciani L, et al. Systematic review and meta-analysis of the effectiveness of chatbots on lifestyle behaviours. *NPJ Digit Med*. Jun 23, 2023;6(1):118. [FREE Full text] [doi: [10.1038/s41746-023-00856-1](https://doi.org/10.1038/s41746-023-00856-1)] [Medline: [37353578](https://pubmed.ncbi.nlm.nih.gov/37353578/)]
222. Geoghegan L, Scarborough A, Wormald JCR, Harrison CJ, Collins D, Gardiner M, et al. Automated conversational agents for post-intervention follow-up: a systematic review. *BJS Open*. Jul 06, 2021;5(4):zrab070. [FREE Full text] [doi: [10.1093/bjsopen/zrab070](https://doi.org/10.1093/bjsopen/zrab070)] [Medline: [34323916](https://pubmed.ncbi.nlm.nih.gov/34323916/)]
223. Alruwaili MM, Shaban M, Elsayed Ramadan OM. Digital health interventions for promoting healthy aging: a systematic review of adoption patterns, efficacy, and user experience. *Sustainability*. Dec 02, 2023;15(23):16503. [doi: [10.3390/su152316503](https://doi.org/10.3390/su152316503)]
224. Liang M, Cui J, Fan X, Zhang J, Liu X, Liu D. The effect of digital health interventions in older adults with frailty: a systematic review and meta-analysis. *Int J Nurs Stud Adv*. Jun 2026;10:100470. [FREE Full text] [doi: [10.1016/j.ijnsa.2025.100470](https://doi.org/10.1016/j.ijnsa.2025.100470)] [Medline: [41551860](https://pubmed.ncbi.nlm.nih.gov/41551860/)]
225. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare. *Cureus*. Oct 2023;15(10):e46454. [FREE Full text] [doi: [10.7759/cureus.46454](https://doi.org/10.7759/cureus.46454)] [Medline: [37927664](https://pubmed.ncbi.nlm.nih.gov/37927664/)]
226. Rahamtalla BM, Medani IE, Abdelhag ME, Eltigani SA, Rajan SK, Falgy E, et al. The AI-powered healthcare ecosystem: bridging the chasm between technical validation and systemic integration—a systematic review. *Future Internet*. Nov 29, 2025;17(12):550. [doi: [10.3390/fi17120550](https://doi.org/10.3390/fi17120550)]
227. Artsi Y, Sorin V, Glicksberg BS, Korfiatis P, Nadkarni GN, Klang E. Large language models in real-world clinical workflows: a systematic review of applications and implementation. *Front Digit Health*. Sep 30, 2025;7:1659134. [FREE Full text] [doi: [10.3389/fdgh.2025.1659134](https://doi.org/10.3389/fdgh.2025.1659134)] [Medline: [41098649](https://pubmed.ncbi.nlm.nih.gov/41098649/)]
228. Lim PC, Lim YL, Rajah R, Zainal H. Usability questionnaire for standalone or interactive mobile health applications: a systematic review. *BMC Digit Health*. Apr 01, 2025;3(1):11. [doi: [10.1186/s44247-025-00150-y](https://doi.org/10.1186/s44247-025-00150-y)]
229. Sharma S, Kumar BA. A systematic review of user-based usability testing practices in self-care mHealth apps. *Digit Health*. Aug 29, 2025;11:20552076251374184. [FREE Full text] [doi: [10.1177/20552076251374184](https://doi.org/10.1177/20552076251374184)] [Medline: [40893169](https://pubmed.ncbi.nlm.nih.gov/40893169/)]
230. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. *BMC Med Inform Decis Mak*. Dec 02, 2022;22(1):317. [FREE Full text] [doi: [10.1186/s12911-022-02064-5](https://doi.org/10.1186/s12911-022-02064-5)] [Medline: [36461017](https://pubmed.ncbi.nlm.nih.gov/36461017/)]
231. Wechsung I, Weiss B, Kühnel C, Ehrenbrink P, Möller S. Development and validation of the conversational agents scale (CAS). In: *Proceedings of the INTERSPEECH*. Grenoble, France. ISCA; 2013. Presented at: INTERSPEECH; August 25-29, 2013:1106-1110; Lyon, France. [doi: [10.21437/interspeech.2013-298](https://doi.org/10.21437/interspeech.2013-298)]
232. Israfilzade K. Conversational marketing as a framework for interaction with the customer: development validation of the Conversational Agent's Usage Scale. *Journal of Life Economics*. Oct 31, 2021;8(4):533-546. [doi: [10.15637/jlecon.8.4.12](https://doi.org/10.15637/jlecon.8.4.12)]
233. Guerino G, Silva W, Coleti T, Valentim N. Assessing a technology for usability and user experience evaluation of conversational systems: an exploratory study. In: *Proceedings of the ICEIS 2021*. Setúbal, Portugal. Scitepress; 2021. Presented at: ICEIS 2021; Online:463-473; April 26-28, 2021. [doi: [10.5220/0010450204630473](https://doi.org/10.5220/0010450204630473)]
234. Faruk LID, Pal D, Funilkul S, Perumal T, Mongkolnam P. Introducing CASUX: a standardized scale for measuring the user experience of artificial intelligence based conversational agents. *International Journal of Human-Computer Interaction*. Jun 03, 2024;41(9):5274-5298. [doi: [10.1080/10447318.2024.2359206](https://doi.org/10.1080/10447318.2024.2359206)]

Abbreviations

AES: Acceptability E-Scale
CAUSS: Critical Assessment of Usability Studies Scale
CSUQ: Computer System Usability Questionnaire
CUQ: Chatbot Usability Questionnaire
HCA: health care conversational agent
MAUQ: mHealth App Usability Questionnaire
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSSUQ: Post-Study System Usability Questionnaire
SUS: System Usability Scale
UEQ: User Experience Questionnaire
UEQ-S: User Experience Questionnaire—Short Version
UMUX-Lite: Usability Metric for User Experience—Lite Version
USE: Usefulness, Satisfaction, and Ease of Use Questionnaire

Edited by A Parush; submitted 29.May.2026; peer-reviewed by M Chatzimina, L Kettle; comments to author 30.Jun.2026; revised version received 18.Jul.2026; accepted 20.Jul.2026; published 30.Jul.2026

Please cite as:

Pavão J, Bastardo R, Silva AG, Rocha NP

Usability and User Experience Assessment of Health Care Conversational Agents Using Validated Subjective Instruments: Systematic Review and Comparative Analysis

JMIR Hum Factors 2026;13:e102911

URL: <https://humanfactors.jmir.org/2026/1/e102911>

doi: [10.2196/102911](https://doi.org/10.2196/102911)

PMID: [42530974](https://pubmed.ncbi.nlm.nih.gov/42530974/)

©João Pavão, Rute Bastardo, Anabela Gonçalves Silva, Nelson Pacheco Rocha. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 30.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.