Received: 19 May 2026 | Revised: 23 September 2026 | Accepted: 16 November 2026 | Published online:27 November 2026

RESEARCH ARTICLE

Medinformatics
2026, Vol. 00(00) 1–9
DOI: 10.47852/bonviewMEDIN620210454

BON VIEW PUBLISHING

Conversational AI Agents in Physical Rehabilitation: A Scoping Review with Implications for Physiotherapy Practice in Greece

Savina Mariettou[1], Sam Shahbandi[1,2], Vasilis Triantafyllou[1] and Constantinos Koutsojannis[2,*]

[1] *Electrical and Computer Engineering Department, University of Peloponnese, Greece*

[2] *Department of Physiotherapy, University of Patras, Greece*

Abstract: Conversational artificial intelligence (AI), ranging from rule-based chatbots to agents built on large language models (LLMs), is increasingly being applied in physical rehabilitation, yet little synthesized evidence exists on how these systems are used clinically and what their adoption would require within specific national health systems such as that of Greece. This scoping review, conducted according to the PRISMA-ScR guidelines, mapped the architectures, interaction modalities, clinical use cases, and reported the effectiveness of conversational AI in physical rehabilitation and considered the relevance of these findings for Greek physiotherapy practice. PubMed, IEEE Xplore, ScienceDirect, arXiv, and Google Scholar were searched for studies published between 2010 and 2026 that evaluated conversational AI in physical therapy or musculoskeletal care. Nine primary empirical studies were included, while national policy documents and health regulations were reviewed separately to assess feasibility in the Greek context. The included studies examined embodied virtual coaches, mobile text chatbots, LLM-based advisors, and robot-integrated conversational systems, which were used mainly for exercise coaching, postoperative follow-up, adherence monitoring, and patient reassurance. Feasibility, user satisfaction, and short-term engagement were consistently high, but rigorous long-term trials of clinical efficacy were scarce. In Greece, deployment would require adaptation to the Greek language, professional governance, and alignment with EOPYY reimbursement structures and the EU AI Act. Conversational AI therefore appears to be a promising adjunct to physical rehabilitation, whose responsible adoption should proceed in phases with linguistic localization and clear regulatory safeguards.

Keywords: conversational artificial intelligence, physiotherapy, rehabilitation, large language models, digital health

1. Introduction

Digital health technologies have become increasingly integrated into physiotherapy and rehabilitation practice, driven by rising chronic disease prevalence, workforce constraints, and growing demand for patient-centered care. Tele-rehabilitation platforms, wearable sensors, and mobile health applications have expanded access to rehabilitation services while supporting self-management and continuity of care. Within this evolving landscape, artificial intelligence (AI) has emerged as a key enabler of personalized, scalable, and data-driven rehabilitation interventions [1].

Physiotherapy is particularly well suited to digital augmentation due to its emphasis on education, exercise prescription, behavioral adherence, and long-term engagement. However, many existing digital tools rely on static content or limited interactivity, which may constrain patient motivation and responsiveness [2]. Conversational AI agents, systems capable of engaging users through natural language dialogue, offer a novel approach to addressing these limitations by providing interactive, adaptive, and context-aware support [23].

Conversational AI encompasses a range of technologies, including rule-based chatbots, natural language understanding systems, and, more recently, large language models (LLMs). Inception: Early health care chatbots used fixed rule-based decision trees and canned responses, restricting their scalability. With developments in deep learning and transformer-based models [3], LLMs can now create relevant, human-like responses that are appropriate to the context.

The use of LLM-based conversational agents in healthcare has been investigated for clinical documentation, patient education, triage, and clinical decision support [24, 25]. Rehabilitation uses of conversational agents can include exercise teaching, symptom assessment, motivational coaching, and explaining treatment strategies. Voice-enabled conversational agents can also help accommodate people who may have greater difficulty using screens, such as older people and those with movement or visual impairments.

***Correspondence:** Constantinos Koutsojannis, Department of Physiotherapy, University of Patras, Greece. Email: ckoutsog@upatras.gr

© The Author(s) 2026. Published by BON VIEW PUBLISHING PTE. LTD. This is an open access article under the CC BY License (https://creativecommons.org/licenses/by/4.0/).

However, issues exist for response accuracy, hallucinations, privacy, and safety. In rehabilitation, inaccurate information can result in injury and dysfunctional learning, prolonging recovery, and increasing cost. Responsible assessment and management are therefore necessary, before widespread adoption in clinical practice [4, 5].

These challenges manifest in a certain way in Greece, where the situation for the implementation of conversational AI in physiotherapy looks different, given the unique Greek setting. The Greek health system is a mixture of public and private provision, and physiotherapists constitute fundamental health professionals for musculoskeletal and neurorehabilitative therapies. Types of the linguistic particularities, cultural demands for clinician-patient relationship, and legal duties according to the European Union law affect the capacity and acceptability of AI-based interventions [5, 6].

Until now, most conversational AI studies have been carried out within the context of the English-speaking or northern European countries, with the localization or Greek language performance and congruency with the national rehabilitation pathway still not receiving much attention. Thus, a targeted synthesis is needed to advance future research and implementation in Greek physiotherapy [7–10].

Therefore, this scoping review sought to systematically explore the available literature on conversational AI agents that have been implemented in the field of physical rehabilitation and physiotherapy, to describe the system design features and domains of clinical application, and the gaps, ethical issues, and research priorities that are most relevant to physiotherapy practice in Greece.

The rest of this paper is structured as follows. Section 2 details the review methods: the reasons for the choice of the scoping review, the screening criteria, the databases and search terms, and the study selection and data charting procedures. Section 3 reports the results, beginning with the screening of the studies and their general features, then the conversational agent architectures and interaction modes, the clinical purposes, the evidence on their efficacy and safety, and the implications for Greek physiotherapy. Section 4 contextualizes our findings within the broader literature on health chatbots and LLMs, discusses the organizational, linguistic, and regulatory conditions for adoption in the Greek healthcare system, and the limitations of this review. Finally, Section 5 distills the takeaways and lays out the challenges for research and implementation.

2. Methods

2.1. Study design

The scoping review design was selected because the publication of the literature regarding conversational AI in the context of rehabilitation has been recent, is heterogeneous, and spans clinical, engineering, and computer science literatures. The objective was to characterize the types of systems, applications, and outcomes reported, as well as to identify the relevant knowledge gaps in the literature but not to estimate a pooled effect across studies, which was not feasible among the heterogeneous designs. Three questions guided the review: (1) which types of conversational AI agents have been developed and/or evaluated in the context of physical rehabilitation; (2) for which purposes of care and for which patient populations they have been used and with what reported outcomes; and (3) which technical, ethical, and organizational factors may influence their implementation in Greek physiotherapy. The population, concept, and context of the review were determined prior to the search and were used to establish the eligibility criteria, which are reported in section 2.2. The review protocol was not registered.

2.2. Eligibility criteria

Studies: The inclusion criteria were as follows:

Population: People undergoing physical rehab, physiotherapy, or musculoskeletal rehab, or healthy cohorts testing interface prototypes.

Intervention/Concept: Conversational AI agents used in digital health (e.g., rule-based chatbots, embodied conversational agents, voice assistants, and LLM-based advisory systems).

Outcomes: System design specifications, technical usability, effectiveness in a clinical setting, engagement, adherence, or feasibility measures.

Peer-reviewed primary clinical trials, prospective/retrospective cohort studies, usability studies, technical system evaluation studies, and preprints.

Exclusion criteria comprised:

Non-interactive AIs (such as computer vision apps for posture tracking, not conversational dialogue).

Theoretical review articles, commentaries, or conceptual publications without an empirical component or usability assessment.

Non-rehabilitation applications.

2.3. Information sources and search strategy

The systematic search was performed on 15th January 2026, searching five major online bibliographic databases: PubMed, IEEE Xplore, Science Direct, arXiv, and Google Scholar. The published literature was limited to between January 1 st 2010 and 15 th January 2026, to include dialogue systems from the early days of rule-based systems to current implementations of LLMs. Both Greek and English language texts were included in the search.

Search strings consisted of a combination of Medical Subject Headings (MeSH) and free-text words across Title, abstract and metadata: ("conversational AI" OR "chatbot" OR "large language model" OR "virtual agent" OR "embodied agent") AND ("physical therapy" OR "physiotherapy" OR "rehabilitation" OR "musculoskeletal care").

Database searches resulted in 259 total identified articles: PubMed (n = 35), Google Scholar (n = 64), IEEE Xplore (n = 50), Science Direct/MDPI (n = 60), and arXiv (n = 50). Additional citation chaining (Google Scholar and SciSpace) and searches of gray literature were done post hoc to identify newly available articles. Duplicates were removed using EndNote X9 and then manually checked, leaving 114 articles for title and abstract screening. SM and SS independently screened titles and abstracts for eligibility, followed by full-text screening, with discrepancies resolved through group discussion with a senior author (CK).

2.4. Study selection, data charting, and synthesis

Records were imported into EndNote X9 for automated deduplication, followed by manual review. Two independent reviewers (SM and SS) screened titles and abstracts against inclusion criteria. Potentially eligible full-text articles were independently assessed by both reviewers. Disagreements were resolved through discussion or consultation with a senior investigator (CK).

A dual-pathway synthesis was applied:

Primary Scoping Synthesis: Extracted characteristics from included empirical studies (n = 9), mapping country, design, sample, system architecture, clinical application, key findings, and limitations.

Contextual Discussion Synthesis: Analyzed Greek health system literature, EOPYY benefit frameworks, and EU regulations to evaluate implementation barriers and workforce implications.

In accordance with PRISMA-ScR guidelines for scoping reviews, critical appraisal of individual source evidence was omitted due to the heterogeneous mix of technical prototypes and feasibility trials.

The table entries map agent classes to concrete studies, so readers can locate examples for each type [12–18].

3. Results

Findings are organized into four sections. A description of how the studies were selected (Section 3.1) is followed by characterization of the included studies according to system design (Section 3.2) and type of agent (Section 3.3), then according to their intended clinical use (Section 3.4). The evidence relating to efficacy and safety is then summarized (Section 3.5), before the implications for physiotherapy practice in Greece are considered, with a brief summary of the evidence (Section 3.6).

3.1. Study selection and empirical synthesis

The initial search of databases yielded 259 articles: PubMed (n = 35), Google Scholar (n = 64), IEEE Xplore (n = 50), ScienceDirect and MDPI (n = 60), and arXiv (n = 50). Duplicates (n = 145) were subsequently excluded, leaving 114 unique titles and abstracts to be screened. Of these, 95 were removed because they did not fit the scope of the review: 58 were outside the scope, 22 did not describe conversational systems, and in 15 the application was not relevant to rehabilitation. Of the 19 full papers assessed, 10 were removed: 3 because they did not involve an interactive system, 4 because the system had not been evaluated empirically, and 3 because rehabilitation was not the target. Nine primary studies met the inclusion criteria and were included in the synthesis (Figure 1).

3.2. Design characteristics of conversational AI agents

The systems used were very different in design. Agents used earlier and in clinical settings relied on dialogue trees and handwritten scripts, covering every imaginable question and response in detail. This was the system used for the postoperative chatbots after hip arthroscopy [13] and total knee replacement [14], for example. They are predictable and relatively easy to validate, because all messages can be checked by clinicians before use; however, they are not capable of responding well to questions outside the script. The later systems used language models of varying sizes, and the LLM-based physiotherapy advisor, which used its language model and a retrieval component to query curated sources, to avoid unsupported answers [17]. This reduces, but does not eliminate, hallucinated information and means the system must be thoroughly evaluated before clinical deployment. Other systems combined behavioral theory with dialogue, such as the Self-Determination Theory-based chatbot for Parkinson's disease [19], or used reinforcement learning for coaching feedback [20].

In addition, the systems varied in user interaction. The systems were either text-based chatbots using SMS, instant messaging, or smartphone apps [13–15], or embodied agent systems that integrated an animated avatar with speech and text [12, 16]. Others had an associated mobile sensor, wearable, or robotic platform so that the agent could observe the exercise performance and provide feedback [18].

Interaction using voice proved to be of particular interest since it allows patients to keep their hands free while exercising and could help make the system easier to operate for both older adults and those with visual and/or motor impairments. However, speech recognition can prove more difficult to implement well for these populations, for people with speech impairments following a stroke, or in noisy domestic settings, issues which were not tested very often (unless potentially implicitly through exercise performance). Multilingual support was minimal: the systems tested supported English, German, Spanish, or Portuguese languages, but not Greek. This is a major consideration for Greek users, since both speech recognition and language understanding require available training data in the relevant language.

3.2.1. Types of agents

Figure 1 can classify these agents based on their forms and interaction modes. These classes are embodied virtual coaches, text chatbots and SMS-based agents, mobile-based conversational therapy apps, embodied cognitive rehabilitation conversational agents, LLM-based physiotherapy consultation system, and robotic-based AI coaches (Table 1). The forms of interactions, typical use cases, and representative papers for each class can be seen in Table 1. Although the six classes are not mutually exclusive, there are systems that have capabilities in several classes. For example, the virtual coach (i.e., system for measuring the user's physical activity by a wearable device and summarizing this information into understandable motivations) may incorporate behavioral dialogue models; a robot-based coach also includes a conversational interface and motion sensors; and several systems cover multiple interaction modes.

3.3. Clinical applications in physical rehabilitation

In addition to exercise provision, conversational agents were used at various points in the rehabilitation pathway, encompassing prehabilitation, postoperative recovery, discharge planning, and long-term follow-up. In all the studies included, applications were grouped according to major functional categories (Table 2):

Home exercise compliance: Agents provide cues, instruct exercises, and monitor completion to maintain engagement outside clinics [12, 14].

Postoperative follow-up and reassurance automated SMS/chat agents are able to assist with postprocedural questions and reduce clinic contacts by escalating if needed [13].

Motivation, behavior change: Motive-Centered Designs that use motivational theories (e.g., Self-determination Theory) are aimed at fostering autonomy, competence, and relatedness [16, 19].

Cognitive and motor rehabilitation: Conversational programs offer cognitive structured activities and flexible series of exercises for people after stroke or brain damage [15, 16].

First assessment and onward referral: Advisors based on LLMs can provide symptom analysis, recommend exercises, and give realistic advice on managing issues [17].

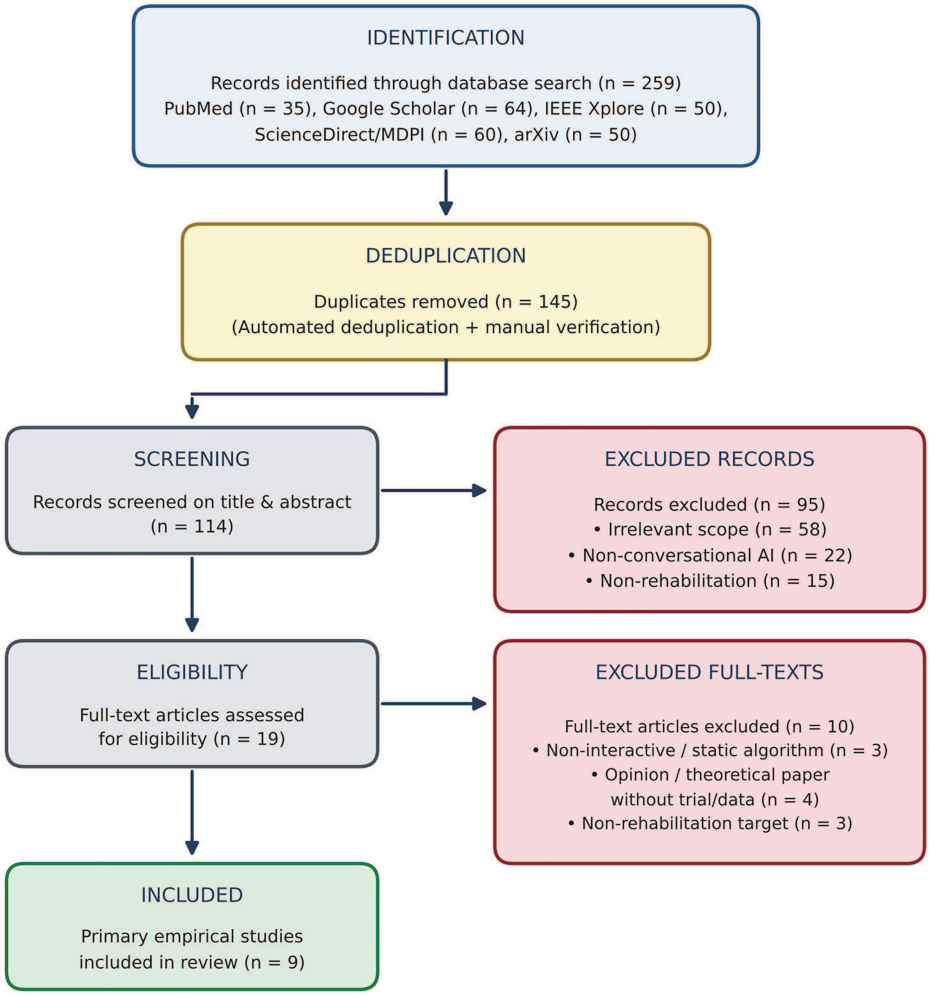

Figure 1. PRISMA-ScR Flow Diagram

Table 1. Types of Conversational AI Agents in Rehabilitation

Agent class	Form and interaction mode	Typical use cases	Example study
EMBODIED VIRTUAL COACH	Avatar or animated agent with voice and/or text, multimodal interaction	Home exercise guidance, motivation, long-term adherence	Virtual coaching for older adults with QoL gains [12]
TEXT CHATBOTS AND SMS AGENTS	Rule-based or ML chat flows delivered by SMS, mobile app, or web chat	Postoperative follow-up, appointment triage, adherence prompts	Hip arthroscopy SMS agent "Felix" [13]; knee replacement chatbot feasibility [14]
MOBILE CONVERSATIONAL THERAPY APPS	Multimodal app with conversational UI for cognitive and motor exercises	Stroke cognitive exercises, tailored therapy routines	Mobile stroke rehabilitation conversational app [15]
EMBODIED CONVERSATIONAL AGENTS FOR COGNITIVE REHAB	Dialogue systems tailored to cognitive impairments with personalization	Brain injury cognitive and motivational support	Motivational ECA for brain injury rehabilitation [16]
LLM-BASED PHYSIOTHERAPY ADVISORS	Large language models with knowledge grounding and retrieval	Initial triage, exercise recommendations, citing sources	Physio LLM-based advisor prototype [17]
ROBOT-INTEGRATED AI COACHES	Conversational front end combined with sensors/robotics and automated monitoring	Real-time exercise guidance and performance feedback	Design and stakeholder evaluation of AI/robotic coaches [18]

Table 2. Summary of Included Studies in Conversational AI for Physical Rehabilitation

Study	Country	Study Design	Population & Sample Size (n)	Agent Architecture & Modality	Clinical Application	Main Findings	Limitations
Schlieter et al. (2024)	Germany	Feasibility / Pilot Cohort	Older adults in home rehab ($n = 24$)	Avatar-based Virtual Coach (Voice/Text)	Home exercise coaching & adherence	>10% QoL increase in experimental group; high user engagement (>4.0/5.0).	Small sample size; short follow-up window.
Dwyer et al. (2023)	USA	Prospective Cohort	Post-hip arthroscopy patients ($n = 26$)	Automated SMS Chatbot ("Felix")	Postoperative follow-up & triage	80% rated helpfulness as good/excellent; 79% patient queries resolved without clinician escalation.	Single-center design; non-randomized.
Blasco et al. (2022)	Spain	Feasibility Trial	Post-total knee replacement ($n = 30$)	Rule-based Mobile Chatbot	Post-TKR home rehabilitation	High system usability; significantly higher home exercise adherence vs. control.	Unblinded pilot study.
Griol & Callejas (2019)	Spain	Technical & Usability	Post-stroke rehabilitation cohort ($n = 12$)	Mobile Conversational Agent	Cognitive & motor therapy exercises	System achieved 82% dialogue accuracy; highly adaptable interface for disabled users.	Small sample size; early functional prototype.
Hocking & Maeder (2021)	Australia	Qualitative / Design Study	Traumatic brain injury cohort ($n = 15$)	Embodied Conversational Agent (ECA)	Motivational cognitive support	Improved intrinsic motivation scores using Self-Determination Theory framework.	Lack of clinical control group.
Almeida et al. (2024)	Portugal	Prototype Evaluation	Clinicians & healthy subjects ($n = 18$)	LLM + RAG Knowledge Grounding ("Physio")	Triage & exercise advice	4.5/5.0 expert safety/relevance score; RAG architecture eliminated safety-critical hallucinations.	Evaluated on simulated clinical vignettes rather than live patients.
Lee et al. (2021)	USA	Iterative Design	Therapists ($n = 4$) & post-stroke survivors ($n = 5$)	Robot-Integrated AI Conversational Coach	Real-time movement coaching	Strong stakeholder acceptance; demonstrated feasibility of voice-guided physical rehab.	Restricted to laboratory environment.
Macedo et al. (2024)	Portugal	Technical Validation	Parkinson's disease cohort ($n = 20$)	SDT-driven Chatbot ("PANDORA")	Exercise empowerment & motivation	100% domain coverage in technical tests; high user satisfaction ($4.2 - 4.9/5.0$).	Short exposure period.

(*Continued*)

Table 2. (*Continued*)

Study	Country	Study Design	Population & Sample Size (*n*)	Agent Architecture & Modality	Clinical Application	Main Findings	Limitations
Siva kumar (2024)	Online / Global	Experimental Technical	Musculoskeletal rehabilitation cohort ($n = 25$)	Conversational AI + Reinforcement Learning	Adaptive exercise progression	Reinforcement learning dialogue adaptation increased exercise completion by 18%.	Heterogeneous cohort; lack of clinical endpoint validation.

Therapist support and monitoring: Sensor-linked and robot-integrated devices combine automated observation with conversational exchange to deliver immediate feedback [18].

Together, the functional use cases suggest that agents are supportive adjuncts that augment and sustain rehabilitation in the outside world. Agents, therefore, are not intended to be central clinical decision-makers [12–20].

3.4. Evidence of effectiveness and safety

Evidence supporting clinical effectiveness remains mixed and methodologically heterogeneous. Many studies remain small-scale feasibility evaluations, and robust long-term randomized data across conditions are lacking. Few investigations employed adequately powered randomized controlled designs, and outcome measures varied substantially across studies.

Empirical findings can be grouped into several thematic categories.

Quality of life and engagement: A multistage study of a full-featured virtual coach for older patients reported engagement above UX thresholds and QoL improvements of >10% between experimental and control groups [12].

Technical and usability evaluation: A Portuguese Parkinson's co-learning bot (PANDORA) had 100% domain coverage and coherence in technical evaluation and was rated 4.2 – 4.9/5 on user satisfaction in a small user study which may reflect greater engagement, but impact on long-term exercise participation is currently unknown [19].

Postoperative feasibility and satisfaction: SMS chatbot for patients with hip arthroscopy; 80% of users rated helpfulness as good/excellent, and 79% of patient questions were independently managed in a prospective cohort for 6 weeks [13].

An implementation and feasibility study of a chatbot for home management of patients after total knee replacement showed that it is feasible to support adherence monitoring and automated communication in the postoperative period, with a larger study currently in progress [14].

Apps for cognition: Mobile conversational systems for stroke rehabilitation were developed to customize cognitive exercises for individuals and accessibility issues, with prototypes shown to be feasible for seniors and disabled users [15].

Stakeholder acceptance/concerns: Therapists and post-stroke patients welcomed possible gains for systematic management and to support motivation, raising practical issues, such as interaction for users with cognitive impairment and system breakdown [16, 18].

Safety issues were inconsistently addressed. While the risks of errors arising from misguided advice, overreliance on automated solutions, and data privacy were recognized, the issues were seldom empirically tested and followed by a mitigation plan.

Overall, the present indication looks encouraging in terms of feasibility and engagement, but the literature is still quite experimental, focused on pilot-scale and condition-specific work, and on usability rather than on large-scale high-quality randomized clinical trials [12–19]. This area of work seems to be coming of age but not yet mature as a clinical science [12–19].

3.5. Implications for Greek physiotherapy practice

For Greek health care, the results point to an opportunity and a challenge. Conversational AI agents could contribute to maintaining service quality despite workforce shortages and geographical differences, especially when it comes to providing follow-up to rehabilitation outside of urban clinical services. To be viable, implementation would need to consider Greek language conventions, professional standards, and European legislation such as GDPR and the AI Act.

Interestingly, the data processing techniques and technological architectures behind those systems are characterized by different degrees of complexity and clinical risk to ensure their safe integration into Greek physiotherapy setting. Overall, the conversational agents for rehabilitation in all the scientific sources incorporated both classic dialogue management devices and modern AI reasoning, sensing, and behavioral modeling.

It is possible to discern several key technological pathways:

Rule-based dialogues and finite state managers: Many of the chatbots deployed rely on predefined scripts or rule-based flows of more predictable follow-up and information interactions to guarantee safety and reliability [13, 14].

Motivational and behavioral models: The nature of conversation is generally influences by behavior change models; PANDORA had been set up with Self-Determination Theory in mind [19].

Embodiment and multimodality: Virtual coaches and embodied ECAs incorporate avatars, speech, and gestures to improve engagement and adapt to sensory or motor disabilities [12, 15, 16].

LLMs with grounding and retrieval: LLM-based advisors are already being prototyped with external knowledge conditioning and citation mechanisms to mitigate hallucinations and suggest OTC symptom relief exercises [17].

Sensor integration and robotics: Systems use conversational interfaces with motion sensors or robotic devices to verify exercising and give interactive feedback [18].

Adaptive policies and reinforcement learning: Studies in this area have suggested reinforcement learning methods to adapt policies of coaching and dialogue [20].

Within the existing literature, there are various examples hybridization that utilize LLMs, such as multimodal virtual coaches consisting of behavioral dialogue models [12], motivational SDT-driven chatbots for Parkinson's exercise [19], and LLM-based advisors conditioned on source they can verify [17].

Overall, the technological considerations point to the need for a phased and monitored implementation in Greece, whereby the initial introduction involves rule-based and low-risk applications and the more flexible, adaptive AI-based systems are introduced only once proper controls, validations, and safeguards have been implemented and adequate oversight has been ensured.

3.6. Summary of evidence

In brief, conversational agents for physical rehabilitation span embodied virtual coaches, mobile chatbots, LLM virtual coaches, and connected sensor/robot systems. Most of the nine studies included in our review reported that the systems were acceptable and easy to use and showed a high level of engagement, and one study [12] demonstrated quality of life improvements. Few of the trials were long term or controlled, safety has not been tested directly, and functional outcomes including mobility or pain are not commonly assessed. In conclusion, the evidence supports further research and pilot studies, not routine clinical use.

4. Discussion

4.1. Empirical evidence base vs. contextual implementation

This scoping review finds a dynamic field with a growing body of literature on conversational AI agents for physical rehabilitation. It incorporates nine studies using embodied virtual coaches [12, 16], text chatbots [13, 14, 19], LLM advisors [17], and robot connected systems [18]. While quality evidence supports feasibility, user satisfaction, and short-term engagement across the different patient cohorts examined [12–15, 19], clinical trial quality evidence has been limited [12, 14].

Empirical evidence: there is a lack of clinical evidence. These systems are still in their infancy with regard to clinical readiness [12–20]. Most studies were pilot feasibility or technical validation studies [15, 17, 19], indicating the need for large-scale randomized controlled trials prior to implementation [12, 14].

These results are in line with previous work on health chatbots. A systematic review of AI chatbots in health behavior change concluded that "the vast majority of interventions appeared to be acceptable and showed promising results but that the study designs were weak and follow-up periods were short" [23]. LLMs raise additional issues. Participants rated the model's answer as more empathetic and clearer than the doctors' in some cases [25]. As incorrect loads or instructions about pain could prevent a full recovery or cause an injury in physiotherapy, these models should be designed for use as aids to and under the supervision of clinicians, not as autonomous providers of advice.

4.2. Contextual analysis for the Greek healthcare system

As no primary research was carried out in Greek healthcare contexts, the desirability of a conversational AI solution for Greek physiotherapy can only be evaluated from policy and practice perspective:

Health system & reimbursement systems: In Greece, the public health system is organized through the National Organization for the Provision of Health Services (EOPYY). EOPYY reimburses structured physiotherapy sessions, according to the Unified Health Benefits Regulation. However, the regulation does not include digital therapeutics/automated conversational follow-up [6]. In order to prescribe AI-based rehabilitation software in Greece, a route for regulatory approval of this software, equivalent to that used in DiGA (Digital health applications – digital health applications) in Germany, would be required.

Linguistic Localization & Cultural Nuance: Existing commercial LLMs and conversational agents are trained in English. The application of such a tool in Greece would necessitate exhaustive localization for the Greek language, medical/clinical language, speech, and semantics of Greek patients. Greek is also much less represented than English in the data used to develop and evaluate these models, so errors in terminology or in the interpretation of patients' descriptions may be more frequent, and their performance in Greek should be tested directly rather than assumed.

Machines that are used to automate the delivery of triage or exercise instructions as in this example are classified as medical software and regulated under EU Regulation 2017/745 (MDR) and the EU AI Act (Regulation 2024/1689). This is likely to be high risk, requiring explanation of the algorithms, GDPR compliance, and deployment of monitoring of safety (including harms by false or fabricated advice) [4, 5].

Workforce integration: Greek physiotherapists have demonstrated a growing interest towards digital health, especially since the broadening of telerehabilitation during the pandemic [22]. Digital education and national guidelines should be in place in order to avoid "overload" of clinicians and to provide proper supervision.

Taken together, these four factors support the conclusion that the primary issues in Greece are organizational and legal, not technical, and therefore incremental application is the most viable route. Pilot trials of rule-based agents in public and private rehab centers could be a starting point, eventually testing them on low-stakes interactions in Greek, such as exercise assistants or more simple post-surgical queries. This would provide the first Greek data on usability, compliance, and strain for physiotherapists, and the capabilities of LLMs in Greek would need to be well-surveyed beforehand for the agent to be put to use with patients.

Training is also important, because physiotherapists will need the skills to evaluate an agent's treatment recommendation, and to show their patients how to use it. Survey results show that Greek physiotherapists are generally positive about telerehabilitation but also reveal major practical hurdles [22]. Lastly, sustainability will require funding. If there is no reimbursement mechanism through EOPYY, digital tools will probably exist in research projects, or in private practice, which risks increasing the urban-rural divide. A national system for the assessment and reimbursement of digital rehabilitation tools will make sure that patients outside big cities will also benefit.

4.3. Limitations

There are several limitations to this review that should be borne in mind when interpreting its findings. Firstly, the number of studies included was small (n = 9) and study designs were heterogeneous, ranging from technical prototype evaluations and testing to prospective cohorts and feasibility studies. This heterogeneity precluded meta-analysis or determination of effect sizes,

so the results should be considered as a map of existing evidence rather than evidence of effectiveness.

Secondly, there was no formal critical appraisal of studies, following the PRISMA-ScR guidance. Consequently, lower- and higher-quality studies were given equal consideration in the analysis. Several articles were preprints or conference proceedings that had not been peer-reviewed in their entirety, which could impact the trustworthiness of the results.

Third, despite a comprehensive search of five databases and the inclusion of published work in either English or Greek, it is possible that relevant work was not identified. Work published in other languages, consumer company developers' internal evaluations, and unpublished negative findings are unlikely to have been found, and publication bias may overestimate feasibility and user satisfaction.

The search was finalized on 26 January 2026, and more recent systems and evaluations that became available after this date are not covered here. Some of the technical limitations cited in the included studies have perhaps been partly addressed by more recent models.

In the last place, we did not find any Greek primary studies. Therefore, the discussion on the Greek context is based on policy documents, regulatory documents and survey data regarding the attitudes of physiotherapists towards the use of telerehabilitation [22], rather than based on the evaluation of the way conversational agents are used in the Greek clinical practice. Its conclusions should be considered as a starting point for future research.

5. Conclusions

The aim of this scoping review was to summarize the applications of conversational AI agents in physical rehabilitation and to explore its implications for physiotherapy practice in Greece. The nine included studies demonstrate a field that is evolving rapidly but remains in its infancy. The set of identified systems included rule-based text chatbots following hip [13] and knee surgeries [14], embodied virtual coaches [12, 16], LLM-based autonomous advisors [17], and robot-assisted coaching platforms [18]. Overall, most of these systems were well received by both patients and clinicians, with some reporting high usability and engagement, although evidence for clinical benefit remains limited. Participant numbers were often small, and follow-up duration short, with only some employing a control group, making it premature to determine whether function improves in the long term.

Many more steps must be taken before conversational agents can be incorporated into routine rehabilitation. Large randomized trials with common outcome measures (such as validated scales or objective adherence data) are needed, so that comparisons can be made between different interventions. Safety must be incorporated into the design of any agent: it should be able to monitor for red flags and seek help from a clinician if needed, and it should not give advice on topics outside its area of expertise. This is a particular concern for LLM-based systems where hallucination and ideation are big concerns [7]. Ultimately, such tools must augment existing therapy rather than supplant it [21], and they should be incorporated into a clear, professionally supported care pathway.

For Greece, the limitations of the language itself make this advancement crucially different. No reimbursement pathway exists for digital rehabilitation devices within EOPYY, and the support for the Greek language is too limited, and physicians and other clinicians lack training in digital health [22]. The way forward could be first by starting with relatively low-risk applications such as exercise prompts, home program guidance, and standardized postoperative questions using rule-based systems and pilot demonstrations in a Greek clinical setting. More advanced LLM-based interventions could then follow on once their Greek language accuracy is benchmarked and the requirements of the EU AI Act and the Medical Device Regulation are clarified [5]. In these conditions, conversational AI could serve as an adjunct to physiotherapy in Greece, especially for those living in remote areas or requiring structured support in between physiotherapy sessions.

Ethical Statement

This study is a scoping review of previously published literature and publicly available policy documents. It does not contain any studies with human or animal subjects performed by any of the authors, and no new human or animal data were collected. Therefore, ethical approval and informed consent were not required. All findings were derived from published studies, which reported their own ethical approvals where applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Savina Mariettou: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Sam Shahbandi:** Methodology, Investigation, Data curation, Formal analysis, Writing – review & editing. **Vasilis Triantafyllou:** Formal analysis, Validation, Writing – review & editing. **Constantinos Koutsojannis:** Conceptualization, Validation, Supervision, Project administration, Writing – review & editing.

References

[1] Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., … Albekairy, A. M. (2023). Revolutionizing healthcare: The role of artificial intelligence in clinical practice. *BMC Medical Education*, *23*, 689. https://doi.org/10.1186/s12909-023-04698-z

[2] Seron, P., Oliveros, M.-J., Gutierrez-Arias, R., Fuentes-Aspe, R., Torres-Castro, R. C., Merino-Osorio, C., … Sanchez, P. (2021). Effectiveness of telerehabilitation in physical therapy: A rapid overview. *Physical Therapy*, *101*(6), pzab053. https://doi.org/10.1093/ptj/pzab053

[3] Clusmann, J., Kolbinger, F. R., Muti, H. S., Carrero, Z. I., Eckardt, J.-N., Ghaffari Laleh, N., … Kather, J. N. (2023). The future landscape of large language models in medicine. *Communications Medicine*, *3*, 141. https://doi.org/10.1038/s43856-023-00370-1

[4] Gilbert, S., Harvey, H., Melvin, T., Vollebregt, E., & Wicks, P. (2023). Large language model AI chatbots require approval as

medical devices. *Nature Medicine*, *29*(10), 2396–2398. https://doi.org/10.1038/s41591-023-02412-6

[5] European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. L 2024/1689

[6] Ministry of Health of the Hellenic Republic. (2018). *Unified Health Benefits Regulation (EKPY) of the National Organization for the Provision of Health Services (EOPYY)* (Government Gazette FEK 4898/B/2018)

[7] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., … Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571730

[8] Meng, X., Yan, X., Zhang, K., Liu, D., Cui, X., Yang, Y., … Tang, Y.-D. (2024). The application of large language models in medicine: A scoping review. *iScience*, *27*(5), 109713. https://doi.org/10.1016/j.isci.2024.109713

[9] Meskó, B., & Topol, E. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digital Medicine*, *6*, 120. https://doi.org/10.1038/s41746-023-00873-0

[10] Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature Medicine*, *29*(8), 1930–1940. https://doi.org/10.1038/s41591-023-02448-8

[11] Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

[12] Schlieter, H., Gand, K., Weimann, T. G., Caprino, M., Corbo, M., Seregni, A., … Busnatu, S. (2024). Virtual coaching for home rehabilitation: Evidence of an empirical study. *Studies in Health Technology and Informatics*, *316*, 147–151. https://doi.org/10.3233/SHTI240365

[13] Dwyer, T., Hoit, G., Burns, D., Higgins, J., Chang, J., Whelan, D., … Chahal, J. (2023). Use of an artificial intelligence conversational agent (chatbot) for hip arthroscopy patients following surgery. *Arthroscopy, Sports Medicine, and Rehabilitation*, *5*(2), e495–e505. https://doi.org/10.1016/j.asmr.2023.01.020

[14] Blasco, J.-M., Roig-Casasús, S., Igual-Camacho, C., Díaz-Díaz, B., & Pérez-Maletzki, J. (2022). Conversational chatbot to promote adherence to rehabilitation after total knee replacement: Implementation and feasibility. *Archives of Physical Medicine and Rehabilitation*, *103*(12), e125. https://doi.org/10.1016/j.apmr.2022.08.764

[15] Griol, D., & Callejas, Z. (2019). Mobile conversational agents for stroke rehabilitation therapy. In *Proceedings of the IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)* (pp. 513–518). https://doi.org/10.1109/CBMS.2019.00104

[16] Hocking, J., & Maeder, A. (2021). Motivational embodied conversational agent for brain injury rehabilitation. *Studies in Health Technology and Informatics*, *277*, 37–46. https://doi.org/10.3233/SHTI210026

[17] Almeida, R., Sousa, H., Cunha, L. F., Guimarães, N., Campos, R., & Jorge, A. (2024). https://doi.org/10.48550/arXiv.2401.01825 Physio: An LLM-based physiotherapy advisor. *arXiv Preprint: 2401.01825*

[18] Lee, M. H., Siewiorek, D. P., Smailagic, A., Bernardino, A., & Bermúdez i Badia, S. (2024). Enabling AI and robotic coaches for physical rehabilitation therapy: Iterative design and evaluation with therapists and post-stroke survivors. *International Journal of Social Robotics*, *16*(1), 1–22. https://doi.org/10.1007/s12369-022-00883-0

[19] Macedo, P., Madeira, R. N., Santos, P. A., Mota, P. R., Alves, B. R., & Pereira, C. (2025). A conversational agent for empowering people with Parkinson's disease in exercising through motivation and support. *Applied Sciences*, *15*(1), 223. https://doi.org/10.3390/app15010223

[20] Sivakumar, R. D. (2024). Conversational AI and reinforcement learning for virtual coaching in physiotherapy. *Cuestiones de Fisioterapia*, *53*(3), 1141–1152. https://doi.org/10.48047/jpc2x745

[21] OECD & European Observatory on Health Systems and Policies. (2023). *Greece: Country health profile 2023*. *State of Health in the EU*. OECD Publishing.

[22] Tsekoura, M., Fousekis, K., Lampropoulou, S., Xergia, S., Bania, T., Tsepis, E., & Billis, E. (2022). Physiotherapists' perceptions and willingness to use telerehabilitation in Greece: A cross-sectional study. *Cureus*, *14*(12), e32317. https://doi.org/10.7759/cureus.32317

[23] Aggarwal, A., Tam, C. C., Wu, D., Li, X., & Qiao, S. (2023). Artificial intelligence–based chatbots for promoting health behavioral changes: Systematic review. *Journal of Medical Internet Research*, *25*, e40789. https://doi.org/10.2196/40789

[24] Haupt, C. E., & Marks, M. (2023). AI-generated medical advice—GPT and beyond. *JAMA*, *329*(16), 1349–1350. https://doi.org/10.1001/jama.2023.5321

[25] Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., … Smith, D. M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine*, *183*(6), 589–596. https://doi.org/10.1001/jamainternmed.2023.1838

How to Cite: Mariettou, S., Shahbandi, S., Triantafyllou, V., & Koutsojannis, C. (2026). Conversational AI Agents in Physical Rehabilitation: A Scoping Review with Implications for Physiotherapy Practice in Greece. *Medinformatics*. https://doi.org/10.47852/bonviewMEDIN620210454