Interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 15 (2026) ISSN 1929-073X

Contents

ViewPoint

Reviews

XSL•FO
RenderX

Original Papers

Case Report

XSL•FO
RenderX

ViewPoint

Evidence-Based Self-Management Strategies for Fibromyalgia: Foundations for Digital Therapeutic Applications

Thomas Lebeau Foustoukos[1,2], MSc; Isabelle Leclercq[2], MSc; Marc Blanchard[2], PhD; Thomas Hügle[2], MD, PhD

[1]College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, United Kingdom

[2]Department of Rheumatology, University Hospital of Lausanne, Lausanne, Vaud, Switzerland

Corresponding Author:
Thomas Hügle, MD, PhD
Department of Rheumatology
University Hospital of Lausanne
Av. Pierre-Decker 4
Lausanne, Vaud, 1011
Switzerland
Phone: 41 21 314 14
Email: thomas.hugle@chuv.ch

Abstract

Fibromyalgia is a prevalent musculoskeletal pain condition that causes major personal, social, and societal burden. Pharmacological therapies often provide only limited benefit, making multimodal approaches and self-management the cornerstones of care. Such strategies, spanning lifestyle modification, physical activity, psychoeducation, and cognitive-behavioral approaches, target the biopsychosocial complexity of fibromyalgia and promote sustainable coping. In parallel, digital health technologies are transforming how these interventions can be delivered and coordinated in the form of digital therapeutics. This viewpoint draws on a multiphase investigation to appraise the current and future landscape of fibromyalgia self-management in the digital era. Its objective is to present an evidence-based framework and recommendations to guide the development of a mobile health self-management program for patients with fibromyalgia. In phase 1, we conducted a review of international guidelines and randomized controlled trial–based systematic reviews addressing nondigital self-management interventions for fibromyalgia and related nociplastic pain conditions. In phase 2, we analyzed the content and certification status of currently available mobile and virtual health applications for fibromyalgia. In phase 3, we convened a multidisciplinary focus group of rheumatologists, patients, and digital health developers to identify priorities for translating evidence-based self-management content into mobile health formats. Collectively, we suggest that effective digital self-management for fibromyalgia should evolve beyond single-domain interventions toward validated, personalized, and interactive multimodal platforms. Virtual care may increasingly function at the point of care, linking monitoring, education, and behavioral support in one continuum.

(Interact J Med Res 2026;15:e67523) doi:10.2196/67523

KEYWORDS

chronic pain; cognitive behavior therapy; digital therapeutics; eHealth; fibromyalgia; mind-body therapies; mobile application; physical activity; self-management

Introduction

Fibromyalgia is defined as a syndrome characterized by moderate to severe symptoms, including widespread pain, fatigue, sleep disturbance, cognitive complaints, and an increase in somatic complaints [1]. Despite ongoing efforts to clarify the definition of fibromyalgia, accurately assessing its diagnosis, prevalence, management, and its personal and societal impacts remains challenging [2]. Fibromyalgia can be accompanied by psychiatric, rheumatological, gastrointestinal, or other comorbidities, which further complicate the management and contribute to its overall personal and societal burden [3].

Proposing a clear treatment strategy for fibromyalgia remains challenging. The focus is on nonpharmacological therapies, particularly physical interventions, psychological treatments, and mind-body approaches that aim to promote self-management. Patient education plays a crucial role in enhancing their understanding of the disease and fostering their active participation in its management. Pharmacological treatments have shown limited benefits [4-8]. Multimodal treatment programs are recommended for treatment-resistant

XSL•FO
RenderX

fibromyalgia patients, as they combine different treatment modalities and offer group effects. However, such programs require significant resources that are not necessarily considered cost-effective.

Self-management programs play a crucial role in supporting individuals with chronic conditions [9] and are recommended by management fibromyalgia guidelines [4-7]. The fundamental concept of self-management emphasizes the active involvement of patients [9,10]. Self-management considers the multidimensional nature of the illness and combines multiple interventions such as psychological interventions, exercise, and nutritional aspects. It offers a relevant approach to address the challenges in fibromyalgia management and shows its efficiency to improve fibromyalgia patients' well-being [11,12] and a reduced tendency of patients to seek repeated consultations with health care professionals. A plethora of these programs have shown an improvement in pain intensity, functionality, cognitive behavior, emotions, and quality of life (QoL). However, the quality of these studies is heterogeneous [13-17]. When programs are detailed, their references are often unclear, relying on peer opinions or interviews [18-24], or they may offer only one component, such as psychoeducation or exercises [25-28].

The digital transformation is continuously shaping the health care landscape. It encompasses various digital innovations such as online programs, apps, gaming, virtual reality, or, more recently, generative artificial intelligence (AI) and chatbots [29]. In a clinical context, mobile health (mHealth) apps have demonstrated several key functionalities, including diagnostics and clinical decision support, behavior change interventions, and the delivery of disease-related education and communication support. The latter 3 functionalities are particularly relevant for delivering self-management interventions, empowering patients to monitor and manage their own health conditions, and providing an additional avenue for health care professionals to support their patients' well-being [30]. In 2019, a law was passed in Germany that digital health applications are reimbursable by health insurance companies under certain conditions. Currently, 8 digital health applications for musculoskeletal indications are registered; 2 of those are for chronic pain syndromes, including fibromyalgia [31]. There was a second acceleration of eHealth during the COVID-19 pandemic, notably concerning telemonitoring and video consultations. In the meantime, eHealth interventions have been shown to enhance patients' QoL and treatment experiences [32,33].

Online programs and mobile apps have demonstrated promising benefits for individuals with chronic pain or musculoskeletal conditions [33,34]. The use of mHealth technologies offers an accessible and available 24/7 solution without geographical constraints, empowering patients to become more self-reliant in managing their chronic conditions [35,36]. Notwithstanding, there are several caveats and limitations in existing solutions. A substantial proportion of currently available online fibromyalgia programs are not evidence-based or lack proven, evidence-based content. They are usually not personalized and cannot be updated, as this would alter the certification status of the device. On the other hand, recent digital tools such as large language models (LLMs) and chatbots offer novel approaches that could address some of these limitations.

Aim

The objective of this work is to present an evidence-based framework and recommendations to guide the development of an mHealth self-management program for patients with fibromyalgia. A comprehensive literature review was conducted to evaluate the effectiveness of self-management interventions on symptoms in patients with fibromyalgia (phase 1). Where applicable, relevant practical characteristics of the interventions were also reported. In phase 2, existing therapeutic fibromyalgia apps and online programs were reviewed, and in phase 3, general aspects and user preferences were discussed in focus groups with patients. Finally, we aimed to identify strategies to optimize online self-management interventions and explore their potential for further improvement.

Ethical Considerations

This study is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Phase 1: Review of Existing Evidence

Overview

MEDLINE and Cochrane databases, encompassing English and French articles from the inception of the databases up until 2023, were searched to identify guidelines and systematic reviews of randomized controlled trials for fibromyalgia self-management. Data from 30 systematic reviews that demonstrated a high or moderate level of confidence based on the AMSTAR 2 evaluation and 8 guidelines with an AGREE II–GRS score higher than 18 were considered to categorize and summarize the evidence on self-management interventions (Figure S1 in Multimedia Appendix 1).

Categorization of Self-Management Interventions

Self-management interventions were categorized into 3 groups: physical exercise, psychoeducative interventions, and mind-body interventions. Additionally, a separate category called multimodal interventions was identified, which encompassed interventions combining 2 or more of the previously mentioned self-management approaches. The European, American, Canadian, and Italian guidelines covered physical exercise, psychoeducative interventions, and mind-body self-management interventions as part of the overall management of patients with fibromyalgia, alongside other treatment modalities [4,6,7,37]. The German guideline divided its recommendations into 4 separate studies: general recommendations [7], recommendations on physical interventions [38], psychological interventions [39], and other interventions such as mind-body therapies [40]. In total, 8 studies focused on physical activities [41-48]; 14 studies on psychoeducative interventions [25,49-60]; 4 studies on mind-body therapies [61-64]; 1 study on multicomponent interventions [11]; 1 study that investigated both psychoeducative and mind-body interventions [65]; and 2

XSL•FO
RenderX

studies that examined physical activities, psychoeducative interventions, mind-body therapies, and multicomponent interventions [66,67]. The studies used various types of control groups, which encompassed the following options: treatment as usual, waiting list or no treatment, alternative pharmacological or nonpharmacological interventions, active or attention control, and placebo or sham interventions. Designing a control group for complex self-management interventions was considered challenging due to the inherent difficulty of blinding participants, which can introduce measurement and interpretation bias [68].

Evidence for Specific Self-Management Interventions

A wide range of variables were assessed both before and after the interventions: pain (both qualitative and quantitative measures), QoL, psychological functions such as depression, anxiety, self-efficacy, acceptance, catastrophizing, and fear-avoidance, physical functions such as strength and disability, sleep quality, fatigue, health care use indicators such as visits, sick leaves, return to work, and care-seeking behavior, as well as trial withdrawal rates and adverse effects. The Fibromyalgia Impact Questionnaire (FIQ) was used as a multidimensional measurement tool.

Physical Interventions

All of the guidelines strongly endorse the inclusion of aerobic exercises and resistance (or strength) exercises in self-management interventions for fibromyalgia [4,6,7,37,38]. On the other hand, flexibility and stretching exercises are either not recommended [6,7] or recommended with weak or insufficient evidence of benefit [37,38]. Exercise interventions were found to be superior to control groups in improving pain, FIQ scores, sleep quality, fatigue, and depression [67]. The characteristics of the exercises were extracted based on the prescription guidelines provided by the American College of Sports Medicine, which include frequency, intensity, time, and type of exercises [69]. Additionally, the duration of the exercises, delivery modalities, whether they were tailored or standardized, and safety considerations were also considered. Aerobic exercises had positive effects on pain, QoL, and depression [42,45,46]. Aerobic exercises demonstrated a high effect on reducing the FIQ [41]. Aerobic exercises probably improve stiffness and slightly improve physical function and cardiovascular function [42]. According to Couto et al [46], resistance exercises yielded positive effects on pain and QoL but did not show a significant effect on depression. On the other hand, Albuquerque et al [41] found that resistance exercise demonstrated a moderate effect in reducing the FIQ scores. Resistance exercise has the potential to improve fatigue and sleep [49]. Although flexibility exercise did not show superiority over control groups (other interventions or no intervention), the quality of evidence is insufficient to draw definitive conclusions regarding the effectiveness of flexibility exercise on fibromyalgia symptoms [41,45,46,48]. However, there is evidence suggesting that physical exercise may help decrease muscle stiffness [48]. Interventions that combine aerobic exercises, resistance exercises, and flexibility exercises have shown greater effectiveness in improving FIQ scores compared to single-type exercise interventions [41]. However, Bidonde et al [43] state that no definitive conclusion can be drawn regarding the optimal proportion, synergy, or specific characteristics of each exercise type [43]. Exergames have demonstrated promising benefits in improving pain, disability, and physical function in both the short and long term [66]. There is a lack of available information regarding other exercises, such as Pilates and motor control exercises. Exercise (aerobic, resistance, and flexibility) characteristics are shown in Table S1 in Multimedia Appendix 1.

Psychoeducative Interventions

Consensus guidelines recommend the inclusion of psychological and educational interventions, such as cognitive behavioral therapy (CBT) and health education [4-7,37]. These interventions can encourage self-management, enhance self-efficacy, and reduce maladaptive thoughts and behaviors [6]. It is particularly important to consider these interventions when patients have comorbid mental disorders [39]. For hypnosis and guided imagery, recommendations are conflicting [4,7,39]. Although other psychological interventions, such as relaxation, therapeutic writing, Roger therapy, family therapy, psychodynamic therapy, and psychoanalytic therapy, are available, they should not be proposed as stand-alone treatments due to the lack of evidence or limited recommendations [39]. CBT incorporates education on the physiopathology of pain, self-management skills, cognitive reappraisal, pacing activities, and problem-solving techniques to help patients modify their behaviors, thoughts, and emotions, ultimately reducing pain, improving functioning, and enhancing mood. Additional components, such as sleep hygiene, may be included. Homework assignments are given to encourage the practice of skills in everyday life [60]. Acceptance and commitment therapy (ACT) and mindfulness-based therapy are extensions of CBT. ACT comprises 2 main core components: mindfulness and acceptance, as well as commitment and behavior change. It involves several processes, including acceptance, cognitive defusion, present-moment awareness, self-as-context, values clarification, and committed action [70]. ACT interventions often incorporate CBT processes, making it difficult to distinguish between CBT and ACT [51]. Evidence suggests that CBT can reduce pain, disability, and negative mood in patients with chronic pain or fibromyalgia, both immediately after treatment and at long-term follow-up (6 and 12 months), although the effect size is small or very small [68,79. Enomoto et al [52] recommend offering CBT for insomnia, or at least CBT for insomnia and pain, to patients with chronic pain who also experience insomnia as a comorbidity. CBT therapy for insomnia has been effective in improving sleep, pain, disability, and depression. ACT has also demonstrated benefits for patients with chronic pain, improving pain acceptance, QoL, pain intensity, functioning, and mood [57,59]. For fibromyalgia, ACT has shown significant improvements in patient functioning in both the short and long term, based on a meta-analysis with moderate-quality evidence [51]. ACT can be considered to enhance patients' psychological flexibility and, subsequently, their functioning in pursuing valued activities [51,59]. Education serves as both an intervention and a CBT tool [58]. Joypaul et al [55] defined education as providing instructions to inform participants, making it applicable to various interventions as an instructional

XSL•FO
RenderX

tool [55]. Therapeutic pain neuroscience education, also referred to as pain neuroscience education by some authors, aims to enhance patients' knowledge and understanding of pain neurophysiology to improve their cognitive and behavioral skills related to pain [56,71]. Therapeutic pain neuroscience education has demonstrated benefits for patients with chronic pain in reducing fear of movement, pain intensity, pain disability, and pain catastrophizing [58]. Therapeutic pain neuroscience education, in the context of chronic musculoskeletal pain, improved pain intensity, pain knowledge, disability, maladaptive thoughts and behaviors, physical function, and health care use, even up to 1 year post treatment [56]. Other educational content has been studied, including lifestyle components focusing on areas such as nutrition, sexuality, social coping strategies, and the regulation and adjustment of everyday life [11]. Education should be an integral part of a multidisciplinary approach to chronic pain, alongside graded physical activities, graded exposure, and pacing, for example, and not offered as a stand-alone treatment [54-56,58,67]. CBT education characteristics were gathered in Table S1 in Multimedia Appendix 1.

Mind-Body Interventions

The classification of mind-body interventions was based on the Medical Subject Headings terms of the National Library of Medicine. They encompass various modalities such as meditative movement therapies (MMTs), respiration exercises, hypnosis and autogenic training, meditation, and relaxation [72]. They focus on exploring the interconnectedness between the brain, body, mind, and behavior, and how emotional, mental, social, spiritual, experiential, and behavioral factors can directly impact health [73]. Guidelines recommend the use of MMTs with confidence [4,7,40]. However, hypnosis, guided imagery, relaxation, or meditation are not recommended, and it is advised not to propose them as stand-alone interventions [7,40]. The Canadian guideline suggests not discontinuing these interventions but informing the patient about their lack of evidence and potential side effects [6]. Therapies such as Qi Gong, Tai Chi, and Yoga, as examples of MMTs, were included [63]. MMTs have shown improvements in outcomes such as physical functioning, pain, and mood for patients with fibromyalgia [65]. Specifically, Tai Chi has been found to significantly reduce FIQ scores, pain intensity, sleep disturbance, fatigue, and depression, while increasing QoL for individuals with fibromyalgia [61]. Tai Chi and Yoga are weakly recommended, whereas the evidence did not allow for a recommendation regarding Qi Gong for chronic pain management [63]. Lee et al [64] did not provide a recommendation for meditation as a treatment for chronic pain due to the limited evidence available. While no specific recommendation was provided in the systematic review, Lee et al [64] suggested that mindfulness-based interventions could potentially offer benefits for chronic pain management.

Multicomponent Interventions

Multicomponent interventions were defined as combinations of at least 2 components, including psychological, physical activity, medical education, and mind-body therapies, based on the self-management program developed by Miles et al [74] and further refined by Geraghty et al [11]. All guidelines recommend not implementing interventions in isolation but rather combining them with other approaches. Multicomponent intervention should include physical and psychoeducational interventions at a minimum [4-7,75]. Multicomponent interventions have demonstrated effectiveness in improving various aspects such as FIQ, pain, sleep, and depression, with greater effects compared to exercise alone, education alone, or psychological interventions alone [61,62]. The positive effects of multicomponent interventions typically last for an average of 14 weeks, and it is recommended to conduct follow-up assessments every 3 months to review and reinforce the treatment strategy [62]. Considering the variations across studies, multicomponent interventions have shown improvements in physical function, pain, FIQ, fatigue, mood, and QoL in both the short and long term [11].

Phase 2: Review of Self-Management Apps

In this part of the work, we investigated which type of self-management is integrated in existing apps for fibromyalgia and chronic pain syndromes (Table 1). Online fibromyalgia applications were identified via PubMed, Google, and the German Digitale Gesundheitsanwendung (DiGA) registry [76], as well as with ChatGPT (OpenAI). Available mobile apps targeting fibromyalgia self-management reveal a strong emphasis on psychoeducational content, particularly CBT, ACT, and pain neuroscience education. Out of the 12 reviewed apps, 10 integrate psychoeducational modules, often presented through structured lessons, interactive exercises, journaling tools, or virtual coaching. Physical activity interventions, such as aerobic or resistance exercise guidance, were integrated in 6 apps, typically via instructional videos, activity logging, or synchronization with wearable devices. Mind-body techniques, including guided meditation, breathing exercises, and mindfulness-based practices, were found in about 5 apps. This distribution highlights a predominant focus on psychological and cognitive strategies in current digital interventions, with fewer apps addressing physical or somatic components in depth.

XSL•FO
RenderX

Table 1. Existing online self-management programs for fibromyalgia.

App name	Therapeutic mechanisms	Regulatory status	Notes
Stanza	ACT[a], CBT[b], mindfulness, journaling, reminders	FDA[c]-cleared in the United States	12-week structured digital therapy, clinical trial support
HelloBetter Chronic Pain	CBT, ACT, mindfulness	DiGA[d]-approved in Germany	Structured 12-week course, reimbursed by German health insurance
Selfapy Chronic Pain	CBT, mindfulness, education	DiGA approval in Germany withdrawn	Evidence-based course, reimbursed via DiGA
Manage My Pain	Pain tracking, journaling, self-monitoring	Available on app stores	Focuses on tracking pain trends for self-awareness and provider communication
Quell Fibromyalgia	Neuromodulation via wearable, symptom tracking	FDA-authorized medical device	App supports wearable neuromodulation therapy for fibromyalgia
FibroMapp	CBT-based tools, symptom tracking, medication management	Available on app stores	Combines tracking and CBT tools; user-driven format
MoreGoodDays	CBT, mindfulness, education	Available on app stores	Structured CBT modules, lifestyle education, mindfulness tools
FibroMinder	Reminders, symptom tracking, task management	Available on app stores	Simple utility app for scheduling and symptom tracking
PainScale	Pain tracking, education, CBT elements, community support	Available on app stores	Community-focused platform with integrated education and logging
FibroTrack	Symptom tracking, lifestyle management	Available on app stores	Focuses on tracking and lifestyle planning features
Curable	Pain neuroscience education, CBT, guided meditation, writing exercises	Available on app stores	Comprehensive app targeting pain beliefs and coping strategies
Fibrowalk	CBT, mindfulness, therapeutic exercise, pain neuroscience education	Not a certified app; delivered via YouTube and email	Multicomponent program with weekly video sessions

[a]ACT: acceptance and commitment therapy.

[b]CBT: cognitive behavioral therapy.

[c]FDA: US Food and Drug Administration.

[d]DiGA: Digitale Gesundheitsanwendung.

Two applications, HelloBetter Chronic Pain and Stanza, are certified digital therapeutics. HelloBetter is DiGA-approved in Germany, meaning they are reimbursable by statutory health insurance and have met criteria for clinical evidence and data security. Stanza, developed in the United States, is US Food and Drug Administration–cleared as a prescription digital therapeutic, offering a 12-week ACT-based program with demonstrated clinical efficacy in reducing fibromyalgia symptoms [77]. The remaining apps, including Curable, MoreGoodDays, Manage My Pain, and FibroMapp, are publicly available but lack medical certification and are typically framed as wellness tools or digital companions rather than regulated treatments. Overall, while several apps offer well-designed psychoeducational modules and basic symptom tracking, fewer integrate evidence-based physical activity interventions or comprehensive mind-body components. Among the 12 fibromyalgia self-management apps analyzed, only 5 (42%) can be considered multicomponent, meaning they integrate at least 2 of the 3 core evidence-based intervention domains: psychoeducational therapies, physical activity, and mind-body techniques. Notably, the certified apps HelloBetter Chronic Pain, Selfapy Chronic Pain, and Stanza fall into this category, offering combinations of CBT or ACT, mindfulness, and, in some cases, activity planning. Noncertified apps such as MoreGoodDays and Curable also include psychoeducation and mind-body interventions, but vary in the depth and clinical rigor of their content. While most of the reviewed apps integrate evidence-based components aligned with current clinical guidelines, such as CBT, ACT, and mindfulness-based techniques, several also include strategies with limited or inconclusive evidence in fibromyalgia. For example, some noncertified apps incorporate generic meditation, unstructured journaling, expressive writing, or broad lifestyle advice without therapeutic framing or individual tailoring.

Phase 3: Patient Preferences and Focus Groups

As previously reported, we investigated patient preferences for mobile app design through 2 online surveys (53 and 33 patients, respectively) and 3 focus groups comprising a rheumatologist, 1-6 fibromyalgia patients, with or without concomitant post-COVID-19 syndrome, as well as an app designer [78]. All patients in focus groups tested a self-developed app designed to assess fibromyalgia symptoms and to deliver an online therapeutic program containing CBT, mind-body techniques, and physical exercise instructions. This program was initially developed during the pandemic for fibromyalgia associated with postviral syndrome and was subsequently adapted as an

XSL•FO
RenderX

experimental chronic pain companion (Pain Organiser and Companion System). The design and content of this app are described in detail elsewhere [78,79]. In the final focus group (n=1), the online program was connected to an LLM-based chatbot and tested by the patient (see Future Directions section).

Participants emphasized that simplicity, clarity, and accessibility are crucial for adherence, given common cognitive fatigue ("fibrofog") and limited digital literacy. Users preferred short, well-structured modules, clear navigation, and scientifically validated content in plain, nontechnical language. Autonomy and empowerment emerged as central themes. Patients valued tools that visualize progress, such as symptom tracking and diaries, helping them recognize links between behavior and symptoms. Personalization was highly desired; users wanted adaptable exercises, favorite lists, and an interactive virtual coach or chatbot that responds to individual needs and previous activity. An empathic, motivating tone was considered essential. Participants appreciated friendly communication and multimedia content, particularly physiotherapy videos and mindfulness modules. Emotional design elements, positive wording, supportive colors, and encouraging feedback were viewed as vital for engagement and adherence. Reliability and trustworthiness were further priorities: users expected apps to be evidence-based, transparent, and secure, with clear data protection and professional validation. Barriers such as poor navigation, cognitive overload, or lack of emotional connection can significantly reduce adherence. The focus group highlighted that fibromyalgia mHealth apps offer the potential to serve as multicomponent platforms, covering the 3 key therapeutic domains discussed above: psychoeducation, mind-body techniques, and physical exercise guidance. Of note, instructions for physical exercises were appreciated by participants in the focus groups, ideally with more individualization (eg, for individuals with obesity).

Key Findings and Lessons Learned

There is proven evidence for psychoeducational interventions, mind-body therapies, and physical activities in fibromyalgia. All 3 are included in most of the current fibromyalgia mHealth applications on the market, several of those with proven efficacy in randomized controlled trials, although the underlying therapeutic content is not openly accessible [75]. CBT remained strongly recommended to assist patients with fibromyalgia in managing their pain, comorbidities, and lifestyle [49] and is the main approach used in existing fibromyalgia apps. Apart from CBT interventions, psychoeducation focuses on pain coping skills, mindfulness, lifestyle modification, and pain neuroscience education. The recommendations regarding mind-body therapies such as MMT, meditation, relaxation, or hypnosis are conflicting but safe and therefore might be integrated in apps as an adjunct. Certain interventions, such as stress reduction, breathing, and relaxation exercises, are notably suitable for integrating into apps as animation can be positive, such as a balloon for breathing exercises.

Half of the existing fibromyalgia apps integrate physical exercises into their program, but mostly only to a lower degree compared to CBT. The content of existing apps contrasts with technically more sophisticated physiotherapy or online rehabilitation apps; for example, integrating computer vision to instruct and monitor a wide range of exercises. In general, it can be postulated that true multicomponent apps integrating all 3 validated pillars of fibromyalgia self-management, psychoeducation, physical activity, and mind-body techniques, guided by clinical standards and codeveloped with patient input, would improve the effect.

From a technical side, both web and native apps are viable platforms for integrating self-management tools and tracking disease activity in fibromyalgia. Native apps are particularly effective at incorporating wearable data directly from mobile devices, enabling continuous monitoring and data capture. Symptom tracking is typically based on patient-reported outcomes or symptom scores, including measures of pain, QoL, depression, anxiety, self-efficacy, acceptance, catastrophizing, fear-avoidance, physical function (eg, strength and disability), sleep quality, fatigue, and health care use indicators such as clinic visits, sick leave, and return-to-work patterns.

These apps often incorporate validated instruments such as the FIQ, the Symptom Severity Score, or the Widespread Pain Index to assess disease activity. Therapeutic content is typically delivered through text, animated videos, or voice formats, providing education, instructions, and exercises. In some cases, apps are integrated with connected devices, such as wearables that deliver millimeter wave stimulation [80]. Animation can be used to enhance certain exercises, such as balloons being used to synchronize breathing frequency and intensity.

Stand-alone apps, particularly those certified as digital therapeutics, often include a structured program guide to track patient progress. This is important for potential reimbursement by health insurers. Increasingly, mHealth apps also integrate chatbots that guide users through the program, enhance flexibility, and potentially improve adherence, one of the main challenges of these interventions. Successful therapeutic apps for this population require a balance between scientific rigor, user-friendly design, and emotionally intelligent interaction. A clinical and evidence-based framework for fibromyalgia interventions in apps can be found in Table 2. Textbox 1 presents practical recommendations for the corresponding app design and implementation roadmap.

XSL•FO
RenderX

Table 2. Evidence-based framework for fibromyalgia online interventions.

Intervention domain	Evidence level and rationale	Implementation in mobile health app	Digital enablement features	Evaluation metrics
Psychoeducational therapies (CBT[a], ACT[b], and PNE[c])	Strong evidence from RCTs[d] and guidelines for reducing pain, distress, and maladaptive thoughts	Modular courses; microlearning sessions; quizzes; journaling; goal setting; and self-efficacy tracking	Chatbots or avatars for guidance; adaptive progression based on symptom and mood data; empathy-based conversational tone; and voice message transcription	Pain, FIQ[e], self-efficacy, catastrophizing, engagement rate, and module completion
Physical activity and graded exercise	Strong evidence for aerobic and resistance training improving function and mood	Video demonstrations; personalized exercise plans; daily movement reminders; wearable integration (steps, HR[f], and HRV[g])	AI[h]-driven motion feedback (computer vision); adaptive load progression; safety alerts; and gamification	Physical function (FIQ physical domain), fatigue, adherence (logged sessions), and HR and HRV trends
Mind-body techniques (mindfulness, Tai Chi, breathing, and relaxation)	Moderate evidence for improving sleep, mood, and QoL[i]	Audio-guided meditations, breathing animations, mindfulness timers, and relaxation music	Adaptive session lengths; stress biofeedback using HRV; sleep tracking linkage; and integration with smartwatches	Sleep quality, stress index, anxiety, HRV, and app usage continuity
Symptom tracking and patient-reported outcomes	Essential for personalization and clinical insight for HCP[j]	In-app FIQ, pain diaries, and fatigue and sleep trackers	Voice or video symptom entry; AI-based summary visualization; and adaptive dashboard	Data completeness, trend accuracy, and correlation with clinical outcomes
Personalization and adaptive design	Increasingly essential for engagement and relevance	Custom goal setting; phenotype-based pathways (eg, obesity-, menopause-, and PTSD[k]-related FM[l])	Machine learning–based tailoring; predictive suggestions for pacing and exercise. Adapted avatars, eg, older people with obesity	User satisfaction, engagement over time, and adaptive accuracy
Behavior change and motivation	Crucial for long-term adherence	SMART[m] goal planning, feedback loops, and progress visualization	Gamification, motivational messaging, positive reinforcement, and social comparison (optional)	Retention rate, adherence index, and self-efficacy gain
Communication and support	Improves adherence and patient safety	Chatbot or professional chat; asynchronous therapist feedback	Hybrid care integration (eg, AI triage + human follow-up); crisis escalation paths	Message frequency, satisfaction, and safety events
Data integration and clinical workflow	Enables clinical supervision and research	Clinician dashboard, FHIR[n]-based interoperability, and data export to EHR[o]	Sidecar EMR integration; secure teleconsultation channel	Clinical uptake, data completeness, and clinician feedback
Accessibility, UX[p], and emotional design	Critical for usability and adherence in cognitive fatigue	Clean interface, large icons, voice navigation, and light and dark modes	Emotionally supportive design (colors, feedback tone), simplified onboarding, and language localization	SUS[q] score, accessibility compliance, and dropout rate
Privacy, certification, and ethics	Required for trust and scalability	CE[r], FDA[s], or DiGA[t] conformity, and transparent data policies	Privacy-by-design architecture; on-device data processing	Certification status, GDPR[u] or HIPAA[v] compliance, and user trust rating

[a]CBT: cognitive behavioral therapy.

[b]ACT: acceptance and commitment therapy.

[c]PNE: pain neuroscience education.

[d]RCT: randomized controlled trial.

[e]FIQ: Fibromyalgia Impact Questionnaire.

[f]HR: heart rate.

[g]HRV: heart rate variability.

[h]AI: artificial intelligence.

[i]QoL: quality of life.

[j]HCP: health care professional.

[k]PTSD: posttraumatic stress disorder.

[l]FM: fibromyalgia.

[m]SMART: specific, measurable, achievable, relevant, time-bound.

[n]FHIR: Fast Healthcare Interoperability Resources.

XSL•FO
RenderX

[o]EHR: electronic health record.

[p]UX: user experience.

[q]SUS: System Usability Scale.

[r]CE: Conformité Européenne.

[s]FDA: US Food and Drug Administration.

[t]DiGA: Digitale Gesundheitsanwendung.

[u]GDPR: General Data Protection Regulation.

[v]HIPAA: Health Insurance Portability and Accountability Act.

Textbox 1. Recommendations for the fibromyalgia mobile health self-management program development.

Recommendation

- Involve key stakeholders (fibromyalgia patients, multidisciplinary professionals, and digital experts) through surveys, interviews, and co-design workshops.
- Conduct a structured needs assessment, addressing core fibromyalgia symptoms and patient priorities (eg, fatigue, pain, and cognitive dysfunction).
- Provide a transparent overview of the content for users, health care providers, and regulators, eg, using knowledge graphs or content maps.
- Tailor content to user characteristics such as socioeconomic background, gender, age, culture, and health literacy.
- Base all interventions on clinical evidence, including physical activity, cognitive behavioral therapy, psychoeducation, and mind-body strategies.
- Include personalized modules for subtypes and comorbidity (eg, perimenopausal symptoms, obesity, anxiety, posttraumatic stress disorder–associated hypermobility, migraine, etc).
- Adhere to established standards and guidelines, such as Xcertia, National Institute for Health and Care Excellence, Haute Autorité de Santé, and World Health Organization, covering content, privacy, usability, and operability.
- Design the app as a complementary tool, not a substitute for face-to-face care.
- Use interoperable data formats, such as Fast Healthcare Interoperability Resources, to enable secure integration with health care systems and devices.
- Connect to wearable data, allowing tracking of physical activity, sleep, or heart rate variability to enrich outcome measures.
- Include a clinician-facing dashboard to enable remote monitoring, triage, and decision support.
- Ensure robust data security and privacy compliance, including user consent, encryption, and data governance.
- Provide a structured and simple onboarding process, guiding users through initial setup and goals.
- Ensure user-friendliness with intuitive design, a responsive user interface, customizable settings, reminders, and accessible language.
- Include motivational features, such as gamification, progress tracking, and positive reinforcement.
- Incorporate core functional modules, including symptom tracking, educational content, communication, and self-assessment tools.
- Implement adaptive, guided interventions rather than static or generic content to boost engagement and outcomes.
- Define and track recognized fibromyalgia end points to support evidence generation and reimbursement (eg, Fibromyalgia Impact Questionnaire, pain scales).
- Follow an iterative development cycle, including continuous evaluation of effectiveness, usability, and safety.
- Plan for scalability, dissemination, and regulatory approval, including Conformité Européenne marking, Digitale Gesundheitsanwendungen eligibility (Germany), or US Food and Drug Administration listing.

Areas of Uncertainty

The literature search was limited to PubMed and Cochrane databases, potentially introducing selection bias, and the cutoff at 2023 excluded recent studies or emerging interventions. Another limitation of this viewpoint concerns the transferability of conventional evidence-based therapies to digital formats. The effectiveness of individual therapeutic elements within apps remains largely untested, and user experience likely plays a decisive role. Adherence is particularly critical, as unguided online interventions often show low engagement. Psychiatric comorbidities, such as depression, trauma, and cognitive fatigue, further affect adherence and suggest that some users may benefit from guided or coach-assisted programs rather than fully automated ones. The human and group-based components of multimodal therapy are difficult to replicate digitally, raising questions about the optimal degree of human involvement in digital therapeutics. Device-based or biofeedback interventions were not included, as they require additional hardware not universally accessible; however, future versions may integrate smartwatch-based or wearable data.

Finally, personalization remains a key challenge. As fibromyalgia is heterogeneous, tailored approaches—whether based on symptom profiles or machine learning–derived phenotypes—are needed. Future studies should also account for contextual factors such as age, culture, digital literacy, and

XSL•FO
RenderX

socioeconomic status, as well as identify which behavior change techniques drive engagement and outcomes. Guided and adaptive interventions appear more effective than static content, highlighting the importance of iterative, patient-centered design in future mHealth development.

Future Directions

As seen in other digital interventions, such as for depression, the inclusion of a channel for direct interaction with a health care professional or, potentially, a bot can increase effectiveness. Voice message transcription features can facilitate semiautomated, asynchronous communication. In this context, agentic AI trained on such interactions may provide supplementary or alternative support. Dashboards play a critical role in integrating fibromyalgia mHealth applications into the clinical workflow, enabling health care professionals to monitor their patients and app activity. Ideally, these dashboards should interface with electronic medical records through sidecar applications [81]. Personalization of content according to fibromyalgia subtypes should be encouraged. However, this increased individualization may make the creation of robust evidence or standardized certification more challenging. As a future perspective, new content formats could be developed and evaluated, such as therapeutic stories or short interventions. For example, therapeutic stories could resemble YouTube Shorts, with AI algorithms applied to tailor and deliver these formats accordingly. Other future digital directions for fibromyalgia apps will move beyond psychoeducation toward multimodal, personalized therapy ecosystems. Integration of exercise modules, CBT, and mindfulness into cohesive platforms is expected, with adaptive interfaces that adjust content to each user's symptom patterns, motivation, and cognitive load. LLMs and AI-powered chatbots will evolve from static text-based assistants to emotionally intelligent digital companions that guide users through daily self-management, provide empathetic feedback, and interface with sensor data from wearables for sleep, mobility, or stress tracking. These systems could enhance adherence, detect flares early, and personalize pacing and exercise recommendations. Meanwhile, interoperability, data privacy, and regulatory validation will remain crucial to ensure clinical reliability and ethical use. Ultimately, the future of fibromyalgia apps lies in hybrid digital care, combining automated AI-driven support with human coaching and clinical oversight. By blending evidence-based therapy, personalization, and human empathy, next-generation digital therapeutics may finally bridge the gap between self-management and sustained symptom improvement.

Conclusion

Evidence-based foundations are needed to inform the development of effective digital therapeutic applications for fibromyalgia. To build impactful digital therapeutics, it is essential to integrate validated self-management strategies, engage key stakeholders, and consider social and cultural determinants. An iterative development process, guided by ongoing assessment of usability, effectiveness, equity, and safety, will ensure that mHealth tools align with both clinical goals and patient needs.

Acknowledgments

We thank Physio Homecare for providing physical exercise videos for the Pain Organiser and Companion System app. Generative artificial intelligence was used for editing the manuscript. Support for content creation is provided in Table 2.

Data Availability

All data generated or analyzed during this study are included in this published article. Quantitative and qualitative data of focus groups can be found in [78].

Authors' Contributions

All authors have contributed to the conception of this article and approved the final version to be published. TLF and TH wrote the manuscript. TLF performed the literature review and was involved in a focus group. MB and IL facilitated the focus groups. MB created the app and led focus groups. TH was involved in app development and led focus groups.

Conflicts of Interest

MB and TH are the board members of Atreon.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File , 181 KB - ijmr_v15i1e67523_app1.docx]

References

1. Wolfe F, Clauw DJ, Fitzcharles M, Goldenberg DL, Häuser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46(3):319-329. [doi: 10.1016/j.semarthrit.2016.08.012] [Medline: 27916278]
2. Arnold LM, Bennett RM, Crofford LJ, Dean LE, Clauw DJ, Goldenberg DL, et al. AAPT diagnostic criteria for fibromyalgia. J Pain 2019;20(6):611-628 [FREE Full text] [doi: 10.1016/j.jpain.2018.10.008] [Medline: 30453109]

XSL•FO
RenderX

3. Lichtenstein A, Tiosano S, Amital H. The complexities of fibromyalgia and its comorbidities. Curr Opin Rheumatol 2018;30(1):94. [doi: 10.1097/bor.0000000000000464]
4. Ariani A, Bazzichi L, Sarzi-Puttini P, Salaffi F, Manara M, Prevete I, et al. The Italian society for rheumatology clinical practice guidelines for the diagnosis and management of fibromyalgia best practices based on current scientific evidence. Reumatismo 2021;73(2):89-105 [FREE Full text] [doi: 10.4081/reumatismo.2021.1362] [Medline: 34342210]
5. Eich W, Häuser W, Arnold B, Bernardy K, Brückle W, Eidmann U, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Fibromyalgia syndrome. General principles and coordination of clinical care and patient education. Schmerz 2012;26(3):268-275. [doi: 10.1007/s00482-012-1167-z] [Medline: 22760459]
6. Fitzcharles MA, Ste-Marie PA, Goldenberg DL, Pereira JX, Abbey S, Choinière M, et al. 2012 Canadian guidelines for the diagnosis and management of fibromyalgia syndrome: executive summary. Pain Res Manag 2013;18(3):119-126 [FREE Full text] [doi: 10.1155/2013/918216] [Medline: 23748251]
7. Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluß E, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis 2017;76(2):318-328 [FREE Full text] [doi: 10.1136/annrheumdis-2016-209724] [Medline: 27377815]
8. Sarzi-Puttini P, Giorgi V, Marotto D, Atzeni F. Fibromyalgia: an update on clinical characteristics, aetiopathogenesis and treatment. Nat Rev Rheumatol 2020;16(11):645-660. [doi: 10.1038/s41584-020-00506-w] [Medline: 33024295]
9. Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. Lancet 2004;364(9444):1523-1537. [doi: 10.1016/S0140-6736(04)17277-2] [Medline: 15500899]
10. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]
11. Geraghty AWA, Maund E, Newell D, Santer M, Everitt H, Price C, et al. Self-management for chronic widespread pain including fibromyalgia: a systematic review and meta-analysis. PLoS One 2021;16(7):e0254642 [FREE Full text] [doi: 10.1371/journal.pone.0254642] [Medline: 34270606]
12. Häuser W, Bernardy K, Arnold B, Offenbächer M, Schiltenwolf M. Efficacy of multicomponent treatment in fibromyalgia syndrome: a meta-analysis of randomized controlled clinical trials. Arthritis Rheum 2009;61(2):216-224. [doi: 10.1002/art.24276] [Medline: 19177530]
13. Cardosa M, Osman ZJ, Nicholas M, Tonkin L, Williams A, Abd Aziz K, et al. Self-management of chronic pain in Malaysian patients: effectiveness trial with 1-year follow-up. Transl Behav Med 2012;2(1):30-37 [FREE Full text] [doi: 10.1007/s13142-011-0095-2] [Medline: 22448204]
14. Cedraschi C, Desmeules J, Rapiti E, Baumgartner E, Cohen P, Finckh A, et al. Fibromyalgia: a randomised, controlled trial of a treatment programme based on self management. Ann Rheum Dis 2004;63(3):290-296 [FREE Full text] [doi: 10.1136/ard.2002.004945] [Medline: 14962965]
15. Glenn B, Burns JW. Pain self-management in the process and outcome of multidisciplinary treatment of chronic pain: evaluation of a stage of change model. J Behav Med 2003;26(5):417-433. [doi: 10.1023/a:1025720017595] [Medline: 14593851]
16. Rod K. Finding ways to lift barriers to care for chronic pain patients: outcomes of using internet-based self-management activities to reduce pain and improve quality of life. Pain Res Manag 2016;2016:8714785 [FREE Full text] [doi: 10.1155/2016/8714785] [Medline: 27445632]
17. Turner BJ, Rodriguez N, Bobadilla R, Hernandez AE, Yin Z. Chronic pain self-management program for low-income patients: themes from a qualitative inquiry. Pain Med 2020;21(2):e1-e8. [doi: 10.1093/pm/pny192] [Medline: 30312459]
18. LeFort SM, Gray-Donald K, Rowat KM, Jeans ME. Randomized controlled trial of a community-based psychoeducation program for the self-management of chronic pain. Pain 1998;74(2-3):297-306. [doi: 10.1016/s0304-3959(97)00190-5] [Medline: 9520244]
19. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999;37(1):5-14. [doi: 10.1097/00005650-199901000-00003] [Medline: 10413387]
20. Perry J, VanDenKerkhof EG, Wilson R, Tripp DA. Development of a guided internet-based psycho-education intervention using cognitive behavioral therapy and self-management for individuals with chronic pain. Pain Manag Nurs 2017;18(2):90-101. [doi: 10.1016/j.pmn.2016.12.004] [Medline: 28262481]
21. Ruehlman LS, Karoly P, Enders C. A randomized controlled evaluation of an online chronic pain self management program. Pain 2012;153(2):319-330 [FREE Full text] [doi: 10.1016/j.pain.2011.10.025] [Medline: 22133450]
22. Stinson JN, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. iCanCope with Pain User-centred design of a web- and mobile-based self-management program for youth with chronic pain based on identified health care needs. Pain Res Manag 2014;19(5):257-265 [FREE Full text] [doi: 10.1155/2014/935278] [Medline: 25000507]
23. Tse M, Li Y, Tang SK, Ng SSM, Bai X, Lee P, et al. An exploration of the effectiveness of a peer-led pain management program (PAP) for nursing home residents with chronic pain and an evaluation of their experiences: a pilot randomized controlled trial. Int J Environ Res Public Health 2020;17(11):4090 [FREE Full text] [doi: 10.3390/ijerph17114090] [Medline: 32521785]

XSL•FO
RenderX

24. Yuan SLK, Marques AP. Development of ProFibro-a mobile application to promote self-care in patients with fibromyalgia. Physiotherapy 2018;104(3):311-317. [doi: 10.1016/j.physio.2018.04.005] [Medline: 30025715]
25. Trindade IA, Guiomar R, Carvalho SA, Duarte J, Lapa T, Menezes P, et al. Efficacy of online-based acceptance and commitment therapy for chronic pain: a systematic review and meta-analysis. J Pain 2021;22(11):1328-1342 [FREE Full text] [doi: 10.1016/j.jpain.2021.04.003] [Medline: 33892153]
26. de Lorena SB, Duarte ALBP, Bredemeier M, Fernandes VM, Pimentel EAS, Marques CDL, et al. Effects of a physical self-care support program for patients with fibromyalgia: a randomized controlled trial. J Back Musculoskelet Rehabil 2022;35(3):495-504. [doi: 10.3233/BMR-191820] [Medline: 34657869]
27. Musekamp G, Gerlich C, Ehlebracht-Kï Nig I, Dorn M, Hï Fter A, Tomiak C, et al. Evaluation of a self-management patient education programme for fibromyalgia-results of a cluster-RCT in inpatient rehabilitation. Health Educ Res 2019;34(2):209-222. [doi: 10.1093/her/cyy055] [Medline: 30689860]
28. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Børøsund E, Schreurs KMG, et al. A user-centered approach to an evidence-based electronic health pain management intervention for people with chronic pain: design and development of EPIO. J Med Internet Res 2020;22(1):e15889 [FREE Full text] [doi: 10.2196/15889] [Medline: 31961331]
29. Borrelli B, Ritterband LM. Special issue on eHealth and mHealth: challenges and future directions for assessment, treatment, and dissemination. Health Psychol 2015;34S:1205-1208. [doi: 10.1037/hea0000323] [Medline: 26651461]
30. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020;3(1):4 [FREE Full text] [doi: 10.1038/s41746-019-0206-x] [Medline: 31970289]
31. Krusche M, Ruffer N, Grahammer M, Knitza J. Mobile apps and their usage in rheumatology. Z Rheumatol 2020;79(6):554-561. [doi: 10.1007/s00393-020-00822-2] [Medline: 32472178]
32. Shetty A, Delanerolle G, Zeng Y, Shi JQ, Ebrahim R, Pang J, et al. A systematic review and meta-analysis of digital application use in clinical research in pain medicine. Front Digit Health 2022:4 [FREE Full text] [doi: 10.1101/2022.03.02.22271773]
33. Valentijn PP, Tymchenko L, Jacobson T, Kromann J, Biermann CW, AlMoslemany MA, et al. Digital health interventions for musculoskeletal pain conditions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022;24(9):e37869 [FREE Full text] [doi: 10.2196/37869] [Medline: 36066943]
34. Pfeifer AC, Uddin R, Schröder-Pfeifer P, Holl F, Swoboda W, Schiltenwolf M. Mobile application-based interventions for chronic pain patients: a systematic review and meta-analysis of effectiveness. J Clin Med 2020;9(11):3557 [FREE Full text] [doi: 10.3390/jcm9113557] [Medline: 33167300]
35. Ehde DM, Dillworth TM, Turner JA. Cognitive-behavioral therapy for individuals with chronic pain: efficacy, innovations, and directions for research. Am Psychol 2014;69(2):153-166. [doi: 10.1037/a0035747] [Medline: 24547801]
36. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
37. Goldenberg DL, Burckhardt C, Crofford L. Management of fibromyalgia syndrome. JAMA 2004;292(19):2388-2395. [doi: 10.1001/jama.292.19.2388] [Medline: 15547167]
38. Winkelmann A, Häuser W, Friedel E, Moog-Egan M, Seeger D, Settan M, et al. Physiotherapy and physical therapies for fibromyalgia syndrome. Systematic review, meta-analysis and guideline. Schmerz 2012;26(3):276-286. [doi: 10.1007/s00482-012-1171-3] [Medline: 22760460]
39. Köllner V, Häuser W, Klimczyk K, Kühn-Becker H, Settan M, Weigl M, et al. Psychotherapy for patients with fibromyalgia syndrome. Systematic review, meta-analysis and guideline. Schmerz 2012;26(3):291-296. [doi: 10.1007/s00482-012-1179-8] [Medline: 22760462]
40. Langhorst J, Häuser W, Bernardy K, Lucius H, Settan M, Winkelmann A, et al. Complementary and alternative therapies for fibromyalgia syndrome. Systematic review, meta-analysis and guideline. Schmerz 2012;26(3):311-317. [doi: 10.1007/s00482-012-1178-9] [Medline: 22760464]
41. Albuquerque MLL, Monteiro D, Marinho DA, Vilarino GT, Andrade A, Neiva HP. Effects of different protocols of physical exercise on fibromyalgia syndrome treatment: systematic review and meta-analysis of randomized controlled trials. Rheumatol Int 2022;42(11):1893-1908. [doi: 10.1007/s00296-022-05140-1] [Medline: 35604435]
42. Bidonde J, Busch AJ, Schachter CL, Overend TJ, Kim SY, Góes SM, et al. Aerobic exercise training for adults with fibromyalgia. Cochrane Database Syst Rev 2017(6):CD012700. [doi: 10.1002/14651858.cd012700]
43. Bidonde J, Busch AJ, Schachter CL, Webber SC, Musselman KE, Overend TJ, et al. Mixed exercise training for adults with fibromyalgia. Cochrane Database Syst Rev 2019;5(5):CD013340 [FREE Full text] [doi: 10.1002/14651858.CD013340] [Medline: 31124142]
44. Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. Cochrane Database Syst Rev 2013;2013(12):CD010884 [FREE Full text] [doi: 10.1002/14651858.CD010884] [Medline: 24362925]
45. Busch AJ, Barber KAR, Overend TJ, Peloso PMJ, Schachter CL. Exercise for treating fibromyalgia syndrome. Cochrane Database Syst Rev 2007(4). [doi: 10.1002/14651858.cd003786]

XSL•FO
RenderX

46. Couto N, Monteiro D, Cid L, Bento T. Effect of different types of exercise in adult subjects with fibromyalgia: a systematic review and meta-analysis of randomised clinical trials. Sci Rep 2022;12(1):10391 [FREE Full text] [doi: 10.1038/s41598-022-14213-x] [Medline: 35725780]
47. da Silva JM, de Barros BS, Almeida GJ, O'Neil J, Imoto AM. Dosage of resistance exercises in fibromyalgia: evidence synthesis for a systematic literature review up-date and meta-analysis. Rheumatol Int 2022;42(3):413-429. [doi: 10.1007/s00296-021-05025-9] [Medline: 34652480]
48. Kim SY, Busch AJ, Overend TJ, Schachter CL, van der Spuy I, Boden C, et al. Flexibility exercise training for adults with fibromyalgia. Cochrane Database Syst Rev 2019;9(9):CD013419 [FREE Full text] [doi: 10.1002/14651858.CD013419] [Medline: 31476271]
49. Bernardy K, Klose P, Busch AJ, Choy EHS, Häuser W. Cognitive behavioural therapies for fibromyalgia. Cochrane Database Syst Rev 2013;2013(9):CD009796 [FREE Full text] [doi: 10.1002/14651858.CD009796.pub2] [Medline: 24018611]
50. de Oliveira Lima L, Saragiotto B, Costa L, Nogueira L, Meziat-Filho N, Reis F. Self-guided web-based pain education for people with musculoskeletal pain: a systematic review and meta-analysis. Phys Ther 2021;101(10):pzab167. [doi: 10.1093/ptj/pzab167] [Medline: 34174081]
51. Du S, Dong J, Jin S, Zhang H, Zhang Y. Acceptance and commitment therapy for chronic pain on functioning: a systematic review of randomized controlled trials. Neurosci Biobehav Rev 2021;131:59-76. [doi: 10.1016/j.neubiorev.2021.09.022] [Medline: 34536462]
52. Enomoto K, Adachi T, Fujino H, Kugo M, Tatsumi S, Sasaki J. Comparison of the effectiveness of cognitive behavioral therapy for insomnia, cognitive behavioral therapy for pain, and hybrid cognitive behavioral therapy for insomnia and pain in individuals with comorbid insomnia and chronic pain: a systematic review and network meta-analysis. Sleep Med Rev 2022;66:101693. [doi: 10.1016/j.smrv.2022.101693] [Medline: 36410076]
53. Gandy M, Pang STY, Scott AJ, Heriseanu AI, Bisby MA, Dudeney J, et al. Internet-delivered cognitive and behavioural based interventions for adults with chronic pain: a systematic review and meta-analysis of randomized controlled trials. Pain 2022;163(10):e1041-e1053. [doi: 10.1097/j.pain.0000000000002606] [Medline: 35121696]
54. García-Ríos MC, Navarro-Ledesma S, Tapia-Haro RM, Toledano-Moreno S, Casas-Barragán A, Correa-Rodríguez M, et al. Effectiveness of health education in patients with fibromyalgia: a systematic review. Eur J Phys Rehabil Med 2019;55(2):301-313 [FREE Full text] [doi: 10.23736/S1973-9087.19.05524-2] [Medline: 30698402]
55. Joypaul S, Kelly F, McMillan SS, King MA. Multi-disciplinary interventions for chronic pain involving education: a systematic review. PLoS One 2019;14(10):e0223306 [FREE Full text] [doi: 10.1371/journal.pone.0223306] [Medline: 31577827]
56. Louw A, Zimney K, Puentedura EJ, Diener I. The efficacy of pain neuroscience education on musculoskeletal pain: a systematic review of the literature. Physiother Theory Pract 2016;32(5):332-355. [doi: 10.1080/09593985.2016.1194646] [Medline: 27351541]
57. Ma TW, Yuen ASK, Yang Z. The efficacy of acceptance and commitment therapy for chronic pain: a systematic review and meta-analysis. Clin J Pain 2023;39(3):147. [doi: 10.1097/ajp.0000000000001096]
58. Romm MJ, Ahn S, Fiebert I, Cahalin LP. A meta-analysis of therapeutic pain neuroscience education, using dosage and treatment format as moderator variables. Pain Pract 2021;21(3):366-380. [doi: 10.1111/papr.12962] [Medline: 33131210]
59. Veehof MM, Trompetter HR, Bohlmeijer ET, Schreurs KMG. Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. Cogn Behav Ther 2016;45(1):5-31 [FREE Full text] [doi: 10.1080/16506073.2015.1098724] [Medline: 26818413]
60. Williams ACC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2020;8(8):CD007407 [FREE Full text] [doi: 10.1002/14651858.CD007407.pub4] [Medline: 32794606]
61. Cheng CA, Chiu YW, Wu D, Kuan YC, Chen SN, Tam KW. Effectiveness of Tai Chi on fibromyalgia patients: a meta-analysis of randomized controlled trials. Complement Ther Med 2019;46:1-8. [doi: 10.1016/j.ctim.2019.07.007] [Medline: 31519264]
62. Langhorst J, Klose P, Dobos GJ, Bernardy K, Häuser W. Efficacy and safety of meditative movement therapies in fibromyalgia syndrome: a systematic review and meta-analysis of randomized controlled trials. Rheumatol Int 2013;33(1):193-207. [doi: 10.1007/s00296-012-2360-1] [Medline: 22350253]
63. Lee C, Crawford C, Schoomaker E. Movement therapies for the self-management of chronic pain symptoms. Pain Med 2014;15(S1):S40-S53. [doi: 10.1111/pme.12411] [Medline: 24734859]
64. Lee C, Crawford C, Hickey A. Mind-body therapies for the self-management of chronic pain symptoms. Pain Med 2014;15(S1):S21-S39. [doi: 10.1111/pme.12383] [Medline: 24734857]
65. Theadom A, Cropley M, Smith HE, Feigin VL, McPherson K. Mind and body therapy for fibromyalgia. Cochrane Database Syst Rev 2015;2015(4):CD001980 [FREE Full text] [doi: 10.1002/14651858.CD001980.pub3] [Medline: 25856658]
66. Hernando-Garijo I, Jiménez-del-Barrio S, Mingo-Gómez T, Medrano-de-la-Fuente R, Ceballos-Laita L. Effectiveness of non-pharmacological conservative therapies in adults with fibromyalgia: a systematic review of high-quality clinical trials. J Back Musculoskelet Rehabil IOS Press 2022;35(1):3-20. [doi: 10.3233/bmr-200282]

XSL•FO
RenderX

67. Kundakci B, Kaur J, Goh SL, Hall M, Doherty M, Zhang W, et al. Efficacy of nonpharmacological interventions for individual features of fibromyalgia: a systematic review and meta-analysis of randomised controlled trials. Pain 2022;163(8):1432-1445. [doi: 10.1097/j.pain.0000000000002500] [Medline: 34813518]
68. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ British Medical Journal 2000;321(7262):694-696 [FREE Full text] [doi: 10.1136/bmj.321.7262.694] [Medline: 10987780]
69. American COSM, Liguori G, Feito Y, Fountaine C, Roy B. ACSM's Guidelines for Exercise Testing and Prescription. 11th ed. Philadelphia: Wolters Kluwer; 2021.
70. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
71. Moseley GL, Butler DS. Fifteen years of explaining pain: the past, present, and future. J Pain 2015;16(9):807-813 [FREE Full text] [doi: 10.1016/j.jpain.2015.05.005] [Medline: 26051220]
72. National Library of Medicine. NCBI Mesh. NCBI NLM NIH. URL: https://www-ncbi-nlm-nih-gov.ezproxy.is.ed.ac.uk/mesh/ [accessed 2023-06-24]
73. Falci L, Shi Z, Greenlee H. Multiple chronic conditions and use of complementary and alternative medicine among US adults: results from the 2012 national health interview survey. Prev Chronic Dis 2016;13:E61 [FREE Full text] [doi: 10.5888/pcd13.150501] [Medline: 27149072]
74. Miles CL, Pincus T, Carnes D, Homer KE, Taylor SJC, Bremner SA, et al. Can we identify how programmes aimed at promoting self-management in musculoskeletal pain work and who benefits? A systematic review of sub-group analysis within RCTs. Eur J Pain 2011;15(8):775.e1-775.11. [doi: 10.1016/j.ejpain.2011.01.016] [Medline: 21354838]
75. Chew MT, Chan C, Kobayashi S, Cheng HY, Wong TM, Nicholson LL. Online pain management programs for chronic, widespread musculoskeletal conditions: a systematic review with meta-analysis. Pain Pract 2023;23(6):664-683. [doi: 10.1111/papr.13227] [Medline: 37051894]
76. DIGAs. Bundesinstitut für Arzneimittel und Medizinprodukte. URL: https://diga.bfarm.de/de [accessed 2024-06-16]
77. Gendreau RM, McCracken LM, Williams DA, Luciano JV, Dai Y, Vega N, et al. Self-guided digital behavioural therapy versus active control for fibromyalgia (PROSPER-FM): a phase 3, multicentre, randomised controlled trial. Lancet 2024;404(10450):364-374. [doi: 10.1016/S0140-6736(24)00909-7] [Medline: 38991582]
78. Blanchard M, Koller CN, Azevedo PM, Prétat T, Hügle T. Development of a management app for postviral fibromyalgia-like symptoms: patient preference-guided approach. JMIR Form Res 2024;8:e50832 [FREE Full text] [doi: 10.2196/50832] [Medline: 38639986]
79. Blanchard M, Venerito V, Ming Azevedo P, Hügle T. Generative AI-based knowledge graphs for the illustration and development of mHealth self-management content. Front Digit Health 2024;6:1466211 [FREE Full text] [doi: 10.3389/fdgth.2024.1466211] [Medline: 39434919]
80. Remedee Labs. URL: https://remedeelabs.com/en/ [accessed 2025-12-13]
81. Schmajuk G, Nasrallah C, Berrean B, Prugh J, Wilson C, Hamblin A, et al. A step-by-step roadmap for the development and deployment of an electronic health record sidecar application that tracks patient outcomes: the RA PRO dashboard. Digit Health 2024;10:20552076241288739 [FREE Full text] [doi: 10.1177/20552076241288739] [Medline: 39421306]

Abbreviations

ACT: acceptance and commitment therapy
AI: artificial intelligence
CBT: cognitive behavioral therapy
DiGA: Digitale Gesundheitsanwendung
FIQ: Fibromyalgia Impact Questionnaire
LLM: large language model
mHealth: mobile health
MMT: meditative movement therapy
QoL: quality of life

XSL•FO
RenderX

Edited by N Cahill; submitted 16.Oct.2024; peer-reviewed by A D'Silva, Q Ye; comments to author 23.Dec.2024; accepted 16.Nov.2025; published 17.Feb.2026.

Please cite as:
Lebeau Foustoukos T, Leclercq I, Blanchard M, Hügle T
Evidence-Based Self-Management Strategies for Fibromyalgia: Foundations for Digital Therapeutic Applications
Interact J Med Res 2026;15:e67523
URL: https://www.i-jmr.org/2026/1/e67523
doi:10.2196/67523
PMID:

©Thomas Lebeau Foustoukos, Isabelle Leclercq, Marc Blanchard, Thomas Hügle. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 17.Feb.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Serious Games in Supporting Postnatal Depression: Narrative Review

Weijie Wang[1], MDes; Erika Penney[2], DClinPsy; Valerie Gay[3], PhD; Jaime Garcia[1], PhD

[1]School of Computer Science, Faculty of Engineering and IT (Information Technology), University of Technology Sydney, 15 Broadway, Ultimo, Sydney, Australia

[2]Graduate School of Health, Faculty of Health, University of Technology Sydney, Sydney, Australia

[3]School of Electrical and Data Engineering, Faculty of Engineering and IT (Information Technology), University of Technology Sydney, Sydney, Australia

Corresponding Author:
Jaime Garcia, PhD
School of Computer Science, Faculty of Engineering and IT (Information Technology), University of Technology Sydney, 15 Broadway, Ultimo, Sydney, Australia

Abstract

Background: Postnatal depression (PND) is a clinical sign of sadness in certain individuals after childbirth. PND affects the mother, the baby, and the whole family. PND is now recognized as a public health concern worldwide. The global prevalence of PND is approximately 17.22%. However, less than half of those affected seek help, which means over 50% of PND cases are left untreated. Current reviews lack focus on digital interventions targeting parents in late pregnancy or postnatal stages. Existing studies prioritize symptom relief over fostering help-seeking behaviors.

Objective: This study aims to identify what serious games have been applied to support the treatment or help-seeking of PND and what gaps are still left.

Methods: Eligibility criteria for this review included full-text papers from 2015 to 2024 from conferences or peer-reviewed journals that were relevant to serious games to support help-seeking behaviors for individuals with depression. Seven research databases and publisher repositories were used. The final search was conducted in March 2025, and a thematic analysis was used to identify and organize recurring themes. As this review adopts a narrative approach, predefined eligibility criteria, a structured search strategy, and review by an interprofessional team were used to reduce selection bias.

Results: Only 2 studies related to PND were identified. After expanding the search string to depression, 13 studies were included in this review, and the studied games were divided into 3 help-seeking categories: promoting knowledge, reducing stigma, and raising awareness. This review identified that gamification, educational messages, and supportive character interactions could enhance engagement, build coping skills, and promote help-seeking in a practical, parent-friendly format. Nonetheless, this paper is limited by the reliance on depression literature due to scarce PND-specific studies, the quality of included studies, the exclusion of non-English language publications, and the use of common but select academic databases. These factors may affect generalizability but also serve to highlight critical gaps for future research and targeted intervention design.

Conclusions: There is a dearth of studies directly related to PND. Existing games commonly use narrative storytelling and interactive scenarios to promote empathy, correct misconceptions, and encourage help-seeking in broad depression. However, few are designed specifically for new parents, whose unique needs—such as time constraints—make mobile platforms the most suitable format for effective engagement. The authors propose that the future interprofessional codevelopment of a mobile serious game tailored to new parents would address the intervention and literature gaps identified in this review. It is argued that key design elements should include an emotionally engaging narrative, meaningful player choices, real-life parenting scenarios, calming visuals, and accessible, low-pressure gameplay. This review contributes to the progression of serious game research, with a focus on addressing the needs of an often underserved and undertreated PND population.

(Interact J Med Res 2026;15:e70777) doi:10.2196/70777

KEYWORDS

serious game; postnatal depression; digital health; mobile phone; perinatal education; mental health

XSL•FO
RenderX

Introduction

Postnatal Depression: Definition, Consequences, and Treatment

Postnatal depression (PND) is a clinical diagnosis characterized by sadness after the birth of a child [1]. Most cases of PND start between 4 and 8 weeks after childbirth [2]. However, some research suggests that health care providers should screen for PND for up to 3 years after giving birth. This finding is in light of evidence that PND can develop after the traditional 6-month mark in some cases [3]. PND can result in various adverse maternal outcomes, impacting physical health, psychological well-being, interpersonal relationships, and risky behaviors [4]. Moreover, women who experience PND face an increased likelihood of experiencing clinical depression in the future, not just following repeated pregnancies but also throughout other periods of their lives [5].

PND affects not only the mother but the baby and family as a whole [6]. A mother experiencing PND may have difficulty bonding with her infant, providing proper care such as breastfeeding, and playing other maternal roles [1]. The potential consequences of this situation may extend to the child's physical, cognitive, language, emotional, social, and behavioral development and even impact the dynamic within the family unit over an extended period of time [7].

Perinatal mental health issues are now recognized as a significant public health concern all over the world. The global prevalence of PND was found to be approximately 17.22% in the largest meta-analysis of PND to date [8]. It is estimated that perinatal mental health disorders cost the Australian economy Aus $877 million per year (a currency exchange rate of Aus $1=US $0.66 was applicable) [9], the United States economy US $14.20 billion per year [10], and the United Kingdom £8.10 billion each year (a currency exchange rate of UK £1=US $1.34 was applicable) [11].

Health-Seeking Behaviors and PND

However, less than half of those affected seek help, which means over 50% of PND is left untreated [12]. The barriers to help-seeking for PND are various, including social and cultural stigma, lack of knowledge and awareness, and lack of accessibility to the health care system [13]. As highlighted, PND can have a detrimental effect on the entire family. Thus, further research is needed on the most effective ways to empower expectant new parents to seek help earlier and maintain positive mental health, ensuring this condition does not remain a silent struggle.

Digital Interventions and Management of PND

Digital games are interactive, computer-based experiences in which players interact with a virtual environment through input devices such as keyboards, controllers, or touchscreens. They range from simple, text-based applications to complex, immersive worlds powered by advanced graphics and artificial intelligence. Digital games can be played on multiple platforms, including personal computers, gaming consoles, smartphones, and virtual reality (VR) systems.

Game development is a complex process that often leverages specialized tools to streamline workflows and enhance functionality. Two critical components in this ecosystem are middleware solutions and game engines. Middleware provides developers with prebuilt modules to handle specific functionalities, while game engines offer comprehensive frameworks for building games.

Serious games play a crucial role in various domains by blending entertainment with purposeful outcomes, making them valuable tools for education, training, health care, and social awareness [14]. Unlike traditional video games designed solely for recreation, serious games leverage interactive and immersive elements to enhance learning, skill development, and behavioral change. Engaging users in simulated environments provides a risk-free space for experimentation and decision-making, leading to deeper comprehension and improved information retention. This kind of approach makes them particularly effective in fields such as education, where gamified learning experiences can make complex subjects more accessible and engaging for students. In professional training, serious games are widely adopted for skill development and simulation-based learning [15]. Industries such as health care, aviation, and military training use serious games to replicate real-world scenarios, allowing trainees to practice in a controlled setting. For example, medical simulation games help doctors and nurses refine surgical techniques and patient care by providing a safe space rather than performing procedures on real patients [16]. Similarly, corporate training programs integrate serious games to make the traditional learning experience more engaging and dynamic, thereby improving employee knowledge retention and job performance [17].

Beyond education and training, serious games significantly impact mental health and social awareness. Games designed to address depression can provide therapeutic benefits and be valuable and effective in reducing symptoms of depressive disorders [18,19]. Additionally, serious games are used to promote awareness about global challenges, such as climate change and poverty, by immersing players in experiences that foster empathy and drive real-world action [20]. By engaging storytelling, decision-making challenges, and interactive problem-solving, serious games go beyond traditional education and advocacy methods, making complex issues more relatable and actionable for diverse audiences.

In conclusion, serious games are an essential innovation in digital technology, bridging the gap between entertainment and meaningful engagement. Their ability to educate, train, and raise awareness while keeping users involved makes them a powerful tool across various industries. As technology continues to evolve, the potential for serious games to drive positive change in education, health care, and social development will only continue to grow. This narrative review aims to provide a comprehensive and interpretative narrative synthesis of the existing literature on serious games and digital interventions for PND. The goal is to critically evaluate current trends and gaps to explain the need for targeted research in this essential but often overlooked area.

XSL•FO
RenderX

Knowledge Gap and Objective

There is literature from the past 10 years reviewing digital interventions for symptoms of PND. In 2018, van den Heuvel et al [21] published a systematic review of the current literature on eHealth developments in pregnancy to assess this new generation of perinatal care. Studies that reported using eHealth during prenatal, perinatal, and postnatal timeframes include 71 studies covering 6 domains. These domains are information and eHealth use (eg, patients' use of the internet for pregnancy information [22]), lifestyle factors such as gestational weight gain, exercise, and smoking cessation (eg, technology-supported diet and lifestyle interventions [23]), gestational diabetes (eg, telemedicine for diabetes in pregnancy [24]), mental health (eg, therapist-supported internet-based cognitive behavior therapy among postnatal women [25]), low- and middle-income countries (eg, mobile health [mHealth] interventions for prenatal, birth, and postnatal period in low- and middle-income countries [26]), and telemonitoring and teleconsulting (eg, wireless antepartum maternal-fetal monitoring [27]). There were 16 studies related to screening and treatment for PND (van den Heuvel et al [21], 2018).

Hussain-Shamsy et al [28], 2020, aimed to understand the extent, range, and nature of mHealth tools for prevention, screening, and treatment of perinatal depression and anxiety in order to identify gaps and inform opportunities for future work. Compared to the van den Heuvel et al [21] review, it targeted interventions delivered through mobile phones, including apps and text message–based interventions with prevention, screening, and treatment purposes. Tools were for prevention (10/22, 45%), screening (6/22, 27%), and treatment (6/22, 27%). Interventions delivered included psycho-education (16/22, 73%), peer support (4/22, 18%), and psychological therapy (4/22, 18%); however, interventions that started in pregnancy and continued into the postnatal period were rare (2/22, 9%).

In the same year, Dosani et al [29] investigated existing uses of mobile phone technologies for perinatal depression in low- and middle-income countries, finding improved depressive symptoms after the interventions. Similarly, Wan Mohd Yunus et al [30], 2022, found positive outcomes for digitalized cognitive behavioral therapy (CBT) interventions for depression symptoms during pregnancy. However, it should be noted that the studies within this systematic review demonstrated relatively high risks of bias and some missing outcome data. Lau et al [31] (2022) also conducted a review synthesizing 18 randomized controlled trials and demonstrating support for the effectiveness of digital CBT for perinatal psychology symptoms (ie, depression, anxiety, and stress symptoms) in high-income countries; however, again, the included studies were found to be of limited quality. It is argued that the literature is promising in its indicative efficacy for digital and mobile interventions for perinatal mental health; however, further robust, high-quality, larger sample size, and interprofessionally co-designed randomized controlled trial research between health care professionals and application developers is required in this field.

Beyond the more common digital and mobile interventions, VR has been explored for its utility in this field. In 2024, Fallon et al [32] conducted a scoping review on the use of VR to support parents during birth and in the first year postbirth in different settings, finding that across these studies, VR was found to be effective in improving both physiological and psychological outcomes. Furthermore, mothers reported positive experiences of using VR, which could indicate VR's acceptability in this population.

Literature has looked not only at interventions targeting symptoms but also at prevention. Some have assessed web- and mobile-based psychological interventions' role in preventing depression during the perinatal period [33], as well as reviews on the current state of diagnostic and screening apps for perinatal mental health [34]. These review studies highlight important considerations for future research, including the need for comprehensive digital assessment tools, ensuring data protection and safety of the intended app use, and improving data sharing features between users and health care professionals for timely support.

Other prevention-focused research has aimed at increasing partner support (Pilkington et al [35], 2015). Several prevention studies that include a partner component have demonstrated some benefits [35]. However, not all of these interventions were delivered to both mothers and fathers, and research evaluating their effects on paternal mental health is lacking. Thus, it is argued that future research needs to focus more on developing active interventions for both partners.

In the past decade, numerous reviews have focused on digital interventions for perinatal mental health, particularly PND. Each of these reviews had distinct aims, ranging from examining eHealth developments during pregnancy [21] to evaluating the effectiveness of mHealth tools [28] and digital CBT interventions [30]. Others, such as Fallon et al [32], explored emerging technologies such as VR for supporting parents during the postnatal period. These reviews covered a variety of tools and populations, including both high- and low-income countries, and targeted different stages of perinatal care. Efficacious tools are important. However, those with perinatal mental health symptoms have demonstrated delays in seeking those interventions [36]. To the best of the authors' knowledge, there are limited review papers focused on how digital interventions can improve help-seeking behaviors in perinatal depression, highlighting a significant literature gap with important clinical and economic implications for undertreatment.

There is a paucity of research that explores how the games are designed for PND. However, there is a game-specific exploration of depression symptoms more generally, which is of relevance given that PND and major depression share identical symptoms except for onset [37]. In 2014, there were 2 relevant review papers. One is the systematic review by Fleming et al [38] of the evidence for serious games in the treatment or prevention of depression. The current data suggest that it is possible to develop serious games for depression, that young people are willing to try them, and that available adherence and impact data, whilst limited, are promising. The other systematic review is Li et al [39] systematically examining the effectiveness of game-based digital interventions for depression, investigating psycho-education and training, VR Exposure Therapy, exercise, and entertainment.

XSL•FO
RenderX

Psycho-education and training and VR Exposure Therapy were identified as the most popular types of game applications for depression. Given the demonstrated potential gains for global access, it becomes even more crucial to meet the needs of those reluctant to seek help through current methods.

To the best knowledge of the authors, there is only one relevant review from 2018 (Dias et al [40]). Dias et al [40] aimed to identify through a systematic review how gamification and serious games support depression treatment and identified two significant literature gaps, including the limited study of (1) the effectiveness of gamification and serious games for depression and (2) hazard assessment of the side effects of using gamification and serious games in depression treatment. The technologies identified in this review included mobile, computer, wearable, and web applications that were applied in gamification, serious games, VR, and speech analysis. This area has not been revisited since this 2018 review, as far as the authors are aware, yet in recent years, there has been an increasing interest in serious games for depression and health problems more broadly [41].

A systematic review of casual video games (CVGs) by Pine et al [42] provided a systematic review of the effects of CVGs on treating anxiety, depression, stress, and low mood. This work includes 9 CVGs, of which 6 studies aimed at reducing anxiety, 2 studies examined effects for depression, and 4 studies investigated the impact of CVGs on treating stress or low mood. CVGs are becoming increasingly popular, and people report playing them for various reasons, such as relieving stress and relaxing, which may provide some initial benefit for those who need additional support or are waiting to access more comprehensive treatment.

In 2021, a total of 4 relevant works on games for depression and border mental health concerns were published in Australia, Ireland, Austria, and Taiwan. King et al [43] narratively reviewed indie—or independent—games that address mental health, trauma, and grief, highlighting developers' difficulties due to insufficient game design and research studies when creating, developing, and evaluating serious games. Indie games were initially defined as games made on low budgets by small teams that were published outside more mainstream channels and used by larger companies. There are many challenges when developing serious games, including limited funding, small teams, a need for broad expertise, and limited examples for developing both games and studies. Serious mental health games that target those with mental illness directly face additional challenges, such as mental health disorder symptoms that can affect the present ability to engage with content. This challenge means developing mental health serious games presents unique challenges to developers, who must decide how to best address their intended audience.

Differing from previous work that targeted independent games, Kowal et al [41] focused on the mental health benefits associated with playing commercial video games to address symptoms of depression and anxiety. In light of the current research, we conclude that commercial video games show great promise as inexpensive, readily accessible, internationally available, practical, and stigma-free resources for mitigating some mental health issues in the absence of, or in addition to, traditional therapeutic treatments. Aside from the game target general public, Martinez et al [44] investigated serious games published between 2015 and 2020 for depression and anxiety in children and adolescents, with a new approach focusing on their applications: awareness, prevention, detection, and therapy. The results of this systematic review show that more awareness and detection games and games with awareness, prevention, detection, and therapy applications are needed. In addition, games for depression and anxiety should equally target all age ranges. For future research, developing and evaluating serious games should be standardized, and the games should always offer support while playing. Meanwhile, Yen and Chiu [45] conducted a review to explore the effectiveness of VR exergames in improving older adults' cognition and ameliorating depressive outcomes. This study suggests that VR exergames can potentially positively influence cognition, memory, and depression in older adults. VR exergames could be an interesting strategy for active aging and a good mental health status.

In 2022, there were 5 key sources. First, Abd-Alrazaq et al [46] aimed to assess the effectiveness of serious games in alleviating depression by summarizing and pooling the results of previous studies. Their findings indicate that exergames are as effective as active interventions, which are usually delivered and supervised by health care providers. Meanwhile, Ruiz et al [18] systematically reviewed the evidence and found video game-based interventions were valuable and effective in reducing symptoms of depressive disorders. Additionally, Kim et al [47] conducted a systematic review, finding that serious games were beneficial in reducing depression in older adults. Regardless of the study setting, serious games appeared to reduce depression. Particularly, serious games, including physical activities, had a significant impact on reducing depression.

Concerning combined depression and anxiety-related symptoms, Townsend et al [19] conducted a review into the effectiveness of gaming interventions for treating either depression or anxiety in individuals aged 12 - 25 years. Preliminary evidence suggests that gaming interventions may be an effective treatment for youth depression but not anxiety. Further research is warranted to establish the utility, acceptability, and effectiveness of gaming interventions in treating mental health problems in young people.

The fifth relevant work of this year was published by Chitale et al [48], reviewing the viability of games and VR for the assessment of anxiety and depression, finding that possible digital correlates or biomarkers of depression and anxiety could help researchers with their design. It is important to emphasize that to ensure safety, efficacy, and privacy at a health care standard. It is argued that game developers and researchers must collaborate interprofessionally with qualified mental health specialists. More clinical data is necessary to further evidence the effective use of video games or VR in assessment methods for anxiety and depression.

In 2023, Gliosci et al [49] systematically reviewed the current array of scientific research on video games used as a therapeutic intervention tool for depression. The latest research on the use of video games as depression treatment methods shows that

XSL•FO
RenderX

there are significant gaps in the sector, particularly across generations. The age group that is most affected by the condition and now consumes the most video games—adults aged 18 - 40 years—has had limited focus in the game and mental health literature. Given that depressed people tend to be more isolated, there may be untapped possibilities, such as the unique way that games can bring people together, that could be used in different types of interventions.

Most recently, in 2025, Gómez-León [50] systematically reviewed the current serious games designed to train emotional regulation skills in children and adolescents, finding that serious games can be effective, acceptable, and feasible for learning emotional regulation strategies and reducing symptoms related to depression, anxiety, and lack of impulse control.

As highlighted, the current reviews have several limitations and clinically significant research gaps that need to be addressed. Existing reviews rarely address digital interventions specifically targeting parents in late pregnancy or postnatal stages. Most studies focus on adolescents, older people, or the general population with depression. Given the unique physiological and psychological needs of a perinatal mental health group and the common reluctance to seek help, it is crucial to develop targeted digital games and help-seeking interventions for new parents, especially during late pregnancy and the postnatal period.

Current reviews focus primarily on symptom relief rather than fostering help-seeking behaviors. For new parents, it is essential to explore how digital games can educate and provide psychological support to help them recognize symptoms of depression and empower them to seek professional help. Crucially, the physiological and psychological changes that occur during pregnancy and the postnatal period differ from those experienced by the general population with depression, often influenced by hormonal shifts and changes in life roles [1]. Existing reviews lack a deep exploration of how these specific mechanisms impact the effectiveness of gamified interventions, and therefore, this will be a particular aim of this review. Game design must consider these unique psychological and physiological changes to support the mental health of new parents effectively.

Although serious games have shown promise in supporting mental health, few have been tailored to the unique needs of new parents facing PND. Existing interventions often overlook practical constraints such as time limitations and emotional vulnerability. Given this need, the current review aims to explore how game-based tools can be thoughtfully designed to provide emotional support, foster help-seeking, and offer accessible education during this critical period.

There is growing recognition that serious games could offer an accessible, flexible, and emotionally supportive solution for broad mental health problems. This narrative review provides a comprehensive and interpretative narrative synthesis of the existing literature on serious games and digital interventions for PND and, where data is limited, extends its review to related literature on depression, given its almost identical diagnostic profile with PND. The aim is to critically assess prevailing trends and gaps and to inform the need for targeted research in key, often overlooked, areas. By addressing these aspects, we aim to make future direction recommendations for research and game developments that could cater to the unique needs of parents in late pregnancy and postnatal stages, particularly in raising awareness of help-seeking behaviors toward PND. This review aims to highlight how digital games can be more effective in supporting this specific underserved population.

Methods

Search Strategy and Study Selection

The literature search was conducted across 7 research databases and publisher repositories: PubMed Central, the ACM Digital Library, Google Scholar, the IEEE Xplore Digital Library, Science Direct, the Springer Library, and Wiley. Keywords included combinations of major terms such as “serious games,” “digital games,” “postnatal depression,” and “help-seeking.”

The resulting search string is as follows and shown in Table 1: ((serious game OR serious games OR gamification OR gamified OR game design OR game based OR game-based OR gaming) AND (postnatal depression OR postnatal depression OR perinatal depression OR maternal depression OR post-birth depression)). For the detailed search strategy, please refer to Multimedia Appendix 1.

Table . Search string.

Major terms	Synonyms
Postnatal depression	postnatal depression OR perinatal depression OR maternal depression OR post-birth depression
Serious game	serious games OR gamification OR gamified OR game design OR game based OR game-based OR gaming

The search covered literature published between 2015 and 2024 and was limited to studies published in English. We chose to only include articles published after 2015, considering relevance and time efficiency.

Given the only 2 game research specific to PND, we also chose to include depression as an analogous disorder for inclusion because, according to the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition) [51], major depressive disorder and PND share core diagnostic symptoms. Thus, the broadened search string is as follows: ((serious game OR serious games OR gamification OR gamified OR game design OR game based OR game-based OR gaming) AND (depression OR sadness OR depressive))

Initial search results were screened based on titles and abstracts by author WW. Then, full-text reviews were conducted by author WW and decided across the team, EP and JG, to

XSL•FO
RenderX

determine their relevance to the research objectives. Reference lists of key articles were also reviewed to identify additional sources. The screening process is shown in the Results section.

Studies were included if they discussed the design, application, or evaluation of serious games related to promoting help-seeking behaviors toward depression or perinatal mental health. Articles focusing solely on maternal health without mention of game-based interventions were excluded.

Sufficiency Statement

The authors acknowledge that their disciplinary backgrounds and research interests in digital mental health and game design may have influenced the interpretation and prioritization of specific themes during the analysis. Author WW is a PhD candidate who has experience in interaction design, data analysis, and software development. Author EP is a senior lecturer in the faculty of health. She has published reviews and empirical research in the field of mental health. She is also a clinical psychologist practitioner working with vulnerable and mentally unwell clients and families, including clients with PND. Author JG is a senior lecturer in games development, an experienced software engineer, and a serious game designer working with diverse clients. A particular emphasis was placed on design elements that support emotional engagement and help-seeking behaviors, which aligns with the authors' focus on developing interventions for new parents experiencing PND. Additionally, the narrative nature of the review allows for interpretive flexibility, which, while beneficial for exploring emerging ideas, may introduce subjective bias in theme identification and synthesis. The selection of literature in English only and a focus on peer-reviewed sources may have also limited the inclusion of culturally diverse perspectives and nontraditional forms of evidence.

Analysis Approach

To guide the synthesis and interpretation of findings across the selected literature, this review adopted a thematic analysis approach. Commonly used in qualitative research, thematic analysis is a systematic method for identifying, analyzing, and interpreting patterns of meaning—referred to as themes—within textual data [52]. It is particularly appropriate for narrative reviews, as it allows for the integration of diverse sources and offers a deeper understanding of complex issues across varied contexts. The analysis followed the 6-step process [53]. Initially, all included texts were read thoroughly by author WW to ensure familiarity with the content. From there, initial codes were generated inductively by author WW to capture recurring ideas, concepts, and relevant patterns. These codes were then reviewed by the team (EP and JG) and grouped into broader themes that reflected meaningful trends across the dataset. Each theme was refined and decided across the team (WW, EP, and JG) to ensure clarity, distinctiveness, and relevance to the review's aims. The final step was author WW weaving the themes into a cohesive narrative that highlighted their significance in relation to the central research questions.

By using thematic analysis, this review was able to go beyond a simple summary, offering a structured yet flexible framework to interpret findings. A thematic narrative approach was considered appropriate due to the heterogeneity of the sources, which varied in methodology, population focus, and design strategy. This flexibility allowed for a deeper exploration of how serious games function as educational and emotional tools, especially in the context of PND.

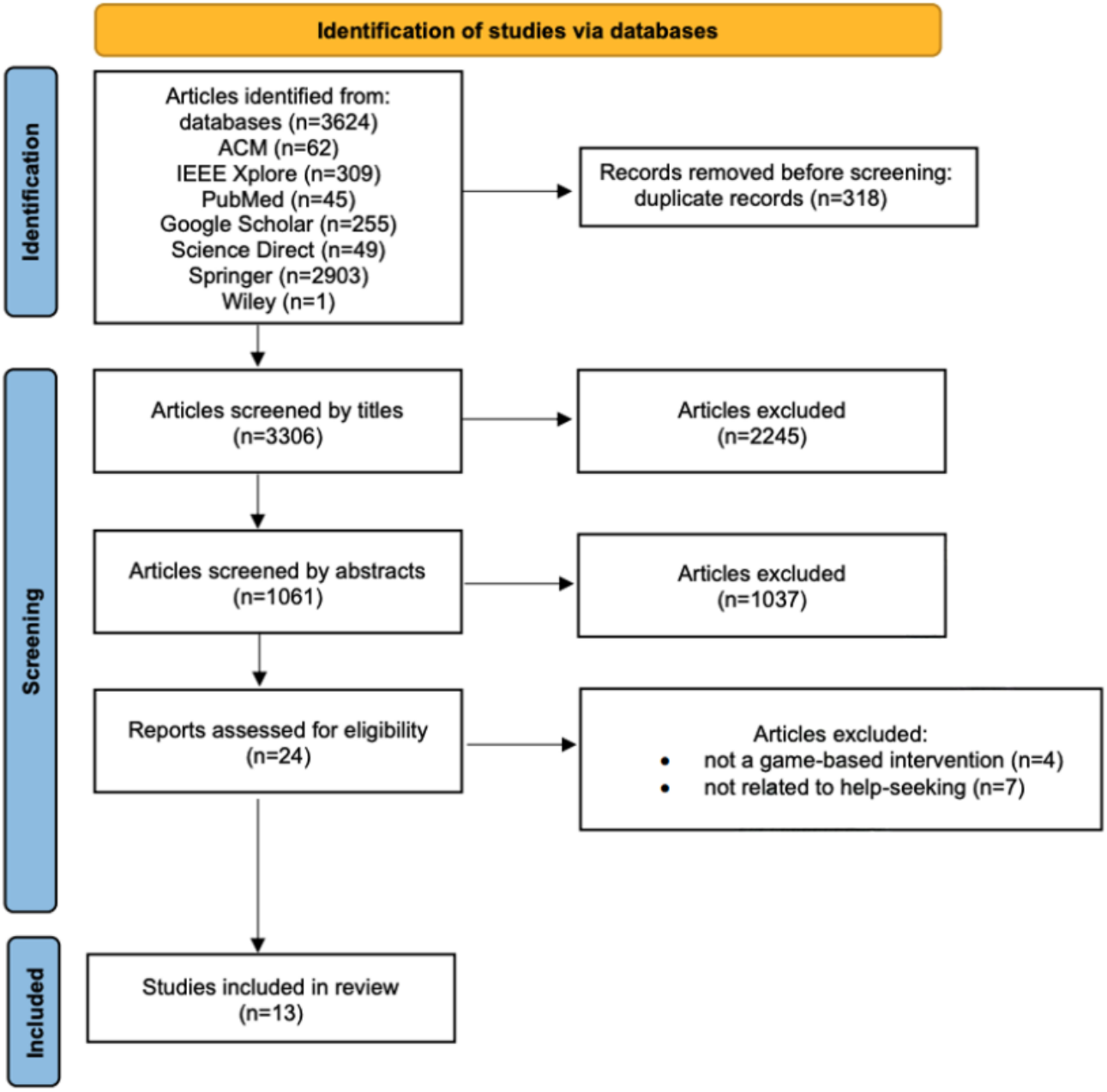
Results

Overview

There are 2 studies identified specifically from the search string PND, and 11 studies included from the search with depression. In total, 13 eligible studies were included in this review. There were 9 games published in academic articles, and the other 4 are from market research. In the past 5 years, there were 5/9 (55.6%) studies published. The screening process is shown in Figure 1. An overview of their characters is shown in Table 2.

XSL•FO
RenderX

Figure 1. PRISMA diagram. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



XSL•FO
RenderX

Table . Summary of included studies.

Name	Year	Country	Study design	Sample size	Intervention length
Above Water	2016	Canada	N/A[a]	N/A	N/A
Stigma-Stop	2017	Spain	Comparison study	552	N/A
Leadership Training in Mental Health Promotion (LMHP)	2017	Germany	Quasi-experimental study	48	3 months
Moving Stories	2019	Netherlands	Cluster randomized controlled trial	185	6 months
MindMax	2020	Australia	Participatory design	40	N/A
MANTRA	2020	United Kingdom	Cocreation process	60	N/A
Hellblade: Senua's Sacrifice	2020	United States	Two-condition, randomized design	198	45 minutes
COEX-IST	2022	Brazil	N/A	N/A	N/A
PPDHero	2024	Bangladesh	Requirement elicitation study	108	N/A

[a]N/A: the data were not clear or not mentioned in the original paper.

This section divides the 13 studied games included into 3 categories based on their perspectives on improving help-seeking behaviors: promoting knowledge, reducing stigma, and raising awareness.

Promoting Knowledge

Overview

An overview of the games that focus on promoting knowledge to help improve depression can be seen in Table 3.

XSL•FO
RenderX

Table . Games focused on promoting knowledge.

Name	Year	Category	Aim	Audience	Game engine	Player mode
Above Water [54]	2016	Strategy games	This game informs people about the available strategies to cope with 2 types of anxiety disorders—generalized anxiety disorder and panic disorder. The game teaches players about existing treatments.	Game players	Physical cards and a mobile application	Multiple players
Moving Stories [55]	2019	RPGs[a]	This game aims to make adolescents have better mental health literacy and endorse fewer stigmatizing attitudes regarding depression.	Adolescents	Not specified	Single player
MindMax [56]	2018	Casual games	This game is to deliver psycho-educational modules and create a web-based community centering on well-being, AFL[b].	Men aged 16 to 35 years who are interested in AFL or video games	Mobile phone	Single player
MANTRA [57]	2020	Casual games	This game is to increase maternal and child health resilience before, during, and after disasters.	Vulnerable low-literacy female audiences in rural Nepal	Mobile phone	Single player
PPDHero [58]	2023	Serious games	This game aims to screen PND[c] and provide education and support to new mothers to identify and manage symptoms and connect with health professionals.	New mothers	Mobile gamification app	Single player

[a]RPG: role playing game.

[b]AFL: Australian Football League.

[c]PND: postnatal depression.

Above Water

In 2016, Wehbe et al [54] released the educational game *Above Water* for anxiety. It is a digital or physical hybrid game to inform people about the available strategies to cope with 2 types of anxiety disorders, which are Generalized Anxiety Disorder and Panic Disorder. It is designed to inspire players to share their experiences and develop their narratives by teaching them existing treatments. *Above Water* is a game that players play with physical cards and digital devices, such as tablets and smartphones. The game is designed to be played with 4 players. They are required to achieve all 6 categories of life goals, including career, education, health, self-improvement, financial stability, and relationships, while managing anxiety. They need to select 4 life goals on their smartphone, which cannot be changed during the play session. On each player's turn, they will draw 1 card to face up, play, or discard. Anxiety cards can be discarded by using treatment cards. They can be taught deep breathing skills, yoga, and stretching poses using mobile applications. The digital implementation does not require downloading an app. It uses a simple HTML5-enabled website. This game has several strengths, particularly in raising awareness and encouraging social support through collaborative gameplay. Its focus on educating players about anxiety and introducing treatment options makes it helpful in fostering an open dialogue about mental health. However, the game falls short in targeting specific populations and providing tailored

XSL•FO
RenderX

interventions. Additionally, it lacks mechanisms for improving long-term behavior change and cultural adaptation, which are critical elements for effectively addressing the unique challenges faced by parents dealing with PND. It could be improved by incorporating specific content for new parents and addressing unique social and cultural barriers.

Moving Stories

Moving Stories is a game-based school program for mental health literacy and stigma regarding depression published by Tuijnman et al [54] in 2019. It is developed by a professional game design company in collaboration with researchers and relevant stakeholders. Moving Stories includes three parts: (1) an introduction lesson, (2) a single-player, mobile, 3D video game, and (3) a contact session with someone with lived experience with a depressive disorder. The program focuses on three perspectives of mental health literacy: (1) recognizing the disorder, (2) having knowledge of help-seeking options and treatment available, and (3) first aid skills to support others who are developing or in a mental health disorder. The video game asks players to help a girl named Lisa, who has almost lost interest in everything. The players can do multiple things for Lisa to accumulate or decrease their relationship scores. As a bridge, the players are guided to share messages and discuss the reflection in a contact session with the school welfare coordinator or school counselor. This game demonstrates several strengths as a mental health intervention tool, particularly in improving mental health literacy, reducing stigma, and promoting help-seeking behaviors. By empowering players with knowledge, its approach to directly confronting myths through gameplay helps break down stigma. It encourages a more open conversation about mental health issues, which sets the foundation for increased understanding and empathy. It is crucial to tackle the stigma that new parents might face regarding PND. A key strength of the game lies in its focus on promoting help-seeking behavior. Through interactive scenarios that involve assisting the character Lisa and guiding her toward appropriate support, *Moving Stories* emphasizes the importance of recognizing mental health issues and seeking help. This game aligns well with the need to address the avoidance of help-seeking behaviors seen in PND, where new parents often struggle to ask for support due to stigma or lack of awareness. Incorporating role-playing elements demonstrates how and when to seek help, which is critical for empowering individuals to take action in real-life situations.

MindMax

MindMax is an app that incorporates gamification, mini-games, and social connection to improve men's mental health and well-being, published by Cheng et al [56] in 2018. This game is based on the original well-being training program that the Australian Football League Players' Association offered players. Thanks to the collaboration with the University of Sydney and the University of Technology Queensland, it aims to deliver a portable, digital version of Australian Football League Players' Association's existing program. The design process includes 3 phases. The first phase is 6 participatory design workshops consisting of 3 stages: discovery, evaluation, and prototype to identify how best to frame the well-being concept discussed in *MindMax* and, more broadly, how to structure a mental health and well-being app for the intended audience. The second phase is knowledge translation. The third phase is user experience testing interviews. This work explained well why and how they implement the participatory design process. The core components of *MindMax* include two parts: (1) multiple educational modules, each around 10 minutes, and (2) a web-based community centering on well-being, sports, and video games. The good point in an educational context is that the audience does not care about what context is presented, but how it was presented. The tone of expression is very important. Regarding the social element, it suggested the app should enable rather than emulate, and sharing personal experiences with potential groups they are conscious of may need careful consideration.

One of the key strengths of *MindMax* is its incorporation of psycho-educational modules that provide users with mental health information in a clear and accessible manner. Each module is designed to be short (approximately 10 min), making it feasible for users with limited time, such as new parents with demanding schedules. These modules cover foundational topics such as mindfulness, emotional well-being, and basic coping skills. For new parents experiencing PND, such content can be particularly beneficial as it provides them with the tools they need to understand their emotional state and learn practical strategies for managing stress. Making these modules easily digestible helps ensure that even those unfamiliar with mental health concepts can engage with the material effectively. Besides, the mini-games and reward-based mechanics used in *MindMax* are particularly useful in maintaining user engagement, which might be a crucial factor for individuals who may be overwhelmed with responsibilities or have limited energy to devote to self-care. Using games as rewards helps create positive reinforcement for completing educational tasks, encouraging continued learning and interaction. For new parents, especially those struggling to make time for themselves, gamification can transform mental health support into something enjoyable and less intimidating, thus promoting consistent use of the intervention. The light-hearted games and achievement-based systems can provide much-needed moments of relaxation and enjoyment, breaking up the often stressful and exhausting routines of caring for a newborn.

MANTRA

MANTRA is a serious mobile game that aims to reach a low-literacy audience in a low-resource setting with knowledge of maternal health, neonatal health, and geohazards. This work was published by Mueller et al [57] in 2020. The game mechanic is relatively simple in this game: picture matching with audio and visual feedback. The project's core goal is to tailor the knowledge to a local cultural context to fit specific settings with low literacy and resources. Throughout the design process, co-design and cocreation were guiding principles. The game design methodology builds on these principles and typical software development processes. The "drag-and-drop" is the only mechanism needed in the game. There is a touchscreen tutorial at first to help users who are not familiar with the screen interface. There are 3 modules in total, each containing 3 levels of complexity. The unsuccessfully answered question will be

XSL•FO
RenderX

pulled into an error pool and repeated up to 3 times throughout the level. The player will fail this level if all 3 times are answered incorrectly. *MANTRA* demonstrates several notable strengths in health education, particularly its accessibility for low-literacy users, cultural localization, and reinforcement of knowledge through repetition. It is designed for low-literacy users. It uses pictograms, audio prompts, and simple, intuitive interactions to ensure that even those with limited literacy skills can engage with and benefit from the game. This setting is particularly valuable for new parents from underserved communities who may lack access to formal health education, effectively addressing knowledge gaps. Second, the game's cocreation and cultural localization enhance its relevance and acceptance. Given that cultural context plays a significant role in the understanding and treatment of PND, incorporating cultural elements into the game helps to establish a deeper connection with users, reduce stigma, and increase engagement. Lastly, *MANTRA* uses repeated learning modules and immediate feedback, which helps reinforce knowledge retention. For new parents needing to master coping strategies and understand when to seek help, this method of repetition is crucial for effectively supporting their mental health outcomes. However, *MANTRA* also has some limitations. First, while using simple mechanics (such as drag-and-drop) makes the game accessible, it may not provide the depth of engagement needed for users facing significant mental health challenges. Second, the game lacks features that actively promote help-seeking behavior. For a digital game aimed at addressing PND, it is essential to include explicit pathways to professional support, such as information on local health care services or in-game guidance on seeking help, to bridge the gap between awareness and action. Lastly, the content of *MANTRA* primarily focuses on mothers without sufficient consideration of the role of fathers or nonbirthing parents in supporting maternal mental health.

PPDHero

PPDHero is a novel web-based system published by Sara et al [58] in 2024 for new mothers to screen PND and provide them education and support to identify and manage symptoms and connect with health professionals if needed. The design process starts with an elicitation study, which involved 6 senior doctors, who were incorporated with surveys from the broader public. The system was developed based on interview and survey data analysis. It includes three modules: (1) relevant content management; (2) screening, assessment, and gamification; and (3) a treatment and support module to enable users to share resources, enroll in screening procedures and treatment plans, register for gamified activities, and involve trusted professionals in their postnatal journey. The gamification tasks, such as quizzes and exercises, are implemented in the screening and assessment session to enhance the assessment process. Its strengths lie in its comprehensive approach, integrating education, screening, and treatment in a single platform. Using gamification elements, such as quizzes and interactive activities, enhances user engagement and makes learning about PND more attractive. Additionally, the platform provides a personalized experience through tailored treatment plans based on screening results, ensuring that users receive targeted support to manage their mental health. However, while the platform provides education and support, it lacks features specifically aimed at reducing stigma, which is an important factor in encouraging new mothers to seek help.

Reducing Stigma

Overview

In Table 4, there is an overview of game-related components that contribute to depression and broad mental health illness by reducing stigma to the game players.

Table . Games focusing on reducing stigma.

Name	Year	Category	Aim	Audience	Game engine	Player mode
Stigma-Stop [59]	2017	Simulation games	This game aims to reduce the stigma toward mental illnesses.	High school students	Unity3D	Single player
Leadership Training in Mental Health Promotion (LMHP) [60]	2017	Serious games	This game is to promote employee mental health and reduce mental illness stigma at work.	Managers	Digital game-based training program	Single player
Hellblade: Senua's Sacrifice [61]	2020	RPGs[a]	This game aiming to let players play video games featuring characters enduring mental illness may ultimately reduce stigma through transportation and identification.	Game players	Unreal Engine 4	Single player

[a]RPG: role playing game.

XSL•FO
RenderX

Stigma-Stop

Stigma-Stop is a 3D serious game to reduce the stigma toward mental illness among high school students, published by Cangas et al [59] in 2017. It is specifically designed to reduce the stigma associated with mental health disorders by educating players about different mental health conditions, such as depression, schizophrenia, bipolar disorder, and panic disorder with agoraphobia. The game provides information about these disorders through interactive dialogues and scenarios where players must choose how to react to individuals with mental health issues. This kind of approach is effective in challenging and correcting misconceptions about mental illness, particularly among young people. The game presents players with scenarios of interacting with characters representing different mental health conditions. Players are asked how they perceive the characters' psychological state, whether they have experienced similar emotions, and how they can help the character. These choices, along with the feedback provided for correct and incorrect responses, allow players to learn more about mental health hands-on, engagingly. However, the game also has notable limitations. Its primary focus on adolescents and broader mental health disorders may make it less directly relevant for new parents dealing specifically with PND. Additionally, the game lacks practical coping strategies or therapeutic guidance, which limits its effectiveness as a support tool aimed at improving the help-seeking behaviors of PND.

Leadership Training in Mental Health Promotion (LMHP)

Leadership Training in Mental Health Promotion (LMHP) is a digital game-based training program for managers to promote employee mental health and reduce mental illness stigma at work, published by Hanisch et al [60] in 2017. It applied various gamification components, such as providing a storyline and clear goals, including the capacity to overcome challenges by learning, giving feedback on performance, showing progress, and reinforcing learning by allocating points rather than simply providing badges for achievements and enabling competition between players, to facilitate an innovative and engaging learning experience. *LMHP* demonstrates several strengths that make it an effective tool for supporting mental health in the workplace and reducing stigma. The program's focus on mental health literacy, reducing stigma, and improving self-efficacy among managers makes it well-suited to foster a supportive work environment. The interactive simulation-based approach helps participants practice these skills in a safe environment, leading to higher engagement and better skill retention. While *LMHP* has strengths in terms of mental health literacy and stigma reduction, it focuses on the workplace environment and specifically targets managers. The scenarios and skills taught are tailored to workplace situations, which may not fully address new parents' specific emotional and psychological challenges in their home environment.

Hellblade: Senua's Sacrifice

Hellblade: Senua's Sacrifice is an action-adventure game accurately portraying psychosis to reduce public mental health stigma, developed and published by Ninja Theory [61] in 2017. It aims to reduce the mental health stigma in 2 ways: by lowering stereotyping and limiting participants' desire for social distance. *Hellblade: Senua's Sacrifice* is set in an age of Vikings. Its story is about Senua's battles for the soul of her departed lover, Hela. Senua endures psychosis and must contend with her mental illness along with the challenges presented by her quest. The game blends gameplay mechanics and concepts such as puzzle solving, psychological horror, and melee combat. Voice acting is an integral part of the game, while its cut scenes combine motion capture by Melina Juergens and live-action performances by other actors.

The game's powerful and immersive portrayal of mental health struggles provides players with a unique opportunity to experience the emotional complexities associated with conditions such as psychosis. Through this kind of realistic depiction, new parents experiencing PND can see their struggles mirrored in the protagonist's experiences, helping them feel validated and understood. Second, the game's high-quality audio design plays a crucial role in enhancing the emotional depth of the experience. Using binaural audio to simulate auditory hallucinations, *Hellblade: Senua's Sacrifice* creates a profoundly immersive atmosphere that draws players into Senua's inner world. This kind of powerful audio experience can help new parents connect more profoundly with the emotional struggles portrayed, fostering empathy and offering a unique perspective on mental health challenges. Lastly, the game's development process involved collaboration with mental health experts and individuals with lived experience of psychosis. This process makes the portrayal of Senua's mental health struggles both vivid and respectful, adding credibility to the game's narrative.

Raising Awareness

Overview

In Table 5, there is an overview of how game-related components contribute to depression and broad mental health illnesses by raising awareness among game players.

XSL•FO
RenderX

Table . Games focused on raising awareness.

Name	Year	Category	Aim	Audience	Game engine	Player mode
Depression Quest [62]	2013	RPGs[a]	This game aims to show other people with depression that they are not alone in their feelings and to illustrate to people who may not understand the illness the depths of what it can do to people.	Game players	Twine engine	Single player
Keep in Mind: Remastered [63]	2018	Simulation games	This game is meant to be a therapeutic experience for those who struggle with mental illness on a journey of reflection and emotional healing.	Those who struggle with mental illness or those who find themselves lost in the dark	Unity	Single player
Before I Forget [64]	2020	Simulation games	This game aims to raise awareness about mental health and communicate to the audience the struggles of people experiencing dementia.	Game players	Unity	Single player
Sea of Solitude [65]	2021	Simulation games	This game aims to reveal the themes of depression, loneliness, and hopelessness, how we unwittingly push people away, and how our actions can negatively affect others.	Game players	Unity	Single player
COEX-IST [66]	2022	Strategy games	This game is meant to increase mental health awareness and focus on self-care as a form of prevention.	People with a wide age range, from teenagers to adults	Not specified	Single player

[a]RPG: role playing game.

Depression Quest

Depression Quest is an interactive fiction game dealing with depression. It was developed by Zoe Quinn [62] using Twine (Interactive Fiction Technology Foundation) and published in 2013. This game aims to spread awareness by showing other people with depression that they are not alone in their feelings. *Depression Quest* is not an easy gaming experience because its goal is precisely to let you experience the difficulties of depressed people. It is a text-based game, and your choice will impact your ending. *Depression Quest* has a straightforward game structure. There are no animations or complex controls. As the story progresses, the player chooses different options to move the story in different directions. The part that *Depression Quest* comes closest to reality is that it tries to make people who do not have experience with depression understand that there are choices that do not exist for people who have depression. In the game, as the depressive condition deepens, the optimistic options are crossed out by the system with a red line—the player can only choose the closed, pessimistic options, and the depressive condition further deepens. There are fewer and fewer options to choose from in the game. It is a vicious circle and one that many people with depressive tendencies face.

This game has several strengths in its design elements. First, the game's narrative-driven depiction of depression provides a deeply empathetic look into the daily struggles of a person experiencing mental health challenges. Second, the game uses a unique mechanic where certain positive choices become restricted based on the protagonist's mental health status. This mechanic reflects the real-life constraints imposed by depression, such as the difficulty in making proactive decisions when feeling overwhelmed. For new parents experiencing PND,

XSL•FO
RenderX

this feature can help explain why even seemingly simple actions might feel impossible. It encourages self-compassion by showing that these limitations are part of the condition, not personal failings. Furthermore, as players engage with the narrative and take actions such as seeking help, they unlock new options, which effectively illustrate the empowerment that can come from actively addressing mental health challenges. Third, *Depression Quest* is also effective in raising awareness and fostering empathy among those who may not have firsthand experience with depression. By taking players through scenarios involving hopelessness, lack of motivation, and interpersonal struggles, this game mechanic could help partners and family members of new parents better understand the challenges their loved ones face. This improved understanding can lead to more supportive relationships and foster a compassionate environment. Finally, the game places a strong emphasis on encouraging help-seeking behavior. Players can pursue therapy or medication throughout the match, impacting the protagonist's situation and improving the choices. By showcasing how seeking help can positively affect mental health, *Depression Quest* actively promotes the idea that professional support is valuable and effective. New parents, who may feel hesitant to seek help due to stigma or uncertainty, could see the benefits of reaching out in this kind of game, which can help motivate them to pursue the support they need.

However, despite these significant strengths, the game has a key limitation. The intense focus on the struggles of depression, combined with the dark and sometimes overwhelming themes, may risk leaving players feeling disheartened without enough positive reinforcement or hope to balance the experience.

Keep in Mind: Remastered

Keep in Mind: Remastered [63] is a heavy story-driven psychological indie game that follows Jonas, a man who struggles with grief, depression, and alcoholism, on a journey of reflection and emotional healing. One night, he awakens to a shadowy mirror world where beasts lurk, and stars do not shine. Lost and scared, Jonas must face the twisted beasts if he ever wishes to return home and learn the truth about his darkness. It was released by Little Moth Games and Akupara Games in March 2018. This game was created for those who struggle with mental illness or those who find themselves lost in the dark. This game is meant to be a therapeutic experience. *Keep in Mind: Remastered* has notable strengths that make it a compelling game for raising awareness about mental health challenges. For new parents, especially those experiencing PND, the portrayal of such internal struggles can help them feel seen and understood. This level of emotional storytelling can help reduce the stigma around PND by showing that struggling with mental health issues is not uncommon and that others share the negative thoughts and emotions they may be experiencing. Besides, the game uses symbolism effectively to represent the different facets of the protagonist's mental health challenges, such as dark creatures and eerie settings. This symbolism allows players to externalize their internal emotional struggles, making them more tangible and easier to understand. However, the game's focus on heavy and potentially triggering themes, lack of practical coping strategies, and absence of integrated help-seeking features might limit its effectiveness as a supportive intervention.

Before I Forget

Before I Forget is a single-player exploration game developed and published by 3-Fold Games [64] in 2020. It is a short game that draws the player to a character experiencing dementia. She is housebound and unable to leave the flat. Only the windows could offer a view of the outside world, with its letterboxes and birdbaths. Otherwise, the exploration all happened in the mind as the character scrabbles to dredge artifacts from her past. In this game, players guide Sunita Appleby, who is a scientist with early-onset Alzheimer disease. As she interacts with now-unfamiliar objects in her house, some of her memories return. The game features puzzle-like elements, such as recalling where mementos have been left and asking players to guide Sunita to the bathroom when she perceives her house as a shifting maze. The game's use of empathetic storytelling and emotional depth provides a powerful connection to the struggles faced by those with mental health challenges. The game fosters understanding and reduces stigma around vulnerability and emotional struggle by allowing players to see the world through the protagonist's perspective. Additionally, the emphasis on relationships and emotional connections highlights the importance of support networks, encouraging players to lean on their loved ones during difficult times. However, the game also has notable limitations; the lack of practical coping strategies limits its effectiveness as a tool for managing day-to-day mental health issues, such as PND.

Sea of Solitude

Sea of Solitude is an adventure video game developed by Jo-Mei Games and published by Electronic Arts [65]. The player controls a young woman named Kay, who endures such strong loneliness that her inner feelings of hopelessness, anger, and worthlessness turn to the outside, and she becomes a monster. As Kay, the player explores a seemingly empty flooded city and interacts with its scaly red-eyed creatures to reveal why she turned into a monster. Her emotions manifest into giant monsters standing in her way, trying to help but also destroy her. She needs to interact with and understand their underlying intentions to overcome the negative effects of those emotions. The game is an inner dialogue of a person trying to reconcile her shortcomings. It provides an insightful look at how mental illness devastates the lives of not just those it affects but also loved ones on the outside. Kay learns a lot about herself by understanding the value of listening, coming to terms with her flaws, and not just empathizing with her family but also accepting that a simple fix is not always possible. The game's powerful narrative effectively depicts the emotional struggles of loneliness, depression, and anxiety, allowing players to relate deeply to the protagonist's journey. Referring to this representation can help new parents normalize their emotions and reduce the stigma around experiencing mental health challenges during the postnatal period. The exploration mechanics and emotional expression throughout the game provide a meaningful way for players to reflect on their experiences. Additionally, the game's focus on empathy and understanding different characters' struggles helps to foster a

XSL•FO
RenderX

deeper connection between players and the themes of mental health. For new parents, this emphasis on empathy can reduce feelings of isolation and highlight the importance of reaching out for support, both for themselves and others. However, *Sea of Solitude* also has notable limitations in addressing PND effectively. The game lacks practical coping strategies that can be directly applied to real-life challenges. For new parents, having evidence-based tools and actionable advice would be instrumental in managing the day-to-day reality around PND. Without such strategies, the game may fail to deliver tangible help beyond emotional understanding. Moreover, the heavy themes of depression, anxiety, and loneliness may be overwhelming for some new parents, particularly those who are already vulnerable. Without sufficient pacing options or content warnings, the game risks triggering or intensifying negative emotions rather than alleviating them.

COEX-IST

COEX-IST is a 3D interactive story decision-making game that aims to grow awareness about depression and social isolation after a pandemic. In 2022, Rodrigues et al [66] developed and published it. The game story is based on 25 undergraduate students' personal experiences. It is a third-person short game based on narratives and point-and-click to interact with the scene's objects and make decisions. The end of the narrative has 2 possible actions for game over: victory over depression or victory of depression. The story's resolution also involves tying up the loose ends of the climax and falling action. Each choice has impact consequences. Both possible ending scenes allow the player to experience and think about depression, growing awareness of this illness. The game's interactive decision-making process, emphasis on empathy, and reflective storytelling structure offer a deep and meaningful engagement that helps players connect with the content and understand the complexities of mental health challenges. First, the interactive decision-making feature allows players to experience the consequences of their actions realistically. This level of interactivity encourages new parents to see the potential impact of small, positive actions in their own lives, thereby fostering a sense of control over their mental well-being. This sense of agency is particularly important for new parents, who may often feel overwhelmed by the responsibilities and pressures of caring for a newborn. Second, the game's emphasis on mental health awareness and emotional empathy is effectively conveyed through storytelling. By representing depression as "Mr. Shadow," *COEX-IST* provides a powerful visualization of the weight and burden of mental illness. This representation helps players externalize and better understand their own struggles. It also reduces the stigma associated with these emotions by demonstrating that they are a shared experience that can be overcome with the right actions. This empathetic storytelling can provide comfort and connection for those who may feel isolated in their struggles. Lastly, the structured storytelling approach gives the narrative a clear and engaging emotional experience that mirrors the journey of dealing with mental health challenges. For new parents, seeing the protagonist's ups and downs depicted in a structured way can help normalize their own experiences. The 2 possible endings—either overcoming depression or being overtaken by it—underscore the importance of the player's choices and illustrate the ongoing nature of mental health management. This setting highlights that while setbacks may occur, progress is possible, which can provide hope and motivation to players during difficult times.

Many of the reviewed games (eg, *Depression Quest* and *Before I Forget*) aimed to raise awareness and build empathy through immersive experiences that replicate the emotional realities of those affected. These games often use narrative-driven storytelling to allow players to live through the struggles of a character facing depression. Through experiencing the emotional highs and lows, players are encouraged to empathize with the character's struggles. This emotional journey is argued to be crucial not only for players who are personally dealing with mental health challenges but also for their support networks—such as partners, family members, and colleagues—who may benefit from a better understanding of what their loved ones are going through. This approach can externalize internal emotional struggles, providing validation for players dealing with similar issues. The intention here is to not only entertain but offer a profound and empathetic experience that helps players feel less isolated.

Some reviewed games (eg, *MindMax* and *Stigma-Stop*) encouraged players to seek professional help for mental health concerns. By embedding choices in the game where players can seek therapy, use support resources, or talk to someone, these games guide players toward understanding the importance of professional help in managing mental health conditions.

Six games targeted adolescents and youth from 15 to 25 years, and one targeted younger children. Six games do not specifically target an age group, but board game players. Examples of these groups include 1 study for men who are interested in AFL games and 1 study for company managers.

Among these games, only 2 games targeted new mothers with a game intervention focused on promoting knowledge. Although few studies directly target new parents, there is arguably some overlap with the depressive symptoms faced by the other groups discussed above. However, new parents are often in a unique and vulnerable position, facing a combination of physical, emotional, and psychological challenges. They usually experience heightened vulnerability, exhaustion, and lack of time [67], which requires an intervention that is not only effective but also easily accessible and low-pressure. Many of the games targeting adults with depression do an excellent job of portraying the emotional complexities of mental health issues. Still, their intensity and focus on complex gameplay may not be suitable for new parents who are experiencing severe fatigue. It is argued that the new parent target audience would benefit more from short, accessible gameplay that can be easily integrated into their busy routines, and casual yet meaningful experiences that provide emotional validation and support.

Given the needs of the new parent target audience, mobile platforms are likely the most suitable choice. Mobile games can be played in short bursts, require minimal setup, and are highly accessible because most people carry smartphones. Mobile accessibility also allows parents to engage with the game whenever they find spare moments, which is crucial given their unpredictable and fatiguing schedules. However, a hybrid

XSL•FO
RenderX

solution that combines mobile and web-based access to ensure that new parents have multiple ways to engage with the game based on their convenience may also be valuable. A web-based version could provide a larger screen and more detail for those moments when new parents can access a computer, while the mobile version allows instant play. WebGL technology further reduces minimum system requirements and will enable games to be accessed from any electronic device. Hybrid accessibility also increases reach, as it accommodates different user preferences and ensures that players have multiple ways to interact with the content.

Console and VR platforms, while highly immersive, are arguably not ideal for most new parents. While VR today is widely "plug and play" and often requires little more than a connection ID to get started, the transition to becoming a parent, particularly in the early months, is marked by extreme time constraints, chronic sleep deprivation, and a constant state of divided attention [68]. New parents often navigate caring for an infant, are adjusting to unpredictable routines, and manage both emotional and physical exhaustion. Therefore, it is proposed that even simple steps such as clearing space, troubleshooting hardware, or remaining engaged in a virtual environment for more than a few minutes can become overwhelming for parents who might be at risk of cognitive and emotional overload.

Discussion

Overview

There is a dearth of evidence directly related to PND, which is only 2 studies in the past 10 years. After expanding the search to depression, we included 11 more. The included studies predominantly aligned with 3 overarching themes: the promotion of knowledge, the reduction of stigma, and the raising of awareness. A core intention of several of the reviewed games (eg, *Above Water* and *Moving Stories*) is to educate players about mental health and reduce the associated stigma. This is achieved by exposing players to the symptoms, challenges, and lived experiences of individuals who are dealing with various mental health conditions. The games use different approaches, such as presenting players with scenarios that require them to make supportive decisions, simulating real-life interactions with people facing mental health challenges, and providing informative feedback. The ultimate goal is to educate the players on mental health literacy, that is, what different symptoms mean, why they matter, and how to effectively respond. This type of education, delivered in an interactive, engaging way, generally helped to demystify mental health issues, correct misconceptions, and encourage empathetic attitudes. This is argued to play a crucial role in reducing stigma by promoting understanding and highlighting the commonality of these experiences, breaking down the barriers that often prevent open conversations about mental health.

Game Mechanics

For depression management in general, several mechanisms are argued to be particularly impactful. Narrative storytelling, as seen in games such as *Hellblade: Senua's Sacrifice* and *Depression Quest*, may provide emotional engagement that helps players feel understood and validated by illustrating shared experiences of mental health struggles. Given the evidence that the recognition and acceptance of an individual's feelings and experiences promote mental well-being, seeing their experiences reflected in the storyline may create an emotional connection that could combat isolation. Choice-based interaction, where players make decisions that affect the storyline, may offer a sense of agency and sense of control, which is often decreased in those experiencing depression [69]. These decisions can illustrate how positive actions, such as seeking help or practicing self-care, might lead to better outcomes, potentially reinforcing a sense of empowerment for new parents. Simulated real-life scenarios could also play a role in helping players practice practical coping skills, such as managing stress or communicating with loved ones in a safe and controlled environment. Combined with resilience mechanics that highlight small wins, these mechanics may help maintain motivation during challenging times. Additionally, low-pressure gameplay is likely important to this target group, as it may reduce the cognitive and emotional needs of interaction, making the game more accessible to new parents who are experiencing fatigue and time constraints.

For educational purposes and stigma reduction, specific game mechanics may help build knowledge and empathy in accessible ways and minimally overwhelm the player. For example, educational micromessaging, as seen in "Match Emoji" and "Stigma-Stop," delivers brief and embedded content that increases awareness about mental health conditions. This mode of delivery may be especially valuable for new parents, who often benefit from concise, low-effort learning formats. Empathy-building mechanisms could also contribute to stigma reduction, an important function within PND given its high association with stigma [13]. Interactive elements that allow players to make supportive choices during character interactions may foster greater understanding and compassion. Role-playing as characters facing mental health challenges may help players consider these issues more personally, potentially encouraging self-compassion and empathy for others. Additionally, real-life scenario simulations may offer players opportunities to engage in realistic support-seeking situations, which could enhance a better understanding of seeking and receiving support.

Given the cultural, social, and personal barriers to seeking help for people with PND [13], certain game mechanics could play a valuable role in improving help-seeking behavior by reducing hesitancy and encouraging engagement with professional help. Choice-based interactions that include options for seeking therapy or speaking to a health care professional can illustrate the potential positive impact of these decisions and may help normalize the idea of reaching out. In-game scenarios that model conversations with health care professionals or trusted family members could also reduce anxiety by providing players with low-risk opportunities to observe or practice supportive dialogue. Additionally, incorporating positive reinforcement, such as visual or narrative cues that reflect improvements in well-being or relationships following help-seeking, may highlight the value of accessing help. For expectant parents who may experience PND after their baby is born, these mechanics may offer gentle encouragement and promote greater confidence

XSL•FO
RenderX

in taking the first step. As shown in Table 6, here is a comparison of game mechanics and different aims.

Table . Comparison of game mechanics and different aims.

	Build knowledge	Build empathy	Build awareness	Reduce stigma
Narrative storytelling	✓	✓		✓
Choice-based interaction		✓		✓
Educational micromessage	✓		✓	
Character interaction	✓	✓		
Real-life scenarios	✓		✓	
Resilience mechanics			✓	

Principal Findings

To effectively support new parents coping with PND and to address the gaps identified in this review, the authors argue that future research should focus on the interprofessional co-development of a serious game with the following important design elements. First, we argue that a relatable and engaging story is important and should reflect both the struggles and victories that new parents face, helping them see their own experiences mirrored in the game. By doing this, the game can foster a deep sense of understanding, connection, and emotional validation, which may be crucial for reducing the isolation [13] often felt during PND. Allowing players to make meaningful choices that influence the direction of the game could give them a sense of agency and control to combat the low sense of agency [69] frequently found in the experience of depression. These decisions should reflect real-life challenges, and the consequences should demonstrate how positive actions, such as seeking help or practicing self-care, can improve well-being. This interactive element can empower new parents to feel more in control of their own mental health journey. Incorporating real-life scenarios is also a key element. These scenarios should include common parenting challenges, such as communicating effectively with a partner, managing stress, or seeking support. Practicing these skills in a safe, controlled game environment can help players build confidence, making it easier to apply what they have learned to real life.

Gamification elements such as badges, progress tracking, and daily challenges should be used to keep players motivated and committed, in line with a positive reinforcement approach to learning and improving intrinsic motivation [70]. These features add an element of fun and reward, encouraging players to engage consistently with the game. The game should also include educational messages that provide important information about PND. These messages could cover recognizing symptoms, understanding when to seek help, and learning practical self-care techniques. Delivering these messages in small, easily digestible portions ensures that players can learn without feeling overwhelmed, and it is in line with research findings that show benefits for adult learners who engage in mobile-based microlearning steps [71].

Character interactions are another essential design element. Players should have opportunities to interact with supportive characters, which can help build empathy and promote positive communication skills. These interactions can model supportive relationships and reinforce behaviors that new parents can use in their own lives to improve their support networks. Celebrating small wins is also important; highlighting moments such as taking time for self-care or successfully reaching out for help can reinforce resilience and make players feel that each positive step, no matter how small, is a significant achievement. Lastly, the game should be designed to be accessible and easy to engage with in short sessions. Low-pressure gameplay is important, as new parents often face exhaustion and have limited time [68]. The game should be flexible, allowing players to pick it up and put it down whenever they have a few moments, ensuring it serves as a helpful, nonstressful tool in their daily lives. This balance of emotional engagement, practical learning, and ease of access is proposed to create a supportive experience that can make a meaningful difference for new parents coping with PND.

Recommendations

Overview

Serving the aim of our proposed project, we recommend focusing on the following subset of game types and mechanics.

Narrative-Driven Storytelling

Incorporate an engaging storyline that reflects the real-life experiences of new parents, including both struggles and successes. This mechanic helps players feel understood and connected, reducing feelings of isolation. The narrative should also include resilience themes to promote hope and perseverance.

Choice-Based Interactions

Implement choices that directly influence the game's storyline, allowing players to make decisions that affect their mental health journey. This empowers players by giving them a sense of control, emphasizing the positive outcomes of seeking help or practicing self-care.

Real-Life Scenario Simulations

Include scenarios that simulate everyday parenting challenges, such as communicating with a partner, managing stress, or seeking professional support. Practicing these skills in a game setting builds players' confidence and provides them with practical tools for real-life application.

Low-Pressure, Casual Gameplay

Design the game for short, manageable sessions that can be played on a mobile device. This mechanic ensures accessibility

XSL•FO
RenderX

for new parents, allowing them to engage with the game without added pressure or stress. Casual mechanics make the game easier to pick up and play whenever parents have a few free moments.

Educational Micromessaging

Integrate educational content subtly throughout the game to provide important information about PND, such as recognizing symptoms and self-care tips. Delivering these messages in a digestible format helps fill knowledge gaps without overwhelming the player.

Character Interactions and Empathy-Building

Allow players to interact with characters representing supportive individuals, such as partners, family members, or health care professionals. These interactions help model healthy communication and build empathy, both for oneself and others.

Gamification Elements

Add elements such as badges, progress tracking, and daily challenges to encourage ongoing engagement. Gamification makes the experience more rewarding and helps players stay motivated, which is especially important for building positive habits.

Celebrating Small Wins

Highlight small victories in the game, such as successfully reaching out for support or completing a relaxation exercise. Positive reinforcement builds resilience and emphasizes that progress, no matter how small, is meaningful.

Strengths and Limitations

This review draws strength from its interdisciplinary synthesis of literature spanning digital mental health interventions and serious games, offering a comprehensive overview that bridges technological and psychological perspectives. Examining a novel and emerging approach to PND care contributes to a growing area of interest that remains underexplored in current research. In instances where PND-specific evidence is limited, the review draws from broader depression literature to extract relevant insights while maintaining awareness of contextual differences. Importantly, it also provides a forward-looking perspective by identifying clear gaps in the existing literature and outlining specific directions for future research. These include the development and tailoring of serious games for PND populations, informed by existing evidence on effective game mechanics in comparable mental health contexts.

However, this review also has several limitations. First, the limited number of studies specifically focused on PND constrained the depth of analysis and made it necessary to draw from broader depression literature. While this allowed for the extraction of relevant insights, it may affect the generalizability of findings to the unique psychological and social experience of individuals with PND. Second, there is a potential for selection bias in the inclusion of literature; however, attempts to reduce this included using a structured search process and adhering to the preagreed scope. Future reviews might consider incorporating elements of systematic review methodology, such as dual screening or formal bias appraisal, to enhance rigor. Third, the review was limited to English-language publications, which may have excluded relevant studies from culturally and linguistically diverse populations, potentially narrowing the applicability of findings across global contexts. Expanding the language scope in future reviews could provide a more inclusive understanding of how serious games may support help-seeking across varied cultural settings. Finally, while a narrative review design can lack a systematic or quantitative appraisal of study quality and findings, in this case, given the heterogeneity of the studies reviewed, it did allow for a comprehensive synthesis of the available literature to better pinpoint specific targeted areas needing additional research with recommendations for targeted and novel game design to address that identified gap.

Future Directions

Based on the findings of this review, the authors plan to develop a game design prototype that selectively incorporates these recommended game mechanics, which will allow for an empirical investigation of utility and efficacy. The game design will have a particular focus on the accessibility and usability needs of new parents. We will begin by conducting user research, including interviews with new parents, to further understand their challenges and preferences. This process will ensure that the game design aligns well with their experiences and provides meaningful support. We will also collaborate interprofessionally with mental health professionals to ensure that the educational content is accurate and that the overall gameplay promotes mental well-being.

Additionally, we aim to develop a hybrid solution that allows for both mobile and web-based accessibility, ensuring that new parents have multiple ways to engage with the game based on their convenience. Once the prototype is ready, we will conduct usability testing and content value testing to assess its possibility and feasibility in engaging new parents and providing mental health support. Feedback from this testing phase will guide further refinements to make the game as supportive and accessible as possible for new parents to improve help-seeking behavior for PND.

Conclusions

The evidence base for PND-related interventions is currently limited and emerging, highlighting a notable gap in future research. Based on the analysis of existing games, many of them aim to educate players on mental health topics, using engaging scenarios that require players to make supportive decisions and learn about the lived experiences of people with mental health issues. This approach encourages players to understand mental health, corrects misconceptions, and promotes empathy, ultimately reducing stigma. Narrative storytelling also plays a significant role, allowing players to experience the emotional highs and lows of a character's journey, fostering a deep connection and validation for players facing similar struggles. Empowering players was also found to be valuable in seeking professional help by embedding scenarios that encourage reaching out to health care professionals and practicing coping skills in a low-risk, game-based environment. It is argued that these mechanics help players feel more comfortable about seeking support and build their confidence in managing their own mental health.

XSL•FO
RenderX

The included games target a range of audiences, including adolescents, adults, and the general public, identifying a literature gap that only a small number are specifically designed for new parents. New parents face unique challenges, such as exhaustion and time constraints, suggesting the ideal game platform is likely to be mobile—rather than immersive—platforms as they offer the flexibility and ease of use necessary for new parents. Finally, this review underscores the importance of expanding serious game research to more effectively reach and support those experiencing PND, a population that remains inadequately treated and frequently stigmatized in modern culture.

Acknowledgments

We thank Grammarly (Superhuman Platform) for assistance with proofreading. We used the generative artificial intelligence tool ChatGPT by OpenAI to help with expanding the search terms in the search process.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Authors' Contributions

Conceptualization: WW (lead), EP (equal), VG (equal), JG (equal)
Data curation: WW
Formal analysis: WW
Investigation: WW
Methodology: WW (lead), EP (equal), VG (equal), JG (equal)
Supervision: JG (lead), EP (equal), VG (supporting)
Writing – original draft: WW
Writing – review & editing: EP (lead), JG (equal)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed search strategies.
[DOCX File, 16 KB - i-jmr_v15i1e70777_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 106 KB - i-jmr_v15i1e70777_app2.pdf]

References

1. Carlson K, Mughal S, Azhar Y, Siddiqui W. Postpartum depression. In: StatPearls: StatPearls Publishing; 2024. URL: https://www.ncbi.nlm.nih.gov/sites/books/NBK519070/ [accessed 2025-12-16]
2. Perinatal depression. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/publications/perinatal-depression [accessed 2025-12-31]
3. Putnick DL, Sundaram R, Bell EM, et al. Trajectories of maternal postpartum depressive symptoms. Pediatrics 2020 Nov;146(5):e20200857. [doi: 10.1542/peds.2020-0857] [Medline: 33109744]
4. Slomian J, Honvo G, Emonts P, Reginster JY, Bruyère O. Consequences of maternal postpartum depression: a systematic review of maternal and infant outcomes. Womens Health (Lond) 2019;15:1745506519844044. [doi: 10.1177/1745506519844044] [Medline: 31035856]
5. Llewellyn AM, Stowe ZN, Nemeroff CB. Depression during pregnancy and the puerperium. J Clin Psychiatry 1997;58 Suppl 15(15):26-32. [Medline: 9427874]
6. Field T. Postpartum depression effects on early interactions, parenting, and safety practices: a review. Infant Behav Dev 2010 Feb;33(1):1-6. [doi: 10.1016/j.infbeh.2009.10.005] [Medline: 19962196]
7. ACOG Committee on Clinical Practice Guidelines–Obstetrics. Treatment and management of mental health conditions during pregnancy and postpartum. Obstet Gynecol 2023;141(6):1262-1288. [doi: 10.1097/AOG.0000000000005202]
8. Wang Z, Liu J, Shuai H, et al. Correction: mapping global prevalence of depression among postpartum women. Transl Psychiatry 2021;11(1):543. [doi: 10.1038/s41398-021-01692-1]

XSL•FO
RenderX

9. The cost of PNDA. Gidget Foundation Australia. 2019. URL: https://www.gidgetfoundation.org.au/health-professionals/the-cost-of-pnda [accessed 2025-12-24]
10. Luca DL, Margiotta C, Staatz C, Garlow E, Christensen A, Zivin K. Financial toll of untreated perinatal mood and anxiety disorders among 2017 births in the United States. Am J Public Health 2020 Jun;110(6):888-896. [doi: 10.2105/AJPH.2020.305619] [Medline: 32298167]
11. Bauer A, Parsonage M, Knapp M, Iemmi V, Adelaja B, Hogg S. The Costs of Perinatal Mental Health Problems: London School of Economics and Political Science; 2021. URL: https://eprints.lse.ac.uk/59885/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_Bauer%2C%20M_Bauer_Costs_perinatal_%20mental_2014_Bauer_Costs_perinatal_mental_2014_author.pdf [accessed 2025-12-16]
12. Fonseca A, Gorayeb R, Canavarro MC. Women s help-seeking behaviours for depressive symptoms during the perinatal period: socio-demographic and clinical correlates and perceived barriers to seeking professional help. Midwifery 2015 Dec;31(12):1177-1185. [doi: 10.1016/j.midw.2015.09.002] [Medline: 26433622]
13. Grissette BG, Spratling R, Aycock DM. Barriers to help-seeking behavior among women with postpartum depression. J Obstet, Gynecol Neonat Nurs 2018 Nov;47(6):812-819. [doi: 10.1016/j.jogn.2018.09.006]
14. Mouaheb H, Fahli A, Moussetad M, Eljamali S. The serious game: what educational benefits? Procedia - Soc Behav Sci 2012;46:5502-5508. [doi: 10.1016/j.sbspro.2012.06.465]
15. Azadegan A, Riedel JC, Baalsrud Hauge J. Serious games adoption in corporate training. Presented at: Serious Games Development and Applications: Third International Conference, SGDA 2012; Sep 26-29, 2012; Bremen, Germany p. 74-85. [doi: 10.1007/978-3-642-33687-4_6]
16. Higham H. Simulation past, present and future-a decade of progress in simulation-based education in the UK. BMJ Simul Technol Enhanc Learn 2021;7(5):404-409. [doi: 10.1136/bmjstel-2020-000601] [Medline: 35515719]
17. Schöbel SM, Janson A, Söllner M. Capturing the complexity of gamification elements: a holistic approach for analysing existing and deriving novel gamification designs. Eur J Inf Syst 2020 Nov 1;29(6):641-668. [doi: 10.1080/0960085X.2020.1796531]
18. Ruiz M, Moreno M, Girela-Serrano B, et al. Winning the game against depression: a systematic review of video games for the treatment of depressive disorders. Curr Psychiatry Rep 2022 Jan;24(1):23-35. [doi: 10.1007/s11920-022-01314-7]
19. Townsend C, Humpston C, Rogers J, Goodyear V, Lavis A, Michail M. The effectiveness of gaming interventions for depression and anxiety in young people: systematic review and meta-analysis. BJPsych Open 2022 Jan 7;8(1):e25. [doi: 10.1192/bjo.2021.1078] [Medline: 34991767]
20. Solinska-Nowak A, Magnuszewski P, Curl M, et al. An overview of serious games for disaster risk management – prospects and limitations for informing actions to arrest increasing risk. Int J Disaster Risk Reduct 2018 Oct;31:1013-1029. [doi: 10.1016/j.ijdrr.2018.09.001]
21. van den Heuvel JF, Groenhof TK, Veerbeek JH, et al. eHealth as the next-generation perinatal care: an overview of the literature. J Med Internet Res 2018;20(6):e202. [doi: 10.2196/jmir.9262]
22. Sayakhot P, Carolan-Olah M. Internet use by pregnant women seeking pregnancy-related information: a systematic review. BMC Pregnancy Childbirth 2016 Dec;16(1):1-10. [doi: 10.1186/s12884-016-0856-5]
23. O'Brien E, Rauf Z, Alfirevic Z, Lavender T. Women's experiences of outpatient induction of labour with remote continuous monitoring. Midwifery 2013 Apr;29(4):325-331. [doi: 10.1016/j.midw.2012.01.014] [Medline: 23159160]
24. Ming WK, Mackillop LH, Farmer AJ, et al. Telemedicine technologies for diabetes in pregnancy: a systematic review and meta-analysis. J Med Internet Res 2016 Nov 9;18(11):e290. [doi: 10.2196/jmir.6556] [Medline: 27829574]
25. Lau Y, Htun TP, Wong SN, Tam WSW, Klainin-Yobas P. Therapist-supported internet-based cognitive behavior therapy for stress, anxiety, and depressive symptoms among postpartum women: a systematic review and meta-analysis. J Med Internet Res 2017 Apr 28;19(4):e138. [doi: 10.2196/jmir.6712] [Medline: 28455276]
26. Sondaal SFV, Browne JL, Amoakoh-Coleman M, et al. Assessing the effect of mhealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLOS ONE 2016;11(5):e0154664. [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
27. Tapia-Conyer R, Lyford S, Saucedo R, et al. Improving perinatal care in the rural regions worldwide by wireless enabled antepartum fetal monitoring: a demonstration project. Int J Telemed Appl 2015;2015(1):794180. [doi: 10.1155/2015/794180] [Medline: 25691900]
28. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020 Apr 13;22(4):e17011. [doi: 10.2196/17011] [Medline: 32281939]
29. Dosani A, Arora H, Mazmudar S. mHealth and perinatal depression in low-and middle-income countries: a scoping review of the literature. IJERPH 2020;17(20):7679. [doi: 10.3390/ijerph17207679]
30. Wan Mohd Yunus WMA, Matinolli HM, Waris O, et al. Digitalized cognitive behavioral interventions for depressive symptoms during pregnancy: systematic review. J Med Internet Res 2022 Feb 23;24(2):e33337. [doi: 10.2196/33337] [Medline: 35195532]
31. Lau Y, Yen KY, Wong SH, Cheng JY, Cheng LJ. Effect of digital cognitive behavioral therapy on psychological symptoms among perinatal women in high income-countries: a systematic review and meta-regression. J Psychiatr Res 2022 Feb;146:234-248. [doi: 10.1016/j.jpsychires.2021.11.012] [Medline: 34774299]

XSL•FO
RenderX

32. Fallon V, Davies SM, Silverio S, Creagh L. Virtual reality interventions designed to support parents during and throughout the first year after birth: a scoping review. Digit Health 2024;10:20552076241245373. [doi: 10.1177/20552076241245373] [Medline: 38655377]
33. Motrico E, Conejo-Cerón S, Martín-Gómez C, Gómez-Gómez I, Fonseca A, Moreno-Peral P. Effectiveness of web-based and mobile-based psychological interventions to prevent perinatal depression: study protocol for a systematic review and meta-analysis of randomized controlled trials. Internet Interv 2021 Dec;26:100471. [doi: 10.1016/j.invent.2021.100471]
34. Spadaro B, Martin-Key NA, Funnell E, Bahn S. mHealth solutions for perinatal mental health: scoping review and appraisal following the mHealth index and navigation database framework. JMIR mHealth uHealth 2022 Jan 17;10(1):e30724. [doi: 10.2196/30724] [Medline: 35037894]
35. Pilkington PD, Whelan TA, Milne LC. A review of partner‐inclusive interventions for preventing postnatal depression and anxiety. Clin Psychol (Aust Psychol Soc) 2015 Jul 1;19(2):63-75. [doi: 10.1111/cp.12054]
36. Liberto TL. Screening for depression and help-seeking in postpartum women during well-baby pediatric visits: an integrated review. J Pediatr Health Care 2012 Mar;26(2):109-117. [doi: 10.1016/j.pedhc.2010.06.012] [Medline: 22360930]
37. What is perinatal depression? American Psychiatric Association. URL: https://www.psychiatry.org/patients-families/peripartum-depression/what-is-peripartum-depression [accessed 2026-01-15]
38. Fleming T, Cheek C, Merry SN, et al. Serious games for the treatment or prevention of depression: a systematic review. Rev Psicopatologia Psicologia Clin 2014;19(3):227-242. [doi: 10.5944/rppc.vol.19.num.3.2014.13904]
39. Li J, Theng YL, Foo S. Game-based digital interventions for depression therapy: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2014 Aug;17(8):519-527. [doi: 10.1089/cyber.2013.0481] [Medline: 24810933]
40. Dias LPS, Barbosa JLV, Vianna HD. Gamification and serious games in depression care: a systematic mapping study. Telematics Inf 2018 Apr;35(1):213-224. [doi: 10.1016/j.tele.2017.11.002]
41. Kowal M, Conroy E, Ramsbottom N, Smithies T, Toth A, Campbell M. Gaming your mental health: a narrative review on mitigating symptoms of depression and anxiety using commercial video games. JMIR Serious Games 2021 Jun 16;9(2):e26575. [doi: 10.2196/26575] [Medline: 34132648]
42. Pine R, Fleming T, McCallum S, Sutcliffe K. The effects of casual videogames on anxiety, depression, stress, and low mood: a systematic review. Games Health J 2020 Aug;9(4):255-264. [doi: 10.1089/g4h.2019.0132] [Medline: 32053021]
43. King M, Marsh T, Akcay Z. A review of indie games for serious mental health game design. Presented at: Serious Games: Joint International Conference, JCS; Jan 12-13, 2022; Virtual Event p. 138-152. [doi: 10.1007/978-3-030-88272-3_11]
44. Martinez K, Menéndez-Menéndez MI, Bustillo A. Awareness, prevention, detection, and therapy applications for depression and anxiety in serious games for children and adolescents: systematic review. JMIR Serious Games 2021;9(4):e30482. [doi: 10.2196/30482]
45. Yen HY, Chiu HL. Virtual reality exergames for improving older adults' cognition and depression: a systematic review and meta-analysis of randomized control trials. J Am Med Dir Assoc 2021 May;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]
46. Abd-Alrazaq A, Al-Jafar E, Alajlani M, et al. The effectiveness of serious games for alleviating depression: systematic review and meta-analysis. JMIR Serious Games 2022;10(1):e32331. [doi: 10.2196/32331]
47. Kim Y, Hong S, Choi M. Effects of serious games on depression in older adults: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022;24(9):e37753. [doi: 10.2196/37753]
48. Chitale V, Baghaei N, Playne D, et al. The use of videogames and virtual reality for the assessment of anxiety and depression: a scoping review. Games Health J 2022 Jul 26(6):341-354. [doi: 10.1089/g4h.2021.0227] [Medline: 35881854]
49. Gliosci R, Silva TBPE. Therapeutic interventions with videogames in treatments for depression: a systematic review. Games Health J 2023 Aug;12(4):269-279. [doi: 10.1089/g4h.2022.0094] [Medline: 36946750]
50. Gómez-León MI. Serious games to support emotional regulation strategies in educational intervention programs with children and adolescents. Systematic review and meta-analysis. Heliyon 2025 Feb;11(4):e42712. [doi: 10.1016/j.heliyon.2025.e42712]
51. First MB. Diagnostic and Statistical Manual of Mental Disorders, 5th edition, and clinical utility. J Nerv Ment Dis 2013 Sep;201(9):727-729. [doi: 10.1097/NMD.0b013e3182a2168a] [Medline: 23995026]
52. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. Med Teach 2020 Aug;42(8):846-854. [doi: 10.1080/0142159X.2020.1755030] [Medline: 32356468]
53. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023 Oct;22. [doi: 10.1177/16094069231205789]
54. Wehbe RR, Watson DK, Tondello GF, et al. Above Water: an educational game for anxiety. 2016 Oct Presented at: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts; Oct 16-19, 2016 p. 79-84. [doi: 10.1145/2968120.2971804]
55. Tuijnman A, Kleinjan M, Hoogendoorn E, Granic I, Engels RC. A game-based school program for mental health literacy and stigma regarding depression (Moving Stories): protocol for a randomized controlled trial. JMIR Res Protoc 2019 Mar 14;8(3):e11255. [doi: 10.2196/11255] [Medline: 30869652]

XSL•FO
RenderX

56. Cheng VWS, Davenport TA, Johnson D, Vella K, Mitchell J, Hickie IB. An app that incorporates gamification, mini-games, and social connection to improve men's mental health and well-being (MindMax): participatory design process. JMIR Ment Health 2018 Nov 19;5(4):e11068. [doi: 10.2196/11068] [Medline: 30455165]
57. Mueller S, Soriano D, Boscor A, et al. MANTRA: development and localization of a mobile educational health game targeting low literacy players in low and middle income countries. BMC Public Health 2020 Jul 28;20(1):1171. [doi: 10.1186/s12889-020-09246-8] [Medline: 32723317]
58. Sara TA, Nazifa SN, Tasneem S, Shakib TU, Islam MN. PPDHero: requirements elicitation and development of a system to empower new mothers on postpartum depression. 2024 Aug Presented at: International Conference on Intelligent Systems Design and Applications; Dec 11-13, 2023 p. 389-398. [doi: 10.1007/978-3-031-64850-2_37]
59. Cangas AJ, Navarro N, Parra JMA, et al. Stigma-Stop: a serious game against the stigma toward mental health in educational settings. Front Psychol 2017;8:1385. [doi: 10.3389/fpsyg.2017.01385] [Medline: 28878702]
60. Hanisch SE, Birner UW, Oberhauser C, Nowak D, Sabariego C. Development and evaluation of digital game-based training for managers to promote employee mental health and reduce mental illness stigma at work: quasi-experimental study of program effectiveness. JMIR Ment Health 2017 Aug 4;4(3):e31. [doi: 10.2196/mental.7600] [Medline: 28778839]
61. Ferchaud A, Seibert J, Sellers N, Escobar Salazar N. Reducing mental health stigma through identification with video game avatars with mental illness. Front Psychol 2020;11:2240. [doi: 10.3389/fpsyg.2020.02240] [Medline: 33013585]
62. Depression Quest. URL: http://www.depressionquest.com/ [accessed 2025-12-16]
63. Keep in Mind: Remastered. Steam. URL: https://store.steampowered.com/app/643200 [accessed 2025-12-16]
64. Before I Forget. Steam. URL: https://store.steampowered.com/app/1126600/Before [accessed 2025-12-16]
65. Sea of Solitude. Wikipedia. URL: https://en.wikipedia.org/wiki/Sea_of_Solitude [accessed 2025-12-16]
66. Rodrigues MAF, de Oliveira TRC, de Figueiredo DL, et al. An interactive story decision-making game for mental health awareness. 2022 Aug Presented at: 2022 IEEE 10th International Conference on Serious Games and Applications for Health(SeGAH); Aug 10-12, 2022; Sydney, Australia p. 1-8. [doi: 10.1109/SEGAH54908.2022.9978592]
67. Taubman-Ben-Ari O, Ben-Yaakov O, Chasson M. Parenting stress among new parents before and during the COVID-19 pandemic. Child Abuse Neglect 2021 Jul;117:105080. [doi: 10.1016/j.chiabu.2021.105080] [Medline: 33930664]
68. Saxbe D, Rossin-Slater M, Goldenberg D. The transition to parenthood as a critical window for adult health. Am Psychol 2018 Dec;73(9):1190-1200. [doi: 10.1037/amp0000376] [Medline: 30525801]
69. Msetfi RM, Kornbrot DE, Halbrook YJ. The association between the sense of control and depression during the COVID-19 pandemic: a systematic review and meta-analysis. Front Psychiatry 2024;15:1323306. [doi: 10.3389/fpsyt.2024.1323306] [Medline: 38414499]
70. Liu Y, Yang Y, Bai X, Chen Y, Mo L. Do immediate external rewards really enhance intrinsic motivation? Front Psychol 2022;13:853879. [doi: 10.3389/fpsyg.2022.853879]
71. Moore RL, Hwang W, Moses JD. A systematic review of mobile-based microlearning in adult learner contexts. Educ Technol Soc 2024;27(1):137-146. [doi: 10.30191/ETS.202401_27(1).SP02]

Abbreviations

CBT: cognitive behavioral therapy
CVG: casual video game
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition
mHealth: mobile health
PND: postnatal depression
VR: virtual reality

Edited by N Cahill; submitted 01.Jan.2025; peer-reviewed by A Gazis, M Evans; revised version received 01.Dec.2025; accepted 01.Dec.2025; published 16.Jan.2026.

Please cite as:
Wang W, Penney E, Gay V, Garcia J
Exploring Serious Games in Supporting Postnatal Depression: Narrative Review
Interact J Med Res 2026;15:e70777
URL: https://www.i-jmr.org/2026/1/e70777
doi:10.2196/70777

© Weijie Wang, Erika Penney, Valerie Gay, Jaime Garcia. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 16.Jan.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

XSL•FO
RenderX

in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Key Predictors of Appropriate Discharge Destinations for Older Inpatients in Acute Care: Scoping Review

Christoph Leinert[1], Dr med; Marina Liselotte Fotteler[1,2], Dr biol hum; Thomas Derya Kocar[1], Dr med; Jessica Wolf[1]; Lisa Beissel[1], BSc; Kathrin Grummich[3], Diplom (FH); Dhayana Dallmeier[1,4], Prof Dr med, PhD; Michael Denkinger[1], Prof Dr med

[1]Institute for Geriatric Research at Agaplesion Bethesda Ulm, Ulm University Medical Center, Zollernring 26, Ulm, Germany

[2]DigiHealth Institute, Neu-Ulm University of Applied Sciences, Neu-Ulm, Germany

[3]ZB MED - Information Centre for Life Sciences, Cologne, Germany

[4]Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States

Corresponding Author:
Christoph Leinert, Dr med
Institute for Geriatric Research at Agaplesion Bethesda Ulm, Ulm University Medical Center, Zollernring 26, Ulm, Germany

Abstract

Background: Postacute care (PAC) services are important to ensure functional recovery and provide adequate care for geriatric inpatients in acute care. The choice between different PAC options can be challenging, and predictors for the most appropriate among diverse discharge options are warranted.

Objective: We conducted a scoping review to identify predictors of appropriate discharge destinations for older adults (≥65 y) in acute care transitioning to different PAC settings and extract the most relevant predictors for different PAC settings as well as a generalizable set of predictor domains.

Methods: The databases of Medline, Embase, Cochrane Central Register of Controlled Trials, PsycINFO, CINAHL, and Emcare were systematically searched for English or German literature published until February 25, 2022. A total of 3 researchers screened, extracted, and categorized the data according to domains, discharge destinations, mean age, and health care systems origin, focusing on predictors that increase the likelihood of a discharge destination (positive predictors). The Jaccard index was calculated to compare the similarity between different possible domain combinations and existing literature.

Results: Of 22,382 records screened, 171 quantitative and 10 qualitative studies were included. After separating combined discharge destinations, we found 1047 predictors for different discharge destinations including nursing home (n=297, 28%), skilled nursing facility (n=223, 21%), inpatient rehabilitation (n=206, 20%), home with (n=97, 9%) or without (n=74, 7%) support, assisted living (n=63, 6%), and early inpatient rehabilitation (n=21, 2%). Of all positive predictors (n=723), age was the most frequently reported predictor (80/723, 11%). Geriatric syndromes were found more often in patients 80 years or older (121/192, 63%) and in non-US studies (174/285, 61%). The top reported predictors for discharge to nursing homes were diagnosed dementia (9/297, 3%) and deficits in instrumental activities of daily living (ADL; 10/297, 3%); for discharge to rehabilitation, the top predictors were longer length of stay (11/205, 5%) and existent cardiopulmonary disease (10/205, 5%); and for back home without support, the top predictors were good ADL (10/74, 14%) and mobility assessments (9/74, 12%). Among 20 predictor domains, 8 were most concordant with the literature: cognitive impairment, ADL, demographics, social support, hospitalization data, multimorbidity, mobility, and primary diagnosis.

Conclusions: This scoping review provides a comprehensive overview of predictors for appropriate discharge decisions in older adults in acute care, stratified by destination, age, study origin, and the predictor domains most concordant with the literature. The results will be valuable to inform the choice of features for clinical decision support systems, including the training of machine learning algorithms.

(Interact J Med Res 2026;15:e76582) doi:10.2196/76582

KEYWORDS

CDSS; clinical decision support system; continuity of care; discharge destination; discharge planning; feature selection; geriatrics; postacute care; predictor

XSL•FO
RenderX

Introduction

Geriatric inpatients are at a high risk of functional decline during acute care treatment, and post-acute care (PAC) services are needed to ensure recovery [1]. Effective discharge planning has been shown to mitigate readmission rates, reduce hospital stays, lower associated costs, and enhance patient satisfaction [2]. Thus, choosing the most appropriate discharge destination is vital within geriatric comanagement, guaranteeing alignment with individualized rehabilitation and care needs. Modern health care systems present diverse PAC options, making selection challenging and highlighting the need for predictors of appropriate discharge destinations [3-6].

The World Health Organization defines continuity of care (COC) as the degree to which health care events are perceived as connected and coherent over time, aligning with patients' health necessities and preferences [7]. COC can be further categorized into relational continuity (patient-provider relationship), informational continuity (communication), and management continuity (coordination) [8]. In the context of geriatric comanagement in acute inpatient care, the choice of the most appropriate discharge destination refers to the management continuity aspect. In addition to COC and discharge planning, various terms encompass this dimension, including integrated care, case management, or transitional care [9,10].

Existing research has focused on predictors differentiating binary outcomes [11-13]. However, there remains a need for predictors supporting decisions among multiple discharge options to not miss relevant features.

To address this gap, we conducted a scoping review to identify potential predictors for the most appropriate discharge destination for older inpatients in acute care transitioning to different PAC options such as outpatient, inpatient or early rehabilitation, skilled nursing facility (SNF), nursing home, assisted living, or home-based care with or without support. As different health care systems offer different discharge destinations and funding options, we also planned to stratify according to the different health care system origins. Our scoping review will inform the feature selection for the development of a machine learning−driven clinical decision support system (CDSS) within the "Supporting Surgery with Geriatric Co-Management and AI" (SURGE-Ahead) project by providing a broad overview of predictive measures for different discharge options [14].

Methods

Conceptualization

We conducted a scoping review adhering to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Scoping Review extension guidelines; Checklist 1) and the Joanna Briggs Institute guidance [15,16]. A review protocol was registered on Open Science Framework and was adapted after a piloting phase to focus on "predictors" instead of "predictors and outcome measures" due to a lack of specific outcome measures for COC identified [17].

Eligibility Criteria

Using the Joanna Briggs Institute's population-concept-context framework, our focus was on older adult inpatients 65 years and older treated in acute surgical or medical hospital departments [18]. We included only participants 65 years and older, verifying this by examining the inclusion criteria and the reported age structure of each study population. Studies were included if they reported an age range with a lower limit of 65 years or above or at least provided a mean or median of 65 years or more. Publications with unclear age demographics were conservatively excluded. We aimed to identify predictors utilized by health care professionals to determine the most appropriate PAC setting. Predictors were defined as discrete, separable parameters associated with a specific discharge destination. When faced with complex models combining multiple parameters, we included the individual parameters whenever feasible but not the complex model itself. In quantitative analyses, we further required evidence of a statistically significant effect ($P \leq .05$). Studies had to involve transitions from acute care to different PAC options. German- or English-language publications were considered, excluding letters, comments, case studies, editorials, and studies primarily addressing health economic issues or involving psychiatry or rehabilitation departments for older adults. We excluded studies that chose to report composite discharge options like "discharge to nursing home or death," as these outcomes should not be combined. Similarly, studies including palliative care services like hospice as part of a composite discharge destination were excluded, as we consider individualized decision-making essential for palliative care.

Search Strategy

Comprehensive search strategies were developed that included different concepts for the successful choice of discharge destinations, including continuity of care, coordination of care, transition of care, integrated health care, case management, discharge planning, or rehabilitation eligibility determination. On February 25, 2022, the databases Medline (OVID interface), Embase (OVID), Cochrane Central Register of Controlled Trials (Wiley), PsycINFO (EBSCOhost), CINAHL (EBSCOhost), and Emcare (OVID) were systematically searched, guided by an experienced information specialist (KG). Search strategies for all databases searched are shown in Multimedia Appendix 1.

Study Selection and Data Extraction

Records obtained through the database search were imported into the Covidence systematic review software (Veritas Health Innovation). Additionally, the primary publications of all reviews identified in the literature search were imported into Covidence. Duplicates were removed using the deduplication functionality of the software, followed by manual verification in unclear cases marked by the software (MLF and CL). Two out of the 4 reviewers (MLF, CL, LB, and JW) independently screened all titles and abstracts for eligibility, followed by full-text screening. Any discrepancies were resolved through adjudication by a third author. Data extraction for each article was conducted by 2 of the 3 authors independently (MLF, CL, and JW) and consolidated by the team. The extracted data

XSL•FO
RenderX

encompassed study and population characteristics, health care systems origin, discharge destination, statistical analysis type, predictor, and quantitative measures of predictive strength (odds ratio, relative risk, confidence interval, *P* value) if available. The identified predictors were categorized into different domains after discussion in the review team. Data extraction for qualitative studies was conducted separately using a thematic evidence synthesis, assigning all qualitative and quantitative predictors, including representative quotes, to the predefined domains [19].

Data Analysis

Predictive directions indicating an increased (positive) or reduced (negative) probability of discharge to a specific environment were determined from odds ratios, relative risks, or manual labeling and joint consensus when necessary (MLF and CL). The negative predictors reported in the included studies often used double negations or lacked a differentiation of discharge alternatives. To avoid misdirected results, we focused our analyses on the positive predictors.

Predictors were also stratified based on discharge destinations (outpatient, inpatient or early rehabilitation, SNF, nursing home, assisted living, or home-based care with or without support, long-term acute care setting, or other acute care setting). In studies that consolidated multiple discharge destinations into a single outcome, these were extracted separately. The direction of effect was added to the top 5 predictors per discharge destination via manual extraction from the literature and joint consensus labeling (MLF and CL). Additionally, we stratified by mean population age (<80 y or ≥80 y) and health care systems origin: predictors from Anglo-European studies with mostly publicly funded health care systems (Europe, Canada, Australia) were also analyzed separately from studies from the United States of America with a relevant proportion of pay-for-service health services [20-22].

Similar predictive factors identified across multiple studies were consolidated, whereas validated assessment tools were treated independently. In some studies, data on predictive measures were reported without using a validated assessment instrument. In these cases, we categorized the results as “no specific score.” As an example, in the activities of daily living (ADL) domain, many studies reported “decreased ADL” or similar but did not use an established assessment like the Barthel index (BI). In these cases, we categorized the reported predictors as “ADL no specific score.”

The Jaccard Index can be used to measure similarity and diversity of sample sets [23]. It was used to determine the concordance between different potential domain combinations and those reported in the reviewed studies. A higher index value signifies better agreement. Based on the 20 defined domains, we juxtaposed approximately 1 million potential combination sets (all possible combinations) to be compared with those found in the literature. The domain set with the highest average Jaccard index value was identified as a potentially generalizable combination of key predictor domains most commonly associated with each other and existing literature.

Descriptive statistics and data manipulation were carried out using Microsoft Excel 2019 (Microsoft). We used Python 3.12.2 for the analysis, using the itertools standard library for generating combinations and set class methods for calculating the Jaccard index. For loading the data from Excel, the pandas 2.2.1 library was used.

Results

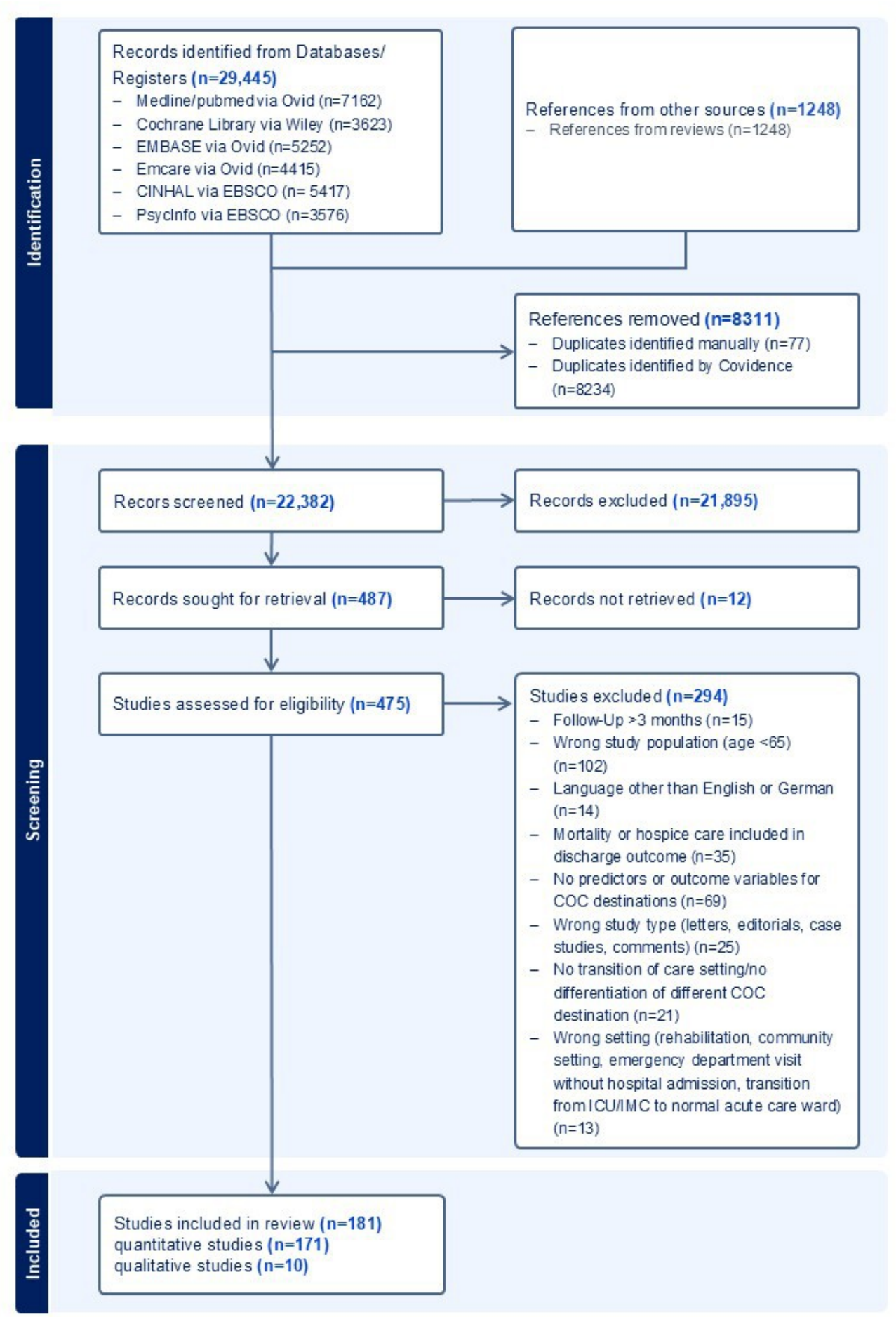
Overview

Our search yielded 22,382 database entries after removing duplicates. After title or abstract screening, we assessed the full texts of 475 studies for eligibility and included 181 studies (n=171 quantitative studies, n=10 qualitative studies). Figure 1 shows the PRISMA flow chart of the scoping review. All included studies are listed in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flowchart. COC: continuity of care; ICU: intensive care unit; IMC: intermediate care unit.

XSL•FO
RenderX



XSL•FO
RenderX

The included studies were published between 1984 and 2022. Most of the 181 studies originated from North America (USA: n=84, 46%; Canada: n=14, 8%). Europe accounted for approximately one-third (n=60, 33%), including studies from the United Kingdom (n=11, 6%), the Netherlands (n=11, 6%), Switzerland (n=8, 4%), Germany (n=7, 4%), Italy (n=5, 3%), and Spain (n=3, 2%). The remaining 15% were conducted in other parts of the world, including Australia (n=14, 8%) and Japan (n=7, 4%).

A total of 275 (in part combined) settings were described, and most studies (n=161, 58.5%) were conducted in surgical settings such as trauma or orthogeriatrics (n=75, 27%) and general or visceral surgery (n=28, 10%). Other prevalent settings were geriatrics (n=29, 11%) and general or internal medicine (n=27, 10%). In 7% (n=19) of settings, no clinical department was specified. The included studies encompass data from 6,357,026 participants with a mean age of 77.42 years.

Quantitative Studies

In the 171 quantitative studies, a total of 856 predictors for various discharge destinations were identified. More than half of these studies were prospective (88/171, 51.5%), while the remaining utilized a retrospective (78/171, 46%) or mixed approach (5/171, 3%). Most reported predictors were positive predictors (723/856, 84.5%). About two-thirds of the studies (116/171, 68%) reported the participants' mean age, with 55 of 171 (32%) studies having a mean age of 80 years or older and 61 (36%) having a mean age younger than 80 years.

Predictors

Age emerged as the most prevalent predictor across all categories. The 10 most frequent positive predictors are listed in Table 1 for all studies, and stratified by mean age <80 years, ≥80 years, US and Anglo-European studies.

XSL•FO
RenderX

Table . Top 10 positive predictors for continuity of care of all the included studies, stratified by mean age (<80 vs ≥80 y) and US and Anglo-European origin[a].

Predictor	All (n=723), n (%)	<80 years mean age (n=274), n (%)	≥80 years mean age (n=192), n (%)	US (n=413), n (%)	Anglo-European (n=285), n (%)
Age	80 (11)	34 (12)	20 (10)	50 (12)	28 (10)
Wound problems	25 (4)	23 (8)	—[b]	25 (6)	—
Cardiopulmonary disease	24 (3)	16 (6)	—	23 (6)	—
ADL[c] NSS[d]	23 (3)	10 (4)	—	19 (5)	—
Length of stay	22 (3)	6 (2)	—	—	14 (5)
Infectious disease	17 (2)	16 (6)	—	16 (4)	—
Number of comorbidities	16 (2)	11 (4)	—	15 (4)	—
ASA[e] score [24] (multimorbidity)	16 (2)	9 (3)	—	13 (3)	—
Falls	14 (2)	—	5 (3)	9 (2)	7 (3)
Frailty index[f]	14 (2)	8 (3)	—	—	—
Female sex	14 (2)	—	6 (3)	9 (2)	—
Mobility NSS	—	7 (3)	5 (3)	9 (2)	—
Caregiver support	—	6 (2)	—	—	—
Deep vein thrombosis diagnosis	—	6 (2)	—	—	—
Source of admission	—	6 (2)	—	—	—
IADL[g] NSS	—	—	9 (5)	—	9 (3)
Problems personal hygiene	—	—	9 (5)	—	7 (3)
Katz score [25] (ADL)	—	—	8 (4)	—	11 (4)
Charlson comorbidity index [26]	—	—	7 (4)	—	—
Dementia diagnosis	—	—	6 (3)	—	7 (3)
Cognitive deficits NSS	—	—	5 (3)	—	—
Living with or without companion	—	—	5 (3)	—	7 (3)
Barthel index [27] (ADL)	—	—	—	—	7 (3)
Injury severity score [28]	—	—	—	—	7 (3)

[a]n: number of extracted positive predictors, % proportion of extracted positive predictors in each group.

[b]Not applicable.

[c]ADL: activities of daily living.

[d]NSS: no specific score.

[e]ASA: American Society of Anesthesiologists.

[f]Frailty index: different assessments possible.

[g]IADL: instrumental activities of daily living.

Discharge Destinations

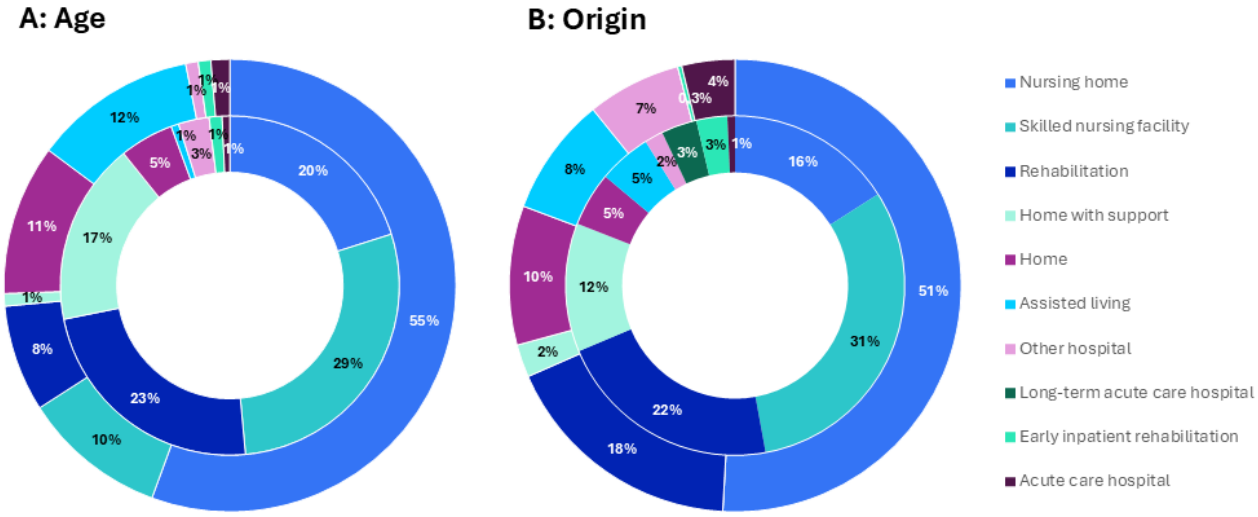
We analyzed predictor-destination sets (n=1047) across 10 different discharge options. A discharge to nursing homes emerged as the most common option (n=297, 28%), followed by SNF (n=223, 21%), inpatient rehabilitation (n=206, 20%), discharge home with support (n=97, 9%) and without support (n=74, 7%), and assisted living (n=63, 6%). Less frequently

XSL•FO
RenderX

found were predictors related to early inpatient geriatric rehabilitation (n=21, 2%), long-term acute care hospitals (n=23, 2%), and other acute care hospitals (n=18, 2%).

Nursing homes were the predominant destination in the oldest old category (≥80 y: 124/223, 55% vs <80 y: 82/408, 20%), and inpatient rehabilitation (<80 y: 95/408, 23% vs. ≥80 y: 17/223, 8%) and SNF (<80 y: 116/408, 29% vs ≥80 y: 23/223, 10%) were the main destination in the younger old category (see Figure 2A). In Anglo-European studies, the most frequently reported discharge destination was nursing home (174/342, 51%) compared to SNFs (221/677, 31%) in US studies (see Figure 2B).

Figure 2. Predictors stratified by discharge destination. (A) Proportion of discharge destinations of studies stratified with mean age <80 years (inner circle, n=408) and ≥80 years (outer circle, n=223). (B) Proportion of discharge destinations of studies originating from the United States (inner circle, n=677) and Anglo-European countries (outer circle, n=342).



Predictors for Different Discharge Destinations

Age emerged as the predominant positive predictor for most discharge options (Table 2). Nevertheless, for a discharge home, the BI was found more often. Because of the scarcity of predictors in the categories early inpatient rehabilitation, acute care hospital, long-term acute care hospital, and other hospitals, no further stratified analysis was done for these destinations.

XSL•FO
RenderX

Table . Top 5 positive predictors for 6 discharge destinations[a].

Predictor	Nursing home (n=297), n (%)	Skilled nursing facility (n=211), n (%)	Rehabilitation (n=205), n (%)	Home with support (n=97), n (%)	Home (n=74), n (%)	Assisted living (n=63), n (%)
Age (older)[b]	33 (11)	24 (11)	25 (12)	10 (10)	—[c]	4 (6)
IADL[d] NSS[e] (dependent on the instrument used)	10 (3)	—	—	—	—	—
Dementia diagnosis (yes)	9 (3)	—	—	—	—	—
Length of stay (longer)	9 (3)	—	11 (5)	4 (4)	—	—
Sex (female)	9 (3)	—	—	—	—	4 (6)
ADL[f] NSS (dependent on the instrument used)	8 (3)	8 (4)	9 (4)	4 (4)	—	—
Cardiopulmonary disease (yes)	—	11 (5)	10 (5)	9 (9)	—	—
Wound problem (yes)	—	11 (5)	10 (5)	6 (6)	—	—
Number of comorbidities (higher)	—	8 (4)	—	4 (4)	—	—
Frailty index (frailer)	—	8 (4)	—	—	—	—
Infectious disease (yes)	—	—	—	6 (6)	—	—
Age (younger)	—	—	—	—	4 (5)	—
Barthel index [27] (higher)	—	—	—	—	6 (8)	—
Katz score [25] (ADL) (higher)	—	—	—	—	4 (5)	—
Cumulated ambulation score [29] (higher)	—	—	—	—	3 (4)	—
De Morton Mobility Index [30] (higher)	—	—	—	—	3 (4)	—
Short physical performance battery [31] (higher)	—	—	—	—	3 (4)	—
ICU[g] treatment (yes)	—	—	—	—	—	4 (6)
Older people's QoL[h] questionnaire [32] (low)	—	—	—	—	—	4 (6)
Anesthesia or ICU treatment procedure (yes)	—	—	—	—	—	3 (5)

[a]n: number of extracted positive predictors, % proportion of extracted positive predictors in each group.

[b]Direction of effect in parentheses.

[c]Not applicable.

[d]IADL: instrumental activities of daily living.

[e]NSS: no specific score.

XSL•FO
RenderX

[f]ADL: activities of daily living.

[g]ICU: intensive care unit.

[h]QoL: quality of life.

Negative predictors were a minority (133/856, 15.5%) and very heterogeneous, and often seemed not useful for the identification of the most appropriate discharge destinations, such as ethnicity or male sex for discharge to rehabilitation. More relevant predictors were the unavailability of caregivers for discharge to a nursing home, a hip fracture diagnosis, or a higher frailty index, not allowing for discharge home.

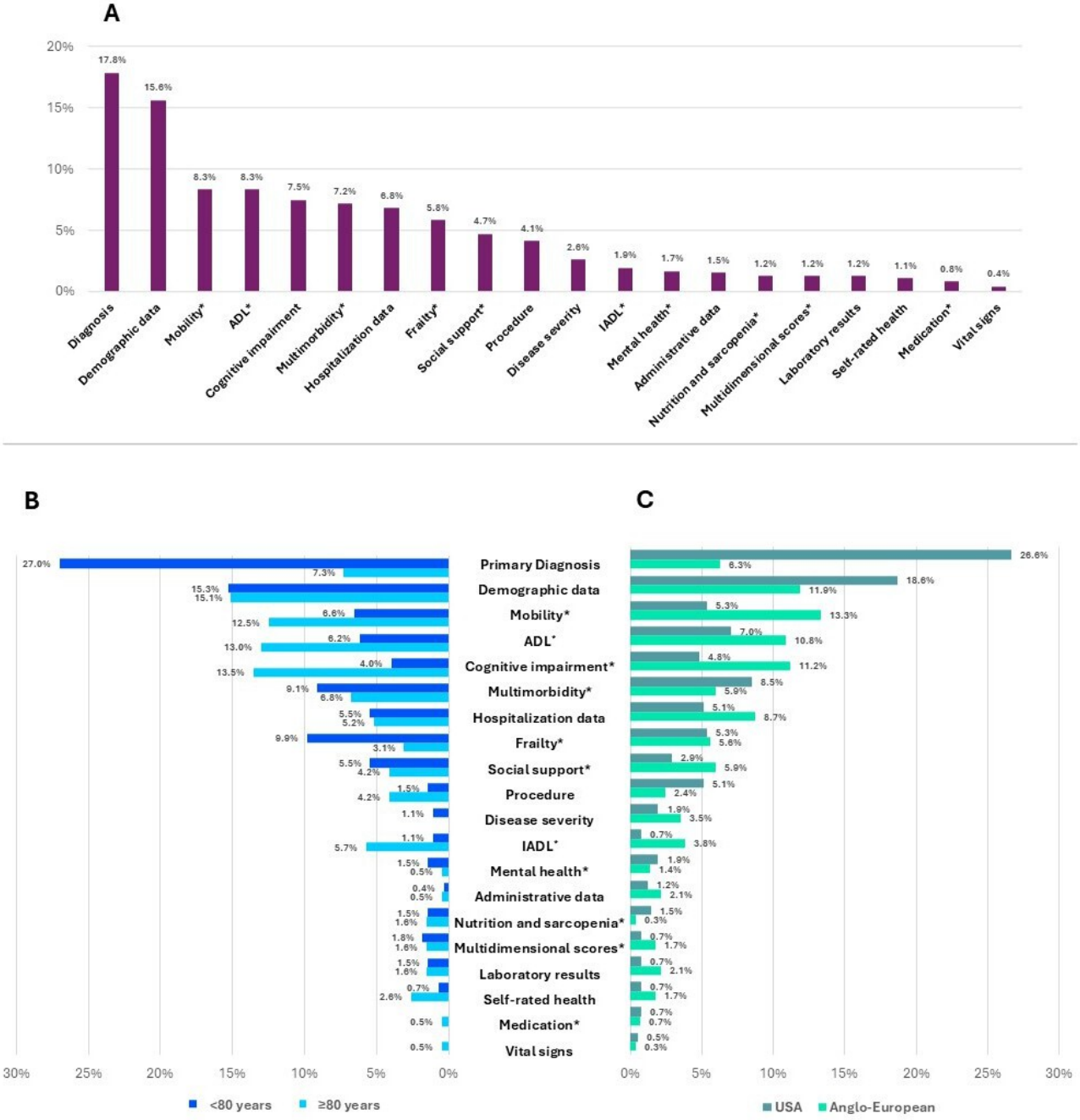
A detailed overview of the extracted evidence stratified by mean age, study type, study origin, and discharge destination is shown in Multimedia Appendix 3 and, according to the domains, positive as well as negative predictors, predictive strength indicators, and corresponding literature references in Multimedia Appendix 4.

Predictor Domains

After initial data inspection, 20 predictor domains representing common health care data and geriatric syndromes were defined, and all the extracted predictors were categorized accordingly. Diagnoses (129/723, 17.8%) and demographic data (113/723, 15.6%) formed the largest proportion. The assessment domains of geriatric syndromes, such as mobility (60/723, 8%), ADL (60/723, 8%), cognitive impairment (54/723, 8%), and frailty (42/723, 6%), were commonly represented as well (Figure 3A). Geriatric syndromes were more often predictive in the oldest old category (all: 352/723, 48.7%; <80 y: 129/274, 47.1%; ≥80 y: 121/192, 63%). In contrast, the frailty domain was a more frequent predictor in the younger old category (<80 y: 27/274, 10% vs ≥80 y: 6/192, 3%), as shown in Figure 3B.

XSL•FO
RenderX

Figure 3. Predictor domains. (A) Dark red columns: proportion of predictors across predictor domains (n=723). (B) Proportion of predictors across all predictor domains stratified by age. Dark blue bars: mean age <80 years (n=274); light blue bars: mean age ≥80 years (n=192). (C) Proportion of predictors across all predictor domains stratified by study origin. Dark green bars: US origin (n=413); light green bars: Anglo-European origin (n=286). ADL: activities of daily living, IADL: instrumental activities of daily living. *Geriatric syndromes.



Geriatric syndromes were more often extracted from Anglo-European compared to US studies (USA: 163/413, 39.5%; Anglo-European: 174/286, 60.8%). Among these, mobility (38/286, 13%), demographic data (34/286, 12%), along with cognitive impairment (32/286, 11%), and ADL (31/286, 11%) emerged as the predominant predictor domains (Figure 3C).

In total, 93 combinations of domain sets could be identified in the literature in our review. Contrasting these sets with all potential predictor combinations of the 20 domains, the greatest concordance with the highest average Jaccard Index score of 0.28 was found for the 8 elements shown in the upper part of Table 3 as a potentially generalizable domain set. As many literature-derived sets comprise solely 1 element, based on an 8-element set, the best Jaccard Index of 2 one-element sets could be 0.125. Thus, achieving a higher index value of 0.28 with an 8-element set across all literature represents a favorable outcome.

XSL•FO
RenderX

Table . Predictor domains most concordant with the literature and potentially generalizable. Upper part: Domain set with the highest mean Jaccard index in comparison to all sets found in the literature, in alphabetical order, including the top 3 predictors of each domain. Lower part: Other domains recognized in the review, in alphabetical order, including the top 3 predictors of each domain.

Domain	Predictor examples
Domain set most concordant with the literature	
ADL[a]	ADL NSS[b], Barthel index [27], Katz score [25]
Cognitive impairment	Dementia diagnosis, confusion assessment method [33], short portable mental status questionnaire [34]
Demographic data	Age, sex (female), ethnicity
Hospitalization data	Length of stay, ICU[c] treatment, source of admission
Primary diagnosis	Wound problems, cardiopulmonary disease, infectious disease
Mobility	Falls, De Morton Mobility Index [30], pre-fracture mobility score [35]
Multimorbidity	American Society of Anesthesiologists score [24], number of comorbidities, Charlson comorbidity index [26]
Social support	Living alone or with companion, preadmission professional support, caregiver support
Other domains	
Administrative data	Level of care hospital, prior hospitalizations, prior ICU admission
Disease severity	Glasgow coma scale [36], injury severity score [28], injury severity NSS
IADL[d]	IADL no specific score, Lawton index [37], InterRAI Acute Care (finances) [38]
Frailty	Frailty index[e], clinical frailty scale [39], Fried frailty scale [40]
Laboratory results	Anemia, fluid and electrolyte disorder, hypoalbuminemia
Medication	Number of medication, vitamin K antagonist therapy, therapeutic anticoagulation
Mental health	Geriatric depression scale [41], depression diagnosis, 8-item Patient Health Questionnaire [42]
Multidimensional scores	Hospital Admission Risk Profile [43], ISAR[f] (≥2) [44] and CGA[g] normal, multidimensional prognostic index (every 0.1 increase) [45]
Nutrition and sarcopenia	BMI [46], Mini Nutrition Assessment short form [47], sarcopenia diagnosis
Procedures	Orthopedic or trauma surgery, anesthesia or ICU treatment procedure, gastrointestinal surgery
Self-rated health	Older people's quality of life questionnaire [32], 3-item brief health literacy screen, Short Form-12 physical component summary [48]
Vital signs	Vital capacity, respiratory rate, systolic blood pressure

[a]ADL: activities of daily living.

[b]NSS: no specific score.

[c]ICU: intensive care unit.

[d]IADL: instrumental activities of daily living.

[e]Frailty index: different assessments possible.

[f]ISAR: identification of seniors at risk.

[g]CGA: comprehensive geriatric assessment.

Qualitative Studies

A total of 10 studies were found that mostly utilized semistructured interviews complemented by observational techniques and reviews of patients' clinical records. Participants encompassed patients (n=8, 80%), health care professionals (n=8, 80%), and informal caregivers or relatives (n=2, 20%).

All study characteristics can be found in Multimedia Appendix 2.

A total of 98 predictors supporting an appropriate discharge decision were extracted. Predictors spanned various discharge destinations, namely discharge in general (n=36, 37%), return to home (n=25, 26%), transfer to rehabilitation (n=15, 15%), SNF (n=14, 14%), nursing home (n=5, 5%), and home with support (n=3, 3%).

XSL•FO
RenderX

These 98 predictors were categorized into 10 domains aligning partially with those identified in the quantitative studies: patient or caregiver involvement (n=27, 28%), organizational structures (n=16, 16%), health status or morbidity (n=16, 16%), communication among health care professionals or providers (n=9, 9%), social support (n=7, 7%), staff education (n=7, 7%), regional aspects (n=6, 6%), administrative health care data (n=4, 4%), hospitalization data (n=4, 4%), and home assessment (n=2, 2%).

Within the domain of patient or caregiver involvement, effective strategies emphasize communication and respect. Notable recurring themes include discussing discharge plans with patients and caregivers in a timely manner, addressing their capability to self-manage post-discharge, providing detailed medication schedules, and honoring patient preferences. As highlighted in 1 study, "Patients reported receiving contradictory information, especially with respect to medications and how to manage their care following discharge" [49], whereas another underscored that "...elders who thought they received enough information about how to manage their care after they left the hospital reported feeling satisfied" [50].

Regarding organizational structures, clear responsibilities, including staff continuity, geriatric comanagement, and standardized discharge procedures, emerged as significant components facilitating discharge decisions. A recent study emphasized the "...lack of recognized decision-making tools or algorithms as a critical issue in practice" [51].

The domain health status or morbidity includes parameters akin to those observed in quantitative studies, including ADL and cognitive impairment due to dementia or delirium. In 1 study, it was cited that "... clinicians noted the clinical challenges of managing delirium and the need for development of patient pathways for those with delirium" [52].

Discussion

Principal Findings

This scoping review identified and synthesized predictors of appropriate PAC destinations for older adults (≥65 y), transitioning from acute care to specific discharge destinations. Based on the analysis of 181 studies (171 quantitative, 10 qualitative), key predictors varied by discharge destination. Diagnosed dementia and deficits in instrumental activities of daily living were frequently associated with discharge to nursing homes, while longer length of stay and cardiopulmonary disease predicted discharge to rehabilitation, and good performance in activities of daily living and mobility assessments favored discharge home without support. Furthermore, the review highlighted the influence of age and geographical origin, with geriatric syndromes being more prominent in those aged 80 years and older and in non-US studies. The 8 predictor domains—cognitive impairment, activities of daily living, demographics, social support, hospitalization data, multimorbidity, mobility, and primary diagnosis—demonstrated the highest concordance with existing literature. These findings provide a comprehensive overview to inform clinical decision-making and the development of clinical decision support systems, including machine learning apps, aimed at optimizing PAC planning for older adults.

Comparison to Prior Work

Age emerged as the most frequent predictor over nearly all subgroups. Older adults ≥80 years tend to be more often discharged to nursing homes rather than rehabilitation facilities [53]. This could be due to the reduced availability and accessibility of facilities as well as misconceptions regarding their potential for improvement through these services. Investigations involving populations with a mean age of ≥80 years demonstrate an increased use of geriatric syndromes as predictors for discharge destinations, identifiable by a comprehensive geriatric assessment (CGA) [54]. Notably, frailty appears more frequently as a predictive factor in studies encompassing individuals with a mean age of <80 years. This could be because frailty was not assessed in these studies or because frailty has been mediated by age or functional parameters such as the BI [55,56]. Qualitative analysis revealed that patient or caregiver involvement was another important predictor for appropriate discharge destination, although hardly used in quantitative studies.

To address the observed heterogeneity in predictors, our review presents a higher-level categorization based on predictor domains. This provides a more standardized and manageable framework for analysis. A CGA is another established method for capturing the heterogeneity of potential limitations across geriatric domains, thereby improving the likelihood of living at home after 3 months [57]. Comparing the predictor domains most concordant with existing literature with a CGA, it stands out that 5 of the 8 predictor domains (mobility, ADL, multimorbidity, social support, and cognitive impairment) are also core elements of a CGA [54]. Therefore, a CGA supplemented by routine measurements such as current hospitalization, primary diagnosis, and demographic data forms a solid foundation for enhancing prediction accuracy of the most appropriate discharge setting and is probably one reason why it ultimately reduces discharge to higher levels of care [58].

Discharge destination choices are inherently contingent upon local health care infrastructure and accessible resources. Therefore, our analysis distinguishes between Anglo-European studies, predominantly characterized by publicly funded health care systems, and US studies with a higher proportion of pay-per-service options. Anglo-European research tends to prioritize geriatric syndromes and length of hospital stay as key predictors, whereas US studies place greater emphasis on diagnostic factors. Variations in health care systems result in disparate PAC options whose definitions diverge across settings [59,60]. For example, we found many predictors for discharge to SNF, uniquely offered in the United States, combining skilled nursing care with rehabilitative interventions, even though similar initiatives are emerging in European countries such as Germany [3,61]. Despite regional differences, parallel options exist for clinically stable yet functionally declined older patients requiring both nursing care and rehabilitation within other health care frameworks, illustrated by examples like the Australian respite residential aged care or early inpatient geriatric rehabilitation in Germany [6,62]. Regrettably, we did not find

XSL•FO
RenderX

any comparative analyses examining various PAC alternatives across countries.

A lack of standardized decision support tools assisting health care professionals when deciding on the best discharge option is one of the issues that were raised in the qualitative studies [51,63,64]. This was also supported by a recent scoping review by Singh et al [65] focusing on digital health solutions facilitating transitions in care. In a systematic review on 35 CDSSs for PAC referral that included mainly studies of non-geriatric adults, Kennedy et al [13] revealed that merely 14% have been integrated into regular clinical practice, potentially hindered by constraints in time resources. Conversely, positive outcomes have emerged from specific implementations, such as a pre- or post-implementation study of a CDSS with a 2-step approach that showed a significant reduction of readmission rates [66]. Also, other binomial prediction models for discharge destinations, for example, for routine versus nonroutine discharge as in Karhade et al [67], showed promising results with an area under the curve of 0.823 in a machine learning model. However, when comparing these findings to one of the few CDSS developed for multinomial differentiation across 6 discharge destinations involving over 14,000 participants, it achieved a lower overall area under the curve of 0.685 [68]. The findings presented in this review will help to develop future multinomial COC or discharge prediction models by identifying eligible features.

Strengths and Limitations

A key strength of our scoping review lies in the detailed differentiation of discrete prediction parameters, which enabled us to identify and statistically analyze a robust set of core predictor domains. However, our review is not without limitations, which can be categorized as pertaining to the included studies themselves and our review methodology.

We acknowledge the following limitations. The first limitation was the predominance of studies conducted in the US health care system with its specific PAC options and at least partially unmet health care needs due to some patients' lack of insurance coverage [69,70]. To address this limitation, we compared US studies with studies from Europe, Canada, and Australia with more similarly conceptualized health care systems, even though we are aware that this simplifies the complex differences between countries in the Anglo-European group that also exist.

The second limitation was the frequent use of nonstandardized assessments that we addressed by summarizing these measures into a "no specific score" predictor. For example, the predictors "problems with personal hygiene" or "ADL no specific score" are included in both the Katz Score and BI that again report comparable information. This indicates a need for a wider use of validated assessment instruments to capture the functional status in older patients in a uniform way. When selecting assessments for the future development of a CDSS tool or compiling assessments for a CGA, these validated and established assessments should be used.

Third, many of the included studies presented composite end points of different discharge destinations with unknown proportions among the different destinations as well as different populations concerning investigated diseases and clinical settings, which limits the validity of the extracted predictors for generalizable discharge destination options. Wherever possible, we tried to extract the singular discharge destination.

Fourth, focusing on the older geriatric population might have led to an omission of other studies mainly including nongeriatric adults that might also have included valuable information on the topic. For example, a review by Kennedy et al [13] found 33 studies presenting CDSS for optimizing discharge destinations, but only 6 studies were conducted in older age populations.

Fifth, our focus on positive predictors to avoid often confusing directions of effect among the negative predictors may have resulted in an undercapture of relevant predictors. With the negative predictors representing only 15.5% of the predictors, we consider the bias of this approach to be low.

Sixth, there is a lack of further details on the extracted predictors, for example, concerning the timepoint of the assessment or the direction of effect of the predictor. These parameters have been found relevant in other studies [66]. We partially addressed this lack of detail by manually extracting the direction of effect of the most frequent predictors stratified according to different discharge destinations. The aim of a scoping review is to cover a broader topic. This approach, however, results in a large heterogeneity among the included studies and a lack of details. Thus, the analysis provided here remains on a descriptive level and offers the opportunity for future research.

Seventh, only 10 studies were included in the qualitative analysis, with most studies from the 1990s and early 2000s, limiting the informative value of the qualitative evidence synthesis.

Eighth, the broad scope of studies screened and the comprehensive analysis undertaken resulted in a literature search limited to publications up to 2022. Nevertheless, we are confident that our review of this relevant, yet relatively stable, topic remains current.

Conclusion and Implications

This scoping review synthesized evidence on predictors of appropriate discharge destinations for older adults, highlighting the heterogeneity of factors influencing care transitions. The identified 8-domain set—including demographic data, activities of daily living, social support, hospitalization data, primary diagnosis, cognitive impairment, mobility, and multimorbidity—underscores the value of comprehensive geriatric assessment in guiding discharge planning. This work will directly inform the feature selection process for a machine learning algorithm within the SURGE-Ahead project, designed to improve discharge recommendations. Recognizing the limitations of existing data and the need for system-specific adaptation, we advocate for continued research and the implementation of evidence-based discharge planning strategies.

XSL•FO
RenderX

Acknowledgments

We thank Prof Dr Karen Andersen-Ranberg for critical discussion of preliminary data.

The SURGE-Ahead Study Group comprises Karen Andersen-Ranberg, Sophia Andres, Sibylle Beck, Christian Bolenz, Simone Brefka, Raffael Cintean, Karen Clauss, Dhayana Dallmeier, Genia Decker, Miriam Deniz, Michael Denkinger, Vanessa Disque, Anna Lena Flagmeier, Marina Liselotte Fotteler, Florian Gebhard, Julia Geiger, Janina Hahn, David Hansen, Thomas Hoffmann, Wolfgang Janni, Hans Kestler, Reinhold Kilian, Jochen Klenk, Thomas Derya Kocar, Aleyna Koese, Johann Kraus, Christoph Leinert, Elena Leinert, Fabia Mangold, Judith Mantz, Annabel Müller-Stierlin, Gabi Müller, Nadir Nasir, Graziano Onder, Desmond O'Neill, Marcin Orzechowski, Carlos Pankratz, Nina Parchmann, Nuh Rahbari, Johannes Schobel, Fabienne Schochter, Konrad Schütze, Tobias Skuban-Eiseler, Vytautas Stasiunaitis, Florian Steger, Walter Swoboda, Adriane Uihlein, Sebastian Voigt-Radloff, Martin Wehling, Pascal Wendel, Felix Wezel, Jana Willems, Jessica Wolf, Philip Wolf, Friedemann Zengerling, and Marco Zeptner.

Generative artificial intelligence was used to improve the quality of writing. It assisted with proofreading and enhancing proficiency in professional English. The model used is available at [71]. The models were run locally on an RTX A6000 GPU.

Funding

This work was funded by the Federal Ministry of Education and Research (BMBF) and is part of the SURGE-Ahead—Supporting Surgery with Geriatric Co-Management and AI project (funding number: 01GY2101). The funder had no role in the study design, data extraction and synthesis, decision to publish, or preparation of the manuscript. CL, MLF, and TDK are employed in the SURGE-Ahead project.

Authors' Contributions

Conceptualization: CL (lead), DD, KG, MD, MLF (equal), TDK (supporting)
Data curation: CL (lead), JW (supporting), KG, LB, MLF (supporting)
Formal analysis: CL (lead), MLF, TDK (supporting)
Funding acquisition: MD
Investigation: CL (lead), JW (supporting), KG, LB, MLF (equal)
Methodology: CL (lead), DD, KG, MD, MLF (equal), TDK (supporting)
Project administration: CL, MD (supporting), MLF (lead), TDK
Resources: CL (lead), MLF (supporting)
Supervision: DD (supporting), MD (lead)
Validation: MD (lead), TDK (supporting)
Visualization: CL (lead), MLF, TDK (supporting)
Writing – original draft: CL (lead), MLF (supporting)
Writing – review & editing: DD, JW, KG, LB (supporting), MD (lead), TDK (equal)
Final approval of the manuscript: all authors.
All the authors fulfill the International Committee of Medical Journal Editors criteria for authorship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings.
[PDF File, 217 KB - i-jmr_v15i1e76582_app1.pdf]

Multimedia Appendix 2
Overview of included studies and reference list. (A) Overview of included quantitative studies ordered by year of publication. P: prospective observational study (eg, cohort study); R: retrospective observational study (eg, population-based study); ES: expert survey. (B) Overview of included qualitative studies ordered by year of publication.
[PDF File, 628 KB - i-jmr_v15i1e76582_app2.pdf]

Multimedia Appendix 3
Predictors positive and negative stratified according to age, origin, study design, and discharge destinations.
[XLSX File, 663 KB - i-jmr_v15i1e76582_app3.xlsx]

Multimedia Appendix 4

XSL•FO
RenderX

Predictors positive and negative, domain, discharge destinations, predictive strength indicators, and corresponding literature references.
[XLSX File, 88 KB - i-jmr_v15i1e76582_app4.xlsx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 122 KB - i-jmr_v15i1e76582_app5.pdf]

References

1. Wang YC, Chou MY, Liang CK, Peng LN, Chen LK, Loh CH. Post-acute care as a key component in a healthcare system for older adults. Ann Geriatr Med Res 2019 Jun;23(2):54-62. [doi: 10.4235/agmr.19.0009] [Medline: 32743289]
2. Gonçalves-Bradley DC, Lannin NA, Clemson L, Cameron ID, Shepperd S. Discharge planning from hospital. Cochrane Database Syst Rev 2022 Feb 24;2(2):CD000313. [doi: 10.1002/14651858.CD000313.pub6] [Medline: 35199849]
3. Heiks C, Sabine N. Long term care and skilled nursing facilities. Dela J Public Health 2022 Dec;8(5):144-149. [doi: 10.32481/djph.2022.12.032] [Medline: 36751604]
4. Bachmann S, Finger C, Huss A, Egger M, Stuck AE, Clough-Gorr KM. Inpatient rehabilitation specifically designed for geriatric patients: systematic review and meta-analysis of randomised controlled trials. BMJ 2010 Apr 20;340(apr20 2):c1718. [doi: 10.1136/bmj.c1718] [Medline: 20406866]
5. Zimmerman S, Sloane PD. Definition and classification of assisted living. Gerontologist 2007;47(suppl 1):33-39. [doi: 10.1093/geront/47.supplement_1.33] [Medline: 18162567]
6. Kwetkat A, Lehmann T, Wittrich A. Early geriatric rehabilitation: an opportunity for the oldest old. Z Gerontol Geriatr 2014 Jul;47(5):372-378. [doi: 10.1007/s00391-014-0660-7] [Medline: 24906436]
7. Continuity and coordination of care: a practice brief to support implementation of the WHO framework on integrated people-centred health services. World Health Organization. 2018. URL: https://iris.who.int/handle/10665/274628 [accessed 2025-12-23]
8. Bahr SJ, Weiss ME. Clarifying model for continuity of care: a concept analysis. Int J Nurs Pract 2019 Apr;25(2):e12704. [doi: 10.1111/ijn.12704] [Medline: 30393894]
9. Uijen AA, Schers HJ, Schellevis FG, van den Bosch W. How unique is continuity of care? A review of continuity and related concepts. Fam Pract 2012 Jun;29(3):264-271. [doi: 10.1093/fampra/cmr104] [Medline: 22045931]
10. Hirschman KB, Shaid E, McCauley K, Pauly MV, Naylor MD. Continuity of care: the transitional care model. Online J Issues Nurs 2015 Sep 30;20(3):1. [Medline: 26882510]
11. Kind AJH, Smith MA, Frytak JR, Finch MD. Bouncing back: patterns and predictors of complicated transitions 30 days after hospitalization for acute ischemic stroke. J American Geriatrics Society 2007 Mar;55(3):365-373. [doi: 10.1111/j.1532-5415.2007.01091.x]
12. Eichler S, Völler H, Reibis R, et al. Geriatric or cardiac rehabilitation? Predictors of treatment pathways in advanced age patients after transcatheter aortic valve implantation. BMC Cardiovasc Disord 2020 Apr 6;20(1):158. [doi: 10.1186/s12872-020-01452-x] [Medline: 32252646]
13. Kennedy EE, Bowles KH, Aryal S. Systematic review of prediction models for postacute care destination decision-making. J Am Med Inform Assoc 2022 Dec 28;29(1):176-186. [doi: 10.1093/jamia/ocab197] [Medline: 34757383]
14. Leinert C, Fotteler M, Kocar TD, et al. Supporting SURgery with GEriatric Co-management and AI (SURGE-Ahead): a study protocol for the development of a digital geriatrician. PLoS One 2023;18(6):e0287230. [doi: 10.1371/journal.pone.0287230] [Medline: 37327245]
15. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
16. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
17. Leinert C, Fotteler M, Kocar TD, Dallmeier D, Denkinger M. Predictors for continuity of care decisions for older inpatients in acute care settings: a scoping review. OSF. 2022 Mar 22. URL: https://osf.io/6qsvb/ [accessed 2025-12-15]
18. Hirshon JM, Risko N, Calvello EJB, et al. Health systems and services: the role of acute care. Bull World Health Organ 2013 May 1;91(5):386-388. [doi: 10.2471/BLT.12.112664] [Medline: 23678202]
19. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8(1):45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
20. Palm W, Webb E, Hernández-Quevedo C, et al. Gaps in coverage and access in the European Union. Health Policy 2021 Mar;125(3):341-350. [doi: 10.1016/j.healthpol.2020.12.011] [Medline: 33431257]
21. Ridic G, Gleason S, Ridic O. Comparisons of health care systems in the United States, Germany and Canada. Mater Socio Med 2012;24(2):112-120. [doi: 10.5455/msm.2012.24.112-120] [Medline: 23678317]
22. Dixit SK, Sambasivan M. A review of the Australian healthcare system: a policy perspective. SAGE Open Med 2018;6:2050312118769211. [doi: 10.1177/2050312118769211] [Medline: 29686869]

XSL•FO
RenderX

23. Ahmadinejad N, Chung Y, Liu L. J-score: a robust measure of clustering accuracy. PeerJ Comput Sci 2023;9:e1545. [doi: 10.7717/peerj-cs.1545] [Medline: 37705621]
24. Knuf KM, Maani CV, Cummings AK. Clinical agreement in the American Society of Anesthesiologists physical status classification. Perioper Med (Lond) 2018;7:14. [doi: 10.1186/s13741-018-0094-7] [Medline: 29946447]
25. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist 1970;10(1):20-30. [doi: 10.1093/geront/10.1_part_1.20] [Medline: 5420677]
26. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
27. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. Md State Med J 1965;14:61-65. [Medline: 14258950]
28. Baker SP, O'Neill B, Haddon W Jr, Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma 1974 Mar;14(3):187-196. [Medline: 4814394]
29. Foss NB, Kristensen MT, Kehlet H. Prediction of postoperative morbidity, mortality and rehabilitation in hip fracture patients: the cumulated ambulation score. Clin Rehabil 2006 Aug;20(8):701-708. [doi: 10.1191/0269215506cre987oa] [Medline: 16944827]
30. de Morton NA, Davidson M, Keating JL. The de Morton Mobility Index (DEMMI): an essential health index for an ageing world. Health Qual Life Outcomes 2008 Aug 19;6(1):63. [doi: 10.1186/1477-7525-6-63] [Medline: 18713451]
31. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
32. Bowling A. The psychometric properties of the older people's quality of life questionnaire, compared with the CASP-19 and the WHOQOL-OLD. Curr Gerontol Geriatr Res 2009;2009:298950. [doi: 10.1155/2009/298950] [Medline: 20168974]
33. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med 1990 Dec 15;113(12):941-948. [doi: 10.7326/0003-4819-113-12-941] [Medline: 2240918]
34. Erkinjuntti T, Sulkava R, Wikström J, Autio L. Short Portable Mental Status Questionnaire as a screening test for dementia and delirium among the elderly. J Am Geriatr Soc 1987 May;35(5):412-416. [doi: 10.1111/j.1532-5415.1987.tb04662.x] [Medline: 3571790]
35. Voeten SC, Nijmeijer WS, Vermeer M, Schipper IB, Hegeman JH, DHFA Taskforce study group. Validation of the fracture mobility score against the parker mobility score in hip fracture patients. Injury 2020 Feb;51(2):395-399. [doi: 10.1016/j.injury.2019.10.035] [Medline: 31668574]
36. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet 1974 Jul 13;2(7872):81-84. [doi: 10.1016/s0140-6736(74)91639-0] [Medline: 4136544]
37. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
38. Gray LC, Beattie E, Boscart VM, et al. Development and testing of the interRAI acute care: a standardized assessment administered by nurses for patients admitted to acute care. Health Serv Insights 2018;11:1178632918818836. [doi: 10.1177/1178632918818836] [Medline: 30618486]
39. Cheung A, Haas B, Ringer TJ, McFarlan A, Wong CL. Canadian study of health and aging clinical frailty scale: does it predict adverse outcomes among geriatric trauma patients? J Am Coll Surg 2017 Nov;225(5):658-665. [doi: 10.1016/j.jamcollsurg.2017.08.008] [Medline: 28888692]
40. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
41. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
42. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
43. Sager MA, Rudberg MA, Jalaluddin M, et al. Hospital Admission Risk Profile (HARP): identifying older patients at risk for functional decline following acute medical illness and hospitalization. J Am Geriatr Soc 1996 Mar;44(3):251-257. [doi: 10.1111/j.1532-5415.1996.tb00910.x]
44. McCusker J, Bellavance F, Cardin S, Trépanier S, Verdon J, Ardman O. Detection of older people at increased risk of adverse health outcomes after an emergency visit: the ISAR screening tool. J Am Geriatr Soc 1999 Oct;47(10):1229-1237. [doi: 10.1111/j.1532-5415.1999.tb05204.x] [Medline: 10522957]
45. Pilotto A, Ferrucci L, Franceschi M, et al. Development and validation of a multidimensional prognostic index for one-year mortality from comprehensive geriatric assessment in hospitalized older patients. Rejuvenation Res 2008 Feb;11(1):151-161. [doi: 10.1089/rej.2007.0569] [Medline: 18173367]
46. Eknoyan G. Adolphe Quetelet (1796-1874)—the average man and indices of obesity. Nephrol Dial Transplant 2008 Jan;23(1):47-51. [doi: 10.1093/ndt/gfm517] [Medline: 17890752]

XSL•FO
RenderX

47. Kaiser MJ, Bauer JM, Ramsch C, et al. Validation of the Mini Nutritional Assessment short-form (MNA-SF): a practical tool for identification of nutritional status. J Nutr Health Aging 2009 Nov;13(9):782-788. [doi: 10.1007/s12603-009-0214-7] [Medline: 19812868]
48. Ware J, Kosinski M, Keller SD. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
49. Bull MJ, Kane RL. Gaps in discharge planning. J Appl Gerontol 1996 Dec;15(4):486-500. [doi: 10.1177/073346489601500406]
50. Bull MJ. Patients' and professionals' perceptions of quality in discharge planning. J Nurs Care Qual 1994 Jan;8(2):47-61. [doi: 10.1097/00001786-199401000-00009] [Medline: 8312594]
51. Bradley G, Baker K, Bailey C. Exploring how occupational therapists and physiotherapists evaluate rehabilitation potential of older people in acute care. Br J Occup Ther 2022 Mar;85(3):199-207. [doi: 10.1177/03080226211011386] [Medline: 40337082]
52. Isbel ST, Jamieson MI. Views from health professionals on accessing rehabilitation for people with dementia following a hip fracture. Dementia (London) 2017 Nov;16(8):1020-1031. [doi: 10.1177/1471301216631141] [Medline: 26843421]
53. Marengoni A, Agüero-Torres H, Timpini A, Cossi S, Fratiglioni L. Rehabilitation and nursing home admission after hospitalization in acute geriatric patients. J Am Med Dir Assoc 2008 May;9(4):265-270. [doi: 10.1016/j.jamda.2008.01.005] [Medline: 18457803]
54. Deutsche Gesellschaft für Geriatrie eV. (DGG). S3 guideline: comprehensive geriatric assessment (CGA) in hospitalized patients, long version 2.3. AWMF Leitlinien-Register. 2024. URL: https://register.awmf.org/de/leitlinien/detail/084-003 [accessed 2024-11-26]
55. Hegendörfer E, Vaes B, Van Pottelbergh G, Matheï C, Verbakel J, Degryse JM. Predictive accuracy of frailty tools for adverse outcomes in a cohort of adults 80 years and older: a decision curve analysis. J Am Med Dir Assoc 2020 Mar;21(3):440. [doi: 10.1016/j.jamda.2019.08.029] [Medline: 31678074]
56. Gordon EH, Peel NM, Hubbard RE, Reid N. Frailty in younger adults in hospital. QJM-Int J Med 2023 Oct 23;116(10):845-849. [doi: 10.1093/qjmed/hcad173] [Medline: 37467071]
57. Ellis G, Gardner M, Tsiachristas A, et al. Comprehensive geriatric assessment for older adults admitted to hospital. Cochrane Database Syst Rev 2017 Sep 12;9(9):CD006211. [doi: 10.1002/14651858.CD006211.pub3] [Medline: 28898390]
58. Eamer G, Taheri A, Chen SS, et al. Comprehensive geriatric assessment for older people admitted to a surgical service. Cochrane Database Syst Rev 2018 Jan 31;1(1):CD012485. [doi: 10.1002/14651858.CD012485.pub2] [Medline: 29385235]
59. van Balen R, Gordon AL, Schols J, Drewes YM, Achterberg WP. What is geriatric rehabilitation and how should it be organized? A Delphi study aimed at reaching European consensus. Eur Geriatr Med 2019 Dec;10(6):977-987. [doi: 10.1007/s41999-019-00244-7] [Medline: 34652774]
60. Sezgin D, O'Caoimh R, O'Donovan MR, et al. Defining the characteristics of intermediate care models including transitional care: an international Delphi study. Aging Clin Exp Res 2020 Nov;32(11):2399-2410. [doi: 10.1007/s40520-020-01579-z] [Medline: 32430887]
61. Klauber J, Wasem J, Beivers A, Mostert C, Scheller-Kreinsen D, editors. Krankenhaus-Report 2024: Strukturreform: Springer; 2024. [doi: 10.1007/978-3-662-68792-5]
62. Mitchell R, Harvey L, Draper B, Brodaty H, Close J. Risk factors associated with residential aged care, respite and transitional aged care admission for older people following an injury-related hospitalisation. Arch Gerontol Geriatr 2017 Sep;72:59-66. [doi: 10.1016/j.archger.2017.05.012] [Medline: 28599139]
63. Burke RE, Lawrence E, Ladebue A, et al. How hospital clinicians select patients for skilled nursing facilities. J Am Geriatr Soc 2017 Nov;65(11):2466-2472. [doi: 10.1111/jgs.14954] [Medline: 28682456]
64. McWilliam CL, Sangster JF. Managing patient discharge to home: the challenges of achieving quality of care. Int J Qual Health Care 1994 Jun;6(2):147-161. [doi: 10.1093/intqhc/6.2.147] [Medline: 7953214]
65. Singh H, Tang T, Gray CS, et al. Recommendations for the design and delivery of transitions-focused digital health interventions: rapid review. JMIR Aging 2022 May 19;5(2):e35929. [doi: 10.2196/35929] [Medline: 35587874]
66. Bowles KH, Ratcliffe SJ, Holmes JH, et al. Using a decision support algorithm for referrals to post-acute care. J Am Med Dir Assoc 2019 Apr;20(4):408-413. [doi: 10.1016/j.jamda.2018.08.016] [Medline: 30414821]
67. Karhade AV, Ogink P, Thio Q, et al. Development of machine learning algorithms for prediction of discharge disposition after elective inpatient surgery for lumbar degenerative disc disorders. Neurosurg Focus 2018 Nov 1;45(5):E6. [doi: 10.3171/2018.8.FOCUS18340] [Medline: 30453463]
68. Sultana I, Erraguntla M, Kum HC, Delen D, Lawley M. Post-acute care referral in United States of America: a multiregional study of factors associated with referral destination in a cohort of patients with coronary artery bypass graft or valve replacement. BMC Med Inform Decis Mak 2019 Nov 14;19(1):223. [doi: 10.1186/s12911-019-0955-0] [Medline: 31727058]
69. Rice T, Rosenau P, Unruh LY, Barnes AJ. United States: health system review. Health Syst Transit 2020 Dec;22(4):1-441. [Medline: 33527901]
70. Kennedy EE, Davoudi A, Hwang S, et al. Identifying barriers to post-acute care referral and characterizing negative patient preferences among hospitalized older adults using natural language processing. AMIA Annu Symp Proc 2023;2022:606-615. [Medline: 37128417]

XSL•FO
RenderX

71. Saama AI Labs. Llama3-OpenBioLLM-70B. Hugging Face. URL: https://huggingface.co/aaditya/Llama3-OpenBioLLM-70B [accessed 2025-12-15]

Abbreviations

ADL: activities of daily living
BI: Barthel index
CDSS: clinical decision support systems
CGA: Comprehensive geriatric assessment
COC: continuity of care
IADL: instrumental activities of daily living
PAC: postacute care
SNF: skilled nursing facility
SURGE-Ahead: Supporting Surgery with Geriatric Co-Management and AI

Edited by N Cahill; submitted 26.Apr.2025; peer-reviewed by K Bowles, M Gagesch; revised version received 30.Sep.2025; accepted 08.Oct.2025; published 22.Jan.2026.

Please cite as:
Leinert C, Fotteler ML, Kocar TD, Wolf J, Beissel L, Grummich K, Dallmeier D, Denkinger M
Identifying Key Predictors of Appropriate Discharge Destinations for Older Inpatients in Acute Care: Scoping Review
Interact J Med Res 2026;15:e76582
URL: https://www.i-jmr.org/2026/1/e76582
doi:10.2196/76582

© Christoph Leinert, Marina Liselotte Fotteler, Thomas Derya Kocar, Jessica Wolf, Lisa Beissel, Kathrin Grummich, Dhayana Dallmeier, Michael Denkinger. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 22.Jan.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effectiveness, Facilitators, and Barriers of Digital Mental Health Services for First Nations People in Australia: Systematic Scoping Review

Siyu Zhai[1]; Andrew Goodman[1,2]; Anthony C Smith[3,4,5,6]; Sandra Diminic[1,7]; Xiaoyun Zhou[1,7]

[1]School of Public Health, The University of Queensland, 266 Herston Road, Herston, Queensland, Australia

[2]Australian eHealth Research Centre (AEHRC), Commonwealth Scientific and Industrial Research Organisation (CSIRO), Herston, Australia

[3]Centre for Online Health, The University of Queensland, Brisbane, Australia

[4]Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark

[5]Centre for Innovative Medical Technology, University of Southern Denmark, Odense, Denmark

[6]Centre for Health Services Research, The University of Queensland, Herston, Queensland, Australia

[7]Queensland Centre for Mental Health Research, Brisbane, Australia

Corresponding Author:
Siyu Zhai
School of Public Health, The University of Queensland, 266 Herston Road, Herston, Queensland, Australia

Abstract

Background: First Nations people in Australia experience inequitable mental health outcomes and service access. Digital mental health (DMH) services, which refer to offering mental health services through digital platforms, are considered potential solutions to address such mental health service inequities and improve the mental health outcomes of First Nations Australians. However, evidence on the effectiveness of DMH services for First Nations people in Australia is yet to be synthesized.

Objective: This systematic scoping review aims to fill this gap and to identify the facilitators and barriers that influence the implementation of DMH services among First Nations people in Australia.

Methods: A systematic search was conducted across 6 academic databases to search for studies related to DMH services for First Nations people in Australia. Search terms relating to First Nations people, geographic terminologies of Australia, mental health, and DMH services were used. Studies were included if they assessed the effectiveness of DMH services or the determinants of the facilitators and barriers of implementing DMH interventions among First Nations people in Australia. Data were extracted based on study design, targeted services, and research findings, and were then synthesized using a thematic analysis framework.

Results: In total, 22 studies met the inclusion criteria. DMH services were used to support and treat First Nations Australians and conduct psychological assessments in these individuals. Evidence of effectiveness was stronger for nonsevere mental health conditions. The determinants of the facilitators and barriers of the implementation of DMH services included the following: (1) organizational and administrative factors; (2) cultural appropriateness; (3) accessibility; (4) integration of DMH services in the existing health system; (5) engagement between clients and service providers; (6) coverage of different conditions and clients; (7) acceptability to DMH services; (8) digital literacy; and (9) efficiency.

Conclusions: Evidence on the use of DMH services for First Nations Australians remains heterogeneous in terms of study design and outcome measurement. DMH services appear to be most effective for managing nonsevere mental health conditions. Successful implementation requires multilevel structural support, including policy and organizational commitment, enhanced digital infrastructure, workforce training and engagement, and the design of culturally responsive DMH models to improve uptake and equitable access to mental health care among First Nations Australians.

Trial Registration: PROSPERO CRD42024612517; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024612517

(Interact J Med Res 2026;15:e80386) doi:10.2196/80386

KEYWORDS

digital mental health; Aboriginal and Torres Strait Islander people; eHealth; mHealth; mobile health; telehealth; mental health services

XSL•FO
RenderX

Introduction

Aboriginal and Torres Strait Islander people in Australia ("First Nations" people used hereafter) view mental health as a part of a holistic view of health across all life stages [1,2]. Maintaining a personal connection to traditional culture, health views, and community serves as an important protective factor for the mental health of First Nations people [1,3-6]. Therefore, reinforcing cultural strength and resilience would foster personal resilience and self-esteem, thereby protecting their individual mental health [7] and reducing the likelihood of developing mental illness [5,8,9]. Moreover, mental health services that acknowledge the right of self-determination and the need for cultural understanding of First Nations people [10], provide culturally appropriate care [11], or employ First Nations health workers [12] are also more likely to reinforce protective factors and achieve better treatment outcomes.

However, First Nations people are facing more complex mental health challenges than the general population in Australia due to external historical and social determinants [13]. Historically, colonization led to intergenerational trauma that worsened their mental health [14]. Currently, First Nations people continue to face daily discrimination [15], along with social inequity in areas such as education and employment [9]. These structural inequities increase the likelihood of mental health issues within these communities [16] and create mental health outcome gaps between First Nations people and the general population. For instance, although the national health survey showed that the prevalence of any mental illness among First Nations people dropped from 29.3% in 2014‐2015 [8] to 24% in 2018‐2019, it was still considered high [17]. Additionally, a survey conducted in regional Australia that applied structured clinical interviews in 2014‐2016 showed that the prevalence of any mental illness was 42.2% (4.2-fold higher than the prevalence in the general population) [5]. The prevalence of anxiety and mood disorders among First Nations people was estimated to be 1.6‐3.3 times the national prevalence in Australia [18]. Furthermore, 16% of deaths of First Nations people were related to suicide in 2001‐2005, which is 10% higher than the proportion in the general population [4].

Despite significant mental health needs and gaps, First Nations people experience barriers to accessing mental health services. First, structural racism, stigma toward mental health conditions, low health literacy, and less comfort in seeking help from health professionals prevent First Nations people from using more mental health services [11,12,19]. Second, a lack of local services and service providers creates barriers to travel for First Nations people [11,20], especially in areas with higher geographical remoteness that have a higher proportion of First Nations populations [21] but lower service access [22]. Moreover, community-controlled, culturally safe services may be unavailable for some First Nations people [23], and non-First Nations health workers who do not receive appropriate cultural safety training also contribute to service access barriers [20,24-26]. In 2018‐2019, only 31% of First Nations adults with high or very high levels of psychological distress received professional mental health services [27]. Moreover, mental health service gaps were mostly reported by First Nations primary health organizations (134 out of 198 organizations) [28]. Furthermore, among suicide cases, First Nations populations were approximately 2 times less likely than non-First Nations populations to have received mental health support (23.8% vs 43.3%) [29]. In the Northern Territory, mental health service utilization of First Nations youth is nearly half that of non-First Nations youth (1.9% vs 4.1%) [30].

Digital mental health (DMH) services refer to providing mental health services through digital technologies [31-33]. DMH services have been implemented for the prevention, screening, intervention, and rehabilitation of mental health conditions [34] with or without mental health professionals' guidance or support [35], and they demonstrate promising results for managing various mental health conditions, including anxiety disorders [36], depressive disorders [37], posttraumatic stress disorder [38], and alcohol and other drug-related mental disorders [39]. Moreover, DMH services have been evaluated among different populations, including children and young people [40], adults [41], and older populations [42].

DMH services have been evaluated among First Nations people worldwide [43-46] and have shown the potential to address the current health outcome gaps faced by them. Evidence from Canada, New Zealand, and the United States indicates that DMH services support improvements in psychological assessment [47-50] comparable to existing face-to-face services. DMH services were also found to improve clinical outcomes [49,51] and continuity in mental health care [52], both as stand-alone and blended care. International evidence also suggests that DMH services can reduce mental health gaps by improving the accessibility and availability of mental health services among First Nations people [23,34-36].

In Australia, evidence on DMH services among First Nations people is mixed. Although digital gaps remain a barrier to DMH implementation [53], government funding has enabled the provision of digital infrastructure in rural and remote First Nations communities [54]. Another service gap is the need for culturally appropriate services [23]. Evidence suggests that DMH services may fail to deliver culturally safe services [55,56] due to the preference of First Nations people for the familiarity and safety of face-to-face services [50]. However, some studies have indicated that digital health may support cultural safety by increasing the involvement of First Nations health workers [43,57] and community-led design services [58], and by allowing patients to see familiar faces via video conferencing [54]. DMH services also enable First Nations people with mental health needs to include family and community members in decision-making, without the need for travel [54].

Currently, most systematic reviews have focused on international evidence regarding the application of DMH services to First Nations people [43-46]. Existing Australian evidence is more likely to review broad and general health conditions rather than focus on mental health [47,54,57,59] or to solely focus on the design and assessment of DMH services for First Nations people rather than the implementation [58]. There is a paucity of evidence on the effectiveness of DMH services and the facilitators and barriers of implementing these services among First Nations people in Australia. Our study

XSL•FO
RenderX

aims to address this research gap by exploring the following research questions: (1) What is the effectiveness of DMH services for First Nations people in Australia? (2) What are the determinants of the facilitators and barriers of implementing DMH services among First Nations people in Australia? Such evidence is critical for policymakers, health providers, and First Nations users to make evidence-informed decisions while funding, implementing, prescribing, and choosing DMH services.

Methods

Reporting and Registration

The reporting of this study followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines (Checklist 1) and was registered through PROSPERO (CRD42024612517).

Acknowledgment to Indigenous Data Sovereignty

This study follows the Maiam Nayri Wingara principles [60]. The authors acknowledge that First Nations people have control of the data ecosystem; all data are available to First Nations people; and this study pays respect to First Nations communities, will be accountable to these communities, and empowers the sustainable self-determination of these communities.

Outcome Definition

Our study outcomes include the effectiveness of DMH services and the determinants of facilitators and barriers. The effectiveness of DMH services is defined as the extent to which DMH services obtain the desired mental health outcomes. For example, for DMH services that are related to psychological diagnosis and assessment [49], the accuracy of diagnosis is considered as effectiveness. For treatments provided through DMH services [51,52,61], effectiveness is defined as the degree to which a treatment service improves mental health symptoms. The determinants of facilitators and barriers are defined as the determinants that can promote and impede, respectively, the implementation of DMH services for First Nations people in Australia.

Eligibility Criteria

Based on the PICO (population, intervention, comparison, outcome) protocol, this review considered studies that (1) included participants who were First Nations Australians; (2) included interventions that were DMH services with various functions; (3) included comparisons that were typical mental health services or did not include comparisons; and (4) targeted outcomes that involved changes in mental health conditions. In addition, studies written in English were included. Studies that were irrelevant to our PICO criteria and those that did not have full text available were excluded. Moreover, there was no eligibility criterion relating to First Nations groups, date of publication, or age of participants. See Multimedia Appendix 1 for detailed inclusion and exclusion criteria and the PICO protocol.

Search Strategy

A systematic search was conducted from September 9, 2024, to September 15, 2024, across 6 online databases: PsycINFO, PubMed, MEDLINE, Embase, Web of Science, and Google Scholar. Considering that more than 10,000 search results were obtained from Google Scholar, searching on this database was limited to the first 30 pages.

The search strategy was developed after consultation among 3 authors (SZ, XZ, and SD). Boolean operators, truncations, and search terms were tailored to each database. The search terms were organized around the following themes: First Nations people, geographic terminologies of Australia, mental health (including specific mental health conditions), and DMH services. Considering that DMH is a relatively new research area, there was no limitation on the publication date of the searched studies. See Multimedia Appendix 2 for details of the search strategy.

Study Selection

Selected studies were imported into EndNote 20 and Covidence, and duplicates were removed at this stage. Afterwards, the titles and abstracts were screened. Studies found to be irrelevant to the review topic based on an assessment of the title and/or abstract were excluded. Studies that had insufficient information in the title and abstract were moved to full-text screening. During full-text screening, studies that met the exclusion criteria were excluded.

One reviewer (SZ) conducted the title and abstract screening. Uncertainty in this process was solved through discussion with another reviewer (XZ), and any disagreements were resolved by involving the third reviewer (SD). Two reviewers completed the full-text screening, where one reviewer (SZ) made initial decisions and the other reviewer (XZ) checked the findings. Therefore, the selection process was not blinded. Disagreements were resolved through discussion or by involving the third reviewer (SD).

Data Extraction

Extracted data were tabulated based on author, year of publication, study design (qualitative, quantitative, or mixed methods), study aim, study population, research setting, type of DMH intervention, outcome (effectiveness, facilitators, and barriers), and key research findings. See Multimedia Appendix 3 for details of the extracted data.

For missing data, if the authors did not specify the reason, the reviewers tried to contact the authors to obtain relevant information. If such information was unavailable or no response was received, the missing data were excluded from the analysis process.

Strategy of Data Synthesis

The synthesis of data followed the thematic analysis method introduced by Thomas and Harden [62] and was guided by the SWiM (Synthesis Without Meta-Analysis) guideline [63].

Reviewer SZ read selected studies completely and focused on the results section of those studies. Results relevant to the targeted outcomes of this review, such as the effect size of interventions or reported user experiences related to these

XSL•FO
RenderX

determinants, were summarized and coded into descriptive themes. Descriptive themes were categorized into subgroups and synthesized based on the type of data (qualitative or quantitative), research design (with or without a control group), and relevant determinants (facilitators, barriers, or effectiveness). Afterwards, reviewers standardized themes using the direction of the effect. Vote counting was conducted based on the direction of the effect. Then, the commonality among descriptive themes was found based on the direction of the effect or the different determinants they related to, which were used to formulate the analytical themes that summarized multiple similar descriptive themes. Analytical themes interpreted the primary results to fit the objectives of this review. To ensure the trustworthiness and validity of thematic analysis, during the generation of themes, reviewer SZ consulted reviewers XZ and SD to resolve any uncertainty. Once reviewer SZ completed generating themes, reviewer XZ checked all themes. Any disagreements were resolved by consulting the third analyst, SD.

Quality Assessment

This study applied the Aboriginal and Torres Strait Islander Quality Appraisal Tool [64] to assess the quality of health research from a cultural sensitivity perspective. Reviewer SZ carefully reviewed the included studies and answered the 14 questions in the Aboriginal and Torres Strait Islander Quality Appraisal Tool based on the content of all selected studies. Moreover, as the instructions for this quality assessment tool suggested, the numbers of studies assessed as "yes," "partially," "no," and "unclear" have been reported. Reviewer SZ applied the Mixed Methods Appraisal Tool (MMAT) [65] for assessing the methodological quality of the included studies. Only the relevant part of the included study was assessed. For instance, if the qualitative part of a mixed-methods study was the only part relevant to this review, the study was appraised as a qualitative study. The MMAT does not recommend scoring each study; hence, this review only reports the performance of the included studies on each criterion and the overall trend of the design of the included studies. All quality assessment results were checked by reviewers XZ and SD. See Multimedia Appendices 4 and 5 and the Results section for the quality assessment results.

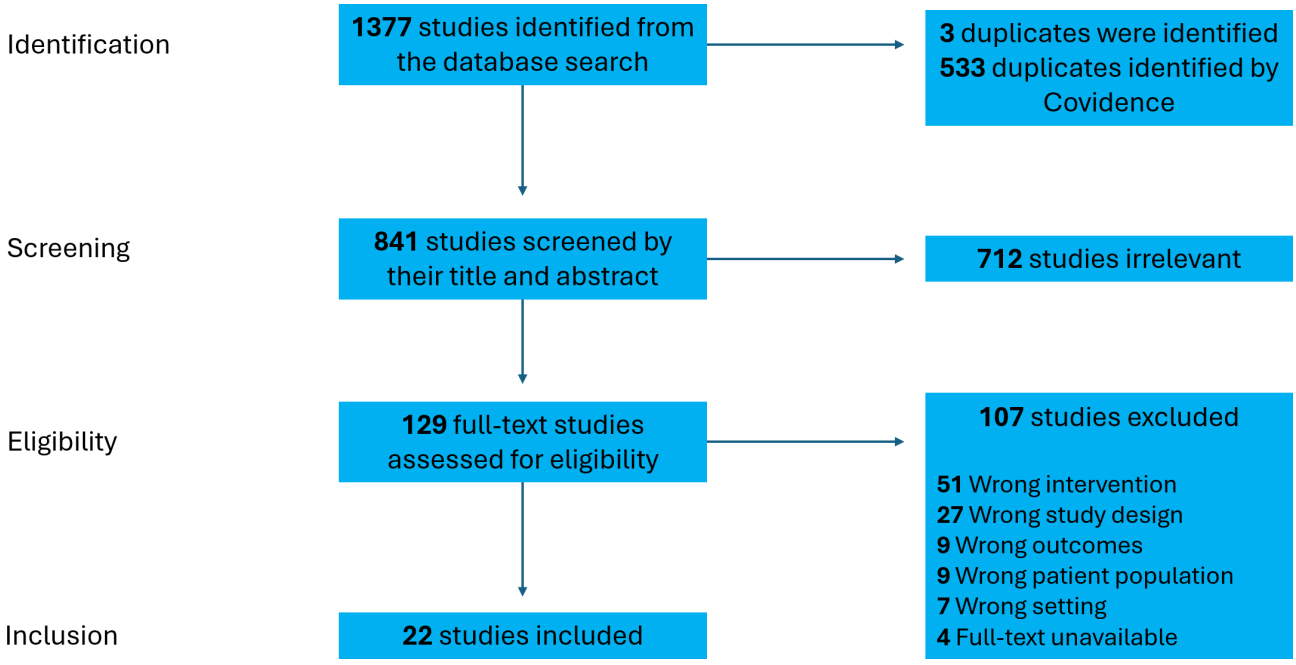
In this systematic scoping review, all studies were included regardless of quality rating. However, the quality appraisal results informed the interpretation of the searched studies. Considering this study was a narrative synthesis and did not include a meta-analysis, no statistical weighting or sensitivity analysis was performed.

Results

General Characteristics of the Studies

A total of 22 articles were included in this review [66-87]. The PRISMA flowchart is presented in Figure 1. See Multimedia Appendix 3 for detailed characteristics. Except for 1 study that solely reported barriers [73] and 4 studies that solely reported effectiveness [72,82,85,87], the rest of the studies reported more than one targeted outcome (ie, facilitators, barriers, and effectiveness). Moreover, 16 studies [67,68,71-74,76,77,79-84,86,87] were published within 5 years, and all included studies were published between 2015 and 2023.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of article screening and inclusion.



Among the included studies, there were 8 qualitative studies [66,67,69,74,75,78,79,83], 8 mixed-method studies [68,70,71,76,77,80,81,84], and 6 quantitative studies [72,73,82,85-87]. Among the quantitative studies, 3 adopted a cross-sectional research design [72,73,87], 1 was a crossover study [82], 1 was a randomized controlled trial [85], and 1 was a prospective cohort study [86].

All studies involved First Nations people as the clients of DMH services. There were 5 studies involving health care providers who provided DMH services to First Nations people

XSL•FO
RenderX

[66,69,78-80], 2 studies involving both clients and service providers [67,74], 12 studies involving adult clients or clients whose ages were unspecified [68,70-73,75,81,82,84-87], and 3 studies involving youth or student clients [76,77,83].

Regarding the types of DMH services, 5 studies were related to general perspectives on DMH services [66,77-80]. Among the studies, 2 involved DMH services delivered through web pages [70,86], 12 involved DMH services delivered through mobile apps [67-69,71,72,74-76,79,80,84,85], 5 involved DMH services delivered through telehealth methods (text or voice-only phone call) [70,73,86,87], 1 involved DMH services delivered through video conferencing [82], and 2 involved DMH services containing both online and offline activities [81,83]. See Table 1 and Multimedia Appendix 3 for the details of DMH service types.

XSL•FO
RenderX

Table . Overview of the findings in relation to the effectiveness of digital mental health services.

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
Qualitative evidence				
Dingwall et al [68], 2023	Qualitative interviews at the end of the 4-week intervention	• Deductive approach for qualitative analysis	Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app	• Participants had mental health improvements or gained help during tough times by using this app, which suggested its effectiveness.
Perdacher et al [74], 2022	Semistructured interviews	• Thematic analysis under the framework provided by the constant comparison method	Stay Strong app (custodial version)	• Both clients and practitioners reported that the app improved the abilities of self-insight and self-reflection of clients and guided the positive thinking of clients. • Both clients and practitioners reported that the app could enhance clients' confidence, view of self, and empowerment.
Povey et al [75], 2016	Three focus groups	• Thematic analysis	AIMhi Stay Strong iPad app and the iBobbly suicide prevention app	• Both apps have the potential to help people with minor mental health conditions overcome difficulties. • For the suicide prevention function, participants thought the app is useful for nonsevere conditions but may not be helpful in preventing suicide due to it being a severe condition.
Raphiphatthana et al [80], 2020	Semistructured interviews	• Thematic analysis that oscillated between deductive and inductive approaches	General digital mental health services (but many of the responses were related to the use of the Stay Strong app)	• Electronic mental health was perceived as potentially beneficial and useful for First Nations communities.
Routledge et al [81], 2022	Paper-and-pen test in classes or an online questionnaire	• Thematic analysis	Web-based lesson package, Strong & Deadly Futures	• Some facilitators reported that students learned the knowledge that the lesson planned to deliver, while students reported that the substance-related knowledge had been learned and understood.

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
Tighe et al [84], 2020	Semistructured interview	• Thematic analysis	iBobbly program	• Participants indicated that an app may not be able to provide enough support for vulnerable people who have severe mental health conditions. • The improved mental health literacy helped participants feel distracted from their thoughts and reduced their distress. • Some participants reported that the program has the effect of suicide prevention.
Quantitative evidence with comparison to typical mental health services or no intervention groups				

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
Lee et al [72], 2019	Intervention group: data from the Grog Survey app; Control group: clinical interview	• Triangulating the Finnish method to compare drinking status and average consumption between 2 groups using descriptive statistics. • Agreement between consumption data/daily alcohol consumption and withdrawal tremors was assessed through the Spearman rank correlation coefficient. • Sensitivity and specificity analyses (unspecified) were used to compare how well the app and clinical interview agreed when classifying drinkers' risk.	Grog Survey app	• The app can accurately test participants' drinking status (n=157), with 94.7% agreement with the clinical interview results. • The alcohol consumption level recorded by the app was usually higher than clinical interviews (drinks per drinking occasion: median 17.0, IQR [10.5, 27.9] vs median 15.4, IQR [9.6, 23.2]). • The app tended to categorize more participants at risk. The app identified 76.7% of drinking males and 69.7% of females as being at risk, while the interview identified 74.4% of males and 65.2% of females as being at risk. • The app appeared to be highly sensitive (92.7% - 97.5% sensitivity) in detecting at-risk drinkers, and when using matched and unmatched reference periods, the app could detect 10 and 37 more risky drinkers, respectively, than clinical interviews. • The average daily consumption recorded by the app and by the interview was moderately correlated (r=.52). Correlations were greater when only data from matched time periods were used (r=.62). Correlations were all statistically significant. • Participants were more likely to report the presence of tremors when using the app than when responding to the interview (17.0% and 7.2%, respectively; r_s=0.48). • There was high test-retest reliability for the app (r_s=0.81 for both participants and drinkers).
Veinovic et al [87], 2023				

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
	Test results of the face-to-face and telehealth versions of both the MMSE[a] and the KICA-screen[b] tests	• The MMSE and KICA-screen test version total scores were compared using Spearman correlations. • Agreement between test versions was examined using Bland-Altman analyses. • Telephone test scores were prorated to equivalent in-person total scores to examine overall performance discrepancies directly.	Telehealth versions of the MMSE (TMMSE) and KICA-screen (TKICA-screen)	• Despite a reduction in effectiveness, the accuracy of the TMMSE was generally acceptable. The median TMMSE results were slightly lower than the MMSE results (24 vs 28.5, IQR 2 for both). There was a moderate correlation and a reasonable agreement between MMSE versions (r_s=0.33; *P*=.20), although the limits of agreement were unacceptably wide (−4.1 and 4.8 point differences). • The TKICA-screen was significantly less effective than the KICA-screen. There was a significant mean difference of 2.17 (95% CI 1.39 - 2.95), with wide limits of agreement (−1.09 and 5.43 point differences) and evidence of proportional bias (B=−1.03, SE=0.32; *P*=.005), with a tendency for poorer agreement at the lower end of the performance.
Tighe et al [85], 2017	Test results from the following psychological instruments: Depressive Symptom Index-Suicidality Subscale (DSI-SS), Patient Health Questionnaire-9 (PHQ-9), Kessler Psychological Distress Scale-10 (K-10), and Barratt Impulsivity Scale-11 (BIS-11)	• Paired *t* test	iBobbly program	

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
				• The effectiveness of the iBobbly app on suicide prevention was insignificant. The pre- and postintervention changes in the DSI-SS score were significant in the iBobbly arm (t=2.40, df=58.1; P=.02), but these differences were not significant compared with the waitlist arm (t=1.05, df=57.8; P=.30). • The iBobbly app was effective in reducing depression. The iBobbly arm showed a statistically significant decrease in PHQ-9 scores compared with waitlist controls. The interaction of the iBobbly arm by time was significant (t=2.79, df=56.9; P=.007; Cohen d=0.71, 95% CI 0.17-1.23), reflecting a substantial effect. • The iBobbly app was effective in reducing general psychological distress measured by the K-10. The interaction of the intervention arm by time was significant (t=2.44, df=57.5; P=.02; Cohen d=0.65, 95% CI 0.12-1.17), reflecting a substantial effect. • However, the PHQ-9 and K-10 scores for the control group (comparing baseline and receiving the iBobbly intervention afterwards) were insignificant. • The iBobbly app was not effective in reducing impulsivity, as measured by the BIS-11. The preintervention scores for the waitlist group were significantly lower than those for the iBobbly group (t=2.05, df=59.2; P=.045). Postintervention means were identical. However, after applying the iBobbly app to the waitlist group, the score reduction was insignificant (t=−1.82, df=29.1; P=.08).

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
Russell et al [82], 2021	KICA-screen results obtained from both face-to-face and online tests	• Liner correlation	KICA-screen delivered via telehealth	• The KICA-screen test delivered through digital platforms did not cause a decrease in effectiveness compared to the face-to-face version, owing to good correlation (Pearson r=0.851; P<.01) and good agreement (intraclass correlation coefficient=0.85; P<.01).
Titov et al [86], 2019	Health records measured by the K-10, PHQ-9, and Generalized Anxiety Disorder Scale-7 (GAD-7)	• A generalized estimation equation (GEE) modeling technique was used to examine the significance of changes in symptom measures over time. The main effects of course and Indigenous status were also examined.	MindSpot service	• For First Nations patients, the MindSpot service was effective at causing a significant decrease in the K-10 score (χ^2=5933.1; P<.001), PHQ-9 score (χ^2=5938.8; P<.001), and GAD-7 score (χ^2=5658.8; P<.001). • There were no differences in treatment outcomes between Indigenous and non-Indigenous patients (K-10: χ^2=0.1; P=.75; PHQ-9: χ^2=0.5; P=.49; GAD-7: χ^2=0.6; P=.44). • There was no difference in the outcome between Indigenous patients who chose the general wellbeing course and those who chose the Indigenous wellbeing course (K-10: χ^2=1.2; P=.27; PHQ-9: χ^2=0.5; P=.48; GAD-7: χ^2=2.9; P=.09).
Quantitative evidence with no comparison				

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
Dingwall et al [68], 2023	Comparing participants' scores of psychosocial instruments at baseline and after the 4-week intervention: K-10, Patient Health Questionnaire-2 (PHQ-2), Generalized Anxiety Disorder-short form (GAD-2), Alcohol Use Disorders Identification Test (AUDIT-C), Drug Use Disorders Test (DUDIT), and Parent-rated Strengths and Difficulties Questionnaire (SDQ-parent rated)	• Paired sample 2-tailed *t* tests for psychological test results with within-group effect sizes (Cohen *d*)	AIMhi-Y app	• The app was significantly effective in reducing psychological distress (t_{29}=4.63; *P*<.001; Cohen *d*=0.85, 95% CI 2.79-7.21) and depressive symptoms (t_{29}=4.09; *P*<.001; Cohen *d*=0.75, 95% CI 0.42-1.25). • Participants presented improvements in reducing anxiety symptoms (t_{29}=1.58; *P*=.13; Cohen *d*=0.28, 95% CI −0.10 to 0.77). • The app was ineffective in reducing alcohol use disorder symptoms (t_{26}=1.55; *P*=.13; Cohen *d*=0.31, 95% CI −0.06 to 0.43), drug use disorder symptoms (t_{26}=1.55; *P*=.13; Cohen *d*=0.31, 95% CI −0.06 to 0.43), and lead dependence (t_{29}=0; *P*>.99; Cohen *d*=0, 95% CI −0.29 to 0.29). • Participants' parents reported that the app was ineffective in improving psychological adjustment (t_3=2.15; *P*=.12; Cohen *d*=1.07, 95% CI −2.03 to 10.53).
Kennedy et al [71], 2021	The uMARS[c] survey comprises 26 questions and 5-point Likert scales	• Calculate descriptive statistics and mean and SD for continuous variables	Multibehavioral change app "MAMA-EMPOWER"	• Participants (n=16) reported acceptable perceived impact (mean 3.55, SD 1.06).
Tighe et al [84], 2020	Record usage of the app and results from psychological instruments: DSI-SS, K-10, and PHQ-9	• Regression analyses	iBobbly program	

XSL•FO
RenderX

Evidence type and study (author, year)	Data collection method	Data analysis method	Digital mental health service type	Major findings
				• The iBobbly app was ineffective for the primary outcome, suicidal ideation, which was measured with the DSI-SS. For every minute spent on the app, DSI-SS scores reduced by 0.013 points (R^2=0.29). • The iBobbly app was ineffective in reducing psychological distress that was measured with the K-10. For every minute spent on the app, K-10 scores reduced by 0.007 points (R^2=0.35). • The iBobbly app was ineffective in reducing depression that was measured using the PHQ-9. For every minute spent on the app, PHQ-9 scores reduced by 0.001 points (R^2=0.268). • According to the authors, the low validity could be caused by the small sample size.
Routledge et al [81], 2022	Paper-and-pen test on classes or an online questionnaire	• Analysis of descriptive data (method unspecified)	Web-based lesson package, Strong & Deadly Futures	• Among First Nations students (n=15), more than half (53.3%) reported that they were likely to use the information and skills taught by the course in their own lives. Nearly half (46.7%) of First Nations students were unsure, and 0% of First Nations students believed they were unlikely to use the information and skills taught by the course. • First Nations students generally thought this program was effective in handling problems related to substance use (73.3% of students supported this idea), stress (66.7%), and peer pressure (80%).

[a]MMSE: Mini-Mental State Examination.

[b]KICA-screen: Kimberly Indigenous Cognitive Examination-short form.

[c]uMARS: user version of the Mobile Application Rating Scale.

XSL•FO
RenderX

Targeted Mental Health Conditions

Five studies did not specify the mental health conditions targeted by DMH services [66,67,70,77-80]. Three studies focused on psychological assessment and diagnosis [82,86,87], and 5 studies addressed the management of distress, depression, and anxiety disorders [71,75-85,86]. Additionally, 5 studies examined the management of substance use [71,72,75,76,83], and another 5 studies were related to the management of suicide or self-harm behaviors [67,73,75,84,85]. See Table 2 for detailed information about the settings of the included studies.

XSL•FO
RenderX

Table . Settings of the included studies.

Study (author, year)	Study aim	Participants	Research setting	Digital mental health (DMH) service type	Targeted mental health issue
Bennett-Levy et al [66], 2017	Report the barriers and enablers of e-mental health uptake among health providers who provide services to First Nations people	50 providers of different mental health services for First Nations people	Two locations (Lismore and Tweed Heads) in northern New South Wales	General e-mental health services (did not specify service type)	Unspecified
Brown et al [67], 2020	Understand how mobile apps support suicide prevention gatekeepers in First Nations communities	12 participants (Indigenous health workers or community members)	A regional city in South West Queensland called Toowoomba	Suicide prevention efforts supported by mobile apps	Suicide and self-harm behaviors
Dingwall et al [68], 2023	Assess the feasibility, acceptability, and use of the Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app and determine the feasibility of study procedures for future assessments	30 First Nations young people aged 12 - 25 years	Darwin, Northern Territory (NT)	AIMhi-Y app	Unspecified
Dingwall et al [69], 2015	Develop and determine the acceptability, feasibility, and appropriateness of the Australian Integrated Mental Health Initiative (AIMhi) Stay Strong App for service providers working with First Nations people	15 service providers, including health professionals, managers, program coordinators, and an Aboriginal elder	Rural and remote primary health care service settings in the NT	AIMhi Stay Strong app	General mental health area (unspecified)
Fletcher et al [70], 2017	Test the acceptability and feasibility of developing the “Stayin’ on Track” website, which offers tailored support and information to young Aboriginal fathers, and adapt and test the mobile phone–based text-messaging and mood-tracker program (SMS4dads program) that provides mental health support to young Aboriginal fathers	20 young Aboriginal fathers in the assessment of the “SMS4dads” smartphone program; 20 young Aboriginal fathers and an uncertain number of community members in the yarn-up session on the “Stayin’ on Track” website	One regional city and two rural towns in New South Wales	“SMS4dads” smartphone program and “Stayin’ on Track” website	General mental health area (unspecified)
Kennedy et al [71], 2021	Describe the development (stages 1 and 2) and pretest (stage 3) of a prototype multibehavioral change app (MAMA-EMPOWER)	Stage 1: 8 First Nations women; stage 2: 6 First Nations women; stage 3: 16 First Nations women	Newcastle and Tamworth, New South Wales	Multibehavioral change app “MAMA-EMPOWER”	Substance use and psychological distress

XSL•FO
RenderX

Study (author, year)	Study aim	Participants	Research setting	Digital mental health (DMH) service type	Targeted mental health issue
Lee et al [72], 2019	Validate a survey app (Grog Survey app) that explores the alcohol consumption status among First Nations people	20 nondrinkers, 40 nondependent drinkers, and 40 dependent drinkers who self-identified as First Nations people and were aged ≥16 years	An Indigenous primary health care service and surrounding community in urban Queensland, and a regional Aboriginal community–controlled health service and a remote Aboriginal community–controlled drug and alcohol day center (a drop-in service) in South Australia	Grog Survey app	Substance use (alcohol consumption)
Ma et al [73], 2022	Understand the expectations of different communities for telehealth crisis support services in Australia	1300 adults living in Australia, including 2.4% First Nations participants	National setting in Australia	Lifeline Australia	Suicide and self-harm behaviors
Perdacher et al [74], 2022	Determine the feasibility of the Stay Strong app as a digital well-being and mental health tool for use by First Nations people in prison	27 clients and 10 health practitioners	3 high-security prisons in Queensland, Australia	Stay Strong App (custodial version)	General mental health area (unspecified)
Povey et al [75], 2016	Explore First Nations community members' experiences of using 2 e-mental health apps and identify the factors that influence acceptability	9 First Nations adults	Darwin, NT	AIMhi Stay Strong iPad app and the iBobbly suicide prevention app	AIMhi Stay Strong app: general mental health issues and substance use; iBobbly app: psychological distress, depressive symptoms, and suicide prevention
Povey et al [76], 2022	Present an in-depth account of the second phase of participatory design in the development of the AIMhi-Y app	75 First Nations youth	Across Australia	AIMhi-Y app (under design)	General mental health area (unspecified)
Povey et al [77], 2020	Report the result of the formative stage of the AIMhi-Y app development process that engaged First Nations youth in the co-design of the new culturally informed AIMhi-Y app	45 First Nations youth	NT, Australia	General e-mental health services (did not specify service type)	General mental health area (unspecified)
Puszka et al [78], 2016	Understand stakeholder perspectives on the requirements for implementing DMH services in regional and remote health services for First Nations people	32 stakeholders who provide mental health services for First Nations people	NT, Australia	General e-mental health services (did not specify service type)	General mental health area (unspecified)
Raphiphatthana et al [79], 2020	Evaluate the process and effectiveness of the e-mental health program	66 participants from First Nations primary care organizations	Remote NT communities, Darwin, Alice Springs, and Adelaide	Not specified, but most participants used the Stay Strong app	General mental health area (unspecified)

XSL•FO
RenderX

Study (author, year)	Study aim	Participants	Research setting	Digital mental health (DMH) service type	Targeted mental health issue
Raphiphatthana et al [80], 2020	Understand service providers' perspectives and experiences of electronic mental health adoption	57 service providers working with Aboriginal and Torres Strait Islander people who had undergone electronic mental health training workshops	Darwin, Alice Springs, and remote NT communities	General DMH services (but many of the responses were related to the use of the Stay Strong app)	Unspecified
Routledge et al [81], 2022	Assess the acceptability and feasibility of Strong & Deadly Futures, a culturally inclusive alcohol and drug prevention program for First Nations secondary students	19 First Nations students out of 281 students	Two independent urban schools and two rural state schools	Strong & Deadly Futures, a 6-lesson, web-based package combining online illustrated storylines with interactive classroom activities	Substance use (both alcohol and tobacco)
Russell et al [82], 2021	Examine the utility of using a culturally appropriate dementia screening tool (KICA-screen[a]) in a telehealth setting	33 medically stable First Nations inpatients/outpatients	Two local rural health care settings	KICA-screen delivered via telehealth	Cognitive assessment
Snijder et al [83], 2021	Describe the development of Strong & Deadly Futures, a web-based substance use education class	41 First Nations students and 36 non-First Nations students	Four schools in New South Wales	Strong & Deadly Futures, a 6-lesson, web-based package combining online illustrated storylines with interactive classroom activities	Substance use (both alcohol and tobacco)
Tighe et al [85], 2017	Evaluate the effectiveness of a self-help mobile app (iBobbly) targeting suicidal ideation, depression, psychological distress, and impulsivity among Indigenous youth in remote Australia	61 First Nations Australians aged 18 - 35 years	Remote and very remote communities in the Kimberley region of North Western Australia	iBobbly program	Suicidal ideation, depression, psychological distress, and impulsivity
Tighe et al [84], 2020	Discover pilot usage and acceptability data from the iBobbly suicide prevention app, an app distributed through a randomized controlled trial	13 First Nations Australians aged 18 - 35 years	Remote and very remote communities in the Kimberley region of North Western Australia	iBobbly program	Suicidal ideation, depression, psychological distress, and impulsivity
Titov et al [86], 2019	Report on Aboriginal and Torres Strait Islander (Indigenous) users of MindSpot, a national service for the remote assessment and treatment of anxiety and depression	780 First Nations participants	National setting across Australia	MindSpot service, including diagnosis and treatment courses	General mental health area, but this study focused on psychological distress, depression, and anxiety

XSL•FO
RenderX

Study (author, year)	Study aim	Participants	Research setting	Digital mental health (DMH) service type	Targeted mental health issue
Veinovic et al [87], 2023	Evaluate mental state examinations delivered face-to-face (MMSE[b] and KICA-screen) and as telehealth versions (TMMSE[c] and TKI-CA-screen[d]) among older First Nations people	20 First Nations people aged 55 - 69 years	Urban and regional New South Wales, Australia	TMMSE and TKICA-screen	Cognitive assessment

[a]KICA-screen: Kimberly Indigenous Cognitive Examination-short form.

[b]MMSE: Mini-Mental State Examination.

[c]TMMSE: telehealth version of the MMSE.

[d]TKICA-screen: telehealth version of the KICA-screen.

Quality Assessment Results

According to the results from the Aboriginal and Torres Strait Islander Quality Appraisal Tool, the included studies presented an acceptable connection to First Nations values and principles, with at least 16 studies having "Yes" or "Partially" (provided sufficient information) answers to 11 out of 14 questions. However, most studies (n=14-18) did not properly negotiate a formal agreement with First Nations people to access and protect their knowledge. Moreover, some studies were deemed less culturally engaging for First Nations people, such as a study that researched DMH services designed for the general population [73], and some studies applied quantitative methods that resulted in lower participation in First Nations communities [82,86,87].

In terms of the results from the MMAT, all studies were considered to have high-quality research designs. For example, all qualitative studies, except for 1 study [71], completely matched all MMAT criteria. Moreover, all included studies had at least 82% of the criteria marked as "Yes."

Effectiveness

Twelve studies [68,71,72,74,75,80-82,84-87] reported the effectiveness of DMH services in managing mental health conditions among First Nations people in Australia (Table 1). Their results are summarized in the following paragraphs based on their study designs.

Qualitative Evidence

Six studies used a qualitative study design to explore the perceived effectiveness of DMH services in managing mental health conditions among First Nations people [68,74,75,80,81,85]. Three studies [68,74,75] assessed the youth version, custodial version, or standard version of the Aboriginal and Islander Mental Health Initiative (AIMhi) Stay Strong app for First Nations people. The AIMhi Stay Strong app helped participants achieve mental health improvements, guided positive thinking, and supported participants in overcoming difficulties [68,74,75]. Participants perceived the AIMhi Stay Strong app and general DMH services to be effective for improving their mental health outcomes [80]. Moreover, "Strong and Deadly Futures" online courses were examined among First Nations people to prevent alcohol and tobacco use and reduce stress and peer pressure [81]. Both clients and practitioners reported that the app enhanced clients' self-insight, self-reflection, positive thinking, confidence, and overall self-perception. However, in the study by Povey et al [75], while both the AIMhi Stay Strong and iBobbly apps were viewed as helpful for managing minor mental health difficulties, participants considered them less suitable for addressing severe conditions or preventing suicide. Participants agreed that the efficacy of the iBobbly app was limited for "vulnerable people with strong emotions" [84].

Quantitative Evidence

Eight studies reported quantitatively measured effectiveness [68,71,72,82,84-87], with mental health outcomes measured through self-reported instruments and scales.

Quantitative Evidence Without Comparison

Four studies reported the effectiveness of DMH services in managing mental health conditions among First Nations people without comparing them to typical services or other control groups [68,71,81,84]. Kennedy et al [71] found that the multibehavioral change app "MAMA-EMPOWER" was effective in managing substance use and psychological distress and was acceptable (average score of 3.55 on a 5-point Likert scale) for First Nations mothers. Meanwhile, Routledge et al [81] found that Strong & Deadly Futures (a web-based lesson package aiming to improve mental well-being and substance use prevention) was effective in 53.3% of First Nations students, who were likely to use the knowledge they obtained from the course, and no participants reported that they were unlikely to apply that knowledge. Moreover, 73.3%, 66.7%, and 80.0% of First Nations students thought that this program was helpful in managing substance use, stress, and peer pressure, respectively. However, Dingwall et al [68] found mixed therapeutic outcomes for the AIMhi-Y general mental health support app among First Nations youth. More specifically, the app was found to be effective in reducing psychological distress and depressive symptoms (Cohen d=0.85 and 0.75, respectively) and reducing anxiety symptoms (without statistical significance), but was not effective in managing substance use disorders or improving psychological adjustment. Furthermore, Tighe et al [84] found that the iBobbly app was not effective in managing suicidal ideation and reducing psychological distress or depression.

XSL•FO
RenderX

However, all indicators presented a trend of improvement. See Table 1 and Multimedia Appendix 3 for detailed information.

Quantitative Evidence With Comparison

Five studies examined the effectiveness of DMH services in managing [85,86] and assessing [72,82,87] mental health conditions by comparing the effectiveness of DMH services with that of typical mental health services [72,82,87] or waitlist control groups [85], or comparing the effectiveness of DMH services between First Nations people and non-First Nations people [86].

The Grog Survey app [72], which was developed for assessing substance abuse among First Nations people, was found to be effective in detecting the alcohol consumption level (94.7% overall agreement between the app and clinical interviews). The app was also found to be more sensitive than a clinical interview for the number of drinks, the greatest number of drinks, the median of reported drinking, and the daily consumption of standard drinks per drinking occasion (mean difference=1.6, 0.8, 3, and 0.4, respectively). Higher sensitivity was found among those classified as risky drinkers (2.3% and 4.5% more detection among males and females, respectively), as well as in the reporting (17.0% vs 7.2%) and prediction (P=.02 vs P=.44) of tremors. See Table 1 and Multimedia Appendix 3 for detailed information.

A previous study compared the telehealth versions of the Mini-Mental State Examination (TMMSE) and Kimberly Indigenous Cognitive Examination-short form (TKICA-screen), which were used for cognitive screening, with their typical versions [87]. The study found that the median TMMSE scores were slightly lower than the Mini-Mental State Examination (MMSE) scores (median difference=4.5), and the agreement measured by correlation was moderate (r_s=0.33; P=.20) but with unacceptable limits of agreement (the effectiveness may not significantly reduce in the TMMSE). On the other hand, the TKICA-screen showed a significant effectiveness loss compared with the Kimberly Indigenous Cognitive Examination-short form (KICA-screen). For instance, there was a significant mean difference of 2.17, with a tendency for poorer agreement at the lower end of the performance. See Table 1 and Multimedia Appendix 3 for evidence of proportional bias. Russel et al [82] found that the KICA-screen delivered through digital platforms was as effective as the face-to-face version, with a statistically significant correlation (Pearson r=0.851; P<.01) and agreement (intraclass correlation coefficient=0.85; P<.01) between the versions.

Titov et al [86] found that the MindSpot service, aiming to provide general mental health support, was effective in managing psychological distress, depression, and anxiety, with statistically significant decreases in the Kessler Psychological Distress Scale-10, Patient Health Questionnaire-9, and Generalized Anxiety Disorder Scale-7 scores (χ^2=5933.1, 5938.8, and 5658.8, respectively). As a DMH service designed for all Australians, there were no statistically significant differences in treatment outcomes between First Nations patients and non-First Nations patients or between First Nations patients who chose the general well-being course and those who chose the First Nations well-being course.

Tighe et al [85] evaluated the iBobbly app that was designed for suicide prevention. The results suggested that there were no statistically significant differences between the iBobbly app group and the waitlist control group in suicide prevention ($t_{58.1}$=2.40; P>.05). However, it was effective in managing depression when compared with a waitlist control group and preintervention baseline scores (P<.05). Similarly, the DMH service was found to be effective in managing psychological distress ($t_{57.5}$=2.44). See Table 1 and Multimedia Appendix 3 for P values and effect sizes. The iBobbly app was found to not be effective in managing impulsivity.

Determinants of DMH Service Implementation

Overview

Overall, this review identified the determinants of facilitators and barriers that influence the implementation of DMH services. See Table 3 for the extracted themes.

Table 3 presents 9 themes and 26 subthemes that describe the facilitators and barriers influencing the implementation of DMH services for First Nations people. There were 4 subthemes under *organizational factors*, 4 subthemes under *cultural appropriateness*, 6 subthemes under *accessibility*, 1 subtheme under *engagement*, 2 subthemes under *coverage*, 6 subthemes under *acceptability*, 1 subtheme under *integration of DMH services with existing context*, 4 subthemes under *digital literacy*, and 1 subtheme under *efficiency of DMH services*.

XSL•FO
RenderX

Table . Themes and subthemes of the identified studies.

Theme and subtheme	Supporting evidence
Theme 1: Organizational factors within DMH[a] service providers	
Supportive structure and managers within DMH service provider organizations facilitate DMH service uptake	• Supportive structures or working culture within organizations facilitated resource allocation or staff training for DMH service uptake [78-80]. • Enthusiastic managers provided direct support, which increased the willingness of staff and the organization itself for DMH service uptake [66,80].
Policy and structural obstacles and a lack of resources in DMH service provider organizations create barriers for DMH service uptake	• Procedural and administrative obstacles and obstructive IT policies limited the access of staff to technologies and DMH services and created failures in digital resource allocation [66,67,79,80]. • High workloads within the workplace and a lack of digital resources resulted in a lack of DMH training for DMH service providers [66,79,80].
Theme 2: Cultural appropriateness of DMH services	
DMH services could be culturally appropriate owing to cultural safety designs and culturally relevant content in DMH services and empowerment, which can increase the willingness of First Nations people to use DMH services	• Overall cultural safety reported by participants [74]. • DMH services empowered clients to record their own information and interact with service providers [78,80]. • Cultural safety designs were applied in DMH services, which created a safe environment or directly increased the acceptability and interest of clients [69,70,75,76,83,84].
Lack of cultural appropriateness creates barriers for First Nations users to overcome existing negative perceptions and accept DMH services	• Negative perception toward the health system created barriers for First Nations people to use DMH services [67]. • Problematic use of English within DMH services and lack of First Nations languages may cause difficulties in using DMH services [75,83].
Theme 3: Accessibility of First Nations people to DMH services	
First Nations people have mixed perceptions of their accessibility to DMH services	• First Nations clients generally appraised the high accessibility of DMH services, such as reporting that the DMH services they received were accessible [83,86]. • However, general accessibility-related concerns, such as a lack of services and resources, and long response times, were also expressed [67].
Portability and flexibility of DMH services, as well as existing digital infrastructure coverage among First Nations communities, could facilitate the implementation of DMH services	• DMH services were available on multiple digital platforms [75]. • A high coverage of digital infrastructure among communities was presented, which facilitated the uptake of DMH services [67,77]. • DMH services can be operated on portable devices, which can create unique advantages in accessibility [78,84].

XSL•FO
RenderX

Theme and subtheme	Supporting evidence
Many First Nations people report limited access to appropriate digital infrastructure, resulting in low accessibility to DMH services	• In some First Nations communities, a lack of digital infrastructure and devices resulted in difficulties in the implementation of DMH services [69,77-80]. • Even in special environments, such as prisons, access to digital devices and some specific functions of DMH services, such as access to images, was limited [74]. • In remote areas, a lack of digital infrastructure created barriers for First Nations residents to implement DMH services [75,78].
Free DMH apps are highly accessible, while those with download fees limit client access	• Apps with high accessibility were preferred, such as free-to-download apps that can operate on different platforms [76]. • The cost of e-mental health apps negatively influenced their accessibility [75].
DMH services that do not have First Nations language versions create accessibility barriers for First Nations people whose first language is not English	• A lack of DMH services in First Nations languages created accessibility barriers [69,75,78].
Theme 4: Ability of DMH services to promote engagement	
DMH services can promote engagement between clients and service providers by breaking the power imbalance and then creating facilitating effects	• DMH services could help establish positive relationships between health care providers and clients, eventually creating positive experiences for clients or increasing First Nations clients' willingness to use these services [69,74,78].
Theme 5: Ability of DMH services to cover a wide range of settings and First Nations population groups	
DMH services could cover different types of clients, topics, and reference periods	• High applicability across different age groups and mental health conditions [69]. • DMH services could cover sensitive topics due to their nature of indirect engagement [69,78]. • Compared to typical face-to-face mental health services, DMH services could cover longer reference periods on substance use topics [72].
DMH services fail to cover some severe mental health conditions, which is considered a barrier	• DMH services may not be able to cover severe mental health conditions, such as suicide [75,78].
Theme 6: Acceptability of First Nations people to DMH services	

XSL•FO
RenderX

Theme and subtheme	Supporting evidence
Acceptability of First Nations people to DMH services is perceived as high	• Satisfaction and optimism regarding DMH services and willingness to use DMH services to seek help were presented by First Nations clients, indicating high acceptability [68,69,71,74,75,77,81,84,86]. • First Nations people also expressed their idea of not expecting to receive help from DMH services [73].
Good visual design can facilitate participants' acceptance of DMH services, while visual design problems are considered barriers	• Clients reported that the DMH services they received had an attractive or engaging visual design, which increased their acceptability of DMH services [68,69,71,74,78]. • Some visual design problems also raised the attention of clients, such as high text density [70].
The unique nature of DMH services has high acceptability among First Nations people	• First Nations clients preferred the web-based, client-led, and clinician-supported nature of DMH services [74,75,83].
Attractive and useful content included in DMH services facilitates First Nations people to accept DMH services	• First Nations clients perceived that mental health resources in DMH services were generally interactive and useful [71,80]. • Embedded media forms, such as games, videos, and animations, were especially preferred by First Nations clients [68,75]. • Tailored content for First Nations students was welcomed [83].
Inadequate, repetitive, and less engaging content included in DMH services is considered a barrier	• Inadequate and repetitive content in DMH services reduced the acceptability of DMH services [68,71,84].
The unique functions of DMH services increase their acceptability among First Nations clients	• The unique functions of DMH services, such as text messages sent as reminders or data capture functions, were acceptable for First Nations clients and prompted clients to use those services [68,69,74,78].
Some functions are not attractive enough or are difficult to use, resulting in lower acceptability of DMH services	• However, some functions of DMH services were ignored, could hardly be used, or could not fit in special contexts, such as in prison [68,71,74,75].
DMH services with data safety have high acceptability, but the concern about data safety is perceived as a barrier	• DMH service apps with high privacy and data safety were preferred [76]. • Some service providers and clients also considered data unsafety as a potential risk and barrier [67,78].
Theme 7: Ability of DMH services to integrate with existing contexts	
Some DMH services are hard to integrate with existing contexts	• Some service providers felt that it was challenging to integrate DMH services into their routine care because DMH services may not easily fit group clients, pre-existing practices, and local community status [79,80].
Theme 8: Digital literacy of clients and service providers	

XSL•FO
RenderX

Theme and subtheme	Supporting evidence
Clients and service providers who have high digital literacy facilitate the uptake of DMH services	• Young clients presented especially high digital literacy and willingness to receive DMH services [67,80]. • Based on their own enthusiasm for using DMH services or specific DMH training, some service providers obtained high DMH literacy [66,67].
Clients who do not have high digital literacy may face challenges when operating DMH services	• Some clients reported that the DMH app is difficult to operate [70,75,78]. • Relatively lower digital literacy resulted in lower awareness of DMH services [68,75,80].
Service providers who do not have high digital literacy may face challenges when operating or providing DMH services	• Low digital literacy of service providers led to negative perceptions, including overestimating the complexity of technology or thinking DMH services will distract them from their current job [66,67,78]. • Inadequate digital literacy among service providers created challenges in maintaining and using DMH services [67,79,80]. • Overall, the digital literacy level among service providers was highly heterogeneous [78,79].
Theme 9: Efficiency of DMH services	
DMH services have high efficiency and are considered time-saving	• Service providers believed that DMH services could reduce the time cost of recording client notes and waiting for data transmission, and these had advantages for data recording and management compared to typical paper format services [78,80].

[a]DMH: digital mental health.

Organizational Factors

Organizational factors related to the administration and management of the organization (usually service providers) could either facilitate or create barriers to implementing DMH services.

Four studies asked stakeholders and service providers to generally comment on DMH services [66,78-80]. The supportive structure within the organizations was a significant facilitator, which included the regular supervision and review of DMH service utilization [78]; being supportive of allocating resources, time, and training [79]; and building the work culture that welcomes DMH services within the workforce [80]. Direct support from managers was also considered to facilitate the uptake of DMH services [80] by increasing the opportunity and interest of staff in adopting DMH services [66].

Four studies asked mental health service providers about their general perspectives on DMH services [66,79,80] and on using DMH services in suicide prevention [67]. Service providers reported that organizational factors created barriers for the uptake of DMH services. These included a lack of organizational support [80], administrative obstacles, obstructive IT policies within the workplace [66,79], work restrictions [67], a lack of training resources within organizations [79], and time allocated to cope with the extra workload brought on by the implementation of DMH services [66,79,80].

Cultural Appropriateness

Cultural appropriateness was perceived as a critical determinant of the uptake of DMH services within First Nations people in Australia, with high cultural appropriateness facilitating uptake and low cultural appropriateness impeding uptake.

Ten studies reported high culturally appropriate DMH services [69-71,74-76,78,81,83,84]. The users of the AIMhi Stay Strong app [69] and its custodial version [74] reported satisfaction with cultural safety. Similarly, participants in programs, such as the "SMS4dads" mobile app program, "Stayin' on Track" website program [70], AIMhi-Y app program [76], and iBobbly suicide prevention program [84], valued the culturally appropriate visual and content design [76] as well as the private space these tools provide to reduce shame and stigma [75,84]. Stakeholders also highlighted the empowering nature of DMH services that allowed users to record personal information and engage with culturally meaningful content [71,78].

The "Strong & Deadly Futures" web-based substance use prevention program, though designed for the general population, was also praised for its cultural appropriateness owing to its inclusion of First Nations characters and an equitable learning

XSL•FO
RenderX

environment for both First Nations and non-First Nations students [81,83].

Three studies reported a lack of cultural safety within DMH services [67,75,83]. Some First Nations health workers perceived digital suicide prevention tools as culturally inappropriate, reflecting broader concerns about cultural unsafety in the mental health system [67]. Users of the AIMhi Stay Strong app and the iBobbly suicide prevention app expressed uncertainty about whether these tools could address trauma linked to colonization [75], and "Strong & Deadly Futures" participants provided mixed responses for the use of "Aboriginal English" [83].

Accessibility

Accessibility was another key determinant of DMH service implementation among First Nations people. Two studies generally evaluated accessibility [83,86], noting that MindSpot could deliver care when local face-to-face mental health services were unavailable [86], and Strong & Deadly Futures was accessible to users with varying abilities [83]. However, concerns remained about poor communication systems, limited resources, slow response times, and service gaps in digital suicide prevention tools [67].

Five studies reported that improved digital infrastructure supported access to DMH services [67,75,77,78,84]. Apps, such as AIMhi and iBobbly, were considered accessible due to their compatibility with multiple devices and their portability [67,75,77,78,84], which enabled use across locations and levels of remoteness. Nevertheless, limited availability of devices and internet connectivity reduced access [78-80], particularly in remote communities and custodial settings [74,75].

Cost was another barrier, with participants preferring free-to-download apps, and the fees for the AIMhi Stay Strong and iBobbly apps were seen as deterrents [75,76].

Finally, language barriers restricted access, as most apps lacked First Nations language options, making it difficult for some users to understand key terms such as "resilience" [69,75,78].

Integration

The difficulty in integrating DMH services with existing practices was generally perceived as a barrier in the included studies. More specifically, some service providers in 2 studies commented on applying DMH services in general mental health areas [79,80] and reported difficulties in integrating DMH services into usual practices when facing non-one-to-one settings, such as facing a group of clients [79], or the need to fit existing practices and the local community status [80].

Engagement (Between Service Providers and Clients)

Three studies highlighted that DMH services promoting engagement between clients and service providers were well received in general mental health care [69,74,78]. The AIMhi app was reported by health providers to reduce barriers, such as power imbalance, and foster positive relationships between health care providers and clients [69]. Moreover, stakeholders expanded this result to other DMH tools that could provide indirect engagements when solving sensitive issues and supported the building of positive relationships [78]. Improved rapport was also observed in custodial settings through the custodial version of the Stay Strong app [74].

Coverage

Coverage refers to the ability of DMH services to cover different types of clients and various mental health conditions. Three studies found that broader coverage facilitated implementation [69,72,78]. The AIMhi app was reported to reach clients across age groups and symptom profiles, including those who were otherwise difficult to engage [69]. Moreover, clients of the AIMhi app [69] and service providers providing general opinions on DMH services [78] noted that DMH services could address sensitive mental health topics. Furthermore, the Grog Survey app, which aimed to investigate the alcohol dependence status, enabled coverage of longer reporting periods than typical clinical interviews [72].

However, 2 studies reported that clients and service providers felt DMH services, including the AIMhi and iBobbly apps, may be unsuitable for addressing severe mental health conditions such as suicide [75,78].

Acceptability

Acceptability refers to the extent to which DMH services are acceptable to First Nations Australians. Ten studies mentioned overall positive acceptability [68,69,71,73-75,77,81,84,86]. Except for Lifeline Australia, which was rated low due to participants' limited expectations of receiving suicide prevention support [73], all other services, including the AIMhi Stay Strong app and its custodial and youth versions [68,69,74], "MAMA-EMPOWER" [71], AIMhi, iBobbly [75,84], "Strong & Deadly Futures" [81], and MindSpot [86], were generally well received, showing high satisfaction, usability, and acceptability [77].

The nature of DMH services was reported as acceptable by 3 studies [74,75,83]. The custodial version of the AIMhi app was appraised for its client-led nature [74], and the AIMhi app and iBobbly suicide prevention app were praised for their clinician-supported nature that could overcome mental health literacy barriers [75]. The web-based nature of "Strong & Deadly Futures" was preferred by students of web lessons [83].

Two studies emphasized relevance and noted that MAMA-EMPOWER was meaningful to users facing relevant issues, although substance use content was viewed as less relevant by nonusers or younger students [71,81].

Four studies discussed digital functions, identifying useful features, such as text messaging in AIMhi [68], service consistency functions in AIMhi [69], goal-setting in the custodial Stay Strong app [74], and data capture features [78]. However, some functions were ineffective or difficult to use, such as the "stories" page in AIMhi-Y [68], and had limited avatar customization [74] and hard-to-use designs [71,75].

Five studies described content as a major contributor to acceptability [68,71,75,80,83]. Participants valued engaging and informative material, including media components (games, videos, and graphics) and culturally tailored content for First

XSL•FO
RenderX

Nations users [83]. Conversely, inadequate or repetitive content reduced engagement [68,71,84].

Six studies addressed visual design [68-71,74,78]. Most participants praised the cultural appropriateness, color, and user-friendly layout of AIMhi and MAMA-EMPOWER, while some found SMS4dads and Stayin' on Track overly text-heavy [70]

Finally, 3 studies discussed data safety and privacy as concerns [67,76,78]. Youth participants preferred DMH tools with secure data handling [76], but general apprehension about data security, particularly in suicide prevention contexts, was reported as a barrier [67,78].

Digital Literacy

Digital literacy was identified as a key determinant of the implementation of DMH services. Three studies reported that high digital literacy facilitated the use and acceptability of DMH services [66,67,80]. Young First Nations clients with higher digital literacy demonstrated greater acceptance of DMH tools [67,80], and some suicide prevention gatekeepers preferred digital platforms, such as Facebook, for service delivery [67]. Training also improved providers' digital literacy, enhancing DMH adoption [66].

However, low digital literacy hindered engagement for both clients [68,70,75,78,80] and service providers [66,67,78-80]. Clients reported difficulty operating apps, such as SMS4dads, Stayin' on Track, AIMhi, and iBobbly [70,75,78], with some forgetting or being unaware of DMH tools [75]. Among providers, digital competency varied widely, with some describing themselves as "less technologically competent" and struggling with basic functions like downloading apps [67,79,80].

Efficiency

Efficiency was another determinant of the implementation of DMH services. Two studies [78,80] reported that service providers perceived DMH services as improving efficiency. More specifically, DMH services could reduce the time cost of recording client notes and waiting for data transmission [78], offering advantages in data recording and management compared with traditional paper-based systems [80].

Discussion

Principal Findings

This is the first review to synthesize evidence on the effectiveness of DMH services and the determinants of the facilitators and barriers of DMH services for First Nations people in Australia. Our review found that DMH services were effective in assessing mental health conditions, monitoring mental health status, and delivering mental health education, and had the potential to improve mental health conditions among First Nations Australians. However, the effectiveness of these services for severe mental health conditions remains limited. Additionally, this review identified determinants that could facilitate or inhibit the uptake of DMH services among First Nations people in Australia.

Effectiveness of DMH Services

The effectiveness of DMH services for First Nations Australians varied across different mental health outcomes and intervention types. Consistent with a previous government report [88] and similar studies [89-91], qualitative studies included in this review suggested that DMH interventions can support mental health improvement, increase self-reflection, and enhance self-confidence and empowerment [68,74,75,80,81,84].

Considering that most included DMH services applied culturally adapted designs, these findings align with broader evidence that culturally adapted DMH tools can provide valuable support for the mental health of First Nations people [3,10,92]. However, concerns remain about the adequacy of digital interventions for severe mental health conditions, particularly suicide prevention, where some participants expressed doubts about their effectiveness beyond mild to moderate distress [75,84,85]. This highlights the ongoing challenge of delivering digital suicide prevention interventions effectively in First Nations communities, as also noted in previous research [55].

Moreover, evidence supports the effectiveness of DMH interventions for substance use. The Grog Survey app accurately identified at-risk drinkers with high sensitivity and reliability [72], and it was as effective as the web-based lesson package Strong & Deadly Futures in preventing substance use [81]. However, the AIMhi-Y app was ineffective in reducing alcohol and drug use symptoms [68], indicating that while DMH tools could be valuable for monitoring and performing early intervention, they might not have sufficient therapeutic impact for substance use disorders without additional face-to-face support. This aligns with previous research on First Nations people in the United States, who had insignificant therapeutic outcomes [93]. The use of DMH interventions for cognitive assessment has also yielded mixed results [82,87]. These findings highlight the need for careful adaptation of cognitive assessments for telehealth delivery, as the telehealth version of the test reduced the visual cues of the test [87] and resulted in insignificant results, but the same test that was delivered through video conferencing obtained acceptable effectiveness compared with typical in-person services [82]. Finally, our engagement and user perception findings indicated that DMH education and behavioral change interventions can be beneficial but require further refinement. For instance, the Strong & Deadly Futures program was generally acceptable to First Nations students owing to the perceived effectiveness in preventing substance use [81]. However, a substantial proportion of students remained unsure about applying the information in their own lives. This suggests that while digital education programs can improve mental health literacy, their impact on long-term behavioral change may be limited without additional reinforcement, as noted in previous research [94].

This review found that the effectiveness of DMH services for First Nations people in Australia is in line with that for First Nations people from other countries. For example, the effectiveness of DMH services in Australia was found in international evidence, such as a general mental health program that successfully caused positive behavioral change for First Nations people in Canada [95] and a mental health screening

XSL•FO
RenderX

program in New Zealand [96]. This indicates that the experience and knowledge generated from First Nations people can be generalized. Heterogeneity was present in some areas. For example, a program in the United States achieved statistical significance in treating substance use issues for First Nations people [93], while a program with a similar target only obtained insignificant results in Australia [68]. Considering that the situation and historical factors faced by First Nations people in Australia are unique, care should be taken when generalizing the results from this review.

Determinants of the Uptake of DMH Services Among First Nations People in Australia

This review highlights that the implementation of DMH services for First Nations people is shaped by interconnected structural, cultural, and individual factors. Organizational readiness, including leadership support, workforce capacity, and digital infrastructure, remains foundational. Leadership drives the strategic vision, fosters innovation, and ensures resource allocation, which are essential for successful implementation [97,98]. Strong leadership also facilitates stakeholder engagement and addresses resistance to change, which are key factors in digital health uptake [99]. Organizational support further enhances workforce readiness and policy development, sustaining long-term adoption [100]. Without these elements, DMH initiatives risk poor adoption and limited impact [66,67,79,80]. Additionally, cultural appropriateness emerged as a defining determinant of uptake. Culturally safe, co-designed tools enhanced engagement and empowerment, whereas culturally unsafe or linguistically limited designs hindered trust and acceptability. Culturally relevant app designs incorporating First Nations art and metaphors were found to enhance engagement [69,70,75,76,83,84]. This echoes prior research emphasizing the importance of embedding cultural elements to ensure relevance and acceptance within First Nations communities [101].

Accessibility, digital literacy, and efficiency collectively reflect the broader digital divide that continues to shape health equity in Australia. First, inconsistent digital literacy among both clients and practitioners has been previously identified as a key obstacle to effective use of DMH services [66-68,70,75,78-80]. Delivering technology training and ongoing support could improve digital literacy and confidence in using DMH services [102,103], ultimately enhancing their uptake and efficacy among First Nations communities. Second, while improved infrastructure and targeted training enhanced access and usability in some First Nations communities, persistent disparities in internet connectivity, device availability, and digital skills constrained the scalability of DMH interventions, particularly in remote and custodial settings [104]. Despite evidence showing common ownership of digital infrastructure in some First Nations communities [67,77], limited infrastructure and technology access remain significant challenges, particularly in remote communities where access to tablets, Wi-Fi, and reliable technology is often restricted [67,69,75,77-80]. It has been reported that First Nations Australians, especially those living in remote areas, are some of the most digitally excluded populations [105]. This finding highlights the urgent need for targeted investments in digital infrastructure to bridge the technology gap and ensure equitable access to DMH services. To address these challenges, it could be helpful to expand community Wi-Fi programs, provide subsidized digital devices, and implement mobile-based DMH solutions that require minimal bandwidth [88]. The Mob Link initiative by the Institute for Urban Indigenous Health is a strong example [106], as it provides culturally responsive digital health support, including assistance with telehealth and digital access, helping to reduce barriers and improve engagement with health care services. Without sustained investment in digital inclusion and community-based training, DMH services risk reproducing existing inequities rather than closing service gaps [107].

At the service level, DMH tools enhance engagement and communication between clients and providers, supporting self-reflection and rapport. However, limited integration with existing care models, concerns about data privacy, and uncertainty about applicability to severe mental health conditions reduce trust and engagement with DMH services and signal the need for clearer clinical pathways and blended-care frameworks. For instance, privacy concerns and the fear of data misuse have been constantly identified as significant deterrents to the uptake of DMH services [67,78]. It is imperative to address the concerns of First Nations people regarding data sovereignty and the potential misuse of personal health information [108], so that DMH services can be designed with robust privacy protection and transparent data governance frameworks. Furthermore, streamlined data management and integration may be helpful for integrating DMH services and reducing administrative burden, enabling system-level adoption of DMH services for First Nations Australians.

Heterogeneity of DMH Service Types

This review included various DMH service types, which caused differentiation in determinants. For example, the telehealth version of an included cognitive screening tool showed lower effectiveness due to the absence of visual cues in the test [87], while the same test with visual cues obtained acceptable effectiveness [82]. A possible explanation for this difference is that the latter test was delivered through video conferencing and involved a complete version of the assessment containing images and visual cues. Similarly, the visual design and multimedia content of DMH apps [68,69,71,74,75,78,80,83] and the various digital functions in DMH apps [68,69,74,78] were praised for increasing the acceptability of DMH services among First Nations people. However, no similar facilitators were reported by the users of telehealth DMH services, hinting at the potential benefits of designing app- or web-based DMH services compared with telehealth DMH services. Despite such diversities, this review aimed to synthesize all evidence representing technology-assisted mental health services. Therefore, it has provided an integrated understanding of how digital approaches have been applied to support mental health among First Nations people.

Limitations

We acknowledge that this review has several limitations. First, the included studies were highly heterogeneous in terms of intervention types, study designs, and outcome measures. This limited our ability to draw precise or generalizable conclusions

XSL•FO
RenderX

about overall effectiveness. Second, as our search was restricted to academic databases, relevant gray literature may have been missed, potentially contributing to publication bias. Third, many included studies had a moderate to high risk of bias, with small sample sizes and self-reported outcomes, which may reduce the certainty of the evidence.

Conclusion

This review highlights the effectiveness of DMH services in assessing, monitoring, and managing mental health conditions for First Nations Australians. However, information on their impact on severe conditions, including suicide, remains limited. Digital health apps showed promising results in improving general mental health support. Digital health apps and online courses were also perceived to be effective in preventing and assessing substance use, but did not have a significant treatment effect on managing it. For screening mental health conditions, the digital health versions of instruments were comparable to face-to-face versions, while the telehealth versions showed less effectiveness than both the digital health and face-to-face versions, indicating that effectiveness may vary across different types of DMH services.

Long-term funding in DMH services through culturally responsive community services (eg, Institute for Urban Indigenous Health) from the government or organizations should be beneficial, especially for First Nations people facing barriers to accessing in-person mental health services (eg, those who reside in remote areas) and for DMH services that are supported by evidence.

This study also identified key facilitators (eg, strong leadership, culturally relevant designs, clinician-supported tools, and community) and major barriers (eg, digital exclusion, low digital literacy, and privacy concerns). These findings highlight the need for targeted infrastructure investment that can reduce the digital divide for First Nations people, especially in regional and remote communities; digital training for both mental health care providers and First Nations clients; and transparent data governance. However, the evidence should be interpreted with caution, given the heterogeneity of study designs and outcomes, potential publication bias, and moderate to high risk of bias across the included studies. Future research should focus on co-designing DMH services with First Nations communities, rigorously evaluating their effectiveness, and integrating them into existing health care models to ensure accessibility, cultural relevance, and long-term impact.

Acknowledgments

Artificial intelligence (ChatGPT, OpenAI) was used solely to refine grammar and improve expression clarity. All authors reviewed and approved all content after artificial intelligence–assisted editing.

Funding

No funding is involved.

Data Availability

All data generated from this study are publicly available.

Authors' Contributions

Conceptualization: SZ, ACS, SD, XZ
Investigation: SZ, SD, XZ
Methodology: SZ, AG, SD, XZ
Project administration: ACS
Supervision: SD, XZ
Writing – original draft: SZ
Writing – review & editing: AG, ACS, SD, XZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed inclusion and exclusion criteria, and PICO (population, intervention, comparison, outcome) protocol.
[DOCX File, 19 KB - i-jmr_v15i1e80386_app1.docx]

Multimedia Appendix 2
Details of the search strategy.
[DOCX File, 19 KB - i-jmr_v15i1e80386_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Details of the extracted data.
[XLSX File, 34 KB - i-jmr_v15i1e80386_app3.xlsx]

Multimedia Appendix 4
Quality assessment results (Mixed Methods Appraisal Tool).
[XLSX File, 14 KB - i-jmr_v15i1e80386_app4.xlsx]

Multimedia Appendix 5
Quality assessment results (Aboriginal and Torres Strait Islander Quality Appraisal Tool).
[XLSX File, 14 KB - i-jmr_v15i1e80386_app5.xlsx]

Checklist 1
PRISMA checklist.
[DOCX File, 26 KB - i-jmr_v15i1e80386_app6.docx]

References

1. Butler TL, Anderson K, Garvey G, et al. Aboriginal and Torres Strait Islander people's domains of wellbeing: a comprehensive literature review. Soc Sci Med 2019 Jul;233:138-157. [doi: 10.1016/j.socscimed.2019.06.004]
2. Ralph S, Ryan K. Addressing the mental health gap in working with Indigenous youth: some considerations for non‐Indigenous psychologists working with Indigenous youth. Aust Psychol 2017 Aug 1;52(4):288-298. [doi: 10.1111/ap.12287]
3. Farah Nasir B, Brennan-Olsen S, Gill NS, et al. A community-led design for an Indigenous model of mental health care for Indigenous people with depressive disorders. Aust N Z J Public Health 2021 Aug;45(4):330-337. [doi: 10.1111/1753-6405.13115] [Medline: 34097328]
4. Parker R. Australia's Aboriginal population and mental health. J Nerv Ment Dis 2010 Jan;198(1):3-7. [doi: 10.1097/NMD.0b013e3181c7e7bc] [Medline: 20061862]
5. Nasir BF, Toombs MR, Kondalsamy-Chennakesavan S, et al. Common mental disorders among Indigenous people living in regional, remote and metropolitan Australia: a cross-sectional study. BMJ Open 2018 Jun 30;8(6):e020196. [doi: 10.1136/bmjopen-2017-020196] [Medline: 29961007]
6. Social networks and wellbeing. Australian Bureau of Statistics. 2016. URL: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4714.0~2014-15~Main%20Features~Social%20networks%20and%20wellbeing~4 [accessed 2024-08-05]
7. Fatima Y, Cleary A, King S, Solomon S, McDaid L, Hasan MM, et al. Cultural identity and social and emotional wellbeing in Aboriginal and Torres Strait Islander children. In: Baxter J, Lam J, Povey J, Lee R, Zubrick SR, editors. Family Dynamics over the Life Course Life Course Research and Social Policies: Springer; 2022, Vol. 15:57-70. [doi: 10.1007/978-3-031-12224-8_4]
8. Aboriginal and Torres Strait Islander people with a mental health condition. Australian Bureau of Statistics. 2016. URL: https://tinyurl.com/5xw32uzz [accessed 2024-08-05]
9. Determinants of health for First Nations people. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/australias-health/social-determinants-and-indigenous-health [accessed 2026-01-16]
10. Sabbioni D, Feehan S, Nicholls C, et al. Providing culturally informed mental health services to Aboriginal youth: the YouthLink model in Western Australia. Early Interv Psychiatry 2018 Oct;12(5):987-994. [doi: 10.1111/eip.12563] [Medline: 29573565]
11. Hepworth J, Askew D, Foley W, et al. How an urban Aboriginal and Torres Strait Islander primary health care service improved access to mental health care. Int J Equity Health 2015 Jun 6;14(1):51. [doi: 10.1186/s12939-015-0183-x] [Medline: 26048616]
12. Isaacs AN, Pyett P, Oakley‐Browne MA, Gruis H, Waples‐Crowe P. Barriers and facilitators to the utilization of adult mental health services by Australia's Indigenous people: seeking a way forward. Int J Mental Health Nurs 2010 Apr;19(2):75-82. [doi: 10.1111/j.1447-0349.2009.00647.x]
13. Markwick A, Ansari Z, Sullivan M, McNeil J. Social determinants and psychological distress among Aboriginal and Torres Strait islander adults in the Australian state of Victoria: a cross-sectional population based study. Soc Sci Med 2015 Mar;128:178-187. [doi: 10.1016/j.socscimed.2015.01.014] [Medline: 25618607]
14. Sherwood J. Colonisation - it's bad for your health: the context of Aboriginal health. Contemp Nurse 2013 Dec;46(1):28-40. [doi: 10.5172/conu.2013.46.1.28] [Medline: 24716759]
15. Thurber KA, Colonna E, Jones R, et al. Prevalence of everyday discrimination and relation with wellbeing among Aboriginal and Torres Strait Islander adults in Australia. Int J Environ Res Public Health 2021 Jun 18;18(12):6577. [doi: 10.3390/ijerph18126577] [Medline: 34207406]

XSL•FO
RenderX

16. Kirkbride JB, Anglin DM, Colman I, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 2024 Feb;23(1):58-90. [doi: 10.1002/wps.21160] [Medline: 38214615]
17. National Aboriginal and Torres Strait Islander health survey. Australian Bureau of Statistics. 2024. URL: https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/national-aboriginal-and-torres-strait-islander-health-survey/latest-release [accessed 2024-08-11]
18. Page IS, Ferrari AJ, Slade T, Anderson M, Santomauro D, Diminic S. Estimating the difference in prevalence of common mental disorder diagnoses for Aboriginal and Torres Strait Islander peoples compared to the general Australian population. Epidemiol Psychiatr Sci 2022 Jun 21;31:e44. [doi: 10.1017/S2045796022000233] [Medline: 35726614]
19. Coates D, Saleeba C, Howe D. Mental health attitudes and beliefs in a community sample on the central coast in Australia: barriers to help seeking. Community Ment Health J 2019 Apr;55(3):476-486. [doi: 10.1007/s10597-018-0270-8] [Medline: 29589218]
20. Hunter E. Disadvantage and discontent: a review of issues relevant to the mental health of rural and remote Indigenous Australians. Aust J Rural Health 2007 Apr;15(2):88-93. [doi: 10.1111/j.1440-1584.2007.00869.x] [Medline: 17441816]
21. Larson A. Rural health's demographic destiny. Rural Remote Health 2006;6(2):551. [Medline: 16623618]
22. Reifels L, Bassilios B, Nicholas A, et al. Improving access to primary mental healthcare for Indigenous Australians. Aust N Z J Psychiatry 2015 Feb;49(2):118-128. [doi: 10.1177/0004867414562046] [Medline: 25492971]
23. Garay J, Williamson A, Young C, et al. Aboriginal young people's experiences of accessibility in mental health services in two regions of New South Wales, Australia. Int J Environ Res Public Health 2023 Jan 18;20(3):1730. [doi: 10.3390/ijerph20031730] [Medline: 36767098]
24. Kavanagh BE, Corney KB, Beks H, Williams LJ, Quirk SE, Versace VL. A scoping review of the barriers and facilitators to accessing and utilising mental health services across regional, rural, and remote Australia. BMC Health Serv Res 2023 Oct 4;23(1):1060. [doi: 10.1186/s12913-023-10034-4] [Medline: 37794469]
25. McIntyre M, Ehrlich C, Barnett L, Kendall E. The experience of isolated practice for Indigenous mental health and addictions workers. Int J Ment Health 2022 Jan 2;51(1):77-94. [doi: 10.1080/00207411.2021.1984128]
26. Schultz R, Quinn S, Wilson B, Abbott T, Cairney S. Structural modelling of wellbeing for Indigenous Australians: importance of mental health. BMC Health Serv Res 2019 Jul 15;19(1):488. [doi: 10.1186/s12913-019-4302-z] [Medline: 31307436]
27. 3.10 Access to mental health services. Australian Institute of Health and Welfare. URL: https://www.indigenoushpf.gov.au/measures/3-10-access-to-mental-health-services [accessed 2026-01-16]
28. 3.14 Access to services compared with need. Australian Institute of Health and Welfare. URL: https://www.indigenoushpf.gov.au/measures/3-14-access-to-services-compared-with-need [accessed 2026-01-16]
29. Sveticic J, Milner A, De Leo D. Contacts with mental health services before suicide: a comparison of Indigenous with non-Indigenous Australians. Gen Hosp Psychiatry 2012;34(2):185-191. [doi: 10.1016/j.genhosppsych.2011.10.009] [Medline: 22154656]
30. Steering Committee for the Review of Government Service Provision. Report on government services 2020. : Productivity Commission; 2020 URL: https://www.pc.gov.au/ongoing/report-on-government-services/2020/data-downloads/ [accessed 2026-01-16]
31. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of technology-based strategies to promote engagement with digital interventions: a systematic review protocol. JMIR Res Protoc 2015 Apr 28;4(2):e47. [doi: 10.2196/resprot.3990] [Medline: 25921274]
32. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research... and three things we can do about them. Psychiatr Serv 2017 May 1;68(5):427-429. [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
33. Tassnim R, Tawfic Ja'far RA, Al-Tarawneh AMA. 2024 2nd International Conference on Cyber Resilience (ICCR): IEEE; 2024. [doi: 10.1109/ICCR61006.2024]
34. Mendes-Santos C, Nunes F, Weiderpass E, Santana R, Andersson G. Understanding mental health professionals' perspectives and practices regarding the implementation of digital mental health: qualitative study. JMIR Form Res 2022 Apr 12;6(4):e32558. [doi: 10.2196/32558] [Medline: 35412459]
35. Christensen H, Hickie IB. Using e-health applications to deliver new mental health services. Med J Aust 2010 Jun 7;192(S11):S53-S56. [doi: 10.5694/j.1326-5377.2010.tb03695.x] [Medline: 20528711]
36. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11. [doi: 10.1002/wps.20083] [Medline: 24497236]
37. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS ONE 2010 Oct 13;5(10):e13196. [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
38. Spence J, Titov N, Johnston L, Jones MP, Dear BF, Solley K. Internet-based trauma-focused cognitive behavioural therapy for PTSD with and without exposure components: a randomised controlled trial. J Affect Disord 2014 Jun;162:73-80. [doi: 10.1016/j.jad.2014.03.009] [Medline: 24767009]
39. Mental health services. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/mental-health/overview/mental-health-services [accessed 2024-08-12]

XSL•FO
RenderX

40. Bernard A, de Ossorno Garcia S, Salhi L, John A, DelPozo-Banos M. Patterns of engagement in a digital mental health service during COVID-19: a cohort study for children and young people. Front Psychiatry 2023;14:1143272. [doi: 10.3389/fpsyt.2023.1143272] [Medline: 37575580]
41. Ftanou M, Machlin A, Mangelsdorf SN, Morgan A, Bassilios B. Australian digital mental health services: consumer perceptions of their usability and acceptability. J Technol Behav Sci 2024;10(2):320-335. [doi: 10.1007/s41347-024-00431-9]
42. Cheng L, Chan WK, Peng Y, Qin H. Towards data-driven tele-medicine intelligence: community-based mental healthcare paradigm shift for smart aging amid COVID-19 pandemic. Health Inf Sci Syst 2023 Dec;11(1):14. [doi: 10.1007/s13755-022-00198-4] [Medline: 36923686]
43. Dawson AZ, Walker RJ, Campbell JA, Davidson TM, Egede LE. Telehealth and indigenous populations around the world: a systematic review on current modalities for physical and mental health. Mhealth 2020;6:30. [doi: 10.21037/mhealth.2019.12.03] [Medline: 32632368]
44. Goetz CJ, Mushquash CJ, Maranzan KA. An integrative review of barriers and facilitators associated with mental health help seeking among Indigenous populations. Psychiatr Serv 2023 Mar 1;74(3):272-281. [doi: 10.1176/appi.ps.202100503] [Medline: 36065579]
45. Hicks LJ, Toombs E, Lund J, Kowatch KR, Hopkins C, Mushquash CJ. Expanding our understanding of digital mental health interventions for Indigenous youth: an updated systematic review. J Telemed Telecare 2025 Aug;31(7):919-960. [doi: 10.1177/1357633X241239715] [Medline: 38584397]
46. Moecke DP, Holyk T, Beckett M, et al. Scoping review of telehealth use by Indigenous populations from Australia, Canada, New Zealand, and the United States. J Telemed Telecare 2024 Oct;30(9):1398-1416. [doi: 10.1177/1357633X231158835] [Medline: 36911983]
47. Caffery LJ, Bradford NK, Wickramasinghe SI, Hayman N, Smith AC. Outcomes of using telehealth for the provision of healthcare to Aboriginal and Torres Strait Islander people: a systematic review. Aust N Z J Public Health 2017 Feb;41(1):48-53. [doi: 10.1111/1753-6405.12600] [Medline: 27868300]
48. Choukou MA, Maddahi A, Polyvyana A, Monnin C. Digital health technology for Indigenous older adults: a scoping review. Int J Med Inform 2021 Apr;148:104408. [doi: 10.1016/j.ijmedinf.2021.104408] [Medline: 33609927]
49. Weiner MF, Rossetti HC, Harrah K. Videoconference diagnosis and management of Choctaw Indian dementia patients. Alzheimers Dement 2011 Nov;7(6):562-566. [doi: 10.1016/j.jalz.2011.02.006] [Medline: 22055972]
50. Fraser S, Mackean T, Grant J, Hunter K, Towers K, Ivers R. Use of telehealth for health care of Indigenous peoples with chronic conditions: a systematic review. Rural Remote Health 2017;17(3):4205. [doi: 10.22605/RRH4205] [Medline: 28930638]
51. Katapally TR. Smart Indigenous youth: the smart platform policy solution for systems integration to address Indigenous youth mental health. JMIR Pediatr Parent 2020 Sep 25;3(2):e21155. [doi: 10.2196/21155] [Medline: 32975527]
52. Fleming TM, Dixon RS, Merry SN. 'It's mean!' The views of young people alienated from mainstream education on depression, help seeking and computerised therapy. Advances in Mental Health 2012 Feb;10(2):195-203. [doi: 10.5172/jamh.2011.10.2.195]
53. Walker R, Usher K, Jackson D, et al. Connection to... Addressing digital inequities in supporting the well-being of young Indigenous Australians in the wake of COVID-19. Int J Environ Res Public Health 2021 Feb 22;18(4):2141. [doi: 10.3390/ijerph18042141] [Medline: 33671737]
54. Clair MS, Murtagh DP, Kelly J, Cook J. Telehealth a game changer: closing the gap in remote Aboriginal communities. Med J Aust 2019 Apr;210 Suppl 6(S6):S36-S37. [doi: 10.5694/mja2.50036] [Medline: 30927470]
55. Gibson K, O'Donnell S, Coulson H, Kakepetum-Schultz T. Mental health professionals' perspectives of telemental health with remote and rural First Nations communities. J Telemed Telecare 2011;17(5):263-267. [doi: 10.1258/jtt.2011.101011] [Medline: 21824967]
56. Newman L, Bidargaddi N, Schrader G. Service providers' experiences of using a telehealth network 12 months after digitisation of a large Australian rural mental health service. Int J Med Inform 2016 Oct;94:8-20. [doi: 10.1016/j.ijmedinf.2016.05.006] [Medline: 27573307]
57. Caffery LJ, Bradford NK, Smith AC, Langbecker D. How telehealth facilitates the provision of culturally appropriate healthcare for Indigenous Australians. J Telemed Telecare 2018 Dec;24(10):676-682. [doi: 10.1177/1357633X18795764] [Medline: 30343658]
58. Povey J, Raphiphatthana B, Torok M, et al. An emerging framework for digital mental health design with Indigenous young people: a scoping review of the involvement of Indigenous young people in the design and evaluation of digital mental health interventions. Syst Rev 2023 Jul 1;12(1):108. [doi: 10.1186/s13643-023-02262-w] [Medline: 37393283]
59. Moffatt JJ, Eley DS. The reported benefits of telehealth for rural Australians. Aust Health Rev 2010 Aug;34(3):276-281. [doi: 10.1071/AH09794] [Medline: 20797357]
60. Indigenous data sovereignty. : Maiam nayri Wingara Indigenous Data Sovereignty Collective and the Australian Indigenous Governance Institute; 2018 URL: https://www.maiamnayriwingara.org/mnw-principles [accessed 2026-01-16]
61. Liu M, Schueller SM. Moving evidence-based mental health interventions into practice: implementation of digital mental health interventions. Curr Treat Options Psych 2023;10(4):333-345. [doi: 10.1007/s40501-023-00298-2]

XSL•FO
RenderX

62. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8(1):45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
63. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
64. Harfield S, Pearson O, Morey K, et al. Assessing the quality of health research from an Indigenous perspective: the Aboriginal and Torres Strait Islander quality appraisal tool. BMC Med Res Methodol 2020 Apr 10;20(1):79. [doi: 10.1186/s12874-020-00959-3] [Medline: 32276606]
65. Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
66. Bennett-Levy J, Singer J, DuBois S, Hyde K. Translating e-mental health into practice: what are the barriers and enablers to e-mental health implementation by Aboriginal and Torres Strait Islander health professionals? J Med Internet Res 2017 Jan 11;19(1):e1. [doi: 10.2196/jmir.6269] [Medline: 28077347]
67. Brown K, Toombs M, Nasir B, et al. How can mobile applications support suicide prevention gatekeepers in Australian Indigenous communities? Soc Sci Med 2020 Aug;258:113015. [doi: 10.1016/j.socscimed.2020.113015] [Medline: 32559573]
68. Dingwall KM, Povey J, Sweet M, et al. Feasibility and acceptability of the Aboriginal and Islander mental health initiative for youth app: nonrandomized pilot with First Nations young people. JMIR Hum Factors 2023 Jun 7;10:e40111. [doi: 10.2196/40111] [Medline: 37285184]
69. Dingwall KM, Puszka S, Sweet M, Nagel T. "Like drawing into sand": acceptability, feasibility, and appropriateness of a new e‐mental health resource for service providers working with Aboriginal and Torres Strait Islander people. Aust Psychol 2015 Feb 1;50(1):60-69. [doi: 10.1111/ap.12100]
70. Fletcher R, Hammond C, Faulkner D, et al. Stayin' on Track: the feasibility of developing Internet and mobile phone-based resources to support young Aboriginal fathers. Aust J Prim Health 2017 Sep;23(4):329-334. [doi: 10.1071/PY16151] [Medline: 28449728]
71. Kennedy M, Kumar R, Ryan NM, Bennett J, La Hera Fuentes G, Gould GS. Codeveloping a multibehavioural mobile phone app to enhance social and emotional well-being and reduce health risks among Aboriginal and Torres Strait Islander women during preconception and pregnancy: a three-phased mixed-methods study. BMJ Open 2021 Nov 24;11(11):e052545. [doi: 10.1136/bmjopen-2021-052545] [Medline: 34819285]
72. Lee KSK, Conigrave JH, Callinan S, et al. Asking about the last four drinking occasions on a tablet computer as a way to record alcohol consumption in Aboriginal and Torres Strait Islander Australians: a validation. Addict Sci Clin Pract 2019 May 1;14(1):15. [doi: 10.1186/s13722-019-0148-2] [Medline: 31039824]
73. Ma JS, Batterham PJ, Kõlves K, et al. Community expectations and anticipated outcomes for crisis support services—Lifeline Australia. Health Social Care Comm 2022 Sep;30(5):1775-1788. [doi: 10.1111/hsc.13557]
74. Perdacher E, Kavanagh D, Sheffield J, Healy K, Dale P, Heffernan E. Using the stay strong app for the well-being of Indigenous Australian prisoners: feasibility study. JMIR Form Res 2022 Apr 8;6(4):e32157. [doi: 10.2196/32157] [Medline: 35394444]
75. Povey J, Mills P, Dingwall KM, et al. Acceptability of mental health apps for Aboriginal and Torres Strait Islander Australians: a qualitative study. J Med Internet Res 2016 Mar 11;18(3):e65. [doi: 10.2196/jmir.5314] [Medline: 26969043]
76. Povey J, Sweet M, Nagel T, et al. Determining priorities in the Aboriginal and Islander mental health initiative for youth app second phase participatory design project: qualitative study and narrative literature review. JMIR Form Res 2022 Feb 18;6(2):e28342. [doi: 10.2196/28342] [Medline: 35179498]
77. Povey J, Sweet M, Nagel T, et al. Drafting the Aboriginal and Islander mental health initiative for youth (AIMhi-Y) app: results of a formative mixed methods study. Internet Interv 2020 Sep;21:100318. [doi: 10.1016/j.invent.2020.100318] [Medline: 32477884]
78. Puszka S, Dingwall KM, Sweet M, Nagel T. E-mental health innovations for Aboriginal and Torres Strait Islander Australians: a qualitative study of implementation needs in health services. JMIR Ment Health 2016 Sep 19;3(3):e43. [doi: 10.2196/mental.5837] [Medline: 27644259]
79. Raphiphatthana B, Sweet M, Puszka S, Dingwall K, Nagel T. Evaluation of a three-phase implementation program in enhancing e-mental health adoption within Indigenous primary healthcare organisations. BMC Health Serv Res 2020 Jun 23;20(1):576. [doi: 10.1186/s12913-020-05431-y] [Medline: 32576266]
80. Raphiphatthana B, Sweet M, Puszka S, Whitty M, Dingwall K, Nagel T. Evaluation of electronic mental health implementation in Northern Territory Services using the integrated "promoting action on research implementation in health services" framework: qualitative study. JMIR Ment Health 2020 May 26;7(5):e14835. [doi: 10.2196/14835] [Medline: 32452812]
81. Routledge K, Snijder M, Newton N, et al. Acceptability and feasibility of Strong & Deadly Futures, a culturally-inclusive alcohol and drug prevention program for Aboriginal and/or Torres Strait Islander secondary students. SSM - Mental Health 2022 Dec;2:100073. [doi: 10.1016/j.ssmmh.2022.100073]
82. Russell S, Quigley R, Strivens E, et al. Validation of the Kimberley Indigenous Cognitive Assessment short form (KICA-screen) for telehealth. J Telemed Telecare 2021 Jan;27(1):54-58. [doi: 10.1177/1357633X19860309] [Medline: 31284810]

XSL•FO
RenderX

83. Snijder M, Stapinski L, Ward J, et al. Strong and deadly futures: co-development of a web-based wellbeing and substance use prevention program for Aboriginal and Torres Strait Islander and non-Aboriginal adolescents. Int J Environ Res Public Health 2021 Feb 23;18(4):2176. [doi: 10.3390/ijerph18042176] [Medline: 33672190]
84. Tighe J, Shand F, McKay K, Mcalister TJ, Mackinnon A, Christensen H. Usage and acceptability of the iBobbly app: pilot trial for suicide prevention in Aboriginal and Torres Strait Islander youth. JMIR Ment Health 2020 Dec 1;7(12):e14296. [doi: 10.2196/14296] [Medline: 33258782]
85. Tighe J, Shand F, Ridani R, Mackinnon A, De La Mata N, Christensen H. Ibobbly mobile health intervention for suicide prevention in Australian Indigenous youth: a pilot randomised controlled trial. BMJ Open 2017 Jan 27;7(1):e013518. [doi: 10.1136/bmjopen-2016-013518] [Medline: 28132007]
86. Titov N, Schofield C, Staples L, Dear BF, Nielssen O. A comparison of Indigenous and non-Indigenous users of MindSpot: an Australian digital mental health service. Australas Psychiatry 2019 Aug;27(4):352-357. [doi: 10.1177/1039856218789784] [Medline: 30058351]
87. Veinovic M, Hill TY, Lavrencic L, et al. Telephone cognitive screening with older Aboriginal Australians: a preliminary study. Australas J Ageing 2023 Jun;42(2):311-316. [doi: 10.1111/ajag.13147] [Medline: 36330778]
88. Australian Institute of Health and Welfare. Digital Mental Health Resources for First Nations People: AIHW, Australian Government; 2023. URL: https://www.aihw.gov.au/getmedia/28f92ca3-ee7f-478a-828f-cfa2e00916e7/aihw-imh-20-digital-mental-health-resources-for-first-nations-people.pdf [accessed 2026-01-16]
89. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
90. Li J, Brar A. The use and impact of digital technologies for and on the mental health and wellbeing of Indigenous people: a systematic review of empirical studies. Comput Human Behav 2022 Jan;126:106988. [doi: 10.1016/j.chb.2021.106988]
91. Hensel JM, Ellard K, Koltek M, Wilson G, Sareen J. Digital health solutions for Indigenous mental well-being. Curr Psychiatry Rep 2019 Jul 1;21(8):68. [doi: 10.1007/s11920-019-1056-6] [Medline: 31263971]
92. Balaratnasingam S, Chapman M, Chong D, et al. Advancing social and emotional well-being in Aboriginal and Torres Strait Islander Australians: clinicians' reflections. Australas Psychiatry 2019 Aug;27(4):348-351. [doi: 10.1177/1039856218789765] [Medline: 30234370]
93. Campbell ANC, Rieckmann T, Pavlicova M, et al. Culturally tailored digital therapeutic for substance use disorders with urban Indigenous people in the United States: a randomized controlled study. J Subst Use Addict Treat 2023 Dec;155:209159. [doi: 10.1016/j.josat.2023.209159] [Medline: 37690525]
94. Chen Q, Zhao Z, Bao J, Lin J, Li W, Zang Y. Digital empowerment in mental health: a meta-analysis of internet-based interventions for enhancing mental health literacy. Int J Clin Health Psychol 2024;24(3):100489. [doi: 10.1016/j.ijchp.2024.100489] [Medline: 39220119]
95. Sockalingam S, Arena A, Serhal E, Mohri L, Alloo J, Crawford A. Building provincial mental health capacity in primary care: an evaluation of a project ECHO mental health program. Acad Psychiatry 2018 Aug;42(4):451-457. [doi: 10.1007/s40596-017-0735-z] [Medline: 28593537]
96. Goodyear-Smith F, Martel R, Darragh M, Warren J, Thabrew H, Clark TC. Screening for risky behaviour and mental health in young people: the YouthCHAT programme. Public Health Rev 2017;38(1):20. [doi: 10.1186/s40985-017-0068-1] [Medline: 29450092]
97. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629. [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
98. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146. [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
99. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006 Jun 23;8(2):e10. [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
100. Whitelaw S, Mamas MA, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. Lancet Digit Health 2020 Aug;2(8):e435-e440. [doi: 10.1016/S2589-7500(20)30142-4] [Medline: 32835201]
101. Black C, Frederico M, Bamblett M. "Healing through culture": Aboriginal young people's experiences of social and emotional wellbeing impacts of cultural strengthening programs. Child Abuse Negl 2024 Feb;148:106206. [doi: 10.1016/j.chiabu.2023.106206] [Medline: 37183121]
102. Alon N, Perret S, Cohen A, et al. Digital navigator training to increase access to mental health care in community-based organizations. PS (Wash DC) 2024 Jun 1;75(6):608-611. [doi: 10.1176/appi.ps.20230391]
103. Dunn P, Hazzard E. Technology approaches to digital health literacy. Int J Cardiol 2019 Oct 15;293:294-296. [doi: 10.1016/j.ijcard.2019.06.039] [Medline: 31350037]

XSL•FO
RenderX

104. St Clair M, Murtagh D. Barriers to telehealth uptake in rural, regional, remote Australia: what can be done to expand telehealth access in remote areas? Stud Health Technol Inform 2019 Aug 8;266:174-182. [doi: 10.3233/SHTI190791] [Medline: 31397320]
105. Featherstone D, Ormond-Parker L, Ganley L, et al. Mapping the digital gap: 2023 outcomes report. : ARC Centre of Excellence for Automated Decision-Making and Society; 2023 URL: https://apo.org.au/node/324397 [accessed 2026-01-16]
106. Mob Link. Institute for Urban Indigenous Health. URL: https://www.iuih.org.au/our-services/mob-link [accessed 2025-04-07]
107. Zhou X, Smith AC, Nelson C, et al. Disparities in access to and utilisation of Medicare subsidised telehealth mental health services by Aboriginal and Torres Strait Islander Australians in Southeast Queensland. PsyArXiv. Preprint posted online on May 19, 2025. [doi: 10.31234/osf.io/kjs42_v1]
108. Sinclair S, Waruk J, Morrisieau T, Cochrane C, Saulog A. Our data, our truths, our voice: reclaiming First Nations data sovereignty in Manitoba. IJPDS 2024;9(5). [doi: 10.23889/ijpds.v9i5.2731]

Abbreviations

AIMhi: Aboriginal and Islander Mental Health Initiative
DMH: digital mental health
KICA-screen: Kimberly Indigenous Cognitive Examination-short form
MMAT: Mixed Methods Appraisal Tool
MMSE: Mini-Mental State Examination
PICO: population, intervention, comparison, outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TKICA-screen: telehealth version of the Kimberly Indigenous Cognitive Examination-short form
TMMSE: telehealth version of the Mini-Mental State Examination

Edited by N Cahill; submitted 09.Jul.2025; peer-reviewed by O Mbanefo, P Usuemerai; accepted 17.Dec.2025; published 27.Jan.2026.

Please cite as:
Zhai S, Goodman A, Smith AC, Diminic S, Zhou X
The Effectiveness, Facilitators, and Barriers of Digital Mental Health Services for First Nations People in Australia: Systematic Scoping Review
Interact J Med Res 2026;15:e80386
URL: https://www.i-jmr.org/2026/1/e80386
doi:10.2196/80386

© Siyu Zhai, Andrew Goodman, Anthony C Smith, Sandra Diminic, Xiaoyun Zhou. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 27.Jan.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Usability of Continuous Monitoring Devices With Deterioration Alerting Systems in Noncritical Care Units: Scoping Review

Jo-Fan Pan[1], PhD; Dawn Dowding[1], PhD; David Wong[2], PhD; Ashley Scott[1], MSc; Qimeng Zhao[1], MSc

[1]University of Manchester, Jean McFarlane Building, Oxford Road, Manchester, United Kingdom

[2]University of Leeds, Leeds, United Kingdom

Corresponding Author:
Jo-Fan Pan, PhD
University of Manchester, Jean McFarlane Building, Oxford Road, Manchester, United Kingdom

Abstract

Background: Delayed recognition of patient deterioration in a non–intensive care unit (ICU) setting contributes to serious adverse events. Continuous monitoring devices with alerting systems offer real-time data to support early detection, but their effectiveness depends on usability. While prior reviews focus on clinical outcomes, usability—defined by effectiveness, efficiency, and satisfaction—remains underexplored.

Objective: This study aims to scope the evidence related to the usability of continuous monitoring devices with deterioration alerting in noncritical adult care units.

Methods: A scoping review was conducted following the Joanna Briggs Institute methodology and reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. A comprehensive search of MEDLINE, Embase, Emcare, Web of Science, and IEEE Xplore was performed for studies published up to November 2024. Title and abstract screening, full-text review, and data extraction were independently conducted by 2 reviewers. Studies were included if they (1) evaluated the usability—defined as effectiveness, efficiency, or satisfaction—of continuous monitoring devices; (2) focused on adult patients in non-ICU hospital settings; (3) used primary data; (4) were published in English; and (5) described how clinicians received alerts.

Results: The search identified 1284 papers, with 35 included. Most studies focused on postoperative patients in surgical wards, mainly from the United States and the Netherlands. Only 2 studies used mixed methods, and 10 reported clinician characteristics. While effectiveness (71%) and efficiency (74%) were widely studied, satisfaction (46%) and usability barriers (29%) received less attention.

Conclusions: Continuous monitoring devices with deterioration alerts may reduce rapid response team calls and ICU transfers, save time, and maintain acceptable alarm frequencies with high user satisfaction. However, usability challenges persist, including technical issues, alarm fatigue, patient discomfort, and limited training or workflow integration. This review mapped current use, usability, and barriers, categorized key usability factors for improvement, and identified the need for further research on clinician perspectives and broader health care settings to enhance generalizability.

(Interact J Med Res 2026;15:e75713) doi:10.2196/75713

KEYWORDS

continuous monitoring; monitoring; wearable devices; early warning score; vital sign; alerts; alarms; deterioration

Introduction

In 2022 - 2023, the United Kingdom recorded 16.4 million hospital admissions, compared to 33.7 million in the United States and 12.1 million in Australia [1-3]. Many of these patients may experience severe adverse events (SAEs), such as cardiac arrests or intensive care unit (ICU) admissions, during their hospital stay, with estimates ranging from 5% to 10% [4,5]. According to UK data, 23% of in-hospital SAEs resulting in sudden death are due to failures to recognize or respond to patient deterioration [6]. This issue is particularly concerning in non-ICU environments, where higher nurse-to-patient ratios require nurses to divide their attention among multiple patients, making timely identification of patient deterioration more challenging [7]. As van Galen et al [8] reported, 46% of 49 unplanned ICU transfers were linked to insufficient monitoring by nurses, highlighting the urgent need for enhanced monitoring strategies to support timely identification of at-risk patients in general wards.

The early warning score (EWS) is a widely used tool for early detection of patient deterioration in non-ICU settings [9]. It assigns weighted scores to vital signs (eg, respiratory rate, oxygen saturation, blood pressure, heart rate, temperature, and consciousness) to assess severity and guide interventions [9].

XSL•FO
RenderX

Studies have shown that the EWS can effectively identify patient deterioration, potentially reducing mortality rates in general wards [10,11]. The systematic review by Smith et al [11] of 21 retrospective studies in academic hospitals, primarily involving patients from medical and surgical wards, found that applying EWS was associated with a reduction in hospital mortality rates from nearly 50% to 40% and reduced cardiac arrest rates from 50% to 35%. However, studies also highlight that the effectiveness of the EWS can be compromised by inaccuracies in recorded data and its inability to capture immediate, real-time changes in vital signs [12,13]. In many general wards, nurses measure vital signs infrequently due to high patient-to-nurse ratios (especially compared with ICU settings) and intense workloads, which leave significant gaps during which critical signs of patient deterioration may be missed [12,13]. Incomplete and delayed data are associated with 19% of EWS inaccuracies in the United Kingdom and untimely responses in 75% of high-risk cases [14].

To help nurses improve the accuracy and timeliness of patient monitoring, real-time continuous monitoring systems with deterioration alerts have been proposed. Using wearable sensors or bedside monitors, these systems provide continuous data on key physiological parameters and generate alerts when out-of-range values are detected [15,16]. Multiple reviews have examined the effectiveness of these systems in non-ICU settings. Two reviews highlight positive outcomes: Downey et al [15] reviewed 24 studies involving over 40,000 patients and reported benefits such as reduced ICU transfers, shorter hospital stays, and significant cost savings. Similarly, Sun et al [16], in their meta-analysis of 14 studies, found that patients under continuous monitoring had a 39% lower risk of mortality compared to those with intermittent monitoring. However, other reviews have not found that continuous monitoring improves patient outcomes. Cardona-Morell et al [17] conducted a systematic review and meta-analysis of 22 clinical studies, involving sample sizes ranging from 16 to 64,661, and found no significant differences in the proportion of high-risk patients requiring urgent attention, preventing serious adverse events, cardiac arrests, or reducing ICU transfers. Similarly, the meta-analysis by Areia et al [18] reported potential reductions in ICU admissions and complications but found no differences compared to standard monitoring.

The differences in results across these reviews may be attributed to several factors. First, the design and functionality of continuous monitoring devices and their deterioration alerting systems vary, leading to differences in how alerts are generated and how clinicians interact with the devices. For example, Areia et al [18] focused solely on wearable devices, while Cardona-Morrell et al [17] included both wearable devices and bedside monitors. Second, the sample size and composition differed across studies. Downey et al [15] included 24 studies covering 40,274 patients and 59 ward staff across 9 countries, whereas Sun et al [16] analyzed 14 studies involving only 14,880 patients without including staff perspectives, resulting in a smaller and more limited participant base. Finally, these reviews primarily focused on device effectiveness, with limited attention to factors such as patient satisfaction and clinicians' adoption preferences, while all emphasized the need for future research on usability factors to enhance the devices' adoption and overall effectiveness [15-18]. Collectively, these findings highlight the limited focus on usability in current reviews, underscoring the need for further investigation into usability barriers and their impact on device usage.

Usability is a key factor in the adoption and successful use of health care technology, as it directly influences whether systems can achieve their intended goals, such as improving patient outcomes. Notably, these devices do not function in isolation. While they cannot make decisions, they provide essential information that shapes clinical judgment. Therefore, careful consideration of how these devices are integrated into clinical practice is critical. In this sense, the device's usability reflects both its inherent attributes (eg, battery life, sensor design, and interface clarity) and its application within the broader health care context, where clinicians interpret data, make treatment decisions, and adapt workflows accordingly. Poor usability, such as unclear interfaces, false alarms, or complex workflows, can lead to clinician disengagement, reducing trust and limiting the system's impact on clinical decisions and care processes [19].

According to the International Organization for Standardization (ISO), usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [20]. ISO Standard 9241-11:2018 further breaks usability into 3 dimensions: effectiveness, which refers to the accuracy and completeness with which users achieve their goals; efficiency, the resources required to achieve these goals; and satisfaction, the comfort and acceptability of the system to its users [20]. As highlighted in the previous paragraph, studies on continuous monitoring and alerting systems in non-ICU settings often emphasize effectiveness while neglecting other critical factors, such as satisfaction and efficiency [15-18]. Therefore, this review aimed to map existing evidence on the usability of continuous monitoring devices with deterioration alerting systems.

Methods

The review followed the Joanna Briggs Institute methodology for scoping reviews. A scoping review was chosen because it enables mapping of heterogeneous evidence across varied definitions of "usability," ensuring that all aspects of the concept are covered. The review also adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines for transparent reporting.

The objectives of the review were framed using a Population, Concept, and Context framework as follows:

- Participants: the review includes studies involving clinicians working in non-ICU adult care settings, focusing on those who interact directly with continuous monitoring devices with deterioration alerting capabilities.
- Concept: the concept evaluated is the usability (effectiveness, efficiency, and satisfaction) of continuous monitoring devices that monitor vital signs and trigger

XSL•FO
RenderX

alerts. These devices must demonstrate how they improve clinical outcomes, enhance workflow efficiency, and increase user satisfaction.
- Context: the context includes hospital settings where these devices are applied in real-world non-ICU scenarios. Studies must describe how clinicians receive alerts and interact with these systems to ensure a comprehensive usability assessment.

No published protocol or registration was in place before this study.

Search

The search was conducted in published electronic databases up to November 13, 2024, including MEDLINE, Embase, Evidence-Based Medicine Reviews, Web of Science, and IEEE Xplore databases. Gray literature was searched through Web of Science.

The search strategy, incorporating the Population, Concept, and Context framework and MeSH (Medical Subject Headings) terms, was refined using search keywords from previous reviews on continuous monitoring devices with deterioration alerting systems' capabilities [15-18,21].

A sample search strategy for MEDLINE was as follows:

("adult care unit*" OR "general ward*" OR hospitalization OR inpatient) AND (continuous* OR "real-time" OR remote* OR wearable* OR sensor* OR monitor*) AND ("early warning score*" OR "track and trigger" OR "deterioration alert*" OR "deterioration alarm*" OR "deterioration warn*")

The final search terms and database results are provided in Multimedia Appendix 1.

Eligibility Criteria

The review included studies that evaluated the usability (of continuous monitoring devices with deterioration alerting, operated by clinicians in non-ICU hospital settings. Usability, as defined by ISO 9241-11:2018, consists of 3 key dimensions:

- Effectiveness: the accuracy and completeness with which users achieve their goals (eg, impact on ICU transfers, hospital stays, and patient outcomes).
- Efficiency: the resources required to achieve these goals (eg, alarm burden, false alert rates, and nursing workload).
- Satisfaction: the comfort and acceptability of the system to its users (eg, clinician and patient perceptions of comfort and acceptability).

The review focused on wearable sensors and bedside monitors that automatically track vital signs—including heart rate, blood pressure, respiratory rate, oxygen saturation, and temperature—at intervals of no more than 15 minutes. Studies were included if they described how clinicians received alerts, as this helps assess the impact of alerts and devices on clinical practice. Exclusion criteria omitted studies conducted in ICU settings; those in which clinicians—especially nurses, who are primarily responsible for patient monitoring—were blinded to alerts or not included in the research; studies involving participants under the age of 18 years; non-English publications; and review or protocol papers. The full criteria are provided in Textbox 1.

To ensure consistency in applying the inclusion and exclusion criteria during the study selection process, a dual-reviewer approach was used. The second reviewer independently evaluated 33% of studies at each review stage, including title/abstract screening and full-text review. If discrepancies exceeded 25%, the criteria were revised, and the screening process was repeated until the discrepancy rate fell below this threshold. Conflicts of less than 25% were resolved through consensus or consultation with a third reviewer. Additionally, reference lists of included studies were examined to identify further relevant studies meeting the inclusion criteria.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Studies evaluating the usability (effectiveness, efficiency, and satisfaction) of continuous monitoring devices with deterioration alerting in hospital settings.
- Studies demonstrating the use of continuous monitoring devices with deterioration alerting in patient monitoring: alerts must be triggered by monitoring one or more vital signs such as heart rate, blood pressure, respiratory rate, oxygen saturation, or temperature.
- Studies that describe how clinicians receive alerts.

Exclusion criteria:

- Studies not published in English.
- Studies that did not apply continuous monitoring devices with deterioration alerting in non–intensive care unit (ICU) hospital settings or failed to clearly segregate data from non-ICU and ICU settings.
- Studies in which clinicians do not receive or acknowledge alerts: exclude studies where clinicians are kept unaware of alerts from monitoring devices (ie, clinicians were "blinded" to alerts, or the device did not generate alerts, unless clinicians had another way to receive the alerts, such as from ringing notifications or a central station).
- Studies stating that they involved participants younger than 18 years.
- Review or protocol papers: systematic reviews, meta-analyses, narrative reviews, and research protocols are excluded to focus solely on original research and first-hand data regarding device usability.

XSL•FO
RenderX

Data Extraction

A structured data extraction form (Multimedia Appendix 2) was used to collect details on citations (author and year), study characteristics (participants, setting, design, methods, and country), device features (type, function, alerts, and application), and usability aspects (effectiveness, efficiency, satisfaction, and barriers).

First, a pilot data extraction phase involved both reviewers independently reviewing 2 included studies. Second, after independent extraction, both reviewers convened to discuss their findings. Discrepancies were resolved through discussion, and any unresolved issues were referred to a third reviewer. Third, following alignment on the pilot studies, the first reviewer extracted data from the remaining studies. The second reviewer then reviewed these extractions to verify consistency with the initial agreement and the standard definitions discussed. Remaining conflicts were resolved through consensus or consultation with a third reviewer.

Synthesis

Studies were summarized and categorized based on key attributes, including publication year, country, study design, and research aim. First, these categorizations mapped the current research landscape and identified the distribution of studies across various contexts. For example, the percentage of studies conducted in different wards was reported to highlight where these devices are most frequently used. Second, the extracted data were categorized by common characteristics, such as device type and alert type. Finally, a narrative synthesis was conducted, thematically categorizing findings to align with the research question and objectives. Studies were grouped by their focus on usability (effectiveness, efficiency, and satisfaction) and barriers to usability, summarizing their findings, such as the accuracy of devices in detecting patient deterioration, categorized under effectiveness.

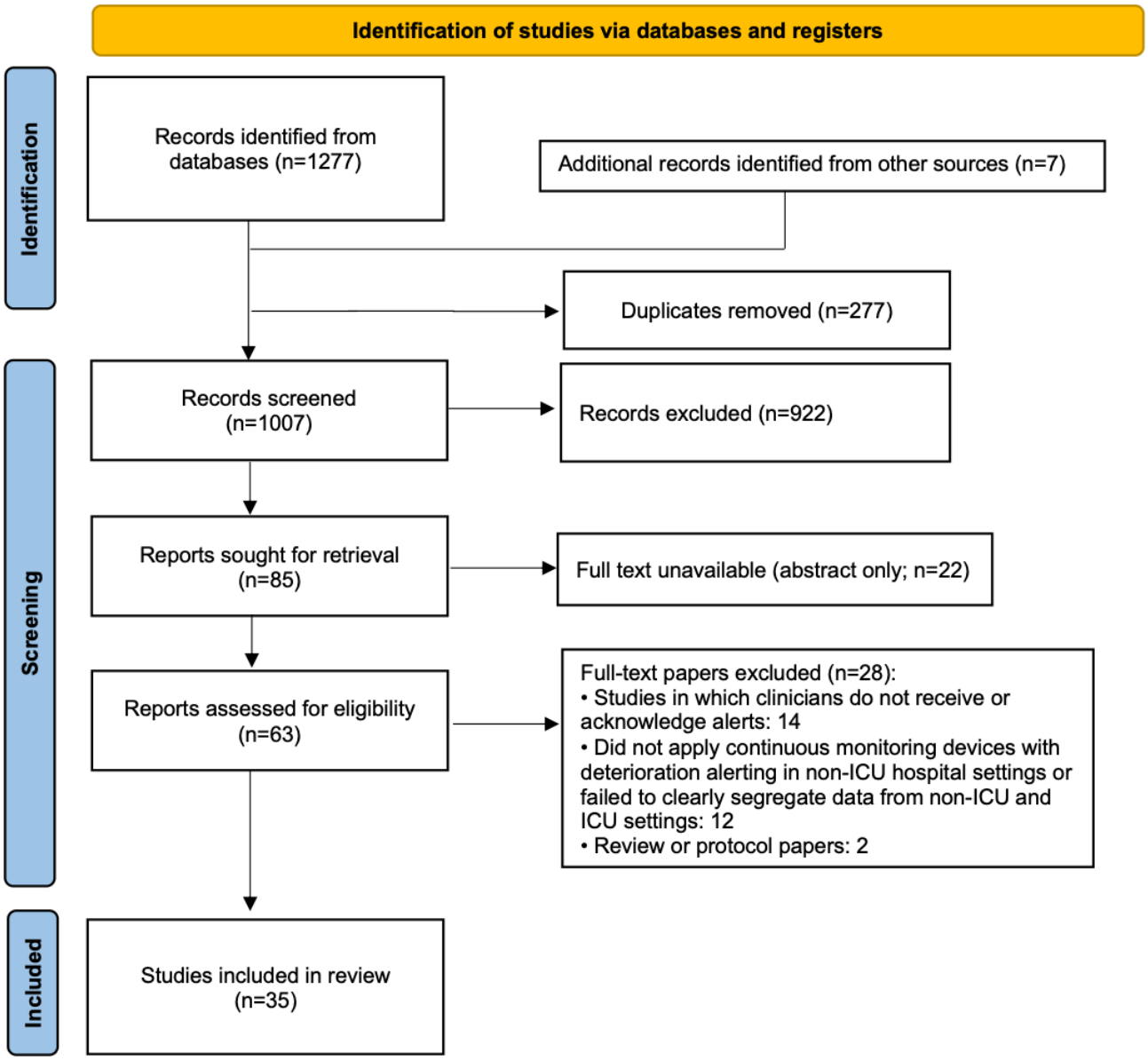
Results

Search Outcome

The search identified 1284 paper citations. After excluding duplicate records, 1007 records were deemed eligible for screening. A total of 63 studies were selected based on abstracts and underwent full-text review. After applying the inclusion and exclusion criteria, 35 studies were selected for this review. Figure 1 provides the study selection process in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (the PRISMA checklist is provided in Checklist 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Overview of Included Studies

An overview of the included studies, including study design, study devices, interventions, comparison groups, and outcomes measured, is summarized in Table 1. Overall, the 35 studies reported outcomes for a total of 65,029 patients and 323 clinicians in 8 countries. Multimedia Appendix 3 provides a summary of the included studies.

XSL•FO
RenderX

Table . Summary of the included studies.

Author	Study aim type	Study design	Sample size	Alert mechanism type	Device type
Becking-Verhaar et al [22]	Implementation and feasibility of continuous monitoring systems	Cross-sectional survey	58 nurses	Threshold alert	Wearable devices
Bellomo et al [23]	Impact on clinical outcomes and patient safety	Observational study	18,305 patients (9617 before intervention and 8688 after intervention)	EWS[a] base alerts	Bedside monitors
Blankush et al [24]	Technological evaluation and alarm strategies	Observational study	133 patients	EWS base alerts	Bedside monitors
Brown et al [25]	Impact on clinical outcomes and patient safety	Randomized controlled trial	Baseline cohort: 1535 (control) and 1433 (intervention); postimplementation cohort: 2361 (control) and 2314 (intervention)	Threshold alerts	Bedside monitors
Downey et al [26]	Nurses' and patients' perspectives and experiences	Randomized controlled trial	226 patients randomized (140 to continuous monitoring and 86 to intermittent monitoring)	Threshold alerts	Wearable devices
Downey et al [27]	Comparison with episodic monitoring	Observational study	12 patients	Threshold alerts	Wearable devices
Downey et al [28]	Implementation and feasibility of continuous monitoring systems	Randomized controlled trial	136 patients	Threshold alerts	Wearable devices
Eddahchouri et al [29]	Comparison with episodic monitoring	Observational study	Baseline cohort: 2466 admissions and intervention cohort: 2303 admissions	Threshold alerts	Wearable devices
Gazarian [30]	Nurses' and patients' perspectives and experiences	Prospective, descriptive, observational study	57 patients observed, 37 on continuous ECG[b] monitoring; 9 nurses	Threshold alerts	Wearable devices
Hravnak et al [31]	Technological evaluation and alarm strategies	Observational study	629 patients (323 in phase I and 306 in phase III)	AI-based[c] alerts	Bedside monitors
Hravnak et al [32]	Technological evaluation and alarm strategies	Observational study	326 patients	AI-based alerts	Bedside monitors
Joshi et al [33]	Technological evaluation and alarm strategies	Observational study	50 patients	Threshold alerts	Wearable devices
Klumpner et al [34]	Technological evaluation and alarm strategies	Observational study	64 monitored rooms	EWS base alerts	Bedside monitor + wearable devices
Kuznetsova et al [35]	Implementation and feasibility of continuous monitoring systems	Observational study	35 (preimplementation: 13 and postimplementation: 22) clinicians	Threshold alerts	Bedside monitors
Leenen et al [36]	Implementation and feasibility of continuous monitoring systems	Observational study	12 nurses	Threshold alerts	Wearable devices
Leenen et al [37]	Implementation and feasibility of continuous monitoring systems	Observational study	30 patients and 23 nurses	Threshold alerts	Wearable devices

XSL•FO
RenderX

Author	Study aim type	Study design	Sample size	Alert mechanism type	Device type
McGrath et al [38]	Impact on clinical outcomes and patient safety	Observational study	2 surgical units and 71 total beds	Threshold alerts	Wearable devices
McGrath et al [39]	Impact on clinical outcomes and patient safety	Observational study	Preimplementation: 4324 patient days and postimplementation: 4382 patient days	Threshold alerts	Wearable devices
McGrath et al [40]	Technological evaluation and alarm strategies	Observational study	General care units, including a 36-bed orthopedics unit and other surgical and medicine units; the exact number of participants was not specified, but the system covered more than 200 inpatient beds.	Threshold alerts	Wearable devices
Mestrom et al [41]	Impact on clinical outcomes and patient safety	Observational study	Control group: 320 patients and intervention group: 274 patients	EWS base alerts	Bedside monitor
Paul et al [42]	Impact on clinical outcomes and patient safety	Randomized controlled trial	Control group: 126 patients and intervention group: 124 patients	Threshold alerts	Wearable devices
Peelen et al [43]	Nurses' and patients' perspectives and experiences	Observational study	1529 patients	AI-based alerts	Wearable devices
Pollack et al [44]	Implementation and feasibility of continuous monitoring systems	Observational study	298 patients	Threshold alerts	Wearable devices
Posthuma et al [45]	Implementation and feasibility of continuous monitoring systems	Observational study	742 patients (515 intermittent monitoring and 227 continuous monitoring)	Threshold alerts	Wearable devices
Sigvardt et al [46]	Nurses' and patients' perspectives and experiences	Observational study	20 patients	Threshold alerts	Wearable devices
Stellpflug et al [47]	Impact on clinical outcomes and patient safety	Observational before-and-after study	547 patients during the intervention period and 27 nurses	Threshold alerts	Wearable devices
Subbe et al [48]	Impact on clinical outcomes and patient safety	Observational before-and-after study	Control: 2139 patients and intervention: 2263 patients	Threshold alert	Bedside monitors + wearable devices
Taenzer et al [49]	Impact on clinical outcomes and patient safety	Observational before-and-after study	Preimplementation: 3118 discharges (intervention unit), 1260 (comparison unit 1), 2628 (comparison unit 2); postimplementation: 2841 discharges (intervention unit), 1162 (comparison unit 1), 2389 (comparison unit 2); 60 nurses	Threshold alert	Wearable devices
Un et al [50]	Implementation and feasibility of continuous monitoring systems	Observational study	34 patients	AI-based alerts	Wearable devices
van Goor et al [51]	Implementation and feasibility of continuous monitoring systems	Observational before-and-after study	209 patients (93 intermittent monitoring and 121 continuous monitoring)	Threshold alert	Wearable devices

XSL•FO
RenderX

Author	Study aim type	Study design	Sample size	Alert mechanism type	Device type
van Rossum et al [52]	Comparison with episodic monitoring	Observational retrospective study	39 patients	Threshold alerts	Wearable devices
Verrillo et al [53]	Comparison with episodic monitoring	Observational study	Preintervention: 427 patients and intervention: 422 patients	Threshold alerts	Wearable devices
Watkins et al [54]	Impact on clinical outcomes and patient safety	Prospective observational study	236 patients and 24 nurses	AI-based alerts	Wearable devices
Weenk et al [55]	Comparison with episodic monitoring	Randomized controlled trial	60 patients	EWS-based alerts	Wearable devices
Weller et al [56]	Impact on clinical outcomes and patient safety	Prospective, observational study	736 patients, 23 nurses, and 20 nursing assistants	Threshold alert	Wearable devices

[a]EWS: early warning score.

[b]ECG: electrocardiogram.

[c]AI: artificial intelligence.

Characteristics of Included Studies

The overview of the studies' characteristics is summarized in Table 2. The aim of the included studies was categorized into 5 main subjects: implementation and feasibility of continuous monitoring systems (n=9, 24%) [22,28,35-37,44,45,50,51], comparison with episodic monitoring (n=5, 14%) [27,29,52,53,55], impact on clinical outcomes and patient safety (n=11, 31%) [23,25,38,39,41,42,47-49,54,56], nurses' and patients' perspectives and experiences (n=4, 11%) [26,30,43,46], and technological evaluation and alarm strategies (n=6, 17%) [24,31-34,40]. The included studies were predominantly from 3 countries: the United States (n=16, 46%) [24,25,30-32,34,35,38-40,44,47,49,53,54,56], the Netherlands (n=10, 29%) [22,29,36,37,41,43,45,51,52,55], and the United Kingdom (n=5, 14%) [26-28,33,48]. The most common study design was observational (n=28, 80%) [23,24,27,29-41,43,44,46-54,56]. Additionally, 20 of the included studies (57%) were categorized as nonrandomized trials [23-25,29,31-33,38,39,41,43,44,47-52,54,56]. Mixed-method studies were the least frequently used, with 2 studies (6%) [22,40]. Seventy-four percent of the included studies had comparison groups (n=26), with before-and-after implementation comparisons (n=11, 31%) [23-25,31,32,35,39,40,49,51,56] and comparisons with intermittent monitoring (n=12, 34%) [26,28,29,34,38,41-43,45,46,48,53] being most commonly used. Sixty-three percent of studies used methods combining electronic health records (EHRs), observations, clinical trials, and experiments (n=22) [23-25,29-34,39,41-43,45,46,48-53,56]. The studies most commonly involved surgical-related wards (n=18, 51%) [22,24-27,29,30,35,36,38-40,42,43,45,49,52,54] and postoperative patients (n=17, 49%) [26-30,33,37-39,41-43,45,49,52-54].

XSL•FO
RenderX

Table . Overview of characteristics of included studies.

Categories	Studies, n (%)	References
Study aim (n=35)		
Implementation and feasibility of continuous	9 (26)	[22,28,35-37,44,45,50,51]
Comparison with episodic monitoring	5 (14)	[27,29,52,53,55]
Impact on clinical outcomes and patient safety	11 (31)	[23,25,38,39,41,42,47-49,54,56]
Nurses' and patients' perspectives and experiences	4 (11)	[26,30,43,46]
Technological evaluation and alarm strategies	6 (17)	[24,31-34,40]
Country (n=35)		
United States	16 (46)	[24,25,30-32,34,35,38-40,44,47,49,53,54,56]
Netherlands	9 (29)	[22,29,36,37,41,43,45,51,52,55]
United Kingdom	5 (14)	[26-28,33,48]
United States, Europe, and Australia	1 (3)	[23]
Others	3 (9)	[42,46,50]
Study design (n=35)		
Observational study (prospective and retrospective before-and-after)	28 (71)	[23,24,27,29-41,43,44,46-54,56]
Randomized control trial	6 (12)	[25,26,28,42,45,55]
Cross-sectional survey	1 (3)	[22]
Comparison group (n=35)		
Before-and-after implementation comparison	11 (31)	[23-25,31,32,35,39,40,49,51,56]
Comparisons with intermittent monitoring	12 (34)	[26,28,29,34,38,41-43,45,46,48,53]
Comparisons with baseline data	1 (3)	[47]
Comparison between the monitor and alert method	2 (6)	[52,55]
No comparison	9 (26)	[22,27,30,33,36,37,44,50,54]
Clinical settings (n=35)		
Surgical-related wards	18 (51)	[22,24-27,29,30,35,36,38-40,42,43,45,49,52,54]
General wards	7 (20)	[23,37,46-48,51,53]
Surgical and internal unit	1 (3)	[55]
Others	9 (26)	[28,31-34,41,44,50,56]
Patient type (n=35)		
Postoperative	17 (49)	[26-30,33,37-39,41-43,45,49,52-54]
General medical, trauma, and surgical patients	1 (3)	[25]
Respiratory includes (COVID-19)	4 (11)	[24,48,50,51]
Other	6 (17)	[31,32,34,46,47,56]
Not specified	7 (20)	[22,23,35,36,40,44,55]
Device type (n=35)		
Wearable devices	26 (74)	[22,26-30,33,36-40,42-47,49-56]
Bedside monitors	5 (14)	[23-25,35,41]
Bedside monitors + wearable devices	4 (11)	[31,32,34,48]
Alert path (n=35)		
Alerts at central stations or system	14 (40)	[22,23,29-32,35,38,41,43,44,47,50,54]

XSL•FO
RenderX

Categories	Studies, n (%)	References
Alert to the central station or system and clinicians' phones or mobile devices	21 (60)	[24-28,33,34,36,37,39,40,42,45,46,48,49,51-53,55,56]
Alert mechanism type (n=35)		
Threshold alerts	25 (71)	[22,25-30,33,35-40,42,44-49,51-53,56]
Early warning score–based alerts	5 (14)	[23,24,34,41,55]
Artificial intelligence–based alerts	5 (14)	[31,32,43,50,54]

Characteristics of Clinicians

Ten (29%) of the included studies reported clinicians' characteristics involved in using continuous monitoring devices with deterioration alerting [22,30,35-37,47,49,53,54,56] (Multimedia Appendix 4). Seven studies reported incomplete characteristics [22,36,47,49,53,54,56], and 4 studies provided sample size and profession [47,49,54,56]. The available data indicated more female than male clinicians, with mean ages between 27 and 30 years. Most clinicians had 5-10 years of work experience, with nurses being the majority of users (Multimedia Appendix 5).

Overview of Study Devices

The devices used can be classified into wearable devices and bedside monitors. Wearable devices, worn directly by the patient, monitor vital signs continuously while allowing patient mobility. The most commonly used wearable devices were the SensiumVital patch (Sensium Healthcare; n=8, 23%) [26-28,33,36,37,45,52] and ViSi Mobile (Sotera Wireless Inc; n=7, 20%) [22,43,51,53-56]. The SensiumVital patch, applied to the chest, monitors heart rate, respiratory rate, and body temperature, transmitting data every 2 minutes to a central station or mobile device, with visual alerts for deviations from preset vital signs [26]. ViSi Mobile monitors heart rate, blood pressure, respiratory rate, body temperature, and oxygen saturation from the upper arm, chest, and wrist, sending visual alerts to a central monitor and the nurse's Wi-Fi phone [56]. Bedside monitors are stationary devices placed near the patient's bed, offering continuous monitoring within the monitor's vicinity. The IntelliVue Guardian Solution (Philips), used in 4 studies, tracks heart rate, respiratory rate, blood pressure, body temperature, and oxygen saturation, providing visual alerts on central and bedside monitors [23,24,41,48]. It can also be used with wireless monitors on the chest, wrist, and upper arm, ensuring continuous monitoring and timely clinical actions [23,24,41,48].

The alerting mechanisms in the included studies were categorized into 3 groups: threshold alerts, EWS-based alerts, and artificial intelligence (AI)–based alerts. Threshold alerts (n=25, 71%) notify health care providers when a monitored vital sign exceeds predefined limits, typically applied to respiratory rate, oxygen saturation, heart rate, and blood pressure (with temperature less consistently included), although only a minority of studies reported explicit numerical cut-off values or which parameters most commonly triggered alarms [22,25-30,33,35-40,42,44-49,51-53,56]. EWS-based alerts (n=5, 14%) are generated using an aggregated score based on multiple vital signs to identify patients at risk of deterioration [23,24,34,41,55]. AI-based alerts (n=5, 14%) use algorithms to analyze vital signs and predict potential clinical deterioration [31,32,43,50,54]. In terms of delivery methods, 20 studies reported systems that sent alerts directly to clinicians' mobile devices, including phones or pagers [24-28,33,34,36,37,39,40,42,46,48,49,51-53,55,56], while also notifying the central nurse station. The remaining studies indicated that alerts were sent exclusively to the central station, highlighting variability in alert dissemination approaches.

Effectiveness, Efficiency, Satisfaction, and Barriers to Usability

The included studies were analyzed for their reported effectiveness, efficiency, satisfaction, and barriers to usability.

Effectiveness

Twenty-five studies (71% of all included studies) provided relevant data on effectiveness [23,25,26,28,29,31-33,36,37,40-45,47-50,52-56]. Eight of these studies (32%) reported on rapid response team (RRT) calls: 2 studies indicated an increase [23,48], 1 study reported no significant change [53], and 5 studies observed a decrease in RRT calls [29,40,47,49,56]. ICU transfer rates were evaluated in 13 (52%) studies, with 3 studies finding no significant change [25,42,45] and 10 studies reporting a decrease [28,29,31,32,40,47-49,53,56]. Mortality rates were assessed in 12 studies, with 5 studies noting an increase in survival rates [23,31,32,48,49] and 7 studies finding no significant change [26,28,29,41,53,56]. The length of hospital stay was examined in 12 studies; 2 studies did not specify results [37,50], 5 studies found no significant change [29,41,45,49,56], and 5 studies reported a decrease [23,25,26,28,47]. Readmission rates were evaluated in 3 studies, with 1 study [41] finding no significant change and 2 studies [26,28] observing a decrease. Serious adverse events were investigated in 6 studies, with 2 studies identifying adverse events [50,52] and 4 studies reporting a decrease [31,32,48,53]. Notably, 5 studies using AI-based alerting systems demonstrated potential benefits [31,32,43,50,54], including reduced mortality and ICU transfer rates [31,32] and improved identification and reduction of SAEs [31,32,50]. However, no significant tendency was observed across other measured outcomes. Of the 4 studies that combined bedside monitors with wearable devices [31,32,34,48], 3 reported strong effectiveness in reducing mortality, ICU transfers, and SAEs [31,32,48]. However, the effectiveness of bedside monitors alone remains unclear due to the limited number of studies.

Overall, more studies suggested that RRT calls and ICU transfer rates decrease after implementing continuous monitoring devices

XSL•FO
RenderX

with deterioration alerting systems. However, the effects on mortality, length of hospital stay, readmissions, and serious adverse events remain inconclusive.

Efficiency

Twenty-six studies (74%) provided insights into efficiency issues, such as alarm frequency, false alert rate, workload impact, and time saving [22-24,30-40,42-44,46,47,49-52,54-56]. Alarm frequency was reported in 15 studies, with 4 studies indicating excessive alarms (over 4 alarms per patient per day) [24,30,36,56] and 11 studies reporting fewer than 5 alarms per patient per day [34,35,37-40,42,43,49,52,54]. The false alert rate was reported in 12 studies; 7 studies reported a high false alert rate (>40%) [24,30,36,37,43,50,52], while 5 studies reported a low false alert rate [31,34,44,49,56]. The impact on workload was investigated in 15 studies; 5 studies were unsure of the impact [22,39,47,51,56], 3 studies reported an increase in workload [24,36,54], and 7 studies observed a reduction in workload [32,34,35,38,40,46,50]. Time savings were evaluated in 8 studies, with 1 study reporting unclear results [55] and 7 studies reporting time savings [22,35,38,44,46,47,56]. Most current studies suggest that continuous monitoring devices save time, although the impact on workload is less clear. Alarm frequency generally remained below 5 alarms per patient per day, but false alarms were reported as a common and significant issue.

Satisfaction

Satisfaction with the devices was reported in 16 (46%) studies. Comfort was reported in 8 studies, focusing only on patient perspectives. Among these, 88% (n=7) [26-28,37,42,44,47] reported that patients found the devices comfortable, while 1 study reported discomfort [22]. Acceptability was reported in 94% of the studies, with 8 studies reporting patient acceptability [22,26-28,37,39,42,49] and 9 studies reporting clinician acceptability [35,36,38,40,44,47,49,56]. High acceptability was found among both patients and clinicians, with 1 study not clearly stating clinician acceptability [53].

Barriers to Usability

Usability barriers were formally reported in 10 studies (29% of all included studies) [22,26,27,30,35,36,39,44,47,49], while 5 studies mentioned them only in their discussion [24,28,40,42,43]. These barriers included technical issues, alarm management, patient comfort and experience, training and knowledge needs, and workflow integration challenges (Table 3).

Table . Summary of barriers to usability (n=10, 29% of all studies).

Studies	Definition	Studies, n (%)
Technical issues [22,39,44]	Issues with device connectivity, battery life, sensor attachment, internet connectivity, and overall device design, along with technical difficulties and artifactual issues, such as poor lead adherence and interference, as well as the cumbersome nature of device use, including removal and reapplication for showers.	3 (20)
Alarm management [26,30,49]	Frequent and high rates of initial false alarms leading to alarm fatigue and inconvenience, inconsistent practices in managing alarms, and excessive alarms from the system, often due to malfunctioning hardware or baseline tachycardia.	3 (20)
Patient comfort and experience [26,27,36,39]	Discomfort and skin reactions from the patch, concerns about practicalities such as showering, trust issues with technology reliability, patient refusals due to discomfort, confusion, or personal reasons, and discontinued monitoring due to various factors, such as contact allergies or initiation of palliative care.	4 (27)
Training and knowledge needs [22,26,35,36,39]	The need for patient education, ongoing training and coaching for health care providers, and challenges including alarm fatigue, accuracy and trust issues, and insufficient training for secondary users are impacting the patient experience.	5 (33)
Workflow and integration issues [27,36,47,49]	Challenges with integrating the monitoring system into clinical workflows, managing alarm burdens, interpreting vital sign trends, ensuring seamless integration with hospital systems, and the potential reduction in face-to-face nursing contact.	4 (27)

Technical issues were identified in 3 studies and included problems with device connectivity, battery life, sensor attachment, internet connectivity, and overall device design, along with artefactual issues such as poor lead adherence (ie,

XSL•FO
RenderX

leads detaching frequently and affecting continuous monitoring), interference, and the cumbersome nature of device use, including removal and reapplication for showers [22,39,44]. Alarm management was a concern in 3 studies [26,30,49], with frequent and high rates of initial false alarms leading to alarm fatigue and inconvenience, inconsistent practices in managing alarms, and excessive alarms often resulting from malfunctioning hardware or baseline tachycardia. Unlike efficiency measures, which assess the total number of alarms and false alarm rates quantitatively, these studies highlighted usability-related challenges, such as how clinicians respond to alarms, their perceived reliability of alerts, and whether excessive alarms led to desensitization or delayed responses to actual critical alerts. Additionally, alarm issues were often linked to hardware malfunctions or patient baseline conditions (eg, persistent tachycardia triggering unnecessary alerts), which further complicated clinical workflows and increased clinician frustration. Patient comfort and experience were highlighted in 4 studies [26,27,36,39], discussing discomfort and skin reactions from the patch, practical concerns such as showering, trust issues with technology reliability, patient refusals due to discomfort, confusion, or personal reasons, and discontinued monitoring due to contact allergies or the initiation of palliative care. Training and knowledge needs were emphasized in 5 studies [22,26,35,36,39], underscoring the need for patient education, ongoing training, and coaching for health care providers, with challenges including alarm fatigue, accuracy and trust issues, and insufficient training for secondary users impacting the patient experience. Workflow and integration issues were reported in 4 studies [27,36,47,49], noting challenges with integrating the monitoring system into clinical workflows, managing alarm burdens, interpreting vital sign trends, ensuring seamless integration with hospital systems, and the potential reduction in face-to-face nursing contact. Finally, only 6 studies systematically surveyed usability barriers from clinicians' perspectives, collecting data through interviews, surveys, or by directly including clinicians as participants [22,30,35,36,47,49].

Discussion

Principal Findings

This scoping review aimed to map the evidence on the usability of continuous monitoring devices with deterioration alerting in non-ICU settings, according to the ISO standard. It focused on effectiveness, efficiency, satisfaction, and the barriers affecting usability. Current review findings suggest that while most research supports the effectiveness and efficiency of these devices, evidence regarding satisfaction and barriers to usability remains limited, with usability barriers receiving the least attention. Through the limited evidence, 5 key barriers to usability were identified in this review: (1) technical issues (eg, connectivity and battery limitations), (2) alarm management challenges (eg, false alarms and alarm fatigue), (3) patient comfort concerns (eg, skin irritation), (4) training gaps for clinicians, and (5) workflow integration difficulties. Considering the impact of barriers to usability on effectiveness, efficiency, and satisfaction, a critical gap in the literature is highlighted [19]. Future research should prioritize investigating usability barriers by examining patient and clinician experiences and developing interventions to overcome implementation challenges.

The findings of this review suggest that continuous monitoring devices with deterioration alerting are associated with reductions in RRT calls and ICU transfer rates in some studies, supporting previous reviews. Cardona-Morrell et al [17] and Sun et al [16] also reported significant reductions in cardiac arrest calls and rescue events associated with continuous monitoring. However, mortality outcomes remain inconclusive, likely due to variability in study designs, small sample sizes, and differences in patient populations or clinical settings. Similarly, Cardona-Morrell et al [17] and Areia et al [18] found no significant impact on mortality, while Sun et al [16] reported a 39% reduction in mortality risk, making it the only review with a significant finding. Interestingly, previous reviews all reported nonsignificant reductions in ICU transfers, which contrasts with the findings of this review [15-18]. These findings collectively suggest that continuous monitoring devices may improve patient outcomes; however, further research is needed to provide more evidence supporting their comprehensive benefits.

The efficiency of continuous monitoring devices with deterioration alerting is supported by evidence indicating time savings and manageable alarm frequencies. Most studies report fewer than 5 alarms per patient per day, which is generally acceptable to clinicians [56]. However, false alarms remain a significant concern, with some studies reporting rates exceeding 40% following device implementation. This issue has been highlighted in previous systematic reviews, which emphasize the need for improved alert accuracy, as false alarms are consistently identified as a usability barrier requiring further research and refinement [15,16,18]. Additionally, the impact of these devices on workload remains inconclusive, as findings vary across studies—some report a reduction in workload, others an increase, while some remain uncertain. Downey et al [15] noted that nurses who received proper training and felt confident using the technology experienced less workload strain, whereas a lack of familiarity led to disengagement and a perceived increase in workload. Given these inconsistencies, further research is needed to establish clearer trends regarding the workload impact of continuous monitoring devices.

Satisfaction with continuous monitoring devices is notably high, with both patients and clinicians reporting positive experiences. Most studies indicate that patients perceive these devices as comfortable and acceptable. While previous reviews have focused less on satisfaction metrics, Downey et al [15] provide valuable insights, noting that patients and clinicians recognize the clinical benefits of these devices and express willingness to adopt them due to enhanced patient safety. These findings align with the current review's conclusions regarding high satisfaction levels.

The barriers to usability associated with continuous monitoring devices with deterioration alerting systems in non-ICU settings are multifaceted, which prior reviews rarely address. Technical issues found in this review, such as unreliable Wi-Fi connectivity and sensors detaching from patients (triggering nonactionable alarms), align with findings from Leenen et al [21], whose review of 13 wearable devices highlighted design

XSL•FO
RenderX

limitations—including partial wiring in supposedly "wearable" systems that restrict patient mobility and generate clinically irrelevant alerts. In addition, patient discomfort and experience, manifested as skin irritation, mobility restrictions, and distrust in device reliability, have been widely recognized as a barrier in existing literature [15-18]. Furthermore, alarm management challenges, particularly false alarms, exacerbate clinician workload and desensitization, a concern echoed across previous review studies [15-18].

Notably, this review identifies training gaps and workflow integration challenges as critical yet underreported usability barriers that were not emphasized in earlier literature [15-18]. Inadequate education for both patients and clinicians—particularly regarding alarm interpretation, device operation, and troubleshooting—amplifies usability challenges and undermines system effectiveness. These findings align with Chaniaud et al [57], who suggest that even basic, widely adopted home devices (eg, blood pressure monitors and pulse oximeters) require robust user education to achieve their full potential. This highlights the universal importance of training programs tailored to device complexity and user expertise. Similarly, poorly integrated systems disrupt clinical workflows. For example, devices that lack seamless connectivity with existing EHR systems force clinicians to manually reconcile data, increasing staff burden and reducing time for direct patient care [49]. Collectively, these barriers highlight the necessity of holistic solutions that address not only technical performance but also human-centered design (eg, prioritizing patient comfort and clinician workflow efficiency) and system interoperability to maximize the potential of continuous monitoring.

One key aspect of the adoption and use of continuous monitoring systems is the critical role of clinicians who are responsible for implementing and managing these technologies. According to these findings, nurses make up the majority of front-line users, managing continuous monitoring, interpreting alerts, and overseeing these technologies in non-ICU environments. This highlights the need to prioritize nurses' needs in the design and integration of continuous monitoring with deterioration alert systems. However, a significant limitation in the literature is the disproportionate focus on patient experiences compared to clinician-related factors. This imbalance is problematic because clinicians' acceptance and effective use of technology are pivotal to integrating these systems successfully into clinical workflows. Few studies have examined usability barriers from the clinicians' perspective or provided detailed data on factors such as familiarity with devices, willingness to adopt new technologies, and demographic characteristics (eg, age, gender, and experience). In fact, clinicians' attributes can significantly influence usability, user behavior, and overall adoption of these systems [58]. Or et al [58] demonstrate that clinicians who trust and are willing to use technology are more likely to adopt EHR systems in their practice. These findings align with Carayon and Hoonakker [59], who argue that clinicians—as both implementers and end users—are as critical as system designers in determining health IT effectiveness. Their review underscores that neglecting clinician-specific barriers risks poor adoption and suboptimal outcomes. Addressing these gaps is essential for ensuring the long-term viability of continuous monitoring technologies and highlights the urgent need for studies that prioritize clinician-centered usability metrics.

Finally, this review adds several key points compared to previous reviews. First, AI-based alerting systems remain underused and are not yet well tested in these studies, indicating a need for further research to establish their effectiveness for future applications. A systematic review by Muralitharan et al [60] suggests that AI-based alerting methods perform better than threshold alerts and EWS-based alerts; however, this scoping review found limited practical application of AI-based alerting, underlining the necessity for additional clinical trials. Second, most studies on continuous monitoring devices with deterioration alerting are concentrated in the United States and are primarily conducted in surgical units with postoperative patients. This suggests that the application of these devices is currently limited to specific clinical settings, patient populations, and geographic regions. However, continuous monitoring systems have been highlighted as promising remote patient monitoring solutions that can save clinicians time by alleviating staffing burdens and improving patient safety—particularly in light of staffing shortages in both limited-, low-, and middle-income countries [61]. Third, implications for smaller community hospitals and lower-resource wards also warrant consideration. Most of the implementations identified in this review were in larger, well-resourced centers, whereas prior work has highlighted that higher nurse-to-patient ratios in general wards and human-related monitoring failures are important contributors to delayed recognition of deterioration [7,8]. In settings where 1 nurse cares for more patients and technical support is limited, the additional alarms and infrastructure required for continuous monitoring may therefore have different consequences for workload, alarm fatigue, and value for money; although earlier reviews have described potential cost savings in selected high-resource wards, economic outcomes were rarely reported in this review's included studies, underscoring the need for formal economic evaluations—particularly in community and resource-constrained hospitals—before large-scale implementation [15,61]. These gaps highlight the importance of expanding research to assess the generalizability of continuous monitoring technologies by exploring diverse clinical environments and rigorously evaluating emerging AI-based alerting systems.

Limitations

This review is limited to studies published in English, potentially excluding relevant non-English publications and unpublished studies. The focus on usability (effectiveness, efficiency, and satisfaction) excludes studies that did not mention or examine these aspects. Only studies providing primary data were included, potentially omitting those reporting secondary data, such as systematic reviews and meta-analyses, or unpublished studies. Due to significant variations in methodology, objectives, and reported data among the included studies, a meta-analysis was not feasible. The search strategy excluded studies conducted in the ICU or nonhospital services, focusing deliberately on noncritical adult care units to investigate usability in these settings. This may have omitted relevant studies that include both ICU and non-ICU settings. Furthermore, as a scoping

XSL•FO
RenderX

review, this study aimed to include a broad range of studies without excluding any based on quality. Finally, usability was measured with heterogeneous, study-specific items, and no study used a validated instrument (eg, System Usability Scale), precluding cross-system benchmarking or pooled usability scores. Alert parameters and thresholds were inconsistently reported, limiting the comparability of alarm burden across devices and protocols.

Conclusion

Continuous monitoring devices with deterioration alerting systems are increasingly recognized as valuable tools for preventing patient deterioration in non-ICU settings. However, their successful implementation hinges on a comprehensive understanding of usability (encompassing effectiveness, efficiency, and satisfaction) and the barriers influencing real-world adoption. This review indicates that these devices can reduce RRT calls and ICU transfers, save time, and maintain manageable alarm frequencies while achieving high user satisfaction. However, significant usability barriers remain, including technical issues, alarm management challenges, patient discomfort, and insufficient training and workflow integration. Moreover, most existing studies focus on effectiveness and efficiency, leaving satisfaction and broader usability factors understudied. Additionally, research has predominantly focused on patient perspectives, often neglecting clinician insights and has been limited to specific clinical contexts, patient populations, and geographic regions, which raises concerns about the generalizability of these findings. Future studies should prioritize usability factors and expand to the clinicians' usage perspective and diverse health care settings to ensure these technologies deliver equitable, scalable improvements in patient safety and optimize care delivery in non-ICU environments.

Acknowledgments

This research was supported by a scholarship grant from the Government of Taiwan, awarded to JFP. The views expressed in this publication are those of the authors and do not necessarily reflect the official policy or position of their affiliated institutions.

Funding

No external financial support or grants were received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search results up to November 2024.
[DOCX File, 26 KB - i-jmr_v15i1e75713_app1.docx]

Multimedia Appendix 2
Data extraction sheet.
[DOCX File, 27 KB - i-jmr_v15i1e75713_app2.docx]

Multimedia Appendix 3
Summary of the included studies.
[DOCX File, 78 KB - i-jmr_v15i1e75713_app3.docx]

Multimedia Appendix 4
Characteristics of the included studies.
[DOCX File, 433 KB - i-jmr_v15i1e75713_app4.docx]

Multimedia Appendix 5
Summary of the clinicians' characteristics in the included studies.
[DOCX File, 23 KB - i-jmr_v15i1e75713_app5.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 104 KB - i-jmr_v15i1e75713_app6.pdf]

References

1. Hospital admitted patient care activity, 2022-23. NHS England. 2023. URL: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2022-23 [accessed 2026-02-04]

XSL•FO
RenderX

2. Total number of hospital admissions in the US from 1946 to 2022. Statista. URL: https://www.statista.com/statistics/459718/total-hospital-admission-number-in-the-us/?__sso_cookie_checker=failed [accessed 2026-01-30]
3. Admitted patient care. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports-data/myhospitals/intersection/activity/apc [accessed 2026-01-30]
4. Massey D, Chaboyer W, Anderson V. What factors influence ward nurses' recognition of and response to patient deterioration? An integrative review of the literature. Nurs Open 2017 Jan;4(1):6-23. [doi: 10.1002/nop2.53] [Medline: 28078095]
5. Churpek MM, Yuen TC, Park SY, Gibbons R, Edelson DP. Using electronic health record data to develop and validate a prediction model for adverse outcomes in the wards*. Crit Care Med 2014 Apr;42(4):841-848. [doi: 10.1097/CCM.0000000000000038] [Medline: 24247472]
6. Donaldson LJ, Panesar SS, Darzi A. Patient-safety-related hospital deaths in England: thematic analysis of incidents reported to a national database, 2010-2012. PLoS Med 2014 Jun;11(6):e1001667. [doi: 10.1371/journal.pmed.1001667] [Medline: 24959751]
7. Bray K, Wren I, Baldwin A, et al. Standards for nurse staffing in critical care units determined by: The British Association of Critical Care Nurses, The Critical Care Networks National Nurse Leads, Royal College of Nursing Critical Care and In-flight Forum. Nurs Crit Care 2010;15(3):109-111. [doi: 10.1111/j.1478-5153.2010.00392.x] [Medline: 20500648]
8. van Galen LS, Struik PW, Driesen B, et al. Delayed recognition of deterioration of patients in general wards is mostly caused by human related monitoring failures: a root cause analysis of unplanned ICU admissions. PLoS ONE 2016;11(8):e0161393. [doi: 10.1371/journal.pone.0161393] [Medline: 27537689]
9. Gerry S, Bonnici T, Birks J, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. BMJ 2020 May 20;369:m1501. [doi: 10.1136/bmj.m1501] [Medline: 32434791]
10. Haegdorens F, Monsieurs KG, De Meester K, Van Bogaert P. An intervention including the national early warning score improves patient monitoring practice and reduces mortality: a cluster randomized controlled trial. J Adv Nurs 2019 Sep;75(9):1996-2005. [doi: 10.1111/jan.14034] [Medline: 31012124]
11. Smith MEB, Chiovaro JC, O'Neil M, et al. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. Ann Am Thorac Soc 2014 Nov;11(9):1454-1465. [doi: 10.1513/AnnalsATS.201403-102OC] [Medline: 25296111]
12. Williams TA, Tohira H, Finn J, Perkins GD, Ho KM. The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: a systematic review. Resuscitation 2016 May;102:35-43. [doi: 10.1016/j.resuscitation.2016.02.011] [Medline: 26905389]
13. Zhang K, Zhang X, Ding W, et al. National Early Warning Score does not accurately predict mortality for patients with infection outside the intensive care unit: a systematic review and meta-analysis. Front Med (Lausanne) 2021;8:704358. [doi: 10.3389/fmed.2021.704358] [Medline: 34336903]
14. Kolic I, Crane S, McCartney S, Perkins Z, Taylor A. Factors affecting response to National Early Warning Score (NEWS). Resuscitation 2015 May;90:85-90. [doi: 10.1016/j.resuscitation.2015.02.009] [Medline: 25703784]
15. Downey CL, Chapman S, Randell R, Brown JM, Jayne DG. The impact of continuous versus intermittent vital signs monitoring in hospitals: a systematic review and narrative synthesis. Int J Nurs Stud 2018 Aug;84:19-27. [doi: 10.1016/j.ijnurstu.2018.04.013] [Medline: 29729558]
16. Sun L, Joshi M, Khan SN, Ashrafian H, Darzi A. Clinical impact of multi-parameter continuous non-invasive monitoring in hospital wards: a systematic review and meta-analysis. J R Soc Med 2020 Jun;113(6):217-224. [doi: 10.1177/0141076820925436] [Medline: 32521195]
17. Cardona-Morrell M, Prgomet M, Turner RM, Nicholson M, Hillman K. Effectiveness of continuous or intermittent vital signs monitoring in preventing adverse events on general wards: a systematic review and meta-analysis. Int J Clin Pract 2016 Oct;70(10):806-824. [doi: 10.1111/ijcp.12846]
18. Areia C, Biggs C, Santos M, et al. The impact of wearable continuous vital sign monitoring on deterioration detection and clinical outcomes in hospitalised patients: a systematic review and meta-analysis. Crit Care 2021 Dec;25(1). [doi: 10.1186/s13054-021-03766-4]
19. Coiera E. Assessing technology success and failure using information value chain theory. Stud Health Technol Inform 2019 Jul 30;263:35-48. [doi: 10.3233/SHTI190109] [Medline: 31411151]
20. Bevan N. International standards for HCI and usability. Int J Hum Comput Stud 2001 Oct;55(4):533-552. [doi: 10.1006/ijhc.2001.0483]
21. Leenen JPL, Leerentveld C, van Dijk JD, van Westreenen HL, Schoonhoven L, Patijn GA. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: systematic review. J Med Internet Res 2020 Jun 17;22(6):e18636. [doi: 10.2196/18636] [Medline: 32469323]
22. Becking-Verhaar FL, Verweij RPH, de Vries M, Vermeulen H, van Goor H, Huisman-de Waal GJ. Continuous vital signs monitoring with a wireless device on a general ward: a survey to explore nurses' experiences in a post-implementation period. Int J Environ Res Public Health 2023 May 11;20(10):5794. [doi: 10.3390/ijerph20105794] [Medline: 37239523]

XSL•FO
RenderX

23. Bellomo R, Ackerman M, Bailey M, et al. A controlled trial of electronic automated advisory vital signs monitoring in general hospital wards. Crit Care Med 2012 Aug;40(8):2349-2361. [doi: 10.1097/CCM.0b013e318255d9a0] [Medline: 22809908]
24. Blankush JM, Freeman R, McIlvaine J, Tran T, Nassani S, Leitman IM. Implementation of a novel postoperative monitoring system using automated Modified Early Warning Scores (MEWS) incorporating end-tidal capnography. J Clin Monit Comput 2017 Oct;31(5):1081-1092. [doi: 10.1007/s10877-016-9943-4] [Medline: 27766526]
25. Brown H, Terrence J, Vasquez P, Bates DW, Zimlichman E. Continuous monitoring in an inpatient medical-surgical unit: a controlled clinical trial. Am J Med 2014 Mar;127(3):226-232. [doi: 10.1016/j.amjmed.2013.12.004]
26. Downey C, Randell R, Brown J, Jayne DG. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. J Med Internet Res 2018 Dec 11;20(12):e10802. [doi: 10.2196/10802] [Medline: 30538086]
27. Downey CL, Brown JM, Jayne DG, Randell R. Patient attitudes towards remote continuous vital signs monitoring on general surgery wards: an interview study. Int J Med Inform 2018 Jun;114:52-56. [doi: 10.1016/j.ijmedinf.2018.03.014] [Medline: 29673603]
28. Downey CL, Croft J, Ainsworth G, et al. Trial of remote continuous versus intermittent NEWS monitoring after major surgery (TRaCINg): a feasibility randomised controlled trial. Pilot Feasibility Stud 2020 Nov 23;6(1):183. [doi: 10.1186/s40814-020-00709-8] [Medline: 33292669]
29. Eddahchouri Y, Peelen RV, Koeneman M, Touw HRW, van Goor H, Bredie SJH. Effect of continuous wireless vital sign monitoring on unplanned ICU admissions and rapid response team calls: a before-and-after study. Br J Anaesth 2022 May;128(5):857-863. [doi: 10.1016/j.bja.2022.01.036]
30. Gazarian PK. Nurses' response to frequency and types of electrocardiography alarms in a non-critical care setting: a descriptive study. Int J Nurs Stud 2014 Feb;51(2):190-197. [doi: 10.1016/j.ijnurstu.2013.05.014] [Medline: 23810495]
31. Hravnak M, Devita MA, Clontz A, Edwards L, Valenta C, Pinsky MR. Cardiorespiratory instability before and after implementing an integrated monitoring system. Crit Care Med 2011 Jan;39(1):65-72. [doi: 10.1097/CCM.0b013e3181fb7b1c] [Medline: 20935559]
32. Hravnak M, Edwards L, Clontz A, Valenta C, Devita MA, Pinsky MR. Defining the incidence of cardiorespiratory instability in patients in step-down units using an electronic integrated monitoring system. Arch Intern Med 2008 Jun 23;168(12):1300-1308. [doi: 10.1001/archinte.168.12.1300] [Medline: 18574087]
33. Joshi M, Ashrafian H, Arora S, et al. A pilot study to investigate real-time digital alerting from wearable sensors in surgical patients. Pilot Feasibility Stud 2022 Jul 6;8(1):140. [doi: 10.1186/s40814-022-01084-2] [Medline: 35794669]
34. Klumpner TT, Kountanis JA, Langen ES, Smith RD, Tremper KK. Use of a novel electronic maternal surveillance system to generate automated alerts on the labor and delivery unit. BMC Anesthesiol 2018 Jun 26;18(1):78. [doi: 10.1186/s12871-018-0540-6] [Medline: 29945569]
35. Kuznetsova M, Kim AY, Scully DA, et al. Implementation of a continuous patient monitoring system in the hospital setting: a qualitative study. Jt Comm J Qual Patient Saf 2024 Apr;50(4):235-246. [doi: 10.1016/j.jcjq.2023.10.017] [Medline: 38101994]
36. Leenen JPL, Dijkman EM, van Hout A, Kalkman CJ, Schoonhoven L, Patijn GA. Nurses' experiences with continuous vital sign monitoring on the general surgical ward: a qualitative study based on the Behaviour Change Wheel. BMC Nurs 2022 Mar 14;21(1):60. [doi: 10.1186/s12912-022-00837-x] [Medline: 35287678]
37. Leenen JPL, Dijkman EM, van Dijk JD, et al. Feasibility of continuous monitoring of vital signs in surgical patients on a general ward: an observational cohort study. BMJ Open 2021 Feb 17;11(2):e042735. [doi: 10.1136/bmjopen-2020-042735] [Medline: 33597138]
38. McGrath SP, Perreard IM, Garland MD, Converse KA, Mackenzie TA. Improving patient safety and clinician workflow in the general care setting with enhanced surveillance monitoring. IEEE J Biomed Health Inform 2019 Mar;23(2):857-866. [doi: 10.1109/JBHI.2018.2834863] [Medline: 29993903]
39. McGrath SP, Pyke J, Taenzer AH. Assessment of continuous acoustic respiratory rate monitoring as an addition to a pulse oximetry-based patient surveillance system. J Clin Monit Comput 2017 Jun;31(3):561-569. [doi: 10.1007/s10877-016-9884-y] [Medline: 27142098]
40. McGrath SP, Taenzer AH, Karon N, Blike G. Surveillance monitoring management for general care units: strategy, design, and implementation. Jt Comm J Qual Patient Saf 2016 Jul;42(7):293-302. [doi: 10.1016/s1553-7250(16)42040-4] [Medline: 27301832]
41. Mestrom E, De Bie A, Steeg MVD, et al. Implementation of an automated early warning scoring system in a surgical ward: practical use and effects on patient outcomes. PLoS ONE 2019;14(5):e0213402. [doi: 10.1371/journal.pone.0213402] [Medline: 31067229]
42. Paul JE, Chong MA, Buckley N, et al. Vital sign monitoring with continuous pulse oximetry and wireless clinical notification after surgery (the VIGILANCE pilot study)-a randomized controlled pilot trial. Pilot Feasibility Stud 2019;5(1):36. [doi: 10.1186/s40814-019-0415-8] [Medline: 30858986]

XSL•FO
RenderX

43. Peelen RV, Eddahchouri Y, Koeneman M, Melis R, van Goor H, Bredie SJH. Comparing continuous with periodic vital sign scoring for clinical deterioration using a patient data model. J Med Syst 2023 May 8;47(1):60. [doi: 10.1007/s10916-023-01954-z] [Medline: 37154986]
44. Pollack CV. Wireless cardiac event alert monitoring is feasible and effective in the emergency department and adjacent waiting areas. Crit Pathw Cardiol 2009 Mar;8(1):7-11. [doi: 10.1097/HPC.0b013e3181980f8b] [Medline: 19258832]
45. Posthuma LM, Breteler MJM, Lirk PB, Dijkum EJ, Visscher MJ, Breel JS, et al. Surveillance of high-risk early postsurgical patients for real-time detection of complications using wireless monitoring. Front Med (Lausanne) ;10. [doi: 10.3389/fmed.2023.1295499]
46. Sigvardt E, Grønbaek KK, Jepsen ML, et al. Workload associated with manual assessment of vital signs as compared with continuous wireless monitoring. Acta Anaesthesiol Scand 2024 Feb;68(2):274-279. [doi: 10.1111/aas.14333] [Medline: 37735843]
47. Stellpflug C, Pierson L, Roloff D, et al. Continuous physiological monitoring improves patient outcomes. Am J Nurs 2021 Apr 1;121(4):40-46. [doi: 10.1097/01.NAJ.0000742504.44428.c9] [Medline: 33755624]
48. Subbe CP, Duller B, Bellomo R. Effect of an automated notification system for deteriorating ward patients on clinical outcomes. Crit Care 2017 Dec;21(1). [doi: 10.1186/s13054-017-1635-z]
49. Taenzer AH, Pyke JB, McGrath SP, Blike GT. Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers. Anesthesiology 2010 Feb;112(2):282-287. [doi: 10.1097/ALN.0b013e3181ca7a9b] [Medline: 20098128]
50. Un KC, Wong CK, Lau YM, et al. Observational study on wearable biosensors and machine learning-based remote monitoring of COVID-19 patients. Sci Rep 2021 Feb 23;11(1):4388. [doi: 10.1038/s41598-021-82771-7] [Medline: 33623096]
51. van Goor HMR, Eddahchouri Y, van Loon K, et al. Can continuous remote vital sign monitoring reduce the number of room visits to patients suspected of COVID-19: a quasi-experimental study. Int J Nurs Stud 2021 Mar;115:103868. [doi: 10.1016/j.ijnurstu.2020.103868] [Medline: 33465579]
52. van Rossum MC, Vlaskamp LB, Posthuma LM, et al. Adaptive threshold-based alarm strategies for continuous vital signs monitoring. J Clin Monit Comput 2022 Apr;36(2):407-417. [doi: 10.1007/s10877-021-00666-4] [Medline: 33575922]
53. Verrillo SC, Cvach M, Hudson KW, Winters BD. Using continuous vital sign monitoring to detect early deterioration in adult postoperative inpatients. J Nurs Care Qual 2019;34(2):107-113. [doi: 10.1097/NCQ.0000000000000350] [Medline: 30095509]
54. Watkins T, Whisman L, Booker P. Nursing assessment of continuous vital sign surveillance to improve patient safety on the medical/surgical unit. J Clin Nurs 2016 Jan;25(1-2):278-281. [doi: 10.1111/jocn.13102] [Medline: 26531215]
55. Weenk M, Koeneman M, van de Belt TH, Engelen LJLPG, van Goor H, Bredie SJH. Wireless and continuous monitoring of vital signs in patients at the general ward. Resuscitation 2019 Mar;136:47-53. [doi: 10.1016/j.resuscitation.2019.01.017] [Medline: 30685546]
56. Weller RS, Foard KL, Harwood TN. Evaluation of a wireless, portable, wearable multi-parameter vital signs monitor in hospitalized neurological and neurosurgical patients. J Clin Monit Comput 2018 Oct;32(5):945-951. [doi: 10.1007/s10877-017-0085-0] [Medline: 29214598]
57. Chaniaud N, Métayer N, Megalakaki O, Loup-Escande E. Effect of prior health knowledge on the usability of two home medical devices: usability study. JMIR Mhealth Uhealth 2020 Sep 21;8(9):e17983. [doi: 10.2196/17983] [Medline: 32955454]
58. Or C, Tong E, Tan J, Chan S. Exploring factors affecting voluntary adoption of electronic medical records among physicians and clinical assistants of small or solo private general practice clinics. J Med Syst 2018 May 29;42(7):121-112. [doi: 10.1007/s10916-018-0971-0] [Medline: 29845400]
59. Carayon P, Hoonakker P. Human factors and usability for health information technology: old and new challenges. Yearb Med Inform 2019 Aug;28(1):71-77. [doi: 10.1055/s-0039-1677907] [Medline: 31419818]
60. Muralitharan S, Nelson W, Di S, et al. Machine learning-based early warning systems for clinical deterioration: systematic scoping review. J Med Internet Res 2021 Feb 4;23(2):e25187. [doi: 10.2196/25187] [Medline: 33538696]
61. Holtz BE, Urban FA, Oesterle J, Blake R, Henry A. The promise of remote patient monitoring. Telemed J E Health 2024 Dec;30(12):2776-2781. [doi: 10.1089/tmj.2024.0521] [Medline: 39535888]

Abbreviations

AI: artificial intelligence
EHR: electronic health record
EWS: early warning score
ICU: intensive care unit
ISO: International Organization for Standardization
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

XSL•FO
RenderX

RRT: rapid response team
SAE: severe adverse event

Edited by M Balcarras; submitted 09.Apr.2025; peer-reviewed by E Ding, H El-Kechai; revised version received 10.Nov.2025; accepted 12.Jan.2026; published 10.Feb.2026.

Please cite as:
Pan JF, Dowding D, Wong D, Scott A, Zhao Q
The Usability of Continuous Monitoring Devices With Deterioration Alerting Systems in Noncritical Care Units: Scoping Review
Interact J Med Res 2026;15:e75713
URL: https://www.i-jmr.org/2026/1/e75713
doi:10.2196/75713

© Jo-Fan Pan, Dawn Dowding, David Wong, Ashley Scott, Qimeng Zhao. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 10.Feb.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Machine and Deep Learning for Detection of Moderate-to-Vigorous Physical Activity From Accelerometer Data: Systematic Scoping Review

Yahua Zi[1], MSc; Sjors RB van de Ven[2], MSc; Eco JC de Geus[2], PhD; Peijie Chen[1], PhD

[1]School of Exercise and Health, Shanghai University of Sport, Shanghai, China

[2]Department of Biological Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Corresponding Author:
Peijie Chen, PhD
School of Exercise and Health
Shanghai University of Sport
No.399 Changhai Rd
Shanghai, 200438
China
Phone: 86 21 5125 3145
Email: chenpeijie@sus.edu.cn

Abstract

Background: Accurate monitoring of moderate-to-vigorous physical activity (MVPA) is critical for advancing public health research and personalized interventions. Traditional accelerometry methods, reliant on regression-derived intensity cut points, exhibit significant misclassification errors and poor generalizability to the free-living environment. Recent advancements in machine learning (ML) and deep learning (DL) offer promising alternatives for automated MVPA detection.

Objective: This scoping review synthesizes evidence on ML and DL techniques for MVPA estimation and prediction using accelerometer data, focusing on performance, algorithm bias, sensor configurations, and translational potential.

Methods: Following PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines, we conducted a systematic search across PubMed, IEEE Xplore, and Web of Science (February 1995-April 2025), supplemented by snowball citation tracking. Two independent reviewers screened titles, abstracts, and full texts against predefined inclusion criteria. Data from included studies were charted by one reviewer and verified by the other, extracting details on study characteristics, sensor configuration, ML and DL techniques, validation methods, and performance metrics. A narrative synthesis approach was used, guided by 6 research questions, to collate and summarize the findings. The synthesis process was rigorously reviewed by multiple authors to ensure consistency.

Results: Of 1938 screened studies, 40 met the inclusion criteria, with 4 studies added by follow-up manual searches. While traditional ML models (eg, random forest, support vector machine) achieved strong laboratory performance with F_1-score of 87.4%-100% and accuracy of 87.9%-100%, their real-world performance declined by 8.0%-13.3% in F_1-score and 6.6%-12.2% in accuracy, due to environment noise and device heterogeneity. DL architectures (eg, convolutional neural networks, transformers) achieved robust performance by leveraging raw signal dynamics with an F_1-score of 71.9%-79.8% and an accuracy of 87.9%-100% in free-living settings. Hybrid models (eg, convolutional neural networks and long short-term memory) demonstrated state-of-the-art performance (F_1-score 91.4%-98.4%, accuracy 97.7%-99.0%). Wrist-worn sensors dominated studies (30/40, 75%) and matched hip/thigh placements in lab settings (mean F_1-scores: 86.5%-88.6%), but multisensor configurations (wrist + hip) yielded the highest accuracy (89.7%). Key challenges included algorithmic bias reducing applicability in older adult populations, and impaired reproducibility, with only 42.5% (17/40) of studies sharing code and data. Emerging opportunities are noted for edge computing and hybrid models integrating contextual data.

Conclusions: ML and DL significantly enhance MVPA monitoring by automating feature extraction and improving adaptability to free-living variability. However, persistent gaps in generalizability, inconsistent validation protocols, and transparency deficits hinder translation. The findings support the need for future research to prioritize inclusive model training, standardized reporting frameworks, and open science practices to realize the equitable potential of artificial intelligence–driven physical activity assessment.

XSL•FO
RenderX

(Interact J Med Res 2026;15:e76601) doi:10.2196/76601

KEYWORDS

physical activity intensity; raw accelerometer data; wearable sensors; free-living validation; classification; estimation; sensor placement; machine learning; deep learning

Introduction

Moderate-to-vigorous physical activity (MVPA) is defined as activities requiring specific metabolic equivalent of tasks (METs), such as 3 METs or 4 METs [1,2]. It is critical to preventive health, linked to reduced risks of cardiovascular disease [3,4], diabetes [5], and premature mortality [6]. Current guidelines, such as those from the World Health Organization, emphasize MVPA as a priority; for example, children and adolescents are advised to engage in MVPA with an average of 60 minutes per day across the week to improve health [7,8]. Additionally, accurate measurement of physical activity is critical for identifying the individual, environmental, and sociocultural determinants and evaluating the efficacy of intervention strategies. Accelerometer-based motion sensors, owing to their compact design, durability, and low cost, have emerged as the predominant tool for objective physical activity assessment in diverse populations [9-12].

Traditional accelerometry methods, though widely adopted, have historically been underused in research due to reliance on intensity-based cut points derived from linear regression models or receiver operating curves [13,14]. These approaches establish thresholds by predicting energy expenditure from accelerometer counts. However, proprietary count-based thresholds, such as Freedson's cut points [15], exhibit significant misclassification of activity intensity (eg, sedentary, light, moderate, vigorous intensity) of approximately 50% in adults [16] and 28%-45% in children and adolescents [17-20]. Such methods fail to account for biomechanical nuances (eg, energy expenditure differences between walking on flat terrain vs uphill terrain) or uncontrolled variables in free-living environments, such as nonexercise movements [21,22]. The proliferation of conflicting regression-derived cut points has further complicated cross-study comparisons [23]. While these thresholds remain standard for quantifying activity intensity, their inability to accurately predict intensity across diverse activities is increasingly acknowledged [22].

The advent of machine learning (ML) and deep learning (DL) has revolutionized intensity recognition by enabling feature extraction and classification from raw accelerometer signals [24,25]. Compared with traditional cut point methods, ML models (eg, random forests [RFs] and support vector machines [SVMs]) leverage time- and frequency-domain features from high-resolution triaxial data (eg, 30-100 Hz) to reduce energy expenditure errors by 25%-50% in school-age children [19,26]. More recently, DL architectures, such as a convolutional neural network (CNN) for local temporal pattern detection, a long short-term memory network (LSTM) for modeling activity sequences, Transformers for long-range dependency learning, and hybrid models (eg, convolutional neural network and long short-term memory [CNN-LSTM]), have further advanced the field. These models identify MVPA bouts by modeling temporal dependencies in continuous data streams [27].

Three distinct methodological approaches have emerged for MVPA detection: The first one is based on activity classification, which directly identifies MVPA from activity-specific movement patterns [28,29]. The second one is based on energy expenditure prediction from predefined MET thresholds (eg, ≥3 METs) [19,30]. The third one is based on an end-to-end DL architecture that automates hierarchical feature extraction from raw accelerometer signals to classify activity intensity directly or through energy expenditure estimation [31-33]. Hybrid models, such as CNN-LSTM, further enhance performance by integrating spatial feature extraction (via convolutions) with temporal modeling (via recurrent layers) to identify subtle biomechanical patterns (eg, stride variability during running) and contextual transitions between movements [34]. However, over 60% of models remain inaccessible due to unshared code or validation protocols, perpetuating a "new cut-point conundrum" that undermines cross-study comparability and clinical utility [35].

Other shortcomings further undermine progress in MVPA-specific research. First, lab-based findings fail to be generalized to real-world conditions. For example, RF achieves >90% accuracy in lab settings [36,37], but its free-living performance degrades dramatically to around 66% [38]. However, only 10% of studies validate models in the real world [39], limiting translational relevance. Second, disparities in validation protocols, such as settings (laboratory-controlled vs free-living environments), or device placement (hip vs wrist), complicate cross-study comparisons. For instance, models trained on hip-based ActiGraph data often underperform when applied to wrist-worn devices [40]. Third, ethical and reproducibility challenges, such as algorithmic bias against older adults or clinical cohorts, and limited code or data sharing, hinder the translation.

Several systematic reviews have explored the broader field of activity recognition using accelerometers and artificial intelligence (AI). However, a focused synthesis on AI-driven MVPA detection is lacking. Previous reviews have either focused on physical activity type detection in real-life conditions rather than intensity-specific thresholds [41], provided the general methodologies of human activity recognition using wearable sensors and ML without a systematic analysis of performance and bias in MVPA classification [42-44], examined the validation of accelerometer-based monitors using ML but not within the specific context of MVPA's lab-to-real-world gap [27], or highlighted the critical issue of accessibility and reproducibility of novel analytical models but not connected them to the development of equitable MVPA models [35]. Other reviews have focused on specific aspects, such as calibration techniques [45], sport-specific movements [46], or compared DL architectures like CNN and LSTM [47]. While 2 recent

XSL•FO
RenderX

reviews touch on predicting physical activity intensity from smartphones or smartwatches [48,49], they do not encompass the full spectrum of research-grade and wearable sensors, model architectures, and the critical synthesis of translational challenges presented here.

Therefore, this scoping review is the first to systematically scope and synthesize the literature exclusively on ML and DL techniques for MVPA intensity. We uniquely quantify the performance of MVPA detection methods as a function of the sensors used, sensor placement, target populations, feature extraction strategies, model architectures, lab-to-real-world settings, and look at possible algorithmic bias introduced by the restricted age and health status of the tested participants.

Methods

Overview

This scoping review follows the Arksey and O'Malley framework, which includes 5 key stages: identifying the research question (RQ), identifying relevant studies, selecting studies, charting the data, and collating, summarizing, and reporting the results. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) was also consulted to ensure methodological rigor. EndNote X9 (Clarivate Analytics) was used for reference management, deduplication, and the screening process.

Identify the Research Questions

This paper presents a scoping review that synthesizes advancements in ML- and DL-driven MVPA estimation and prediction from accelerometer data. The review aims to answer the following RQs:

- RQ1: What ML and DL techniques have been and are currently used for MVPA detection from accelerometer data?
- RQ2: How do accelerometer specifications (eg, sensor type, sampling rate), body placement (eg, wrist, hip, and thigh), and multisensor configurations influence model performance and generalizability?
- RQ3: What's the magnitude of the performance gap between laboratory-controlled and free-living environments, and what potential factors contribute to this disparity?
- RQ4: How do validation protocols vary across studies, and how do inconsistencies in these protocols limit cross-study comparability and clinical utility?
- RQ5: To what extent do current models exhibit biased results, preventing generalization to older adult or clinical populations?
- RQ6: What proportion of studies adhere to open science practices, and how do transparency gaps hinder reproducibility, scalability, and equitable deployment?

Identify Relevant Studies

To ensure a comprehensive and focused literature search, we used a multi-step process. The search strategy was developed and refined in discussion with all the authors, who have specialized expertise in systematic review methodologies and database search strategies. Initially, we conducted a preliminary manual search to identify eligible studies and determine relevant databases and query terms. The search strategy included the following keywords and their combinations: "artificial intelligence" (eg, "machine learning" and "deep learning"), "accelerometer" (eg, "wearable device," "smartphone," "smartwatch," and "inertial measurement unit (IMU)"), and "moderate-to-vigorous physical activity" (eg, MVPA, "physical activity intensity," and "energy expenditure").

The comprehensive search was conducted across multiple databases, including PubMed, IEEE Xplore, and Web of Science. The search was conducted on April 4, 2025. To further enhance the comprehensiveness of our search, a manual citation search was conducted using reference lists of relevant studies to identify other potentially eligible studies. The detailed search strategies used to find relevant studies for this scoping review are described in Multimedia Appendix 1.

Selection of Eligible Studies

We followed the steps outlined in the PRISMA (the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension guidelines) flow diagram to select eligible studies. The study selection process involved 2 independent reviewers (YZ and SRBVDV) who screened titles and abstracts, followed by full-text assessment. Any discrepancies were resolved through discussion. This process is illustrated in the PRISMA flow diagram. The inclusion criteria were as follows: (1) studies that applied ML or DL technique; (2) studies that used accelerometry, no matter any location, number of sensors, or any type of devices, such as accelerometers or smartphones or smartwatches; (3) studies that estimated of MVPA as the outcome; (4) studies that focused on human, any age group or any health status; and (5) peer-reviewed studies published in English.

The exclusion criteria were as follows: (1) studies that did not involve ML or DL techniques (2) studies that relied on multimodal sensor systems (eg, integrated heart rate monitors with accelerometers) or nonaccelerometric data (eg, video-based estimation); (3) studies that focused on nonhuman or nonphysical activity contexts, such as only differentiating sedentary behavior from nonsedentary activity; (4) studies that focused on general activity recognition or physical activity intensity classification without MVPA-specific analysis; (5) studies that focused on theoretical models without empirical validation; (6) studies that were not peer reviewed or reported in a non-English language; and (7) studies without full text available.

Data Charting

Duplicates were identified and removed using the automated deduplication feature in EndNote X9, which was configured to define duplicates based on matching author, publication year, and title field. This automated process was followed by a manual check to ensure the thoroughness and accuracy of deduplication. Then, guided by the RQs, the following details were extracted from included studies: study characteristics (eg, author, publication year, population characteristics), sensor configuration (brand and model, placement, sampling rate), ML or DL techniques used to detect MVPA (features choosing

XSL•FO
RenderX

strategy, features selected, MVPA classification technique), ground truth validation of MVPA (via indirect calorimetry [IC] or direct observation [DO]), validating setting (lab or free-living conditions), classification performance metrics (F_1-score and accuracy), and code availability. In this review, only the F_1-score and accuracy for MVPA classification were extracted. In cases where these metrics were not explicitly reported in the primary studies, they were inferred from the provided confusion matrices using the standard functions. Accuracy, representing the proportion of total correct predictions, was calculated as:

The F_1-score, the harmonic mean of precision and recall, was calculated as

Where:

In which, TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

Collating, Summarizing, and Reporting the Results

To answer the RQs, the results are organized into 6 sections: evolution of feature engineering and model architectures, task-specific insights, sensor performance, validation practices, algorithmic bias, and reproducibility crisis.

A narrative synthesis approach was used, guided by the predefined RQs. The extracted data were summarized quantitatively (using frequencies and percentages) and qualitatively (identifying key themes and trends). To ensure rigor and trustworthiness, the data charting and initial synthesis were performed by one author (YZ) and critically reviewed by the others (SRBVDV, EJCDG, and PC) for accuracy and consistency.

Ethical Considerations

This scoping review synthesized findings from previously published research involving human participants. No new participants were recruited, and no new primary data were collected for this review. Consequently, separate ethical approval for this specific synthesis was not required. All original studies included in this review were expected to have obtained appropriate ethical approval from their relevant institutional review boards or ethics committees and informed consent from participants, consistent with ethical standards for human participation research involving sensor data. We noted that the majority of included studies explicitly reported ethical approval within their publications. For studies where an explicit ethics statement was not found in the publication, we acknowledge this limitation in reporting transparency. As this review analyzed results reported in published literature and did not involve direct access to or reanalysis of the raw accelerometer data from the original studies, specific data licenses or permissions beyond the published findings were not required.

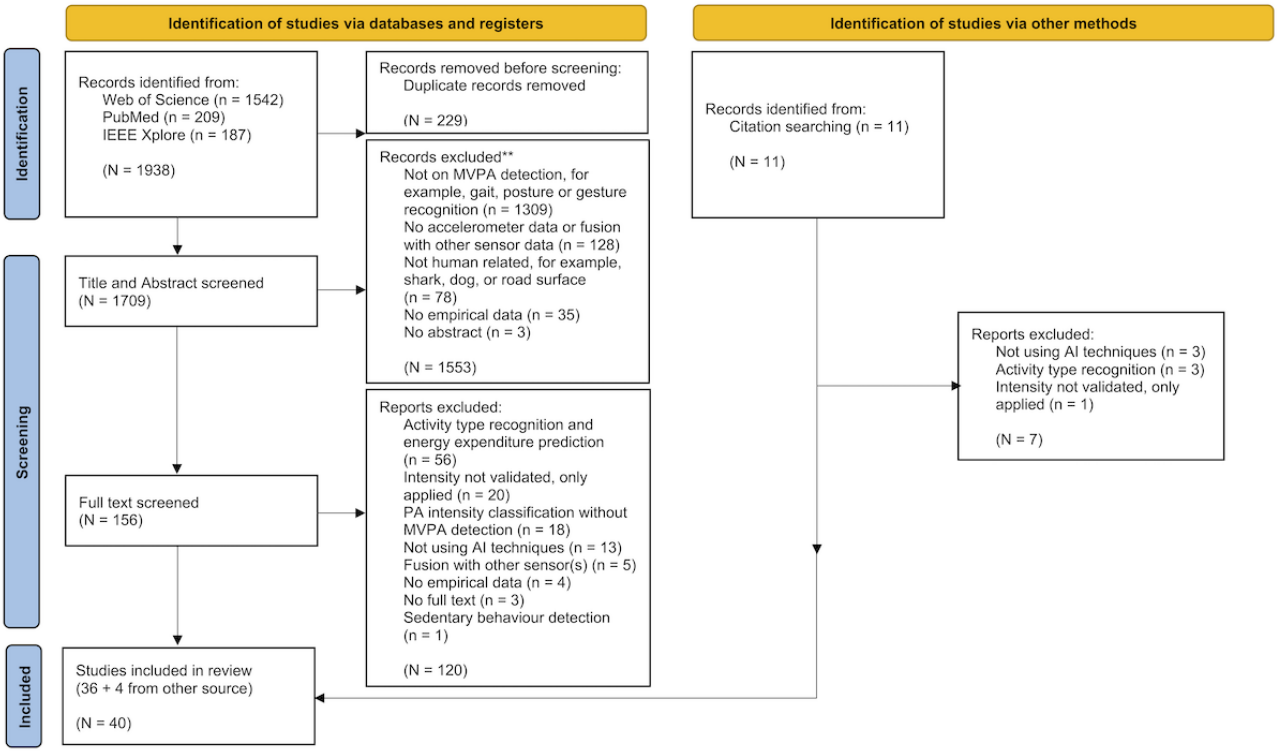
Results

Overview

A total of 1938 articles were identified from PubMed (n=209), IEEE Xplore (n=187), and Web of Science (n=1542). After removing 11.8% (229/1938) of duplicates, 88.2% (1709/1938) of the articles were included in the title and abstract screening phase. After this phase, 156 (8.1%) were screened for eligibility in the full-text screening phase. As a result, 36 articles met the inclusion criteria. In addition, 4 studies were included from manual searches. In total, 40 articles were included in this scoping review, shown in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA flow diagram of study selection. AI: artificial intelligence; MVPA: moderate-to-vigorous physical activity; PA: physical activity.



Overview of Included Studies

Table 1 provides a summary of included studies, including the author's name and publication year, mean age of participants (SD), number of participants, sensor brand and specific model, number of sensors tested, the placement of sensors, MVPA classification techniques used, validation setting (lab or free-living), assessment of MVPA ground truth, and code availability. The information on the country where the test was conducted, sampling rate, feature choosing strategy, window length, feature selected, and F_1-score and accuracy of model detecting MVPA was listed in Multimedia Appendix 2 [19,28,29,31,33,34,36-38,50-81].

XSL•FO
RenderX

Table 1. The summary of included studies (N=40 studies, ranked by health condition and alphabetically by author names).

Reference	Age (years), mean (SD)	Age (years), range	N	Sensor brand/ model	Number of sensors (placement)	MVPA[a] classification technique	Validation setting	Ground truth	Code availability
Healthy condition									
Ahmadi et al [50]	4 (0.9)	—[b]	31	ActiGraph GT3X +	2 (hip, wrist)	RF[c] and SVM[d]	Free-living	DO[e]	No
Ahmadi et al [51]	4 (0.9)	—	31	ActiGraph GT3X +	2 (hip, wrist)	RF	Free-living	DO	Yes
Ahmadi et al [52]	13.9 (3)	—	50	ActiGraph GT3X +	1 (hip)	RF	Lab	PRE[f]	Yes
Ahmadi et al [38]	4 (0.9)	—	31	ActiGraph GT3X +	2 (hip, wrist)	RF	Free-living	DO	Yes
Ahmadi et al [53]	55.8 (12.4)	—	102	MEMS and ActiGraph GT9X	2 (both wrists)	RF	Lab and free-living	DO	No
Ahmadi et al [53]	55.8 (12.4)	—	52	MEMS and ActiGraph GT9X	2 (both wrists)	RF	Lab and free-living	DO, IC[g]	No
Ahmadi et al [53]	—	18-91	151	Axivity AX3	1 (wrist)	RF	Free-living	DO	No
Andò et al [36]	NR[h]	—	NR	LSM9DS1	1 (chest)	k-NN[i] and RF	Lab	PRE	No
Bai et al [54]	72.4 (7.1)	—	247	ActiGraph GT3X-BT	1 (wrist)	XGBoost[j]	Lab	IC	No
Barua et al [34]	29 (0)	18-56	42	Samsung Galaxy S7	3 (pocket, backpack, hand)	1D-CNN-LSTM[k]	Lab	PRE	No
Chen et al [55]	12.3 (1.0)	—	18	smartwatch “mumu”	1 (wrist)	SVM	Free-living	IC	No
Chen et al [55]	24.9 (2.6)	—	24	smartwatch “mumu”	1 (wrist)	SVM	Free-living	IC	No
Davoudi et al [56]	55.2 (17.8)	—	40	ActiGraph GT3X+ and Samsung smartwatch	2 (wrist)	RF	Lab	IC	No
Doherty et al [57]	—	18-91	153	Axivity AX3	1 (wrist)	RF	Free-living	DO	No
Ellingson et al [58]	23.9 (5.3)	—	49	ActiGraph GT3X+ and actiPAL	2 (hip, thigh)	ANN[l] and DT[m]	Lab	IC and PRE	Yes
Ellingson et al [59]	23.5 (4.6)	—	51	ActiGraph GT3X+	1 (hip)	RF	Lab	IC	On request
Farrahi et al [60]	27.5 (11.2); 13.7 (3.1); 27.2 (3.3)	20-30	22; 52; 9; 8	Hookie AM20; Actigraph GT3X+; Colibri inertial measurement unit; Xsens MTx inertial measurement unit	1 (hip); 2 (hip, wrist); 1 (wrist); 2 (wrists)	ANN	Lab	DO	No
Farrahi et al [31]	—	18-91	151	Axivity AX3	1 (wrist)	BiLSTM[n], RF, ANN, SVM, DT, and NB[o]	Free-living	DO	No

XSL•FO
RenderX

Reference	Age (years), mean (SD)	Age (years), range	N	Sensor brand/model	Number of sensors (placement)	MVPA[a] classification technique	Validation setting	Ground truth	Code availability
Freedson et al [61]	38 (12.4)	—	277	ActiGraph GT1M	1 (hip)	ANN	Lab	IC	No
Hagenbuchner et al [29]	4.8 (0.9)	—	11	ActiGraph GT3X +	1 (hip)	ANN	Lab	PRE	No
Hibbing et al [62]	9.4 (2.1)	—	27	ActiGraph GT3X-BT	3 (hip, both wrists)	ANN and DT	Free-living	DO	On request
Hibbing et al [62]	10.0 (2.2)	—	54	ActiGraph GT3X+	2 (hip, wrist)	ANN and DT	Lab	IC	On request
Li et al [63]	4.0 (0.5)	—	34	ActiGraph GT3X-BT	1 (wrist)	k-means[p]	Free-living	Hip cut points	No
Mardini et al [64]	61.7 (17.7)	—	253	ActiGraph GT3X-BT	1 (wrist)	DT, RF, XGBoost, and LASSO[q]	Lab	IC	No
Montoye et al [65]	22.0 (4.2)	—	40	ActiGraph GT3X+ and GENEActiv	4 (thigh, hip, both wrists)	ANN	Lab	PRE	No
Montoye et al [66]	22.0 (4.2)	—	41	activPAL3	1 (thigh)	ANN	Lab	IC	No
Montoye et al [67]	40.8 (19.2)	—	48	ActiGraph GT9X Link	2 (hip, wrist)	RF	Lab and free-living	DO	Yes
Nawaratne et al [68]	45.0 (11.0)	—	119	ActiGraph GT3X +	1 (wrist)	CNN[r]	Free-living	Hip cut points	Yes
Nnamoko et al [69]	69.3 (8.0)	—	33	GENEActiv and ActiGraph	2 (wrist, hip)	Additive regression tree	Lab	IC	No
O'Driscoll et al [70]	44.4 (14.1); 31.9 (10.2)	—	89	ActiGraph GT3-X; SenseWear Armband	2 (wrist, upper arm)	RF, ANN, k-NN, SVM, and gradient boosting	Lab	IC	No
Pober et al [71]	24.8 (4.2)	—	6	Actigraph MTI 7164	1 (hip)	QDA[s] and HMM[t]	Lab	PRE	No
Skjødt et al [72]	80.2 (3.7)	—	67	ActiGraph GT3X +, GENEActiv, and Axivity AX3	6 (both hips, both wrists, thigh, lower back)	RF	Lab	IC	Yes
Staudenmayer et al [73]	35 (0)	21-69	48	Actigraph model 7164	1 (wrist)	ANN	Lab	IC	No
Staudenmayer et al [74]	24.1 (0)	20-39	20	ActiGraph GT3X+	1 (wrist)	RF and DT	Lab	IC	No
Trost et al [19]	11 (2.7)	—	100	ActiGraph GT1M	1 (hip)	ANN	Lab	IC	No
Trost et al [19]	11 (2.7)	—	100	ActiGraph GT1M	1 (hip)	ANN	Lab	IC	No
Trost et al [28]	4.8 (0.9)	—	11	ActiGraph GT3X +	2 (hip, wrist)	RF and SVM	Lab	DO, IC	On request
Tsanas [75]	—	18-91	148	Axivity AX3	1 (wrist)	RF and HMM	Free-living	DO	No
Walmsley et al [76]	—	18-91	152	Axivity AX3	1 (wrist)	RF and HMM	Free-living	DO	No

XSL•FO
RenderX

Reference	Age (years), mean (SD)	Age (years), range	N	Sensor brand/ model	Number of sensors (placement)	MVPA[a] classification technique	Validation setting	Ground truth	Code availability
Wang et al [33]	—	18-91	151	Axivity AX3	1 (wrist)	ViT-BiLSTM[u], CNN-BiLSTM[v], ViT[w], CNN, and BiLSTM	Free-living	DO	No
Wullems et al [77]	73.5 (6.3)	—	40	GENEActiv	2 (both thighs)	RF	Lab	IC	No
Wullems et al [78]	70.0 (12.0)	—	20	GENEActiv	1 (thigh)	RF	Lab	DO, IC	No
Zhou et al [79]	5.0 (0.9)	—	24	Custom inertial measurement unit sensor	1 (arm)	BiLSTM	Lab and free-living	IC	No
Clinical conditions									
Bianchim et al [37]	12.0 (2.8)	—	35^{CF}; 28^{hy}	GENEActiv and ActiGraph	5 (both wrists, waist, both wrists)	k-NN, RF, and XGBoost	Lab	IC	No
Cescon et al [80]	44.9 (5.0)	—	20^{T1Dz}	Empatica E4 wristband	1 (wrist)	RF	Free-living	NR	No

[a]MVPA: moderate-to-vigorous physical activity.
[b]Not applicable.
[c]RF: random forest.
[d]SVM: support vector machine.
[e]DO: direct observation.
[f]PRE: predefined activity schedule.
[g]IC: indirect calorimetry.
[h]NR: not reported.
[i]k-NN: k-nearest neighbor.
[j]XGBoost: extreme gradient boosting.
[k]1D-CNN-LSTM: one directional CNN-LSTM.
[l]ANN: artificial neural network.
[m]DT: decision tree.
[n]BiLSTM: bidirectional long short-term memory.
[o]NB: naive Bayes.
[p]k-means: k-means cluster analysis.
[q]LASSO: least absolute shrinkage and selection operator.
[r]CNN: convolutional neural network.
[s]QDA: quadratic discriminant analysis.
[t]HMM: hidden Markov model.
[u]ViT-BiLSTM: vision transformer bidirectional long short-term memory.
[v]CNN-BiLSTM: convolutional neural network and bidirectional long short-term memory.
[w]ViT: vision transformer.
[x]35^{CF}: 35 participants with cystic fibrosis.
[y]28^{h}: 28 healthy participants in the study.
[z]20^{T1D}: 20 participants with type 1 diabetes.

A total of 40 studies (2006-2025) met the inclusion criteria, with 62.5% (n=25) published between 2020 and 2025, reflecting the growing interest in AI-driven MVPA estimation.

Most studies (37/40 studies, 92.5%) targeted healthy populations, while only 5% (2/40 studies) addressed clinical cohorts, that is, cystic fibrosis [37] and type 1 diabetes [80], with one study (2.5%) did not specify the characteristics of participants [36].

Eleven studies (27.5%, 11/40) focused on children and adolescents [19,28,29,38,50-52,55,62,63,79], 40% (16/40) were

XSL•FO
RenderX

on adults (18-60 years old) [34,53,55,56,58,59,61,65-68,70,71,73,74], and 17.5% (7/40) were on old adults (60 years or older) [54,64,69,72,77,78], and 2 studies reported on the clinical conditions (ie, cystic fibrosis [37] and type 1 diabetes [80]). The remaining 7 studies (17.5%) tested their models using public datasets in adults, such as Capture-24 and Energy-24 (age range 18-91 years) [31,33,53,57,75,76] and a study with multiple datasets, including UOULU (University of Oulu), OSU (Oregon State University), the PAMAP2 Physical Activity Monitoring dataset (the UCI Machine Learning Repository), and the Daily and Sports Activities (the UCI Machine Learning Repository) [60]. Among these, Chen et al [55] covered the analyses both on children and adults; Andò et al [36], though not reporting participant ages, was contextually aligned with older adult research due to its emphasis on age-associated risks of physical inactivity among older adults and its heavy reliance on references related to older adults; Capture-24 and Energy-24 datasets [31,33,53,57,75,76] were grouped into adult, due to the age distributions: 72% of participants were younger than 53 years, with only 22.5% aged 53 years or older [82]; Farrahi et al [60], which included 4 datasets with an average participant age of about 19 years, was classified under adults.

ActiGraph (30/40, 75%) and GENEActiv (6/40, 15%) were the most common sensors using acceleration sensors to identify MVPA, with limited use of consumer wearables, for example, other brands of accelerometers (eg, Axivity AX3, activPAL; 9/40, 22.5%), inertial measurement units (5/40, 12.5%), smartwatches (2/40, 5%), smartphones (1/40, 2.5%), wristbands (1/40, 2.5%), and armbands (1/40, 2.5%).

Lab-controlled validations predominated (30/45 analyses, 66.7%; some studies had multiple analyses); 33.3% of analyses (15/45) were conducted in free-living conditions, while 4 analyses combined lab and free-living validations.

Evolution of Feature Engineering and Model Architectures

A total of 45 analyses from 40 studies used a range of ML and DL techniques for MVPA detection.

Methodological Evolution

The shift from feature-driven ML to end-to-end DL reflects a broader trend toward scalability and generalizability. While traditional ML models excel in interpretability and low computational cost, their dependence on handcrafted features renders them brittle in free-living contexts. In contrast, DL architectures, though data-hungry and opaque, inherently adapt to signal variability through hierarchical abstraction, a critical advantage for real-world deployment [31].

The progression from manual feature engineering to automated DL underscores a paradigm shift toward scalable, context-aware MVPA monitoring. Table 2 synthesizes this evolution, contrasting supervised, unsupervised, and hybrid paradigms. Supervised DL models, particularly those using transfer learning, now dominate, with all studies adopting pretrained CNN or bidirectional long short-term memory (BiLSTM) to mitigate data scarcity [31,33,34,68,79]. Unsupervised approaches, such as the self-organizing map and k-means cluster analysis, remain nascent but offer potential for leveraging unlabeled free-living data [83,84].

XSL•FO
RenderX

Table 2. Task-specific performance comparison.

Task type and methods/ model	Key features	Performance metrics (number of studies)	References
Classification (n=28)			
RF[a] (n=13)	Handcrafted features: time/frequency features (eg, mean, SD, percentiles, lag-1 autocorrelation), ensemble of decision trees	• Lab (n=7): F1-score 91.9%, accuracy 94.0% • Free-living (n=4): F1-score 81.0%, accuracy 87.4% • Lab and free-living (n=2): F1-score 88.1%, accuracy 93.8%	[28,36-38,50-53,64,67,72,74,80]
ANN[b] (n=7)	Handcrafted features: time/frequency features (eg, spectral entropy, signal power), multilayer perceptron	• Lab (n=7): F1-score 88.0%, accuracy 93.1% • Free-living (n=1): F1-score 75.4%, accuracy 82.1%	[19,29,58,60,62,65,73]
SVM[c] (n=4)	Kernal-based classification on RBF[d], advanced cross-correlation metrics (xy, xz, yz)	• Lab (n=1): F1-score 75.4%, accuracy 88.4% • Free-living (n=3): accuracy 86.5%	[28,50,55]
DT[e] (n=4)	Tree-based splits, integrate with ANN outcomes	• Lab (n=3): F1-score 86.6%, accuracy 87.8% • Free-living (n=1): F1-score 75.4%, accuracy 82.1%	[58,62,64,74]
Gradient boosting (n=3)	Gradient boosting framework, handling missing data	• Lab (n=2): F1-score 91.6%	[37,54,64]
HMM[f] (n=3)	Temporal sequence modeling, Viterbi smoothing	• Lab (n=1): F1-score 99.8% • Free-living (n=2): F1-score 73.5%, accuracy 94.0%	[71,75,76]
QDA[g] (n=1)	Quadratic decision boundaries, probabilistic classification	• Lab (n=1): F1-score 100%, accuracy 99.9%	[71]
LASSO[h] (n=1)	L1 regularization, sparse solutions	• Lab (n=1): F1-score 83.6%	[64]
CNN[i] (n=1)	Automated feature extraction via convolutional filters on raw signals	• Free-living (n=1): F1-score 73.4%, accuracy 96.8%	[68]
Estimation (n=10)			
RF (n=6)	Regression trees, bootstrapped subsets of ActiGraph data	• Lab (n=5): F1-score 83.5%, accuracy 86.1% • Free-living (n=1): F1-score 80.0%, accuracy 91.4%	[56,57,59,70,77]
ANN (n=2)	Nonlinear activation functions, raw signal processing	• Lab (n=2): F1-score 91.1%, accuracy 85.7%	[61,66]
SVM (n=1)	Kernal-based regression.	• Lab (n=1): F1-score 90.7%, accuracy 88.7%	[70]
k-NN[j] (n=2)	Instance-based learning, Euclidean distance metrics	• Lab (n=2): F1-score 96.4%, accuracy 95.8%	[37,70]
XGBoost[k] (n=1)	Gradient boosting framework, handling missing data	• Lab (n=1): F1-score 100%, accuracy 100%	[37]
Gradient boosting (n=1)	Iterative error correction, additive regression trees	• Lab (n=1): F1-score 93.2%, accuracy 92.1%	[70]
Deep learning (n=5)			

XSL•FO
RenderX

Task type and methods/ model	Key features	Performance metrics (number of studies)	References
Bi-LSTM[l] (n=3)	Bidirectional temporal modeling, raw signal processing	• Free-living (n=2): F1-score 73.6%, accuracy 93.6% • Lab and free-living: F1-score 53.3%, accuracy 53.7%	[31,33,79]
CNN (n=2)	Automated feature extraction via convolutional filters on raw signals	• Free-living (n=2): F1-score 71.9%, accuracy 94.4%	[33,68]
ViT[m] (n=1)	Self-attention mechanisms for long-range dependencies	• Free-living (n=1): F1-score 79.8%, accuracy 95.0%	[33]
CNN-LSTM[n] or CNN-BiLSTM[o] (n=2)	Hybrid architecture, integrate spatial and temporal learning	• Lab (n=1): F1-score 82.1% • Free-living (n=1): F1-score 91.4%, accuracy 97.7%	[33,34]
ViT-BiLSTM[p] (n=1)	Vision Transformer + BiLSTM, gravity-based acceleration analysis.	• Free-living (n=1): F1-score 98.4%, accuracy 99.0%	[33]

[a]RF: random forest.

[b]ANN: artificial neural network.

[c]SVM: support vector machine.

[d]RBF: radial basis function.

[e]DT: decision tree.

[f]HMM: hidden Markov model.

[g]QDA: quadratic discriminant analysis.

[h]LASSO: least absolute shrinkage and selection operator.

[i]CNN: convolutional neural network.

[j]k-NN: k-nearest neighbor.

[k]XGBoost: extreme gradient boosting.

[l]BiLSTM: bidirectional long short-term memory.

[m]ViT: vision transformer.

[n]CNN-LSTM: convolutional neural network and bidirectional long short-term memory.

[o]CNN-BiLSTM: convolutional neural network and bidirectional long short-term memory.

[p]ViT-BiLSTM: vision transformer bidirectional long short-term memory.

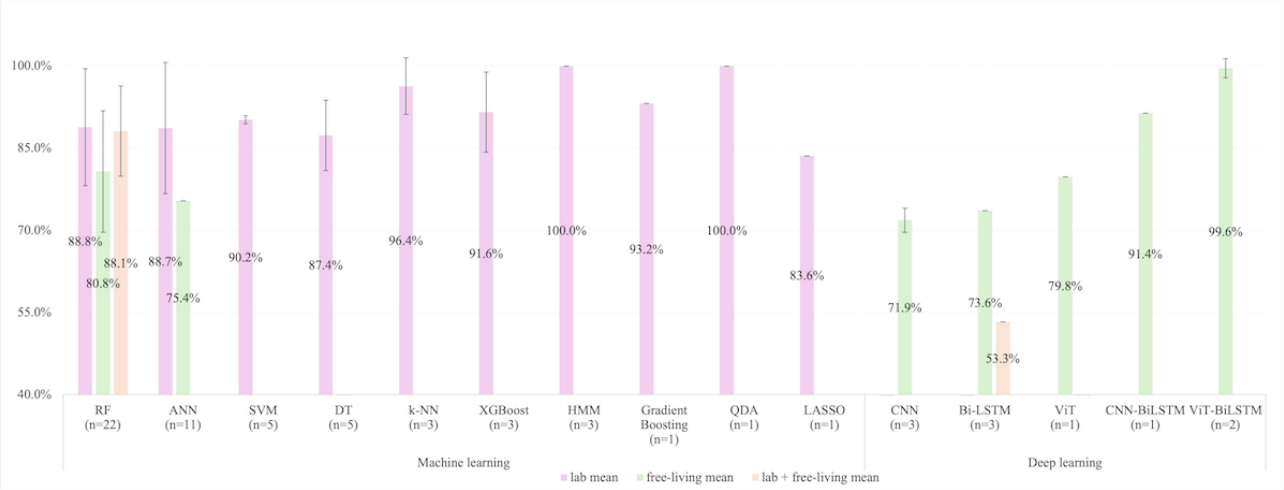
Traditional Machine Learning

Traditional ML techniques have dominated accelerometer-based MVPA detection since 2006 [71], relying on handcrafted features derived from time- and frequency-domain analyses.

Among these, RF emerged as the most prevalent algorithm (22/40 studies, 55%) [28,36-38,50-53,56,57,59,64,67,70,72,74-78,80], achieving mean F_1-scores of 86.6% and mean accuracy of 88.6%. RF's ensemble structure, which aggregates predictions from multiple decision trees (DTs) (usually 100-1000), mitigates overfitting and enhances robustness to noise, a critical advantage in heterogeneous accelerometer datasets [67,76]. Artificial neural network (ANN) followed (11/40 studies, 27.5%), with mean F_1-scores of 87.4% and mean accuracy of 89.5% [19,29,31,58,60-62,65,66,70,73]. ANN used a multi-layer perceptron with input, hidden (3-25 nodes), and output layers, nonlinear activation functions model complex feature interactions. Another often-used method was the SVM (5/40 studies, 12.5%), with an F_1-score of 90.2% and accuracy of 86.5% [28,31,50,55,70]. SVM maps features to a high-dimensional space and constructs optimal hyperplanes using kernel functions (eg, radial basis function). The F_1-scores of models using DT (5/40 studies, 12.5%; F_1-score 87.4%) [31,58,62,64,74] underperformed k-nearest neighbor (3/40 studies, 7.5%; F_1-score 96.4%) [36,37,70], extreme gradient boosting (3/40 studies, 7.5%; F_1-score 91.6%) [54,64,80], hidden Markov model (3/40 studies, 7.5%; F_1-score 100%) [71,75,76], Gradient Boosting (1/40 studies, 2.5%; F_1-score 93.2%) [70], and quadratic discriminant analysis (1/40 studies, 2.5%; F_1-score 100%) [71]. Least absolute shrinkage and selection operator achieved the lowest F_1-score (83.6%) in detecting MVPA among all the ML models [64]. The details are illustrated in Figure 2.

XSL•FO
RenderX

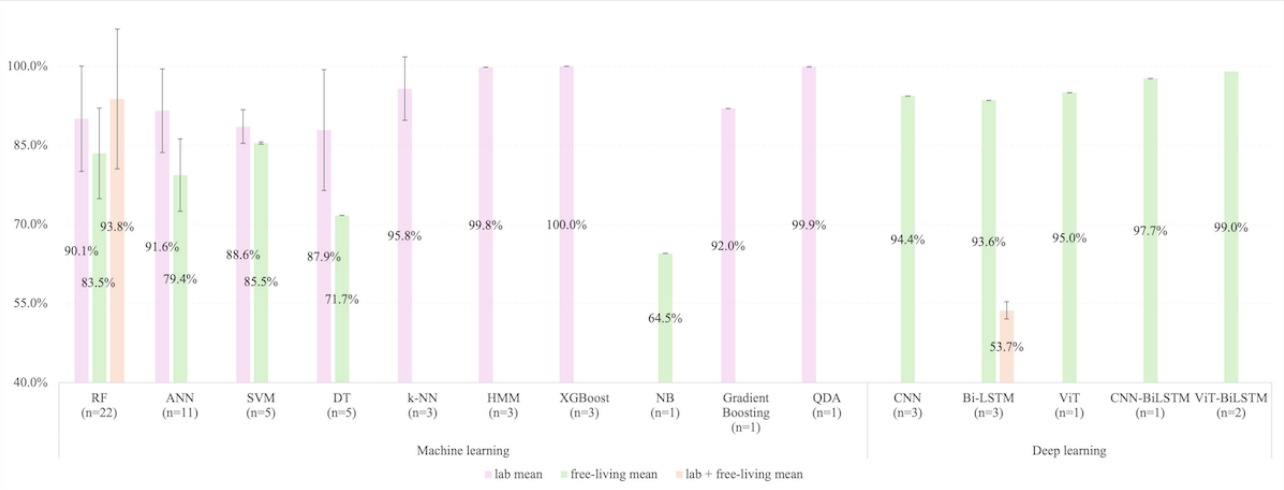
Figure 2. The extracted F1-score for moderate-to-vigorous physical activity from machine learning and deep learning models. Error bars, where applicable, represent SD. "n" means the number of studies using the model. ANN: artificial neural network; BiLSTM: bidirectional long short-term memory; CNN: convolutional neural network; CNN-BiLSTM: convolutional neural network and bidirectional long short-term memory; DT: decision tree; HMM: hidden Markov model; k-NN: k-nearest neighbor; LASSO: least absolute shrinkage and selection operator; QDA: quadratic discriminant analysis; RF: random forest; SVM: support vector machine; ViT: vision transformer; ViT-BiLSTM: vision transformer bidirectional long short-term memory; XGBoost: extreme gradient boosting.



However, performance disparities between lab-controlled and free-living environments underscored inherent limitations. RF models, for instance, exhibited a decline of 8.0% in F_1-score (88.8% lab vs 80.8 free-living) and 6.6% in accuracy (90.1% lab vs 83.5% free-living), attributed to over-reliance on static features (eg, variance, spectral entropy) that fail to generalize to unstructured movement patterns [27]. Similarly, ANN experienced reduced accuracy in free-living contexts (F_1-score 88.7% lab vs 75.4% free-living; accuracy 91.6% lab vs 79.4% free-living), highlighting sensitivity to signal variability introduced by nonexercise movements (eg, gesturing, device placement) [27]. The rest of the algorithms had no free-living validation.

Figures 2 and 3 stratify F_1-score and accuracy by model type, revealing that simpler algorithms like k-nearest neighbor and quadratic discriminant analysis achieved near-perfect lab performance (96.4%-100%) but performed less well in free-living validations.

Figure 3. The extracted accuracy for moderate-to-vigorous physical activity from machine learning and deep learning models. Error bars, where applicable, represent SDs. "n" means the number of studies using the model. ANN: artificial neural network; BiLSTM: bidirectional long short-term memory; CNN: convolutional neural network; CNN-BiLSTM: convolutional neural network and bidirectional long short-term memory; DT: decision tree; HMM: hidden Markov model; k-NN: k-nearest neighbor; NB: naive Bayes; QDA: quadratic discriminant analysis; RF: random forest; SVM: support vector machine; ViT: vision transformer; ViT-BiLSTM: vision transformer bidirectional long short-term memory; XGBoost: extreme gradient boosting.



Deep Learning

DL architectures revolutionized MVPA detection by automating hierarchical feature extraction from raw accelerometer signals, circumventing the manual feature selection bottleneck. CNN was used in 3 studies (3/40 studies, 7.5%), achieving a mean F_1-score of 71.9% (shown in Figure 2) and a mean accuracy of 94.4% (shown in Figure 3) in free-living conditions [33,34,68]. Their layered structure, comprising convolutional filters (64, 128, 256, and 512 filters), pooling layers, and activation functions (such as rectified linear unit), enables granular analysis of signal dynamics.

XSL•FO
RenderX

Recurrent architectures, notably LSTM and BiLSTM, addressed temporal complexity in sustained MVPA bouts (eg, 10 min). BiLSTM, which processes sequences bidirectionally, achieves an average of F_1-score 71.9% and accuracy 80.3% (shown in Figures 2 and 3) in 3 studies (of 40 studies, 7.5%) (2 in free-living settings, and 1 combining both lab and free-living settings) by modeling contextual transitions (eg, walking-to-jogging) [31,33,79]. Transformers, though less prevalent, demonstrated promise in capturing long-range dependencies through self-attention mechanisms. When hybridized with vision, vision transformer (ViT), the accuracy of detecting MVPA is 95.0% (F_1-score 79.8%) in a free-living validating setting in 1 out of 40 studies (2.5%) [33].

Hybrid models (eg, convolutional neural network and bidirectional long short-term memory [CNN-BiLSTM], vision transformer bidirectional long short-term memory), in 2 out of 40 studies (5%), synergized spatial and temporal learning, achieving peak F_1-score (95.5%) [34] and peak accuracy (98.4%) in free-living settings [33].

Task-Specific Insights

Classification

Classification tasks in accelerometer-based MVPA research involve assigning discrete intensity classes, such as sedentary, light, moderate, or vigorous activity. Early methodologies predominantly used traditional ML algorithms, such as RF and SVM, which relied on handcrafted features like signal variance, spectral entropy, and movement counts. For instance, RF achieved robust performance in laboratory settings (F_1-score mean 91.9%, Table 3) by aggregating predictions from DTs trained on bootstrapped subsets of ActiGraph data. However, these models struggled with ambiguity in free-living environments, particularly in distinguishing light-intensity activities (eg, slow walking at 2.5 METs) from MVPA (≥3 METs). SVMs with radial basis function kernels, while effective in lab-annotated running protocols (F_1-score 75.4%) [28], misclassified 9.0% of slow walking (or stroll) bouts as MVPA in unstructured settings due to overlapping signal patterns [50].

Table 3. Taxonomy of machine learning and deep learning technologies for moderate-to-vigorous physical activity (MVPA) detection, categorized by learning paradigm.

Learning paradigm	Key features	Algorithms	Strengths	Limitations	References
Supervised	Require labeled data (activity intensity labels)	RF[a], ANN[b], SVM[c], DT[d], XGBoost[e], HMM[f], QDA[g], LASSO[h], k-NN[i], and gradient boosting	• High accuracy with sufficient labeled data • Interpretable feature importance, Robust to noise and nonlinear patterns	• Dependency on large, labeled datasets • Overfitting risk • Poor generalization to free-living environments	[28,29,33,36-38,50-60, 62,64-67,69-72,74-78,80]
Unsupervised	Work with unlabeled data, focus on clustering or feature learning	k-means, SOM[j], and autoencoders	• No need for labeled data • Identifies hidden patterns in raw signals • Reduces dimensionality	• Limited direct applicability to MVPA classification • Lower accuracy for intensity-specific tasks • Interpretability challenges	[19,29,63]
Hybrid	Combine supervised and unsupervised components, integrates multiple architectures	CNN-BiLSTM[k], ViT-BiLSTM[l], DLEN[m], and multitask learning frameworks	• Capture spatial and temporal dependencies • Improve generalizability • State-of-the-art performance in free-living settings	• High computational complexity • Require large datasets • Synchronization challenges for multisensor data	[29,31,33,34,68,79]

[a]RF: random forest.

[b]ANN: artificial neural network.

[c]SVM: support vector machine.

[d]DT: decision tree.

[e]XGBoost: extreme gradient boosting.

[f]HMM: hidden Markov model.

[g]QDA: quadratic discriminant analysis.

[h]LASSO: least absolute shrinkage and selection operator.

[i]k-NN: k-nearest neighbor.

[j]SOM: self-organizing maps.

[k]CNN-BiLSTM: convolutional neural network and bidirectional long short-term memory.

[l]ViT-BiLSTM: vision transformer bidirectional long short-term memory.

[m]DLEN: deep learning ensemble network.

XSL•FO
RenderX

DL architectures, particularly CNN, addressed these limitations by automating hierarchical feature extraction directly from raw accelerometer signals. By applying convolutional filters to sliding windows of raw data, CNN detected local biomechanical patterns (eg, stride frequency during running), achieving parity with traditional hip-based cut point methods in MVPA detection by the wrist-model [68]. Subsequent advancements, such as Transformer architectures, further improved classification accuracy by using self-attention mechanisms to model long-range dependencies, outperforming CNN by 9.5% in free-living scenarios [33].

Estimation

Estimation tasks focus on predicting energy expenditure metrics (eg, METs) through regression-based models to map accelerometer signals to continuous outcomes. Conventional approaches, such as linear regression-derived cut points (eg, Freedson equations), exhibited significant limitations due to oversimplified assumptions about the relationship between acceleration signals and MET values, especially for free-living activity [65,81]. ML models, such as an additive regression tree, lowered the standard error of estimation by 0.33-22.11 in lab settings using ActiGraph data [69].

DL architectures, like BiLSTM, elevated estimation accuracy by capturing temporal dependencies in accelerometer signals (eg, MET fluctuations during exercise recovery). BiLSTM achieved a mean absolute error of 0.757, with LSTM as the baseline method [79].

Deep Learning

DL frameworks bridge the gap between classification and estimation by unifying feature extraction and task-specific learning with end-to-end frameworks. Multitask architectures, such as AccNet24, integrate BiLSTM layers for activity intensity with fully connected layers for MET prediction, achieving 97.7% accuracy in MVPA detection in free-living settings [31].

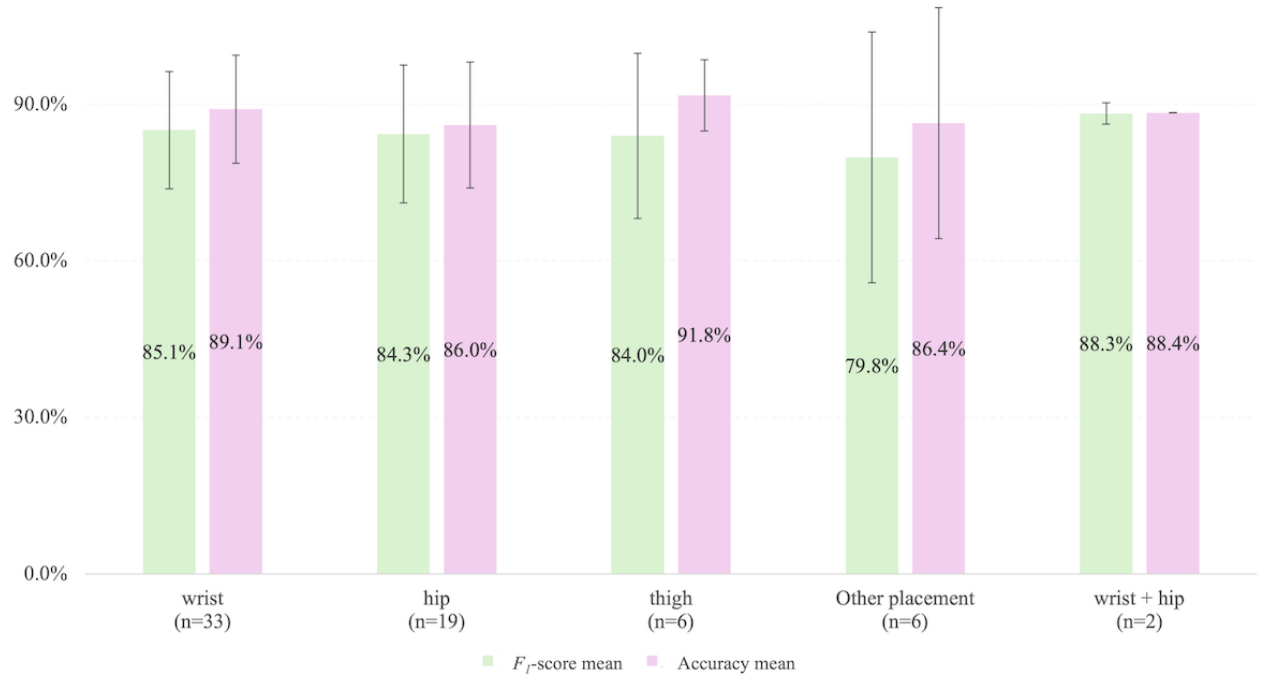
ViTs further optimized task performance via attention mechanisms that dynamically prioritized critical signal regions. For example, ViT allocated closer attention to peak acceleration intervals during jumping, outperforming Bi-LSTM by 6.2% (F_1-score) in free-living MVPA detection [33]. However, these advancements come with trade-offs; hybrid CNN-BiLSTM models require much more training time than traditional RF, limiting real-time deployment on wearables [31].

Sensor Performance

Sensor Placement

The efficacy of accelerometer-based MVPA estimation is significantly influenced by sensor placement. Figure 4 illustrates the averaged performance metrics (F_1-scores and accuracy) across sensor placements.

Figure 4. Mean value of F1-score and accuracy in relation to sensor placement across all validation settings. Error bars, where applicable, represent SDs. Other placements included chest, lower back, backpack, pocket, hand, and upper arm. “wrist + hip” means the multi-sensor configuration using both wrist and hip placements.



Regarding sensors placement, wrist-worn devices dominated in 75% of studies (30/40 studies), followed by hip (19/40, 47.5%), thigh (6/40, 15%), chest (1/40, 2.5%), lower back (1/40, 2.5%), backpack (1/40, 2.5%), pocket (1/40, 2.5%), hand (1/40, 2.5%), upper arm (1/40, 2.5%). Multisensor configurations (eg, wrist + hip) were applied in only 5% (2/40) of studies [28,51]; only the combination of wrist and hip was applied.

Comparative analyses on MVPA revealed that wrist-, hip-, and thigh-worn sensors exhibited comparable mean performance metrics across all validation settings (mean F_1-scores 85.1% vs 84.3% vs 84%, accuracy mean 89.1% vs 86% vs 91.8%) as well

XSL•FO
RenderX

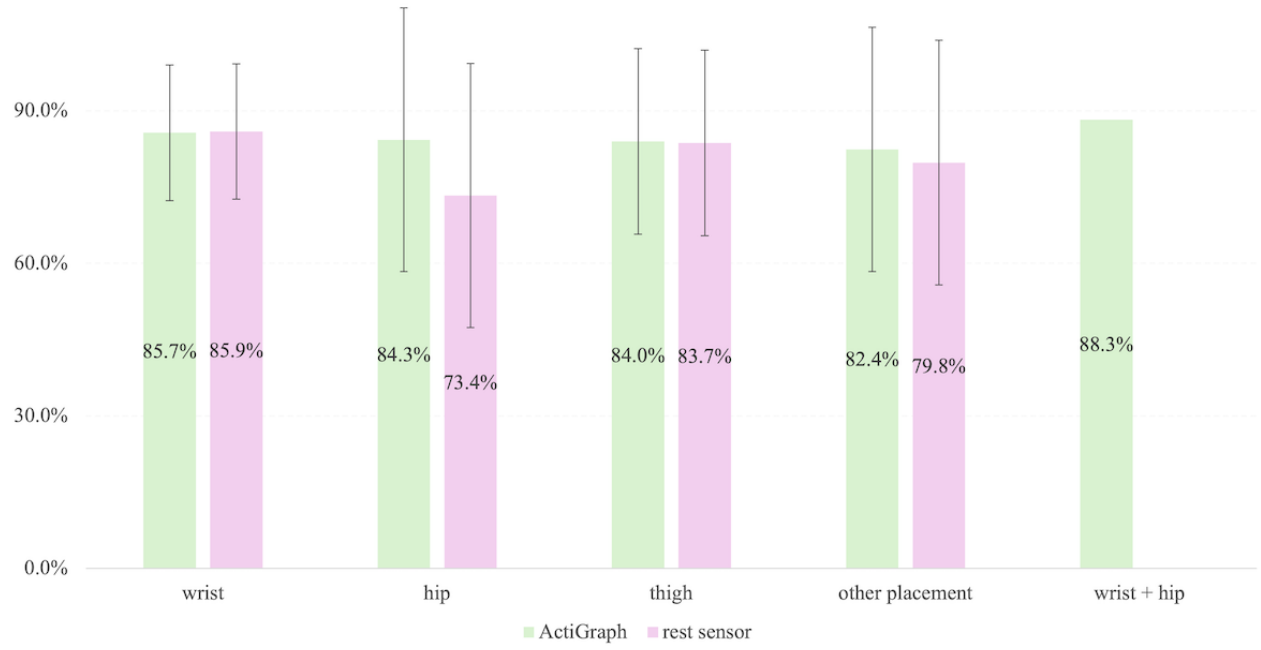
as in laboratory-controlled settings (F_1-score mean 88.6% vs 86.6% vs 86.5%, accuracy mean 91.7% vs 89.6% vs 97.3%). However, disparities emerged in free-living environments. While wrist- and hip-worn sensors demonstrated similar F_1-scores (80.3% vs 79.0%), wrist-worn devices achieved superior accuracy (86.3% vs 70.8%). This discrepancy may stem from the wrist's ability to capture a broader range of upper-body movements associated with MVPA in unstructured environments, such as arm swings during brisk walking or lifting activities, which are less pronounced in hip-worn sensors.

Notably, multisensor configurations (eg, wrist + hip) achieved the highest performance (F_1-score 88.3%, with 89.7% in lab and 86.8% in free-living; accuracy 88.4%, with 88.4% in lab but nonreports in free-living), bridging the gap between controlled and free-living settings. However, practical challenges, including increased participant burden due to multiple devices and synchronization complexities between heterogeneous sensors, limit their widespread adoption.

Sensor Type and Performance Heterogeneity

Device specifications and sensor type further influenced model generalizability. Figure 5 shows the F_1-scores in relation to sensor type (ActiGraph and other types) and sensor placement. While models trained on ActiGraph data achieved a higher mean F_1-score (84.9% vs 83.1%), models using consumer-grade wearables surprisingly achieved a higher mean accuracy (91.8% vs 87.8%). This disparity may arise from the inherent class imbalance in free-living data, where MVPA represents a minority of activities. Accuracy can be inflated by correctly classifying the predominant sedentary and light activities, whereas the F_1-score provides a more balanced measure of performance specifically for the MVPA class. The higher F_1-score associated with ActiGraph models suggests they may be more adept at correctly identifying true MVPA bouts, which is critical for public health monitoring.

Figure 5. Mean value of F1-score in relation to sensor type and placement across all validation settings. Error bars, where applicable, represent SDs.



Validation Practices

Ground Truth Methodologies and Their Implications

Validation of ML and DL models for MVPA detection relies heavily on the specification of ground truth, with IC and DO predominating. IC (21/40 studies, 52.5%), considered the gold standard for energy expenditure measurement, provides MET values through oxygen consumption analysis, enabling precise alignment of accelerometer signals with intensity thresholds (eg, ≥3 METs) [19,28,37,53-56,58,59,61,62,64,66,69,70,72-74,77-79]. However, its laboratory-bound nature limits ecological validity, as structured protocols often fail to replicate free-living movement variability.

In contrast, DO (16/45 analyses, 35.6%) offers real-world applicability by annotating activities in a naturalistic setting but introduces subjectivity, particularly in distinguishing borderline intensities with 2 stages. At the first stage, participant movements were categorized into activity type (eg, sedentary, standing utilitarian tasks, walking, and running) using recordings [28,38,50,51,60,62,65,67,78] and time-stamped images from wearable cameras (eg, combining the usage of a diary in Capture-24) [31,33,53,57,75,76]. At the second stage, physical intensity was coded using references, mainly the Compendium of Physical Activities (sedentary, light, moderate, and vigorous) and Children's Activity Rating Scale (5 categories from stationary/motionless to fast translocation).

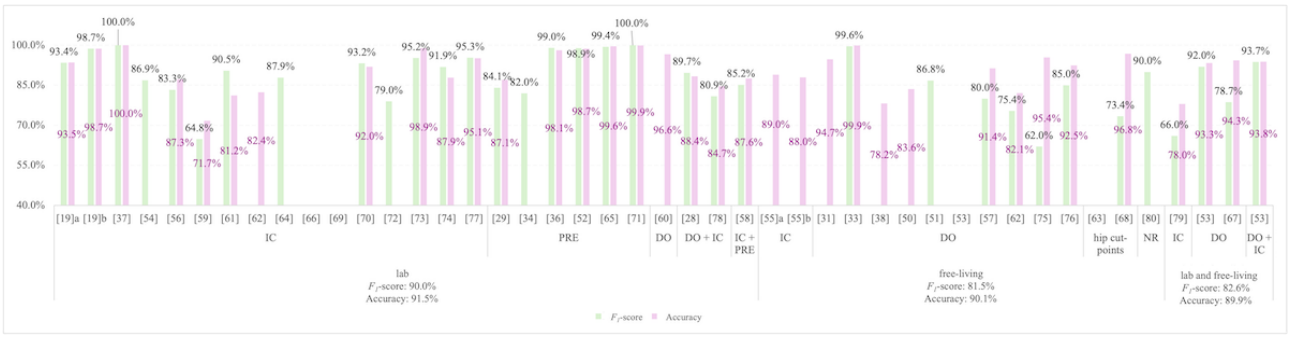
Additionally, reliance on hip-reference cut points as proxies for ground truth (2/45 analyses, 4.4%) perpetuates circular validation, wherein models trained on threshold-based labels inherit the biases of traditional regression methods [63,68]. A total of 15.6% (7/45) of analyses used a predefined activity

XSL•FO
RenderX

schedule in the validation process to define the ground truth [29,36,52,58,71]. Only 11.1% (5/45) of analyses used combined ground truth approaches (eg, IC + DO) [28,53,78], despite evidence that combined methods improve F_1-score by 8.3%-27.7% in free-living compared with IC or DO [53].

Figure 6 illustrates the distribution of ground truth methods across studies, stratified by validation setting. Lab-based studies disproportionately favored IC (19/26 analyses, 73%), while free-living validations leaned on DO (10/15 analyses, 66.7%), with lab validations outperforming free-living (F_1-score mean difference: 8.5%, accuracy mean difference: 1.4%).

Figure 6. F1-score and accuracy metrics for moderate-to-vigorous physical activity (MVPA) classification across studies, stratified by ground truth methods and validation settings. [19]a represents the MVPA classification conducted based on the 10s and 60s window lengths, while [19]b only on the 60s window length. [55]a represents the MVPA classification conducted among the children and adolescents with a mean age of 12.3 (SD 1.0) years, while [55]b represents the MVPA classification conducted among adults aged 24.9 (SD 2.6) years. The mean value of F1-score and accuracy in each group (lab, free-living, and lab and free-living) was shown underneath their names [19,28,29,31,33,34,36-38,50-80]. DO: direct observation; hip: hip reference cut-points; IC: indirect calorimetry; NR: not reported; PRE: predefine activity schedule.



Cross-Validation Protocols and Performance Metrics

The trained models typically undergo rigorous evaluation during the model evaluation and validation phase to verify their generalizability and practical applicability. This process is essential to systematically assess classification accuracy across activity intensities and validate reliability under diverse user scenarios. To mitigate overfitting and ensure model robustness, *k*-fold cross-validation (11/45 analyses, 24.4%) was most commonly implemented. It partitions the dataset into *k* equally sized subsets, iteratively designating one subset as the validation set and the remaining *k*-1 subsets for training. The process is repeated *k* times to ensure all data points contribute to both training and validation. Common configurations include 10-fold and *5*-fold cross-validation, which enhance model generalizability by reducing sensitivity to specific training instances. By systematically evaluating performance across varied data partitions, this method strengthens the activity intensity classification system's reliability and mitigates overfitting, a phenomenon where models memorize training data artifacts rather than learning generalizable patterns.

Leave-one-out cross-validation, which iteratively holds out each individual data point as a test set to evaluate model performance, was also used in 13.3% of analyses (6/45 analyses). Leave-one-subject-out cross-validation, an extension of leave-one-out cross-validation designed for datasets with multiple subjects, iteratively holds out all data from one subject as the test set while training on the remaining participants. This method, used in 24.4% of analyses (11/45), was critical for assessing interindividual generalizability. Its variant, leave-10-subject-out cross-validation, appeared less frequently (1/45 analysis). In contrast, nested cross-validation, which separates hyperparameter tuning from final performance evaluation to prevent data leakage, was sparingly adopted (3/45 analyses, 6.7%) in studies.

Model efficacy was quantified using precision, recall, accuracy, and F_1-score, with metrics calculated iteratively to ensure objective assessment. F_1-score (33/45 analyses, 73.3%) and accuracy (35/45 analyses, 77.8%) were emphasized in this review, though their interpretation varied widely. Notably, studies reporting accuracy exceeding 90.0% often excluded transitional activities (eg, sit-to-stand) [19,37,65,70] or used imbalanced datasets [28,71,80], potentially inflating scores. Conversely, F_1-scores below 75.0% typically correlated with free-living validations, where non-MVPA movements confounded detection [62,68,75].

In the studies not reporting the metrics of MVPA detection, but including a confusion matrix, we calculated F_1-scores and accuracy values from the confusion matrix [28,29,33,36,52,53,57,58,61,62,67,68,70,71,73-78]. Two studies reported the F_1-scores of moderate physical activity (MPA) and vigorous physical activity (VPA) separately without a confusion matrix, so they averaged MPA and VPA to get MVPA [72,79].

Four studies omitted both F_1-score and accuracy [53,63,66,69]. For instance, Li et al [63] reported only the overall accuracy of physical activity intensity classification, while Montoye et al [66] quantified MVPA error (+1.8 min) relative to IC. Ahmadi et al [53] only provided sensitivity (MPA: 80.0%, VPA: 90.0%) and precision (MPA≈75.0%, VPA≈99.0%), and Nnamoko et al [69] reported only the standard error for estimation of personalized cut points.

Algorithmic Bias

The performance of machine and DL models for MVPA detection is inherently tied to the physiological characteristics of the training populations. These have been mostly young, healthy adults. Persistent algorithmic bias induced by using this group can undermine the generalizability of models across the older adult and clinical cohorts.

XSL•FO
RenderX

Figures 7 and 8 stratified MVPA detection performance (F_1-scores and accuracy) by age group (children and adolescents, adults younger than 60 years, adults aged 60 years or older, and clinical populations) and sensor placement (wrist, hip, thigh, other). Among children and adolescents (11/40 studies, 27.5%), both F_1-scores (53.5%-98.9%) and accuracy (52.1%-98.7%) varied widely [19,28,29,38,50-52,55,62,63,79]. Adults under 60 years (20/40 studies, 50%) exhibited consistently high performance (F_1-score mean 85.8%, accuracy mean 91.7%) [31,33,34,53,55-59,61,65-68,70,71,73-76], while older adults (60 years or older) in 15% of studies (6 out of 40 studies) showed relatively reduced score (F_1-score mean 72.3%, accuracy mean 89.9%) [54,64,69,72,77,78]. Clinical populations (2/40 studies, 5%), on the other hand, achieved near-perfect scores (F_1-score 97.6%-100%, accuracy 87.9%-100%), though limited studies (n=2, one in mild cystic fibrosis, and one in type 1 diabetes) necessitate cautious interpretation [37,80].

Figure 7. F1-scores for moderate-to-vigorous physical activity detection across age groups and sensor placements. Brackets ("[]") represent the reference numbers; asterisks ("*") indicate lab-validated results [19,28,29,33,34,36,37,51-54,56-59,61,62,64,65,67,68,70-80]. CF: cystic fibrosis; T1D: type 1 diabetes.

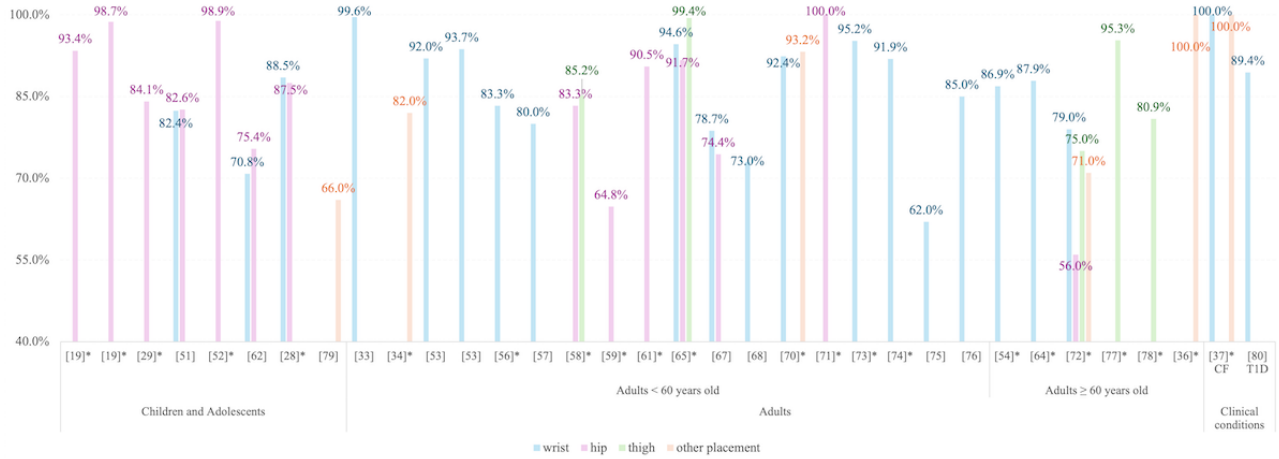
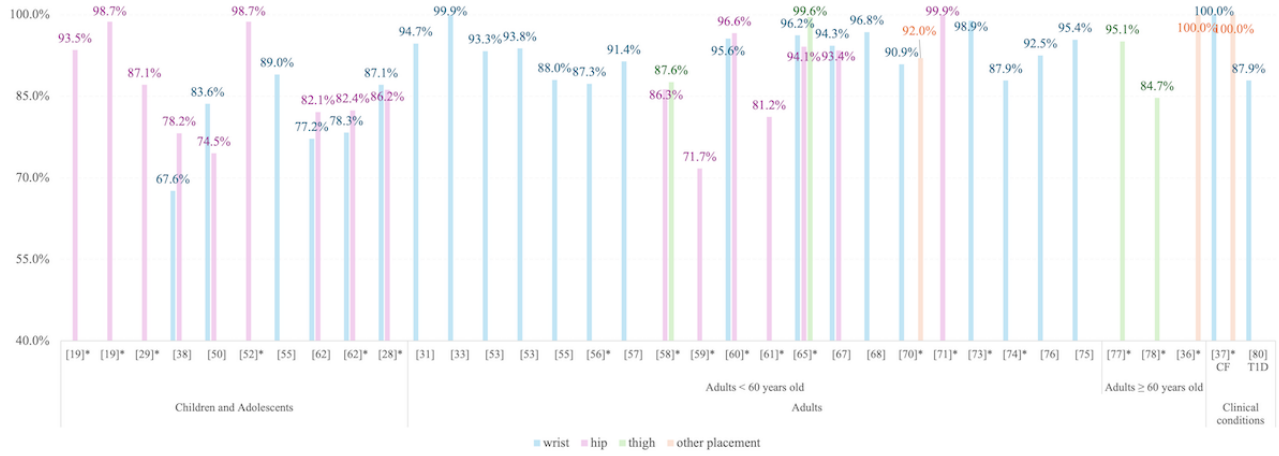


Figure 8. Accuracy metrics for moderate-to-vigorous physical activity detection across age groups and sensor placements. Brackets ("[]") represent the reference numbers; asterisks ("*") indicate lab-validated results [19,28,29,31,33,36-38,50,52,53,55-62,65,67,68,70,71,73-78,80]. CF: cystic fibrosis; T1D: type 1 diabetes.



Reproducibility and Transparency Gaps

Notwithstanding the performance advancement reported, 57.5% of the studies (23/40 studies) failed to disclose code or datasets, and 60% (3/5 studies) of DL studies lacked hyperparameter specifications (eg, learning rates, batch sizes) [34,68,79]. This "black box" methodological opacity mirrors the reproducibility crisis in traditional cut point research, where proprietary algorithms replace opaque thresholds, compromising interpretability. Only 20.0% of studies (8/40) adhered to open science practices by publicly sharing the code of models [38,51,52,58,67,68,72,76], with one study providing only sample code availability [67].

The absence of standardized reporting frameworks exacerbates methodological inconsistencies. For example, window lengths for signal segmentation ranged from 1s to 60s, complicating cross-study comparisons (details shown in Table 1). While 5 studies evaluated multiple window lengths [19,29,33,51,56], optimal performance diverged across populations: 15s or 16s were preferred for adults [33,56] and preschoolers [51], whereas 60s windows were superior to 10s and 30s among preschoolers [29]. Notably, Trost et al [19] found no difference between 10s and 60s among children (mean 11, SD 2.7 years). Similarly, MET thresholds for MVPA classification varied, such as 2.8 METs [38], 3 METs [34,37,59,60,65-67,70,71,74,76,77,79], and 3.9 METs [19,29]. This variability introduces heterogeneity

XSL•FO
RenderX

in intensity categorization, undermining cross-study generalizability.

On the positive side, the proliferation of public datasets (accounting to 25%, 10/40 studies), including Capture-24 [31,33,53,57,75,76], Energy-24 [57], UOULU (University of Oulu) [60], OSU (Oregon State University) [60], the PAMAP2 Physical Activity Monitoring dataset (University of California, Irvine, UCI) [60], and the Daily and Sports Activities (the UCI Machine Learning Repository) [60] has partially mitigated by enabling benchmarking and reducing data dependency.

Discussion

Principal Findings

This systematic scoping review synthesizes advancements in ML and DL techniques for estimating and predicting MVPA from accelerometer data. Traditional ML models (eg, RF, ANN) demonstrated robust lab-based accuracy (F_1-score mean 83.6%-100%) while real-world performance declined by 8.0%-13.3% due to environmental noise and device heterogeneity. DL architectures (eg, CNN, Transformer) achieved superior performance by leveraging raw signal dynamics (F_1-score mean 73.6%-98.4%) in free-living settings, especially with hybrid models (CNN-BILSTM, ViT-LSTM). Wrist-worn devices were most often tested (30/40, 75.0% of studies) and performed comparably to hip/thigh placements in a lab setting (F_1-score mean 84.0%-84.3%, accuracy mean 86.0%-91.8%). Multiaccelerometer configurations (eg, hip + wrist) achieved the best performance (accuracy mean 88.4%) but face practical limitations. Algorithmic bias was seen to disfavor older adult participants, but not clinical populations. However, only a few studies have tested patients, limited to cystic fibrosis and type 1 diabetes.

Methodological Advancements and Challenges

From Feature Engineering to End-to-End Learning

The evolution of MVPA detection methodologies reveals a clear paradigm shift from manual feature engineering to automated DL. Historically, ML models relied on handcrafted features (eg, spectral entropy, variance) derived from time-and frequency-domain analyses. In contrast, DL architectures, such as CNN and Transformers, automate hierarchical feature extraction from raw accelerometer signals, capturing biomechanical nuances (eg, stride variability during running) through convolutional filters and attention mechanisms [33,68]. For instance, hybrid models like CNN-BiLSTM synergized spatial and temporal learning, achieving state-of-the-art accuracy (F_1-score 98.4%-99.6%, accuracy 97.7%-99.0%) in free-living settings [33]. The effectiveness of this architecture is further corroborated by its successful application in related biomechanical modeling tasks, such as predicting ligament fatigue failure risk from complex signal data, highlighting its robust capability to capture critical spatiotemporal patterns [85]. Nevertheless, DL's computational intensity and reliance on high-resolution data (≥100Hz) limited deployment on resource-constrained wearables [33]. Furthermore, while DL reduced manual feature engineering burden, nearly 60.0% of models remained inaccessible due to unshared code, perpetuating reproducibility challenges (section "Reproducibility and Transparency Gaps").

Three key advancements define this evolution: (1) Static to dynamic features, unlike the fixed features of traditional ML, DL architectures dynamically extract nuanced biomechanical patterns from raw signals (section "Traditional Machine Learning" and "Deep Learning") [31,33,34,68,79]. (2) Early studies treated classification and estimation as separate tasks, but modern frameworks like AccNet24 unify these through shared neural pathways, improving efficiency (section "Task-Specific Insights") [33]. (3) Self-supervised learning, pretraining on unlabeled data, reduced annotation costs while maintaining high performance, addressing scarcity of free-living settings (section "Methodological Evolution and Comparative Insights") [33,68].

Lab-to-Real-World Performance Comparison

Although lab-validated models achieved high performance (eg, 87.9%-100% accuracy across ML techniques, section "Evolution of Feature Engineering and Model Architectures"), free-living performance experienced unstructured movement patterns and environment noise. For example, RF accuracy dropped from 90.1% (lab) to 83.5% (free-living), while wrist-based models exhibited superior adaptability to upper-body movements (eg, arm swings) in unstructured settings (section "Sensor Placement"). Notably, only 42.2% of studies validated models in real-world environments (most after 2020), highlighting a critical translational gap.

Two key insights emerge from a lab-to-real-world comparison. (1) There is a 3.1%-16.2% accuracy decline when using ML techniques (section "Evolution of Feature Engineering and Model Architectures"). Context-aware architecture, DL architectures, such as transformers, partially mitigated performance declines by leveraging context-aware attention to movement sequences (eg, detecting walking interruptions), achieving accuracy of 95.0% in free-living scenarios [33]. (2) There is an algorithmic bias across age groups that hinders real-world deployment (section "Algorithmic Bias").

Validation and Reproducibility

A key challenge lies in inconsistent validation protocols. While IC provided precise MET-based thresholds, its lab-bound nature limited ecological validity [58]. Conversely, DO offered real-world applicability but introduced subjectivity in intensity classification [86,87]. Moreover, disparities in metrics reporting (eg, exclusion of transitional activities) [19,37,65,70] and variable parameters (eg, MET threshold: 2.8-3.9, window lengths 1-60s) hindered cross-study comparability (section "Reproducibility and Transparency Gaps"). Compounding these issues, 42.5% of studies adhered to open science practices, perpetuating a "new cut-point conundrum" akin to proprietary regression thresholds.

Our synthesis reveals a vicious cycle underpinning the translational challenges in AI-driven MVPA monitoring. The foundational issue is the lack of standardized validation protocols. Inconsistent MET thresholds and variable data window lengths mean that models are trained and evaluated on

XSL•FO
RenderX

fundamentally different definitions of MVPA. This directly contributes to the lab-to-real-world performance gap, as a model calibrated with one protocol fails to generalize to data collected under another. Furthermore, this inconsistency, when combined with the prevalent lack of code sharing, makes it impossible to audit, replicate, or fairly compare models. Consequently, this opacity hinders the identification and correction of algorithmic bias against underrepresented populations, as the root cause of poor performance, a flawed model versus an incompatible validation method, cannot be discerned. Thus, these challenges are not isolated but are synergistic barriers that collectively impede the development of truly generalizable and equitable models.

Sensor Performance and Device Bias

Device placement and type emerged as critical determinants of model performance, as evidenced in the section "Sensor Performance." For instance, while ActiGraph-trained models achieved high lab accuracy (F_1-score 79.9%, accuracy 90.5%), they underperformed on consumer wearables (eg, Samsung smartwatch, F_1-score mean difference 3.2%) due to differences in sensor calibration and sampling rates (Table 1 and Figure 5) [56]. Additionally, interdevice variability across brands (eg, Axivity vs GENEActiv) exacerbated performance inconsistencies, particularly in free-living settings. Notably, optimal sensor placement (eg, wrist vs hip) influenced adaptability to movement patterns, with wrist-worn devices showing superior capture of upper-body dynamics (eg, arm swings) but struggling with lower-body activities [40,88]. These findings highlight the need for device-agnostic training pipelines to mitigate performance variability across brands and placements.

Translational Opportunities and Challenges

Public Health and Clinical Integration

Wrist-worn devices demonstrated comparable accuracy to hip/thigh placements in lab settings (F_1-score 84.0%-84.3%, accuracy 86.0%-91.8%) and superior adaptability to free-living upper-body movements (Figure 4), supporting their feasibility for scalable monitoring. However, ActiGraph's dominance (n=30, 75.0% of studies) and limited validation on consumer wearables (eg, smartwatches) hinder real-world applicability. Clinically, models achieved high accuracy in controlled settings for cystic fibrosis and type 1 diabetes, but small sample sizes and structured protocols limit ecological validity [37,80]. Expanding validation studies to more diverse clinical populations (eg, mobility impairments) is critical.

Age and Population Disparities

Results revealed systemic biases across the age range (section "Algorithmic Bias"). For instance, models trained on adults misclassified MVPA in children (F_1-score mean difference: –5.7%) due to developmental differences in stride length and metabolic variability [19]. Studies involving preschoolers reported accuracy fluctuations between 53.7% and 88.4%, reflecting challenges in modeling erratic movement patterns typical of young children [28,29,38,50,52,63,79]. Conversely, older adults (60 years or older) exhibited reduced accuracy (F_1-score mean 77.9%) due to slower gait speeds, postural instability, and comorbidities that alter movement signatures [54,64,69,72,77,78]. Wrist-based model, for example, underestimated MVPA in this cohort by 6.0%-16.6% compared with thigh-worn sensors, highlighting the need for age-specific calibration [78].

Emerging Innovations

Hybrid DL models have emerged as a powerful approach. For instance, integrating LSTM with CNN (CNN-LSTM) or ViTs (vision transformer bidirectional long short-term memory) enables the capture of spatial-temporal patterns in accelerometry data [33,34]. Building on this, BiLSTM layers further enhance temporal dependency modeling by analyzing sequences in both forward and backward directions [31].

In parallel, image-based feature extraction methods, such as converting raw accelerometer signals into Gramian angular field images, have improved feature learning by transforming time-series data into visual representations [33]. Additionally, multisensor fusion strategies—combining data from hip, wrist, and thigh placements—address variability in sensor positioning, boosting model robustness [28,34]. Furthermore, transfer learning leverages pretrained architectures like ResNet101, adapting them for accelerometer classification tasks [31].

Another key innovation lies in advanced feature engineering. Autonomous feature extraction via CNN reduces reliance on handcrafted features [68], while time-frequency domain fusion (eg, spectral power) enhances activity discrimination [52]. Notably, real-time and edge computing advancements explore lightweight models through pruning and quantization, enabling deployment on wearable devices [33].

However, significant challenges remain. First, models trained in controlled lab settings often generalize poorly to free-living environments due to uncontrolled variability [52]. Moreover, short and heterogeneous activity bouts, common in populations like preschoolers, result in mixed-activity windows that complicate classification [29]. Another critical challenge is sensor placement variability, as signal patterns differ across body positions [37]. Compounding this, class imbalance from overrepresented sedentary/light activities skews model performance [75]. Additionally, computational complexity limits real-time use, as seen in resource-heavy models like AccNet24 [31]. Finally, distinguishing biomechanically similar activities (eg, climbing vs walking) remains problematic [29].

Future Directions

To address existing gaps, 4 interconnected priorities emerge. First, resolving inconsistencies in ground truth methods, such as variable MET thresholds (2.8-3.9 METs) and window lengths (1-60s), is critical. This requires standardized validation frameworks, including consensus guidelines and open datasets (eg, Capture-24), to harmonize protocols and reduce discrepancies in intensity classification [64].

Second, prioritizing free-living validation is essential to bridge the lab-to-real-world performance gap. For instance, RF models exhibit accuracy declines from 90.1% in lab settings to 83.5% in free-living environments. Concurrently, diversifying training

XSL•FO
RenderX

data to include underrepresented groups, such as the older adult, pediatric, and clinical populations, will improve generalizability and mitigate age-related biases [62,72].

Third, advancing algorithmic fairness through regulatory frameworks and bias audits is imperative. This includes expanding datasets to encompass more and diverse clinical cohorts while addressing disparities in model performance across the age range. Additionally, mandating open science practices, such as code/data sharing and hyperparameter transparency, will enhance reproducibility and resolve the "new cut-point conundrum" plaguing activity intensity thresholds.

Finally, optimizing DL architectures, such as quantized models or hybrid CNN-BiLSTM frameworks, for low-power wearables will enable real-world deployment while maintaining computational efficiency [31,33].

Looking ahead, these priorities align with broader calls for standardization and interpretability. For example, improving the "black-box" nature of DL models [68] and harmonizing evaluation metrics will foster clinical trust. Moreover, lightweight, edge-compatible architectures and multimodal data integration represent promising pathways to overcome current limitations in real-world MVPA monitoring.

Limitations and Methodological Considerations

The strengths of this review include the rigorous adherence to PRISMA-ScR guidelines, a comprehensive search strategy across 3 electronic databases (PubMed, IEEE Xplore, Web of Science, and others via manual citation tracking), and systematic screening of 1938 records. The methodology prioritized transparency through dual-reviewer full-text screening to resolve discrepancies and consultation to ensure methodological rigor. By focusing on peer-reviewed studies, we aimed to synthesize evidence grounded in empirical validation, thereby minimizing inclusion of speculative or opinion-based articles.

However, several limitations warrant consideration. First, the exclusion of gray literature (eg, unpublished trials, industry reports, or conference proceedings) may have omitted insights from ongoing or unsuccessful implementation efforts, particularly those led by technology developers or health care providers. This introduces potential publication bias, as negative results or pragmatic challenges in real-world deployment are often underrepresented in peer-reviewed journals. Second, our decision to exclude non–peer-reviewed studies and prioritize articles reporting empirical implementation in clinical or free-living settings risks overlooking formative research, such as feasibility studies or pilot trials, which could offer valuable lessons for scalable AI integration.

A further limitation arises from our emphasis on the highest reported performance metrics (eg, F_1-scores, accuracy) across studies. While this approach highlights peak algorithmic capabilities, it may overestimate real-world applicability, as optimal configurations (eg, 15-second windows for adults, multisensor placements) often lack generalizability to diverse populations or unstructured environments. For instance, models achieving 99.0% accuracy in lab settings may exhibit significant performance degradation in free-living contexts due to uncontrolled variables like device heterogeneity or nonexercise movements.

Methodologically, while the Arksey and O'Malley framework does not mandate quality appraisal, the inclusion of studies with heterogeneous validation protocols (eg, variable MET thresholds, ground-truth methodologies) complicates cross-study comparisons. Future reviews could strengthen synthesis by incorporating quality assessment tools to evaluate bias risk and methodological consistency. Last, the predominance of studies using young, healthy cohorts limits insights into algorithmic fairness and generalizability for older adult or clinical populations, underscoring the need for more inclusive training datasets.

These considerations do not diminish the review's contributions but highlight critical gaps, such as reproducibility challenges and translational biases, that must be addressed to advance equitable, real-world deployment of AI-driven MVPA monitoring tools.

Conclusions

This systematic scoping review highlights that ML and DL have significantly advanced in the detection of MVPA by using accelerometer data, yet persistent gaps in generalizability and transparency hinder real-world impact. To bridge the lab-to-real-world divide, collaborative efforts across public health and computer science must prioritize reproducibility, inclusive design, and robust validation. By addressing these challenges, AI-driven tools can fulfill their potential as scalable, equitable solutions for advancing global physical activity research and intervention.

Acknowledgments

During the preparation of this work, the authors used DeepSeek to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was supported by the Shanghai Key Lab of Human Performance (Shanghai University of Sport) (grant number 11DZ2261100). The funder provided financial support for the article processing charge of this review.

XSL•FO
RenderX

Authors' Contributions

EJCDG and YZ contributed to the conceptualization of the study. SRBVDV and YZ were responsible for the methodology and formal analysis. PC provided resources and supervised the project, with project administration shared between PC and YZ. YZ prepared the original draft and created the visualizations, while EJCDG, PC, SRBVDV, and YZ contributed to the review and editing of the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 22 KB - ijmr_v15i1e76601_app1.docx]

Multimedia Appendix 2
The full summary of included studies (N=40 studies, ranked by health condition and alphabets of author names).
[DOCX File , 83 KB - ijmr_v15i1e76601_app2.docx]

Multimedia Appendix 3
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 177 KB - ijmr_v15i1e76601_app3.pdf]

References

1. Herrmann SD, Willis EA, Ainsworth BE, Barreira TV, Hastert M, Kracht CL, et al. 2024 adult compendium of physical activities: a third update of the energy costs of human activities. J Sport Health Sci 2024;13(1):6-12 [FREE Full text] [doi: 10.1016/j.jshs.2023.10.010] [Medline: 38242596]
2. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol 1990;13(8):555-565 [FREE Full text] [doi: 10.1002/clc.4960130809] [Medline: 2204507]
3. Lin H, Sardana M, Zhang Y, Liu C, Trinquart L, Benjamin EJ, et al. Association of habitual physical activity with cardiovascular disease risk. Circ Res 2020;127(10):1253-1260 [FREE Full text] [doi: 10.1161/CIRCRESAHA.120.317578] [Medline: 32842915]
4. Ried-Larsen M, Grøntved A, Kristensen PL, Froberg K, Andersen LB. Moderate-and-vigorous physical activity from adolescence to adulthood and subclinical atherosclerosis in adulthood: prospective observations from the European Youth Heart Study. Br J Sports Med 2015;49(2):107-112. [doi: 10.1136/bjsports-2013-092409] [Medline: 23584827]
5. Chen Y, Chen Z, Pan L, Ma Z, Zhang H, Li X, et al. Effect of moderate and vigorous aerobic exercise on incident diabetes in adults with obesity: a 10-year follow-up of a randomized clinical trial. JAMA Intern Med 2023;183(3):272-275 [FREE Full text] [doi: 10.1001/jamainternmed.2022.6291] [Medline: 36716009]
6. Moore SC, Patel AV, Matthews CE, Berrington de Gonzalez A, Park Y, Katki HA, et al. Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. PLoS Med 2012;9(11):e1001335 [FREE Full text] [doi: 10.1371/journal.pmed.1001335] [Medline: 23139642]
7. World Health Organization. WHO Guidelines on Physical Activity and Sedentary Behaviour. Geneva: World Health Organization; 2020.
8. World Health Organization. Guidelines on Physical Activity, Sedentary Behaviour and Sleep for Children under 5 Years of Age. Geneva: World Health Organization; 2019.
9. Cliff DP, Reilly JJ, Okely AD. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. J Sci Med Sport 2009;12(5):557-567. [doi: 10.1016/j.jsams.2008.10.008] [Medline: 19147404]
10. Cain KL, Sallis JF, Conway TL, Van Dyck D, Calhoon L. Using accelerometers in youth physical activity studies: a review of methods. J Phys Act Health 2013;10(3):437-450. [doi: 10.1123/jpah.10.3.437] [Medline: 23620392]
11. Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. J Aging Phys Act 2009;17(1):17-30 [FREE Full text] [doi: 10.1123/japa.17.1.17] [Medline: 19299836]
12. Strath SJ, Pfeiffer KA, Whitt-Glover MC. Accelerometer use with children, older adults, and adults with functional limitations. Med Sci Sports Exerc 2012;44(1 Suppl 1):S77-S85 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399eb1] [Medline: 22157778]
13. Jago R, Zakeri I, Baranowski T, Watson K. Decision boundaries and receiver operating characteristic curves: new methods for determining accelerometer cutpoints. J Sports Sci 2007;25(8):937-944. [doi: 10.1080/02640410600908027] [Medline: 17474047]
14. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med Sci Sports Exerc 2005;37(11 Suppl):S501-S511. [doi: 10.1249/01.mss.0000185660.38335.de] [Medline: 16294113]

XSL•FO
RenderX

15. Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, inc. accelerometer. Med Sci Sports Exerc 1998;30(5):777-781. [doi: 10.1097/00005768-199805000-00021] [Medline: 9588623]
16. Welch WA, Bassett DR, Thompson DL, Freedson PS, Staudenmayer JW, John D, et al. Classification accuracy of the wrist-worn gravity estimator of normal everyday activity accelerometer. Med Sci Sports Exerc 2013;45(10):2012-2019 [FREE Full text] [doi: 10.1249/MSS.0b013e3182965249] [Medline: 23584403]
17. Janssen X, Cliff DP, Reilly JJ, Hinkley T, Jones RA, Batterham M, et al. Predictive validity and classification accuracy of ActiGraph energy expenditure equations and cut-points in young children. PLoS One 2013;8(11):e79124 [FREE Full text] [doi: 10.1371/journal.pone.0079124] [Medline: 24244433]
18. Trost SG, Way R, Okely AD. Predictive validity of three ActiGraph energy expenditure equations for children. Med Sci Sports Exerc 2006;38(2):380-387. [doi: 10.1249/01.mss.0000183848.25845.e0] [Medline: 16531910]
19. Trost SG, Wong W, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. Med Sci Sports Exerc 2012;44(9):1801-1809 [FREE Full text] [doi: 10.1249/MSS.0b013e318258ac11] [Medline: 22525766]
20. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. Med Sci Sports Exerc 2011;43(7):1360-1368. [doi: 10.1249/MSS.0b013e318206476e] [Medline: 21131873]
21. Ellis K, Kerr J, Godbole S, Staudenmayer J, Lanckriet G. Hip and wrist accelerometer algorithms for free-living behavior classification. Med Sci Sports Exerc 2016;48(5):933-940 [FREE Full text] [doi: 10.1249/MSS.0000000000000840] [Medline: 26673126]
22. Trost SG, Brookes DSK, Ahmadi MN. Evaluation of wrist accelerometer cut-points for classifying physical activity intensity in youth. Front Digit Health 2022;4:884307 [FREE Full text] [doi: 10.3389/fdgth.2022.884307] [Medline: 35585912]
23. Bassett DR, Rowlands A, Trost SG. Calibration and validation of wearable monitors. Med Sci Sports Exerc 2012;44(1 Suppl 1):S32-S38 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399cf7] [Medline: 22157772]
24. Wang H, Ma C, Zhou L. A brief review of machine learning and its application. 2009 Presented at: International Conference on Information Engineering and Computer Science; 2009 December 19-20; Wuhan, China p. 19-20. [doi: 10.1109/iciecs.2009.5362936]
25. Mjolsness E, DeCoste D. Machine learning for science: state of the art and future prospects. Science 2001;293(5537):2051-2055. [doi: 10.1126/science.293.5537.2051] [Medline: 11557883]
26. Ruch N, Joss F, Jimmy G, Melzer K, Hänggi J, Mäder U. Neural network versus activity-specific prediction equations for energy expenditure estimation in children. J Appl Physiol (1985) 2013;115(9):1229-1236 [FREE Full text] [doi: 10.1152/japplphysiol.01443.2012] [Medline: 23990244]
27. Farrahi V, Niemelä M, Kangas M, Korpelainen R, Jämsä T. Calibration and validation of accelerometer-based activity monitors: a systematic review of machine-learning approaches. Gait Posture 2019;68:285-299 [FREE Full text] [doi: 10.1016/j.gaitpost.2018.12.003] [Medline: 30579037]
28. Trost SG, Cliff DP, Ahmadi MN, Tuc NV, Hagenbuchner M. Sensor-enabled activity class recognition in preschoolers: hip versus wrist data. Med Sci Sports Exerc 2018;50(3):634-641. [doi: 10.1249/MSS.0000000000001460] [Medline: 29059107]
29. Hagenbuchner M, Cliff DP, Trost SG, Van Tuc N, Peoples GE. Prediction of activity type in preschool children using machine learning techniques. J Sci Med Sport 2015;18(4):426-431. [doi: 10.1016/j.jsams.2014.06.003] [Medline: 25088983]
30. Alinia P, Saeedi R, Fallahzadeh R, Rokni A, Ghasemzadeh H. A reliable and reconfigurable signal processing framework for estimation of metabolic equivalent of task in wearable sensors. IEEE J Sel Top Signal Process 2016;10(5):842-853. [doi: 10.1109/jstsp.2016.2569472]
31. Farrahi V, Muhammad U, Rostami M, Oussalah M. AccNet24: a deep learning framework for classifying 24-hour activity behaviours from wrist-worn accelerometer data under free-living environments. Int J Med Inform 2023;172:105004 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105004] [Medline: 36724729]
32. Ng JYY, Zhang JH, Hui SS, Jiang G, Yau F, Cheng J, et al. Development of a multi-wear-site, deep learning-based physical activity intensity classification algorithm using raw acceleration data. PLoS One 2024;19(3):e0299295 [FREE Full text] [doi: 10.1371/journal.pone.0299295] [Medline: 38452147]
33. Wang L, Luo Z, Zhang T. A novel ViT-BILSTM model for physical activity intensity classification in adults using gravity-based acceleration. BMC Biomed Eng 2025;7(1):2. [doi: 10.1186/s42490-025-00088-2] [Medline: 39891283]
34. Barua A, Jiang X, Fuller D. The effectiveness of simple heuristic features in sensor orientation and placement problems in human activity recognition using a single smartphone accelerometer. Biomed Eng Online 2024;23(1):21 [FREE Full text] [doi: 10.1186/s12938-024-01213-3] [Medline: 38368358]
35. Pfeiffer KA, Clevenger KA, Kaplan A, Van Camp CA, Strath SJ, Montoye AHK. Accessibility and use of novel methods for predicting physical activity and energy expenditure using accelerometry: a scoping review. Physiol Meas 2022;43(9). [doi: 10.1088/1361-6579/ac89ca] [Medline: 35970175]
36. Andò B, Manenti M, Greco D, Pistorio A. An embedded sensing methodology for the classification of activity rate. 2024 Presented at: IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE); 2024 October 21-23; St Albans, United Kingdom p. 21-23. [doi: 10.1109/metroxraine62247.2024.10796795]

XSL•FO
RenderX

37. Bianchim MS, McNarry MA, Barker AR, Williams CA, Denford S, Thia L, et al. A machine learning approach for physical activity recognition in cystic fibrosis. Meas Phys Educ Exerc Sci 2023;28(2):172-181. [doi: 10.1080/1091367x.2023.2271444]
38. Ahmadi MN, Trost SG. Device-based measurement of physical activity in pre-schoolers: comparison of machine learning and cut point methods. PLoS One 2022;17(4):e0266970 [FREE Full text] [doi: 10.1371/journal.pone.0266970] [Medline: 35417492]
39. Clevenger KA, Montoye AHK, Van Camp CA, Strath SJ, Pfeiffer KA. Methods for estimating physical activity and energy expenditure using raw accelerometry data or novel analytical approaches: a repository, framework, and reporting guidelines. Physiol Meas 2022;43(9). [doi: 10.1088/1361-6579/ac89c9] [Medline: 35970174]
40. Gao Z, Liu W, McDonough DJ, Zeng N, Lee JE. The dilemma of analyzing physical activity and sedentary behavior with wrist accelerometer data: challenges and opportunities. J Clin Med 2021;10(24):5951 [FREE Full text] [doi: 10.3390/jcm10245951] [Medline: 34945247]
41. Allahbakhshi H, Hinrichs T, Huang H, Weibel R. The key factors in physical activity type detection using real-life data: a systematic review. Front Physiol 2019;10:75 [FREE Full text] [doi: 10.3389/fphys.2019.00075] [Medline: 30809152]
42. Gupta N, Gupta SK, Pathak RK, Jain V, Rashidi P, Suri JS. Human activity recognition in artificial intelligence framework: a narrative review. Artif Intell Rev 2022;55(6):4755-4808 [FREE Full text] [doi: 10.1007/s10462-021-10116-x] [Medline: 35068651]
43. Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y. Physical human activity recognition using wearable sensors. Sensors (Basel) 2015;15(12):31314-31338 [FREE Full text] [doi: 10.3390/s151229858] [Medline: 26690450]
44. Zhang S, Li Y, Zhang S, Shahabi F, Xia S, Deng Y, et al. Deep learning in human activity recognition with wearable sensors: a review on advances. Sensors (Basel) 2022;22(4):1476 [FREE Full text] [doi: 10.3390/s22041476] [Medline: 35214377]
45. de Almeida Mendes M, da Silva ICM, Ramires VV, Reichert FF, Martins RC, Tomasi E. Calibration of raw accelerometer data to measure physical activity: a systematic review. Gait Posture 2018;61:98-110. [doi: 10.1016/j.gaitpost.2017.12.028] [Medline: 29324298]
46. Cust EE, Sweeting AJ, Ball K, Robertson S. Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance. J Sports Sci 2019;37(5):568-600. [doi: 10.1080/02640414.2018.1521769] [Medline: 30307362]
47. Trujillo-Guerrero MF, Román-Niemes S, Jaén-Vargas M, Cadiz A, Fonseca R, Serrano-Olmedo JJ. Accuracy comparison of CNN, LSTM, and transformer for activity recognition using IMU and visual markers. IEEE Access 2023;11:106650-106669. [doi: 10.1109/access.2023.3318563]
48. Lavanya TV. A machine learning approach for predicting physical activity intensity from wearable sensor data. 2024 Presented at: 7th International Conference on Circuit Power and Computing Technologies (ICCPCT); 2024 August 7-8; Poruvazhy, Kerala, India p. 8-9. [doi: 10.1109/iccpct61902.2024.10673260]
49. Lavanya TV, Sivaraman K. Predictive analytics for monitoring and classifying physical activity intensity using smartphone sensors. 2024 Presented at: 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS); 2024July 10-12; Coimbatore, India. [doi: 10.1109/icscss60660.2024.10625133]
50. Ahmadi MN, Brookes D, Chowdhury A, Pavey T, Trost SG. Free-living evaluation of laboratory-based activity classifiers in preschoolers. Med Sci Sports Exerc 2020;52(5):1227-1234. [doi: 10.1249/MSS.0000000000002221] [Medline: 31764460]
51. Ahmadi MN, Pavey TG, Trost SG. Machine learning models for classifying physical activity in free-living preschool children. Sensors (Basel) 2020;20(16):4364 [FREE Full text] [doi: 10.3390/s20164364] [Medline: 32764316]
52. Ahmadi MN, Pfeiffer KA, Trost SG. Physical activity classification in youth using raw accelerometer data from the hip. Meas Phys Educ Exerc Sci 2020;24(2):129-136. [doi: 10.1080/1091367x.2020.1716768]
53. Ahmadi MN, Hamer M, Gill JMR, Murphy M, Sanders JP, Doherty A, et al. Brief bouts of device-measured intermittent lifestyle physical activity and its association with major adverse cardiovascular events and mortality in people who do not exercise: a prospective cohort study. Lancet Public Health 2023;8(10):e800-e810 [FREE Full text] [doi: 10.1016/S2468-2667(23)00183-4] [Medline: 37777289]
54. Bai C, Wanigatunga AA, Saldana S, Casanova R, Manini TM, Mardini MT. Are machine learning models on wrist accelerometry robust against differences in physical performance among older adults? Sensors (Basel) 2022;22(8):3061 [FREE Full text] [doi: 10.3390/s22083061] [Medline: 35459045]
55. Chen D, Du Y, Liu Y, Hong J, Yin X, Zhu Z, et al. Development and validation of a smartwatch algorithm for differentiating physical activity intensity in health monitoring. Sci Rep 2024;14(1):9530 [FREE Full text] [doi: 10.1038/s41598-024-59602-6] [Medline: 38664457]
56. Davoudi A, Wanigatunga AA, Kheirkhahan M, Corbett DB, Mendoza T, Battula M, et al. Accuracy of samsung gear s smartwatch for activity recognition: validation study. JMIR Mhealth Uhealth 2019;7(2):e11270 [FREE Full text] [doi: 10.2196/11270] [Medline: 30724739]
57. Doherty A, Smith-Byrne K, Ferreira T, Holmes MV, Holmes C, Pulit SL, et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. Nat Commun 2018;9(1):5257 [FREE Full text] [doi: 10.1038/s41467-018-07743-4] [Medline: 30531941]

XSL•FO
RenderX

58. Ellingson LD, Schwabacher IJ, Kim Y, Welk GJ, Cook DB. Validity of an integrative method for processing physical activity data. Med Sci Sports Exerc 2016;48(8):1629-1638. [doi: 10.1249/MSS.0000000000000915] [Medline: 27015380]
59. Ellingson LD, Hibbing PR, Kim Y, Frey-Law LA, Saint-Maurice PF, Welk GJ. Lab-based validation of different data processing methods for wrist-worn ActiGraph accelerometers in young adults. Physiol Meas 2017;38(6):1045-1060. [doi: 10.1088/1361-6579/aa6d00] [Medline: 28481750]
60. Farrahi V, Niemela M, Tjurin P, Kangas M, Korpelainen R, Jamsa T. Evaluating and enhancing the generalization performance of machine learning models for physical activity intensity prediction from raw acceleration data. IEEE J Biomed Health Inform 2020;24(1):27-38. [doi: 10.1109/JBHI.2019.2917565] [Medline: 31107668]
61. Freedson PS, Lyden K, Kozey-Keadle S, Staudenmayer J. Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample. J Appl Physiol (1985) 2011;111(6):1804-1812 [FREE Full text] [doi: 10.1152/japplphysiol.00309.2011] [Medline: 21885802]
62. Hibbing PR, Ellingson LD, Dixon PM, Welk GJ. Adapted sojourn models to estimate activity intensity in youth: a suite of tools. Med Sci Sports Exerc 2018;50(4):846-854. [doi: 10.1249/MSS.0000000000001486] [Medline: 29135657]
63. Li S, Howard JT, Sosa ET, Cordova A, Parra-Medina D, Yin Z. Calibrating wrist-worn accelerometers for physical activity assessment in preschoolers: machine learning approaches. JMIR Form Res 2020;4(8):e16727 [FREE Full text] [doi: 10.2196/16727] [Medline: 32667893]
64. Mardini MT, Bai C, Wanigatunga AA, Saldana S, Casanova R, Manini TM. Age differences in estimating physical activity by wrist accelerometry using machine learning. Sensors (Basel) 2021;21(10):3352 [FREE Full text] [doi: 10.3390/s21103352] [Medline: 34065906]
65. Montoye AHK, Pivarnik JM, Mudd LM, Biswas S, Pfeiffer KA. Validation and comparison of accelerometers worn on the hip, thigh, and wrists for measuring physical activity and sedentary behavior. AIMS Public Health 2016;3(2):298-312 [FREE Full text] [doi: 10.3934/publichealth.2016.2.298] [Medline: 29546164]
66. Montoye AHK, Pivarnik JM, Mudd LM, Biswas S, Pfeiffer KA. Evaluation of the activPAL accelerometer for physical activity and energy expenditure estimation in a semi-structured setting. J Sci Med Sport 2017;20(11):1003-1007. [doi: 10.1016/j.jsams.2017.04.011] [Medline: 28483558]
67. Montoye AHK, Westgate BS, Clevenger KA, Pfeiffer KA, Vondrasek JD, Fonley MR, et al. Individual versus group calibration of machine learning models for physical activity assessment using body-worn accelerometers. Med Sci Sports Exerc 2021;53(12):2691-2701. [doi: 10.1249/MSS.0000000000002752] [Medline: 34310493]
68. Nawaratne R, Alahakoon D, De Silva D, O'Halloran PD, Montoye AH, Staley K, et al. Deep learning to predict energy expenditure and activity intensity in free living conditions using wrist-specific accelerometry. J Sports Sci 2021;39(6):683-690. [doi: 10.1080/02640414.2020.1841394] [Medline: 33121379]
69. Nnamoko N, Cabrera-Diego LA, Campbell D, Sanders G, Fairclough SJ, Korkontzelos I. Personalised accelerometer cut-point prediction for older adults' movement behaviours using a machine learning approach. Comput Methods Programs Biomed 2021;208:106165. [doi: 10.1016/j.cmpb.2021.106165] [Medline: 34118492]
70. O'Driscoll R, Turicchi J, Hopkins M, Duarte C, Horgan GW, Finlayson G, et al. Comparison of the validity and generalizability of machine learning algorithms for the prediction of energy expenditure: validation study. JMIR Mhealth Uhealth 2021;9(8):e23938 [FREE Full text] [doi: 10.2196/23938] [Medline: 34346890]
71. Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of novel techniques to classify physical activity mode using accelerometers. Med Sci Sports Exerc 2006;38(9):1626-1634. [doi: 10.1249/01.mss.0000227542.43669.45] [Medline: 16960524]
72. Skjødt M, Brønd JC, Tully MA, Tsai L, Koster A, Visser M, et al. Moderate and vigorous physical activity intensity cut-points for hip-, wrist-, thigh-, and lower back worn accelerometer in very old adults. Scand J Med Sci Sports 2025;35(1):e70009. [doi: 10.1111/sms.70009] [Medline: 39753998]
73. Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J Appl Physiol (1985) 2009;107(4):1300-1307 [FREE Full text] [doi: 10.1152/japplphysiol.00465.2009] [Medline: 19644028]
74. Staudenmayer J, He S, Hickey A, Sasaki J, Freedson P. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. J Appl Physiol (1985) 2015;119(4):396-403 [FREE Full text] [doi: 10.1152/japplphysiol.00026.2015] [Medline: 26112238]
75. Tsanas A. Accurately inferring physical activity levels and sleep from wrist-worn actigraphy recordings with sample rates as low as 10 hz. IEEE Access 2025;13:27257-27267. [doi: 10.1109/access.2025.3539278]
76. Walmsley R, Chan S, Smith-Byrne K, Ramakrishnan R, Woodward M, Rahimi K, et al. Reallocation of time between device-measured movement behaviours and risk of incident cardiovascular disease. Br J Sports Med 2021;56(18):1008-1017 [FREE Full text] [doi: 10.1136/bjsports-2021-104050] [Medline: 34489241]
77. Wullems JA, Verschueren SMP, Degens H, Morse CI, Onambélé GL. Performance of thigh-mounted triaxial accelerometer algorithms in objective quantification of sedentary behaviour and physical activity in older adults. PLoS One 2017;12(11):e0188215 [FREE Full text] [doi: 10.1371/journal.pone.0188215] [Medline: 29155839]
78. Wullems JA, Verschueren SMP, Degens H, Morse CI, Onambélé-Pearson GL. Concurrent validity of four activity monitors in older adults. Sensors (Basel) 2024;24(3):895 [FREE Full text] [doi: 10.3390/s24030895] [Medline: 38339613]

XSL•FO
RenderX

79. Zhou L, Qu X, Zhang T, Wu J, Yin H, Guan H, et al. Prediction of pediatric activity intensity with wearable sensors and bi-directional LSTM models. Pattern Recognit Lett 2021;152:166-171. [doi: 10.1016/j.patrec.2021.08.030]
80. Cescon M, Choudhary D, Pinsker JE, Dadlani V, Church MM, Kudva YC, et al. Activity detection and classification from wristband accelerometer data collected on people with type 1 diabetes in free-living conditions. Comput Biol Med 2021;135:104633 [FREE Full text] [doi: 10.1016/j.compbiomed.2021.104633] [Medline: 34346318]
81. Lyden K, Keadle SK, Staudenmayer J, Freedson PS. A method to estimate free-living active and sedentary behavior from an accelerometer. Med Sci Sports Exerc 2014;46(2):386-397 [FREE Full text] [doi: 10.1249/MSS.0b013e3182a42a2d] [Medline: 23860415]
82. Chan S, Hang Y, Tong C, Acquah A, Schonfeldt A, Gershuny J, et al. CAPTURE-24: A large dataset of wrist-worn activity tracker data collected in the wild for human activity recognition. Sci Data 2024;11(1):1135 [FREE Full text] [doi: 10.1038/s41597-024-03960-3] [Medline: 39414802]
83. Wang L. Recognition of human activities using continuous autoencoders with wearable sensors. Sensors (Basel) 2016;16(2):189 [FREE Full text] [doi: 10.3390/s16020189] [Medline: 26861319]
84. Ni Q, Fan Z, Zhang L, Nugent CD, Cleland I, Zhang Y, et al. Leveraging wearable sensors for human daily activity recognition with stacked denoising autoencoders. Sensors (Basel) 2020;20(18):5114 [FREE Full text] [doi: 10.3390/s20185114] [Medline: 32911780]
85. Xu D, Zhou H, Jie T, Zhou Z, Yuan Y, Jemni M, et al. Data-driven deep learning for predicting ligament fatigue failure risk mechanisms. Int J Mech Sci 2025;301:110519 [FREE Full text] [doi: 10.1016/j.ijmecsci.2025.110519]
86. Trost SG. State of the art reviews: measurement of physical activity in children and adolescents. Am J Lifestyle Med 2007;1(4):299-314. [doi: 10.1177/1559827607301686]
87. McKenzie TL, van der Mars H. Top 10 research questions related to assessing physical activity and its contexts using systematic observation. Res Q Exerc Sport 2015;86(1):13-29. [doi: 10.1080/02701367.2015.991264] [Medline: 25664670]
88. Cleland I, Kikhia B, Nugent C, Boytsov A, Hallberg J, Synnes K, et al. Optimal placement of accelerometers for the detection of everyday activities. Sensors (Basel) 2013;13(7):9183-9200 [FREE Full text] [doi: 10.3390/s130709183] [Medline: 23867744]

Abbreviations

AI: artificial intelligence
ANN: artificial neural network
BiLSTM: bidirectional long short-term memory
CNN: convolutional neural network
CNN-BiLSTM: convolutional neural network and bidirectional long short-term memory
CNN-LSTM: convolutional neural network and long short-term memory
DL: deep learning
DO: direct observation
DT: decision tree
IC: indirect calorimetry
LSTM: long short-term memory network
MET: metabolic equivalent
ML: machine learning
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
OSU: Oregon State University
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RF: random forest
RQ: research questions
SVM: support vector machine
UOULU: University of Oulu
ViT: vision transformer
VPA: vigorous physical activity

XSL•FO
RenderX

Edited by N Cahill; submitted 27.Apr.2025; peer-reviewed by S Hansun, Y Gu; comments to author 16.Sep.2025; accepted 26.Oct.2025; published 08.Jan.2026.

Please cite as:
Zi Y, van de Ven SRB, de Geus EJC, Chen P
Machine and Deep Learning for Detection of Moderate-to-Vigorous Physical Activity From Accelerometer Data: Systematic Scoping Review
Interact J Med Res 2026;15:e76601
URL: https://www.i-jmr.org/2026/1/e76601
doi:10.2196/76601
PMID:

©Yahua Zi, Sjors RB van de Ven, Eco JC de Geus, Peijie Chen. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 08.Jan.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Treatment of Gender in Research on Intervention Programs Targeting Social Isolation and Loneliness Among Older Adults: Scoping Review

Kenta Nomura[1], PhD; Naoto Kiguchi[2], MA; Eisuke Inomata[3], MA; Takeshi Nakamachi[4], MA; Norikazu Kobayashi[3], PhD

[1]Graduate School of Rehabilitation, Mejiro University, Saitama, Japan

[2]Department of Occupational Therapy, Faculty of Health and Medical Sciences, Ibaraki Prefectural University of Health Sciences, Ibaraki, Japan

[3]Department of Occupational Therapy, Graduate School of Human Health Science, Tokyo Metropolitan University, Tokyo, Japan

[4]Department of Psychology and Social Welfare, Faculty of Psychology and Social Welfare, Seigakuin University, Saitama, Japan

Corresponding Author:
Kenta Nomura, PhD
Graduate School of Rehabilitation
Mejiro University
320 Ukiya
Iwatsuki-ku, Saitama-shi
Saitama, 339-8501
Japan
Phone: 81 487972131
Fax: 81 487972132
Email: nomura@mejiro.ac.jp

Abstract

Background: Social isolation and loneliness have considerable health implications. Research indicates that older men are generally more susceptible to social isolation compared with women, highlighting the need to integrate gender-responsive approaches in the development and implementation of interventions for mitigating social isolation and loneliness in later life.

Objective: This study aimed to conduct a review of intervention programs targeting social isolation and loneliness, focusing on gender-specific considerations. Specifically, it aims to examine the gender composition (male-to-female ratio) of participants in intervention programs and identify and analyze intervention strategies that demonstrate gender-sensitive effectiveness.

Methods: A scoping review was conducted as per the Joanna Briggs Institute manual for evidence synthesis. A comprehensive literature search, including hand searching, was conducted across 6 English-language databases, PubMed, MEDLINE, Cochrane, CINAHL, ScienceDirect, and Web of Science, for papers and reports published in 2013-2023. The authors, country, subjects, research design, intervention method, results, and mentions of gender for each included document were presented.

Results: The study identified 1282 papers and reports, of which 10 were selected for analysis. Only 1 study reported a higher number of male participants compared with female ones; in contrast, all other studies included predominantly female samples. The studies assessed outcomes based on 2 indicators of social isolation, 4 indicators of loneliness, and 29 other indicators. Exercise and workshops proved effective for social isolation and loneliness, while meditation and laughter therapy were effective for loneliness. The intervention with the highest percentage of male participants (264/323, 82%) was a customized meditation program. Conversely, physical activities, social support, and community-based group health classes drew more female participants. In total, 8 studies did not mention gender in the discussion section, and none considered gender-specific issues in formulating research objectives and outcomes.

Conclusions: Research on social isolation and loneliness has generally ignored the influence of gender. The review also indicated a gender bias in participant selection, with women markedly overrepresented in study samples. The study found that women tend to prefer interventions emphasizing conversations, shared experiences, and emotional exchange. In contrast, men showed the highest participation in a meditation program focused on self-dialogue, which required minimal interaction. Importantly, interventions aimed at promoting social interaction or participation are unlikely to succeed without consideration of gender-specific issues. Therefore, systematically identifying conditions necessary for effective interventions that target older men is crucial for guiding future research and program development.

XSL•FO
RenderX

Trial Registration: Open Science Framework 10.17605/OSF.IO/83JQF; hhttps://osf.io/83jqf/overview

(Interact J Med Res 2026;15:e72281) doi:10.2196/72281

KEYWORDS

social isolation; loneliness; intervention; program; gender; men; women; older; review; remote; in-person; scoping review

Introduction

Background

Social isolation is defined as "the objective lack or paucity of social contacts and interactions with family members, friends, or the wider community" [1]. Factors associated with social isolation include advanced age, male gender, presence of depressive symptoms, and low socioeconomic status [2]. The concept of social isolation is closely related to loneliness. According to Valtorta and Hanratty [1], loneliness is "a subjective negative feeling associated with a perceived lack of a wider social network (social loneliness) or the absence of a specific desired companion (emotional loneliness)." Scholars have defined social isolation as a lack of interaction with others and distinguished loneliness from subjective loneliness [3]. Although social isolation and loneliness are considered closely related, their definitions differ. Some studies have found only a weak correlation [4,5]; socially isolated individuals are not necessarily lonely, and vice versa. Despite these distinct definitions and realities, social isolation and loneliness are often studied together as overlapping concepts. Newall [6] also proposed that they should be examined in conjunction.

A growing body of evidence indicates that gender is a crucial factor influencing older adults' experiences of social isolation and loneliness. For example, older men (ie, those aged 65 years and older) in Japan are more likely to experience social isolation compared with their female counterparts [7,8]. In a cross-sectional study, Nomura and Kobayashi [9] discovered that gender can influence the isolation-prevention strategies used by community-dwelling older people in Japan. The authors found that satisfaction with social activities (ie, the degree of satisfaction with activities involving participation in groups and organizations and interpersonal activities with others) and strategies for interactions helped older Japanese women maintain relationships with others. However, for older Japanese men, only satisfaction with social activities contributed to maintaining social relationships. Some studies also identify the need to focus intervention efforts on men [10]. Men are more likely to feel lonely compared with women for one reason: when confronted with challenges or distressing situations, women tend to respond emotionally and engage in conversation with others, whereas men tend to avoid stressors and neglect talking to others about them [11]. These findings underscore the importance of considering gender in the development of interventions targeting social isolation. Nevertheless, few intervention programs have appropriately accounted for gender. Furthermore, a comprehensive overview of findings on the salience of gender in social isolation research is lacking. Moreover, the treatment of gender in research on social isolation is inconsistent, given that researchers adopt different approaches to this issue.

Barreto et al [12] argued that research on loneliness has failed to address gender beyond the binary categories of male and female, thereby obscuring important influences. Other meta-analyses have also indicated a correlation between loneliness and gender. For example, Pollak et al [13] identified gender as a predictor and risk factor for loneliness and functional decline, while Haihambo et al [14] suggested that gender differences exist in both loneliness and social adaptability. However, these studies represent only one dimension of the analysis and fail to specifically focus on gender. Although gender is a critical determinant in understanding social isolation and loneliness, the prevalence of gender-focused interventions remains underexplored.

Several studies have examined intervention programs designed to address social isolation and loneliness. For example, Milligan [15] conducted a scoping review of Men's Sheds and gender interventions to assess their impact on the health and well-being of older men. Men's Sheds are community-based, mutual self-help groups established primarily in Australia and the United Kingdom, among other countries. These initiatives were intended to alleviate social isolation and loneliness among retired men by fostering purpose, companionship, and shared activities [16]. The review identified limited evidence supporting the impact of Men's Sheds and other gender-focused social activities on the mental health and well-being of older men. However, this paper was published in March 2015; therefore, it does not cover research published after 2014. Fakoya et al [17] conducted a scoping review of studies on loneliness and social isolation interventions targeting older adults to identify effective interventions. The sample comprised 33 papers published until 2018. Findings indicated that the individual nature of experiences of social isolation and loneliness may impede the development and implementation of standardized interventions. The study concluded that no universal approach exists for addressing social isolation and loneliness, and interventions must be tailored to specific individuals and groups, as well as the severity of loneliness. Although existing studies encompass diverse interventions, a key limitation is that none explicitly considered gender as an influencing variable. Although a few studies address the potential impact of gender on interventions targeting the alleviation of social isolation and loneliness, the actual prevalence of such gender-focused interventions remains unclear.

Objective

This study aims to comprehensively map intervention programs targeting social isolation and loneliness among older adults, with a particular emphasis on the integration of gender considerations in these programs. Specifically, it intends to (1) identify the male-to-female ratio of participants across studies and (2) determine effective intervention strategies that are responsive to gender differences. The results are expected to

XSL•FO
RenderX

facilitate the development and implementation of new gender-focused intervention programs.

Methods

Research Design

A scoping review was conducted in accordance with the Joanna Briggs Institute (JBI) manual for evidence synthesis (hereafter, JBI framework) [18]. The JBI framework defines the purpose of a scoping review as identifying and analyzing knowledge gaps and examining how research is conducted within a given field [18]. The decision to conduct a scoping review was aimed at mapping the existing evidence and identifying areas that require further exploration rather than synthesizing findings, as customary in a systematic review. The review protocol was registered with the Center for Open Science Framework. No deviations were observed between the registered protocol and the content presented in this study. The inclusion criteria were based on the following definitions of participants, concept, context, and types of sources of evidence.

Participants

The participants were older men and women residing in a community, regardless of whether they reported being socially isolated. The review also focused on older adults but did not strictly define the participants' ages. Thus, age was excluded as an eligibility criterion given the diverse definitions of "elderly" across countries and studies.

Concept

This study focused on the implementation of strategies for addressing social isolation and loneliness. However, studies examining the correlation of social isolation and loneliness with curfew restrictions during the COVID-19 pandemic were excluded. This study focused on cases of social isolation and loneliness in which underlying causes remained unclear and included gender as a variable. The present-day concept of gender is multifaceted; nevertheless, it is often reduced to a binary model that assigns specific biological and behavioral traits to men and women.

Context

Social isolation and loneliness are influenced by various social factors, including geographic location, religious affiliation, and racial identity. This study does not impose any specific limitations regarding such factors; instead, it encompasses diverse community-based interventions.

Types of Sources

This review covered various intervention studies, such as randomized controlled trials, nonrandomized controlled trials, and before-and-after studies. Quantifying the effectiveness of interventions targeting social isolation and loneliness is challenging. Therefore, the review also included websites, reports, and research papers while excluding observational studies, qualitative studies, reviews, and conference abstracts. The rationale for incorporating websites was twofold. First, a significant number of interventions that target social isolation and loneliness have not been disseminated through academic journals. Second, existing research on social isolation and loneliness has predominantly focused on academic publications, with nonacademic sources receiving minimal representation.

Literature Search Strategy

The literature search strategy in this study was developed in consultation with a librarian at Mejiro University. A search was conducted on 6 databases (PubMed, MEDLINE, Cochrane, CINAHL, ScienceDirect, and Web of Science) on May 30, 2023, to identify relevant papers published within an 11-year period (2013-2023). The following terms were searched in combination: intervention, social isolation, loneliness, elderly, older adults, older people, older peoples, sex characteristics, sex, gender, and difference (Textbox 1). In addition to database searches, hand searches were conducted on the websites of 3 journals that frequently publish research on social isolation and loneliness: *Health and Social Care in the Community*, *Journal of Aging and Health*, and *Ageing & Society*. Furthermore, a hand search for gray literature was conducted on 2 websites, the Social Care Institute for Excellence and Connect2Affect, for the terms "social isolation" or "loneliness," with the publication period limited to 2013-2023 (Multimedia Appendix 1).

Textbox 1. Search strategy.

"(intervention)

and (((((social isolation)or(loneliness))

and (((aged)or(elderly))or((""older people"")or(""older peoples""))))

and((sex characteristics)or(((sex)or(gender))and(difference)))))"

Literature Selection

Data extraction was performed by KN. All identified academic studies and relevant texts were uploaded to Rayyan [19]. Duplicate papers were omitted, followed by a screening. KN and EI conducted the initial and full-text screening of the literature, while NK was consulted in cases of disagreement. The sample included English-language papers featuring intervention programs for social isolation and loneliness. The following texts were excluded: conference abstracts; studies without mention of social isolation or loneliness in their objectives, methods, results, or discussion; and studies that focused on COVID-19 countermeasures.

Data Analysis

The selected studies were presented in the form of a table that included the following information for each study: author, country of origin, participants, intervention method, outcome, and gender. An additional table was constructed to illustrate the male-to-female participant ratio per study and to classify the

XSL•FO
RenderX

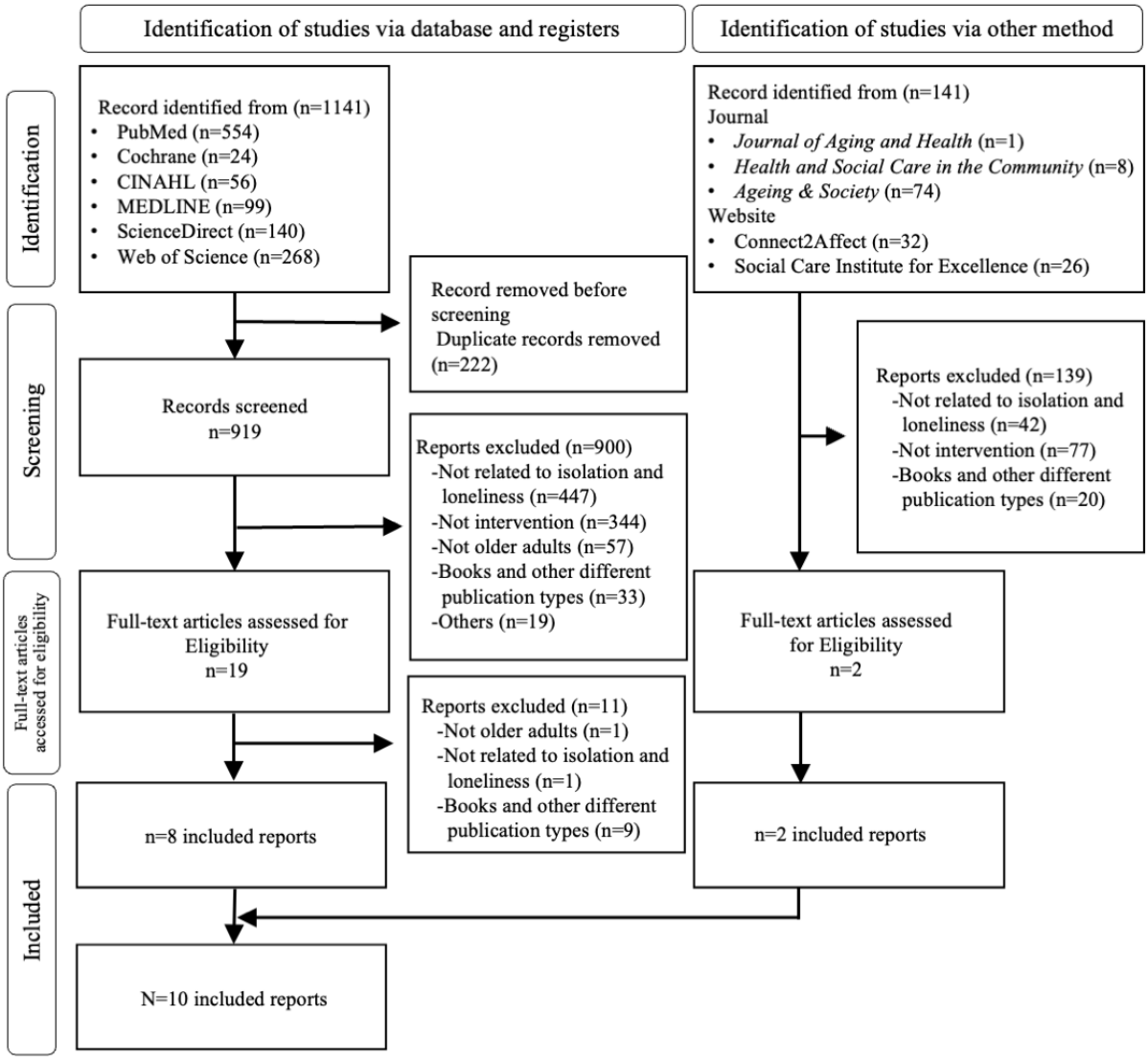
type of intervention (remote or in-person, group-based or individual, or a combination of these formats). Furthermore, the measures and outcome indicators used to evaluate the interventions were presented in tabular form. To enhance clarity and comparability, the primary outcome concepts were systematically categorized and organized according to their thematic relevance.

Results

Characteristics of the Selected Studies

The literature selection process was visualized as a flowchart in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Figure 1; Multimedia Appendix 2) [20]. The database search yielded 1141 studies, of which 222 were duplicates and were therefore excluded. The screening process yielded 19 studies, 8 [21-28] of which were included in the analysis. Of the 900 studies that were excluded during the screening process, 447 were unrelated to social isolation and loneliness, while 344 were not intervention studies. The hand search yielded 141 studies, of which 2 were included (Table 1). A total of 10 studies [21-30] were included in the final analysis. These studies described detailed interventions conducted in the United States (n=3 [23,27,28]), the United Kingdom (n=2 [29,30]), Singapore (n=1 [26]), Turkey (n=1 [25]), Spain (n=1 [21]), Ireland (n=1 [23]), and India (n=1 [22]; Table 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of study selection.



XSL•FO
RenderX

Table 1. Summary of included studies.

First author, date, and country	Participants (final analyzed individuals)	Intervention description	Result	Gender references
Ngiam et al, 2022 [26], Singapore	Older adults who resided in the southeast region of Singapore, aged >55 years, belonging to a lower socioeconomic status. 138 were included for analysis (59 male and 79 female).	Project Wire Up was a volunteer-led, one-on-one, goal-directed, and home-based digital literacy program: (1) equipped with smartphones and internet connection; (2) trained by volunteers for 6 sessions (1 to 2 hours per session) over 3 months that were held in the older adults' homes; and (3) digitally connected to existing social networks.	There were significant improvements in digital literacy scores in the intervention group as compared to controls (mean difference 2.28, 95% CI 1.37-3.20; *P*<.001). There was no statistically significant difference in the University of California, Los Angeles 3-item loneliness scale, Lubben Social Network Scale-6, Personal Wellbeing Score, or EQ-5D Utility and visual analog scale score.	No mention
Santos-Olmo et al, 2022 [21], Spain	The sample (N=68: 18 male, 50 female) was 65 years of age or older; lived alone or lived with other people older than 65 years; had uncovered social and/or health needs; have little or no social support network; refused the assistance offered to cover their needs from the normalized social and/or health services.	The Psychological Support Service for Socially Isolated Elderly People (PSIE) facilitates contact between the elderly person and social and health services (both primary and specialized care) in their area, so that they can receive the care they need on each occasion.	Concerning the total score of the Health and Psychosocial Functioning (Total HoNOS65+, *t*=10.12, *P*<.001, Cohen *d*=1.49) and the Spanish adaptation of the Camberwell Needs Assessment Questionnaire for the Elderly (CANE; average number of unmet needs, *t*=19.99, *P*<.001, Cohen *d*=2.31), significant differences were also observed between pre- and posttreatment. Statistically significant changes were obtained in Global Assessment of Functioning (GAF; *t*=4.06, *P*<.001, Cohen *d*= 0.6) and the Spanish adaptation of the WHO Disability Assessment Short Scale (WHO-DAS-S; A. personal care and survival, *t*=8.82, *P*<.001, Cohen *d*=1.3).	The gender variable does not seem to have an influence on any of the outcome measures studied, except for alcohol use. PSIE has achieved positive results in both men and women, without gender appearing to be a relevant variable in these results.
Dodge et al, 2022 [27], United States	Participants with normal cognition or mild cognitive impairment were recruited from Portland, OR, and Detroit, MI. Key inclusion criteria included (1) age 75 years or older, (2) socially isolated. One hundred eighty-six participants (86 with normal cognition and 100 (52.8%) with mild cognitive impairment) were randomized into the experimental (n=94) or the control group (n=92).	The intervention group had video chats with trained study staff for 30 minutes per day, 4 times per week for 6 months (high dose), and then twice per week for an additional 6 months (maintenance dose). Both intervention and control groups received a phone call once per week (approximately 10 minutes duration) to assess changes in health and social activities.	After the induction period, the experimental group had higher global cognitive test scores (Montreal Cognitive Assessment [primary outcome]; 1.75 points [*P*=.03]). After induction, participants in the experimental group with normal cognition had higher language-based executive function (semantic fluency test [secondary outcome]; 2.56 points [*P*=.03]). At the end of the maintenance period, participants in the experimental group with mild cognitive impairment had higher encoding function (Craft Story immediate recall test [secondary outcome]; 2.19 points [*P*=.04]).	No mention
Pandya et al, 2021 [22], India	Intervention group older adults (IN_2=166, 136 males, 30 females), the control group (CN_2=157, 128 males, 29 females), who underwent no intervention	The key features of the meditation program were (1) postures interspersed with relaxation, (2) slowness in movements, and (3) inner watchful awareness.	There were significant mean differences in the posttest scores on loneliness, well-being, life satisfaction, and contentment outcomes of the intervention group, with high observed effect sizes (Cohen *d* range=2.43-8.78; *P*≤.01). The intervention group older adults reported that they were less lonely and experienced greater well-being, life satisfaction, and contentment posttest (ηp^2 =0.71-0.78; *P*≤.01).	Men, middle class, married, and cohabitating participants, who also comprised a majority of the sample, were less lonely at the pretest phase as compared to women, upper class, single, and living alone. Gender was an important factor, and results showed that retired South Asian men were less lonely and more satisfied pretest as well as posttest.

XSL•FO
RenderX

First author, date, and country	Participants (final analyzed individuals)	Intervention description	Result	Gender references
Mays et al, 2021 [23], United States	Participants (n=382, 63 male, 315 female) were offered enrollment based on the following inclusion criteria: aged 50 years or older, community-dwelling, able to complete questionnaires, able to consent to participate in the study, and able to communicate in English.	Participants met with the program coordinator and selected from 1 of the 4 evidence-based programs: Tai Chi for Arthritis, Enhance Fitness, the Arthritis Foundation Exercise Program, and the Healthier Living Workshop. Group health class lasted 6 to 8 weeks.	Older adults who met with a health coach and participated in a single session of community health programs reported decreased loneliness (ER[a] 0.931, 95% CI 0.895-0.968; *P*<.001) and social isolation (ER 1.033, 95% CI 1.016-1.050; *P*<.001) at 6 months post participation, compared to their baseline scores.	No mention
Zamir et al, 2020 [29], United Kingdom	Twenty-two residents aged ≥65 years (5 male, 17 female) across 3 British care homes	Twenty-two residents engaged with each other using "Skype quiz" sessions with the support of staff once a month over an 8-month trial. Residents met other residents from the 3 care homes to build new friendships and participate in a 30-minute quiz session facilitated by 8 staff members.	Analysis of the field notes revealed 5 themes of: residents with dementia remember faces, not technology, inter- and intraconnectedness, regaining sense of self and purpose, situational loneliness overcome, and organizational issues create barriers to long-term implementation.	No mention
Lawlor et al, 2019 [24], Ireland	Participants (n=40) from older (aged≥50 years) women's groups from 4 different community centers	Intervention consisted of 3 face-to-face group education sessions, encouragement to enlist the support of a buddy (eg, spouse, partner, friend, or a group member), an information pack, and the option of weekly telephone contact. Each education session lasted approximately 20 minutes.	87% (40/46) of women consented to participate, and 78% (31/40) attended all education sessions. Few participants provided valid accelerometer data, but 63% (25/40) completed the HADS[b] questionnaire at all time points. 85% of participants (34/40) were somewhat or very satisfied with their involvement in the study.	No mention
Alıcı et al, 2018 [25], Turkey	The study participants were older adults living in 2 nursing homes set up by foundations located in the capital of Turkey. Their ages were more than 65 years. A total of 50 older adults formed the intervention group (n=20; 9 males, 11 females) and control group (n=30; 14 males, 16 females).	Laughter therapy was conducted by the principal investigator 2 days a week, with one application during each session. The program involved performing yoga, breathing, and physical exercises, as well as laughter therapy. The program continued for 5 weeks for a total of 10 applications. The control group received no intervention.	A statistically significant difference (*P*<.001) between mean De Jong Gierveld Scale scores of the intervention (mean 7.15, SD 1.755) and control groups (mean 15.63, SD 5.027) was observed after the intervention. Median De Jong Gierveld Scale scores were significantly lower in the intervention group than in the control group. After therapy, the social loneliness score was significantly lower in the intervention group (mean 3.10, SD 1.553; *P*<.001) than in the control group (mean 6.90, SD 3.100). Post therapy, the emotional loneliness score was significantly lower (*P*<.001) in the intervention group (mean 4.05, SD 1.538) than in the control group (mean 8.73, SD 2.599).	No mention
Dodge et al, 2015 [28], United States	Eighty-three individuals participated (41 in the intervention group and 42 in the control group). Participants aged 70 years or older were included.	Daily 30-minute face-to-face communications were conducted during a 6-week trial period in the intervention group. The control group received only a weekly telephone interview.	Among the cognitively intact participants, the intervention group improved more than the control group on a semantic fluency test (*P*=.003) at the posttrial assessment and a phonemic fluency test (*P*=.004) at the 18-week assessments. Among those with mild cognitive impairment, a trend (*P*=.04) toward improved psychomotor speed was observed in the intervention group. No difference was found between intervention and control groups in the pre- to posttrial changes in the loneliness score, the secondary outcome.	No mention

XSL•FO
RenderX

First author, date, and country	Participants (final analyzed individuals)	Intervention description	Result	Gender references
Hind et al, 2014 [30], United Kingdom	The eligibility criteria included being aged ≥ 75 years, living independently, and having reasonable cognition. Fifty-six participants (control n=30, intervention n=26) were included in the intention-to-treat analysis group.	Manualized the telephone friendship (TF) with standardized training: (1) one-to-one befriending: 10- to 20-minute calls once per week for up to 6 weeks made by a volunteer befriender, followed by (2) TF groups of 6 participants: 1-hour teleconferences once per week for 12 weeks facilitated by the same volunteer. Control: usual health and social care provision.	The 2 groups were reasonably well matched with respect to baseline demographic characteristics. At 6 months post randomization, the SF-36 mental health mean (SD) scores were 77.5 (18.4) in the intervention group and 70.7 (21.2) in the control group, with a mean difference of 6.5 (95% CI –3.0 to 16.0); after adjusting for age, sex, and baseline scores, the mean difference was 9.5 (95% CI 4.5-14.5). In summary, over the 6-month follow-up period, there was no change in the SF-36 mental health scores in the intervention group, but there was a decline or deterioration in scores in the control group.	No mention

[a]ER: estimated ratio.

[b]HADS: Hospital Anxiety and Depression Scale.

The literature search was conducted using 2 complementary approaches: database and hand searches. Database searches were performed across the 6 abovementioned databases, yielding 1141 records, of which 8 studies [21-28] met the inclusion criteria. Hand searches were conducted across 3 journals and 2 websites, identifying 141 records, of which 2 [29,30] studies were eligible for inclusion.

In total, 10 studies [21-30] were included in the final analysis.

Table 1 presents a summary of the 10 papers included in the final analysis. All studies were published between 2013 and 2023. The United States and the United Kingdom were the countries with the highest representation in terms of intervention programs. Participants were generally older adults aged ≥50 years, with the majority being ≥65 years and predominantly female.

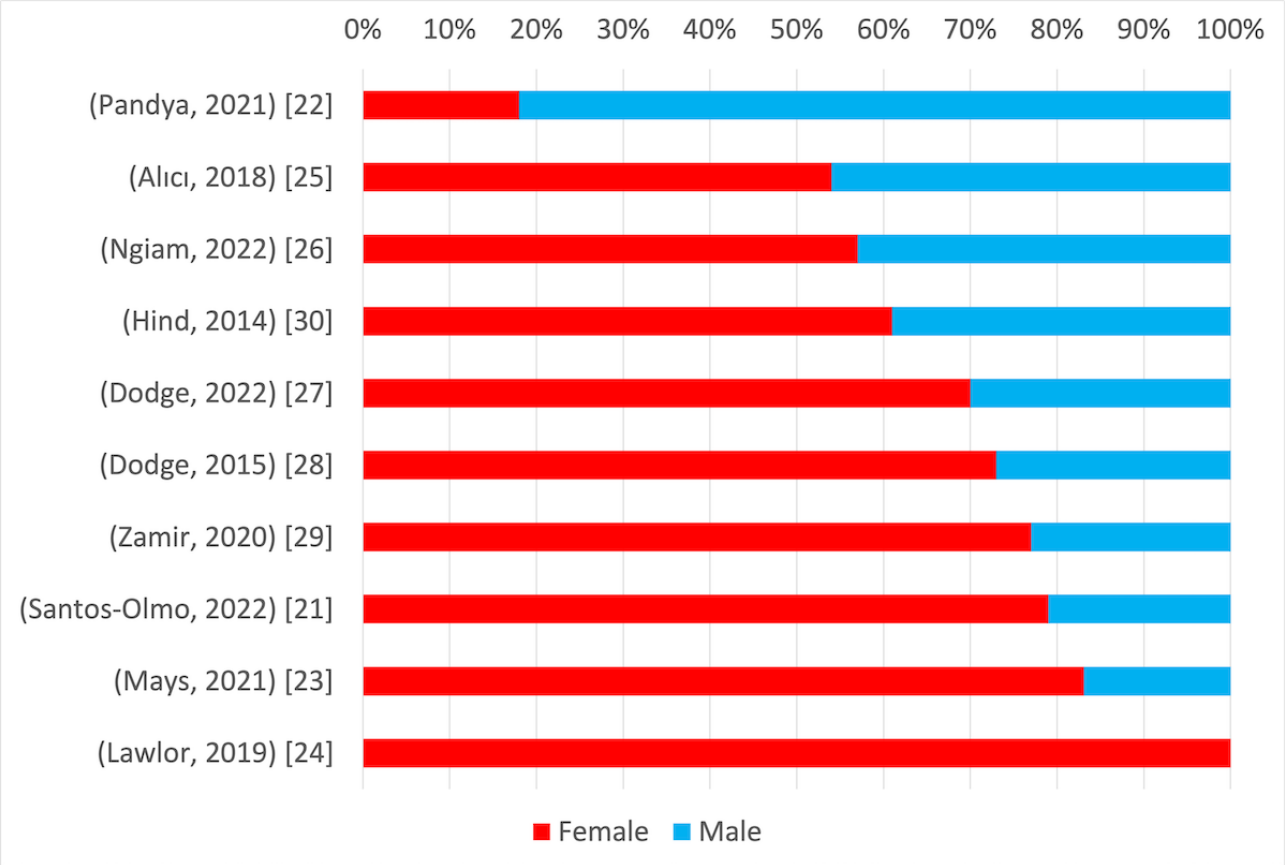
Characteristics of Participants

The 10 included studies [21-30] involved 1343 older adult participants, with sample sizes ranging from 22 to 382 (mean 134.3, SD 119.2; median 73, IQR 51.5-174). The proportion of female participants varied per study: approximately 10%-20%, 50%-60%, 60%-70%, 70%-80%, 80%-90%, and 100% women participated in 1 [26], 2 [21,27], 1 [22], 4 [23-25,29], 1 [28], and 1 [30] study, respectively (Figure 2 [21-30]). Overall, 8 documents incorporated the term “social isolation” or “loneliness” in their objectives, methods, or results. However, these documents did not explicitly address gender in the discussion section. In contrast, a few documents incorporated gender-related content in the discussion section. Studies referencing gender examined its impact on intervention effectiveness [21] and loneliness levels [22]. Among the included studies, the most common intervention method for those with the highest proportion of male participants (82%) was a personalized meditation program [22]. Conversely, studies involving women used different methods, including community-based group health classes [23] and physical activity and social support [24].

Only 1 [22] study reported a higher number of male participants compared with female ones; the remaining studies had a majority of female participants. One study [24] exclusively targeted older women, whereas no study focused solely on older men.

XSL•FO
RenderX

Figure 2. Gender ratio of participants [21-24, 26-30].



Intervention Methods and Means

The studies used 4 types of study designs: randomized controlled trials (n=5), nonrandomized control trials (n=1), pre-post outcome studies (n=3), and action research (n=1). They assessed 5 types of in-group programs: a program intended to enhance access to social and health services [21], a meditation program [22], an exercise and workshop [23], a physical activity program [24], and a laughter therapy program [25]. Overall, 3 studies discussed individualized remote programs, including a digital literacy program [25], a video chatting program [27], and a face-to-face communication program [28]. Another study assessed a remote group program that used a Skype (Microsoft) quiz format [29]. Furthermore, 1 study evaluated a remote individual program and a remote group program based on telephone friendship [30] (Table 2). The exercise workshop [23] effectively alleviated social isolation and loneliness indicators; meanwhile, the meditation [22] and laughter therapy programs [25] effectively mitigated loneliness (Table 1).

The interventions were nearly evenly divided between “remote” (n=5) and “in-person” (n=5) formats. Notably, no individual in-person programs were reported; instead, group-based in-person interventions were the most common. In-person programs typically involved strategies such as facilitating access to social resources, meditation, exercise, and workshops. Remote programs included activities such as digital literacy training, video-based communication, and online conversational sessions.

XSL•FO
RenderX

Table 2. Intervention methods.

Remote	Face-to-face
Individual	
• Digital literacy program (Ngiam et al [26], 2022) • Video chats (Dodge et al [27], 2022) • Face-to-face communications (Dodge et al [28], 2015)	—[a]
Group	
• Skype quiz (Zamir et al [29], 2020)	• Improving access to social and health services (Santos-Olmo et al [21], 2022) • Meditation program (Pandya [22], 2021) • Physical activity (Lawlor et al [24], 2019) • Laughter therapy (Alıcı et al [25], 2018)
Individual and group	
• Telephone friendship (Hind et al [30], 2014)	• Exercise and workshops (Mays et al [23], 2021)

[a]Not available.

Measurement of Effectiveness

The 10 studies [21-30] assessed outcomes based on 2 indicators of social isolation, 4 indicators of loneliness, and 29 other indicators (Table 3). Among these, outcomes with positive effects were observed for the Duke Social Support Index [23] for social isolation, the 3-item loneliness scale of the University of California, Los Angeles [23], the De Jong Gierveld Loneliness Scale [25], and the 6-item De Jong Gierveld Loneliness Scale [22].

The selected studies presented relatively few measures for assessing social isolation and loneliness; however, they used a broad range of indicators to evaluate cognitive function, well-being, satisfaction, and other psychosocial domains. For outcomes related to social isolation or loneliness and cognitive function, the number of studies that reported significant effects was approximately equal to those reporting no effects.

XSL•FO
RenderX

Table 3. Outcomes and effectiveness.

Categories and outcomes	Effective	Ineffective
Social isolation		
LSNS-6[a]	• —[b]	Ngiam et al [26], 2022
DSSI[c]	• Mays et al [23], 2021	—
Loneliness		
UCLA-3[d]	• Mays et al [23], 2021	Ngiam et al [26], 2022
The De Jong Gierveld Loneliness Scale score	• Alıcı et al [25], 2018	—
3-item loneliness scale	—	Dodge et al [27], 2015
DJGLS-6[e]	• Pandya et al [22], 2021	—
Other		
Cognition		
Trail making test	—	Dodge et al [28], 2015
Mini mental state examination	—	Dodge et al [28], 2015
Cogstate computerized tests	—	Dodge et al [28], 2015
The Stroop test	—	Dodge et al [28], 2015
MoCA[f]	• Dodge et al [27], 2022	—
Craft Story immediate recall test	• Dodge et al [27], 2022	—
The Consortium to Establish a Registry for Alzheimer Disease word list learning	—	Dodge et al [28], 2015
The Consortium to Establish a Registry for Alzheimer Disease word list delayed recall	—	Dodge et al [28], 2015
CAMCI[g]	—	Dodge et al [28], 2015
Fluency		
Semantic fluency test	• Dodge et al [27], 2022) • Lawlor et al [24], 2019)	—
The composite of verbal fluency for letters (F, A, and S)	• Dodge et al [28], 2015)	—
Verbal fluency for the category animals	• Dodge et al [28], 2015)	—
Qualitative method		
WHO-DAS-S[h]	• Santos-Olmo et al [21], 2022	—
Thematic analysis	• Zamir et al [29], 2020	—
framework analysis	• Lawlor et al [24], 2019)	—
Well-being		
PWS[i]	—	Ngiam et al [26], 2022
WEMWBS[j]	• Pandya et al [22], 2021	—
QoL[k]		
EQ-5D	—	Ngiam et al [26], 2022
SF-36[l]	• Hind et al [30], 2014	—
Psychosocial functioning		

XSL•FO
RenderX

Categories and outcomes	Effective	Ineffective
HoNOS65+[m]	• Santos-Olmo et al [21], 2022)	—
GAF[n]	• Santos-Olmo et al [21], 2022)	—
Satisfaction		
SWLS[o]	• Pandya et al [22], 2021	—
CLAS[p]	• Pandya et al [22], 2021	—
Emotion		
TDAS[q]	• Alıcı et al [25], 2018	—
NIHTB-EB[r]	• Dodge et al [27], 2022	—
Digital literacy		
Digital literacy score	• Ngiam et al [26], 2022	—
Needs		
CANE[s]	• Santos-Olmo et al [21], 2022	—
Depression		
Hospital Anxiety and Depression Scale	—	Lawlor et al [24], 2019
fMRI[t]		
fMRI analyses	• Dodge et al [27], 2022	—

[a]LSNS-6: Lubben Social Network Scale-6.
[b]Not available.
[c]DSSI: The Duke Social Support Index.
[d]UCLA-3: University of California, Los Angeles 3-item loneliness scale.
[e]DJGLS-6: The De Jong Gierveld Loneliness Scale (six-items).
[f]MoCA: Montreal Cognitive Assessment.
[g]CAMCI: Computer assessment of mild cognitive impairment.
[h]WHO-DAS-S: World Health Organization Brief Disability Assessment Scale.
[i]PWS: Personal Well-being Score.
[j]WEMWBS: Warwick Edinburgh Mental Well-being Scale.
[k]QoL: quality of life.
[l]SF-36: 36-Item Short Form Health Survey.
[m]HoNOS65+: Home Health Outcome Scales for People Over 65.
[n]GAF: Global Assessment of Functioning.
[o]SWLS: The Satisfaction with Life Scale.
[p]CLAS: The Contentment with Life Assessment Scale.
[q]TDAS: The Turkish Death Anxiety Scale.
[r]NIHTB-EB: National Institutes of Health-toolbox emotional battery.
[s]CANE: Camberwell Needs Assessment Questionnaire for the Elderly.
[t]fMRI: functional magnetic resonance imaging.

Discussion

Principal Findings

In this review, a total of 1282 papers were identified, of which only 10 [21-30] were eligible for analysis. Studies were excluded primarily because they did not specifically examine social isolation or loneliness or were merely cross-sectional investigations that emphasized the need for interventions. Of the included studies, East Asia was represented only by Singapore; notably, no research was included from Japan despite its status as a super-aged society.

Regarding the interventions, the methods were evenly distributed between remote and in-person methods. In total, 5 [26-30] of the 10 (50%) studies [21-30] used remote interventions (eg, digital literacy and video chats), while the other 5 studies [21-25] used in-person interventions (eg, exercise, meditation, and

XSL•FO
RenderX

laughter therapy). However, none of them combined remote and in-person approaches.

A significant gender bias was observed; in 9 studies [21-24,26-30], most participants were women. Among studies in which women accounted for ≥80% or more of the participants, interventions frequently emphasized communication, shared emotional expression, and group-based activities. Conversely, the study with the highest proportion of male participants (82%) featured a meditation program characterized by self-dialogue and individual reflection.

Importantly, only 2 [21,22] of the 10 studies [21-30] discussed the impact of gender on intervention outcomes, whereas no study focused on gender-specific issues as a primary research objective. Furthermore, the use of specific indicators for social isolation and loneliness was limited; indicators were typically substituted by measures of cognitive function or quality of life, depending on the study's specific objectives.

Comparison With Prior Work

The finding that 50% of the studies used remote interventions represents a significant increase compared with the result of a previous scoping review conducted by Fakoya et al [17]. The authors found that only 15% (5 of 33) of interventions cited in papers published between 1984 and 2018 were remote. This trend aligns with the growing efficacy and development of remote tools [31,32], likely accelerated by the global COVID-19 pandemic. However, while Sen et al [32] recommended the use of online and offline resources to enhance efficacy, this review found no studies adopting such a hybrid approach.

The predominance of female participants is consistent with previous literature emphasizing that research on social isolation and loneliness is biased toward women [33,34]. This finding contradicts evidence suggesting that older men are more prone to social isolation compared with women [7,8].

Regarding intervention preferences, the findings support existing theories on gender differences. Female participants' preference for exchanging experiences in supportive settings mirrors the findings of Venter et al [35]. In contrast, the male preference for the meditation program aligns with research proposing that men gravitate toward task-oriented activities, autonomy, and less intimate social networks, as seeking help or emotional disclosure can be perceived as vulnerability [36].

The lack of gender-specific analysis in the included studies is in contrast with the recommendations of Santos-Olmo et al [21] and Wen et al [10], who argued that future research needs to prioritize gender as a variable for predicting the effectiveness of interventions. The absence of studies conducted in Japan is also a concern, given the cultural specificity of isolation [13,37], thus highlighting a gap between Japan's urgent policy needs [38] and the available evidence base.

Strengths and Limitations

Strengths

The primary strength of this study is its explicit focus on gender. To the best of our knowledge, this is the first literature review on social isolation and loneliness that considers gender as a central variable. Additionally, relevant studies were identified using a rigorous methodology based on multidisciplinary databases.

Limitations

First, the review may not have captured all relevant studies, partially because the study period overlapped with the COVID-19 pandemic, potentially causing delays in publication. Second, the literature search concluded in May 2023, excluding subsequent studies. Third, despite manual searches, some high-quality sources or community-based practices that indirectly address isolation (eg, neighborhood association activities) may have been missed. Finally, as a scoping review, no quality assessment of the included papers was conducted; thus, the quality of evidence may vary.

Future Directions

Gender-specific interventions: a clear need exists to develop and implement new gender-focused intervention programs.

For older women: interventions should prioritize group-based activities that foster dialogue, mutual understanding, and emotional connection.

For older men, future efforts must rigorously identify criteria for interventions targeting men. Programs should incorporate activities that enable self-expression without overt emotional disclosure, such as those emphasizing skill use, personal autonomy, and a sense of contribution.

Research design: future studies must move beyond merely reporting gender demographics. Studies should include gender as the primary variable in order to analyze its impact on intervention outcomes. A reexamination of previous studies that failed to show significant effects could reveal whether the lack of gender consideration was a contributing factor.

Hybrid approaches: given the increased popularity of remote interventions and the established value of in-person interaction, future interventions should aim to combine remote and in-person methods. Leveraging the experience of online-based interventions to create hybrid models could enhance the precision and efficacy of countermeasures against social isolation and loneliness.

Conclusions

This scoping review, which aimed to examine the composition of participants by gender and gender-sensitive strategies, confirms that research on social isolation generally underestimates gender. This finding reveals a bias in which women are markedly overrepresented in study samples, contradicting evidence that older men are more susceptible to isolation. Critically, no study considered gender-specific issues in its objectives, undermining the efficacy of general interventions. We found gender-based differences in preferences: women favored conversation and emotional exchange, while men exhibited the highest proportion of participation in self-dialogue meditation programs. Therefore, a clear need exists to develop new gender-focused intervention programs and systematically identify conditions for effective interventions that specifically target older men.

XSL•FO
RenderX

Acknowledgments

The authors extended their gratitude to the librarians at Ibaraki Prefectural University of Health Sciences for their invaluable assistance in the design of the search strategy. No text generated by artificial intelligence was used in the writing of this study.

Funding

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grants-in-Aid for Scientific Research (grant JP23K12637).

Data Availability

All data generated or analyzed during this study are included in this published paper.

Authors' Contributions

Conceptualization: KN (lead) and N Kiguchi (supporting)
Data curation: KN
Formal analysis: KN (lead) and EI (supporting)
Funding acquisition: KN
Investigation: KN (lead) and EI (supporting)
Methodology: KN (lead) and N Kiguchi (supporting)
Project administration: KN
Resources: KN
Supervision: KN (lead), TN (supporting), and N Kiguchi (supporting)
Validation: KN
Visualization: KN
Writing – original draft: KN
Writing – review and editing: KN (lead), TN (supporting), and N Kobayashi (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 15 KB - ijmr_v15i1e72281_app1.docx]

Multimedia Appendix 2
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 555 KB - ijmr_v15i1e72281_app2.pdf]

References

1. Valtorta N, Hanratty B. Loneliness, isolation and the health of older adults: do we need a new research agenda? J R Soc Med 2012;105(12):518-522 [FREE Full text] [doi: 10.1258/jrsm.2012.120128] [Medline: 23288086]
2. Kino S, Stickley A, Arakawa Y, Saito M, Saito T, Kondo N. Social isolation, loneliness, and their correlates in older Japanese adults. Psychogeriatrics 2023;23(3):475-486 [FREE Full text] [doi: 10.1111/psyg.12957] [Medline: 36966745]
3. Dickens AP, Richards SH, Greaves CJ, Campbell JL. Interventions targeting social isolation in older people: a systematic review. BMC Public Health 2011;11:647 [FREE Full text] [doi: 10.1186/1471-2458-11-647] [Medline: 21843337]
4. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. Proc Natl Acad Sci U S A 2013;110(15):5797-5801 [FREE Full text] [doi: 10.1073/pnas.1219686110] [Medline: 23530191]
5. Tanskanen J, Anttila T. A prospective study of social isolation, loneliness, and mortality in Finland. Am J Public Health 2016;106(11):2042-2048. [doi: 10.2105/AJPH.2016.303431] [Medline: 27631736]
6. Newall NEG, Menec VH. Loneliness and social isolation of older adults: why it is important to examine these social aspects together. J Soc Pers Relat 2017;36(3):925-939. [doi: 10.1177/0265407517749045]
7. Saito M, Shmizu Y, Yamaguchi M, Takei S. Prevalence and characteristics of the socially isolated elderly in a large urban area. Japanese Journal of Social Welfare 2009;50(1):110-122 [FREE Full text]

XSL•FO
RenderX

8. Miura KW, Sekiguchi T, Otake-Matsuura M, Sawada Y, Tanaka E, Watanabe T, et al. Patterns of social relationships among community-dwelling older adults in Japan: latent class analysis. BMC Geriatr 2022;22(75):1-9 [FREE Full text] [doi: 10.1186/s12877-022-02748-7] [Medline: 35078419]
9. Nomura K, Kobayashi N. Gender gaps in strategies for maintaining the social participation and interaction of older adults with people in a local community. Health Soc Care Community 2024;2024(1). [doi: 10.1155/2024/3572577]
10. Wen Z, Peng S, Yang L, Wang H, Liao X, Liang Q, et al. Factors associated with social isolation in older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2023;24(3):322-330.e6. [doi: 10.1016/j.jamda.2022.11.008] [Medline: 36549651]
11. Tamres LK, Janicki D, Helgeson VS. Sex differences in coping behavior: a meta-analytic review and an examination of relative coping. Pers Soc Psychol Rev 2002;6(1):2-30. [doi: 10.1207/s15327957pspr0601_1]
12. Barreto M, Doyle DM, Maes M. Researching gender and loneliness differently. Ann N Y Acad Sci 2025;1544(1):55-64. [doi: 10.1111/nyas.15283] [Medline: 39761171]
13. Pollak C, Verghese J, Blumen H. Loneliness and functional decline in aging: a systematic review. Res Gerontol Nurs 2023 Jul;16(4):202-212. [doi: 10.3928/19404921-20230503-02] [Medline: 37159388]
14. Haihambo N, Layiwola D, Blank H, Hurlemann R, Scheele D. Loneliness and social conformity: a predictive processing perspective. Ann N Y Acad Sci 2025;1547(1):5-17. [doi: 10.1111/nyas.15324] [Medline: 40173107]
15. Milligan C, Neary D, Payne S, Hanratty B, Irwin P, Dowrick C. Older men and social activity: a scoping review of Men's Sheds and other gendered interventions. Ageing Soc 2015;36(5):895-923 [FREE Full text] [doi: 10.1017/S0144686X14001524]
16. Kelly D, Steiner A, Mason H, Teasdale S. Men's Sheds: a conceptual exploration of the causal pathways for health and well-being. Health Soc Care Community 2019 [FREE Full text] [doi: 10.1111/hsc.12765] [Medline: 31206945]
17. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020;20(1):129 [FREE Full text] [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
18. JBI manual for evidence synthesis. JBI. URL: https://synthesismanual.jbi.global [accessed 2025-06-27]
19. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
20. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
21. Santos-Olmo A, Ausín B, Muñoz M. People over 65 years old in social isolation: description of an effective community intervention in the city of Madrid (Spain). Int J Environ Res Public Health 2022;19(5) [FREE Full text] [doi: 10.3390/ijerph19052665] [Medline: 35270358]
22. Pandya SP. Meditation program mitigates loneliness and promotes wellbeing, life satisfaction and contentment among retired older adults: a two-year follow-up study in four South Asian cities. Aging Ment Health 2021;25(2):286-298. [doi: 10.1080/13607863.2019.1691143] [Medline: 31755300]
23. Mays AM, Kim S, Rosales K, Au T, Rosen S. The leveraging exercise to age in place (LEAP) study: engaging older adults in community-based exercise classes to impact loneliness and social isolation. Am J Geriatr Psychiatry 2021;29(8):777-788 [FREE Full text] [doi: 10.1016/j.jagp.2020.10.006] [Medline: 33268235]
24. Lawlor ER, Cupples ME, Donnelly M, Tully MA. Promoting physical activity among community groups of older women in socio-economically disadvantaged areas: randomised feasibility study. Trials 2019;20(1):234 [FREE Full text] [doi: 10.1186/s13063-019-3312-9] [Medline: 31023329]
25. Alıcı NK, Bahceli PZ, Emiroğlu ON. The preliminary effects of laughter therapy on loneliness and death anxiety among older adults living in nursing homes. Int J Older People Nurs 2018;13(4):e12206. [doi: 10.1111/opn.12206] [Medline: 30004172]
26. Ngiam N, Yee WQ, Teo N, Yow KS, Soundararajan A, Lim JX, et al. Building digital literacy in older adults of low socioeconomic status in Singapore (project wire up): nonrandomized controlled trial. J Med Internet Res 2022;24(12):e40341 [FREE Full text] [doi: 10.2196/40341] [Medline: 36459398]
27. Dodge H, Yu K, Wu C, Pruitt P, Asgari M, Kaye JA, et al. Internet-based conversational engagement randomized controlled clinical trial (I-CONECT) among socially isolated adults 75+ years old with normal cognition or mild cognitive impairment: topline results. Gerontologist 2024 Apr 01;64(4) [FREE Full text] [doi: 10.1093/geront/gnad147] [Medline: 37935416]
28. Dodge HH, Zhu J, Mattek NC, Bowman M, Ybarra O, Wild KV, et al. Web-enabled conversational interactions as a means to improve cognitive functions: results of a 6-week randomized controlled trial. Alzheimers Dement (N Y) 2015 May;1(1):1-12 [FREE Full text] [doi: 10.1016/j.trci.2015.01.001] [Medline: 26203461]
29. Zamir S, Hennessy C, Taylor A, Jones R. Intergroup 'Skype' quiz sessions in care homes to reduce loneliness and social isolation in older people. Geriatrics (Basel) 2020 Nov 11;5(4) [FREE Full text] [doi: 10.3390/geriatrics5040090] [Medline: 33187242]
30. Hind D, Mountain G, Gossage-Worrall R, Walters SJ, Duncan R, Newbould L, et al. Putting Life in Years (PLINY): a randomised controlled trial and mixed-methods process evaluation of a telephone friendship intervention to improve mental well-being in independently living older people. Public Health Research 2014:1-225. [Medline: 25642518]

XSL•FO
RenderX

31. Czaja S, Boot W, Charness N, Rogers W, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018;58(3):467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]
32. Sen K, Prybutok G, Prybutok V. The use of digital technology for social wellbeing reduces social isolation in older adults: a systematic review. SSM Popul Health 2022 Mar;17 [FREE Full text] [doi: 10.1016/j.ssmph.2021.101020] [Medline: 35024424]
33. Chen YR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016;18(1):e18 [FREE Full text] [doi: 10.2196/jmir.4596] [Medline: 26822073]
34. Jarvis M, Padmanabhanunni A, Balakrishna Y, Chipps J. The effectiveness of interventions addressing loneliness in older persons: an umbrella review. Int J Afr Nurs Sci 2020 May 17;12:100177-100118 [FREE Full text] [doi: 10.1016/j.ijans.2019.100177]
35. van de Venter E, Buller A. Arts on referral interventions: a mixed-methods study investigating factors associated with differential changes in mental well-being. J Public Health (Oxf) 2015;37(1):143-150. [doi: 10.1093/pubmed/fdu028] [Medline: 24839293]
36. Shepherd G, Astbury E, Cooper E, Dobrzynska W, Goddard E, Murphy H. The challenges preventing men from seeking counselling or psychotherapy. Ment Health Prev 2023:1-10 [FREE Full text] [doi: 10.1016/j.mhp.2023.200287]
37. Meehan DE, Grunseit A, Condie J, HaGani N, Merom D. Social-ecological factors influencing loneliness and social isolation in older people: a scoping review. BMC Geriatr 2023 Nov 09;23(1):726 [FREE Full text] [doi: 10.1186/s12877-023-04418-8] [Medline: 37946155]
38. You are not alone. The Office for Policy on Loneliness and Isolation. URL: https://www.notalone-cao.go.jp/english/ [accessed 2025-08-21]

Abbreviations

JBI: Joanna Briggs Institute
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by N Cahill, M Balcarras; submitted 07.Feb.2025; peer-reviewed by C Chikwe, VS Silva, J Scott; comments to author 14.Aug.2025; accepted 29.Dec.2025; published 18.Feb.2026.

Please cite as:
Nomura K, Kiguchi N, Inomata E, Nakamachi T, Kobayashi N
Treatment of Gender in Research on Intervention Programs Targeting Social Isolation and Loneliness Among Older Adults: Scoping Review
Interact J Med Res 2026;15:e72281
URL: https://www.i-jmr.org/2026/1/e72281
doi:10.2196/72281
PMID:

©Kenta Nomura, Naoto Kiguchi, Eisuke Inomata, Takeshi Nakamachi, Norikazu Kobayashi. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 18.Feb.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Trends, Predictors, and Outcomes of Monitored Acute Care Unit Admissions in Older Adults: 10-Year Retrospective Analysis

Ludovic van Delden[1], MD; Stephan von Düring[2,3], MD, MSc; Frederic Rouyer[3], MD; Anne Rossel[4], MD; Christophe Marti[4,5], MD; Sara Cereghetti[1], MD; Xavier Roux[1,6], MD; Thibaut Desmettre[2,3], MD, PhD; Olivier Grosgurin[2,3,4,5], MD; Christophe A Fehlmann[2,3], MD

[1]Division of Critical Care Medicine, Department of Acute Medicine, Geneva University Hospitals, Geneva, Switzerland

[2]Department of Anaesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland

[3]Division of Emergency Medicine, Department of Acute Medicine, Geneva University Hospitals, Rue Gabrielle Perret Gentil 4, Geneva, Switzerland

[4]Division of Internal and General Medicine, Department of Internal Medicine, Geneva University Hospitals, Geneva, Switzerland

[5]Department of Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland

[6]Division of Internal Medicine for the Elderly, Department of Rehabilitation and Geriatrics, Geneva University Hospitals, Geneva, Switzerland

Corresponding Author:
Christophe A Fehlmann, MD
Department of Anaesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Abstract

Background: Global population aging places an increasing burden on health care systems. This is driven by multimorbidity, frailty, and polypharmacy. Older adults, particularly those aged 65 years or older, use emergency departments (EDs) more frequently and experience poorer outcomes. In this population, decisions regarding admission to monitored acute care units—intensive care units, intermediate care units, and operating rooms—are frequent and complex. While ED and intensive care unit use are well documented, data on monitored acute care units as a whole remain limited. Evidence on admission trends, patient characteristics, and outcomes in older adults is scarce.

Objective: This study aimed to describe temporal trends in monitored acute care unit admissions, identify predictors of such admissions, and assess outcomes following these admissions.

Methods: We conducted a retrospective cohort study using routinely collected electronic health record data. We included patients aged 65 years or older who visited the EDs of the Geneva University Hospitals, Switzerland, between 2009 and 2019. The primary outcome was admission to a monitored acute care unit. The secondary outcomes were hospital length of stay, 7-day mortality, and 1-year mortality. Logistic regression models were used to identify factors associated with monitored acute care unit admission and to assess the association between age and mortality.

Results: During the 10-year period, 701,838 ED visits were recorded. Annual visits increased from 56,944 to 76,368 (+34.1%). The increase was greater among patients aged 65 years or older (+56.1%) than among younger patients (+26.5%). A total of 180,189 older patients presented to the ED. Of these, 887 (0.5%) died in the ED, 97,238 (54.0%) were discharged home, 63,025 (35.0%) were admitted to a ward, and 19,039 (10.6%) were admitted to a monitored acute care unit. Monitored acute care unit admissions increased from 1379 (10.3%) in 2009 to 2240 (11.1%) in 2019. This represented an absolute increase of 62.4% and a relative increase of 0.8%. Predictors of monitored acute care unit admission included younger age, male sex, ambulance arrival, higher triage level, being married or in a relationship, not residing in a nursing home, and French as the primary language. Among patients admitted to a monitored acute care unit, mortality was 5.8% (1105/19,039) at 7 days and 22.3% (4251/18,039) at 1 year. Older age was associated with higher 7-day mortality (adjusted odds ratio 1.55, 95% CI 1.14 - 2.10) and 1-year mortality (adjusted odds ratio 1.28, 95% CI 1.08 - 1.51).

Conclusions: Admissions to monitored acute care units among older patients increased over time. These findings indicate a growing demand for high-level care in this population. Hospitals should adapt infrastructure and resource allocation to address the needs of an aging population.

(Interact J Med Res 2026;15:e80629) doi:10.2196/80629

KEYWORDS

older patients; intensive care units; intermediate care units; critical care; hemodynamic monitoring; risk assessment

XSL•FO
RenderX

Introduction

The global population is aging rapidly, exerting substantial pressure on health care systems [1,2]. With increasing life expectancy, older adults present with a higher burden of chronic illnesses, polypharmacy, and biological frailty. However, only limited data exist on the use of monitored acute care units within this population.

Individuals aged between 65 and 79 years exhibit the highest prevalence of comorbidity, while those aged 80 years and older have a 30% likelihood of presenting with at least one chronic condition [3]. This demographic shift places considerable strain on health care services. In the United States, studies have shown that patients aged older than 65 years account for 15% of all emergency department (ED) admissions, with these consultations increasing by 25% over the past decade [4,5]. These patients typically present with more severe emergencies, requiring extensive hospital resources [6]. Moreover, they experience poorer outcomes following admission compared to younger patients, particularly in the intensive care unit. In-hospital mortality rates are markedly higher for older patients during the postintensive phase [7,8], with the majority of deaths occurring within 3 months after intensive care unit discharge [9]. Furthermore, survivors frequently experience a substantial decline in quality of life, with age and frailty being crucial factors that affect the recovery of functional capacity [10].

These age-associated challenges necessitate careful consideration when evaluating older patients for admission to high-level care. Understanding these aspects is essential for health care providers in developing strategies that effectively address the specific needs of older patients, ensuring that the allocation of high-level care resources is both efficient and beneficial. While extensive research exists on ED use and intensive care unit admissions in older patients, there is a notable gap in data regarding the trends in monitored acute care unit admissions over time. Furthermore, research on monitored acute care unit admissions, encompassing intensive care units, intermediate care units, and operating rooms, as a whole entity, remains limited. Finally, intermediate care units have only recently been introduced in health care systems, and their impact on patient disposition and role as acute care providers remains to be determined.

This study aimed to address this gap by analyzing ED and monitored acute care unit admissions among patients aged 65 years or older over the past decade in a Swiss tertiary care institution. Our primary objective was to describe the trends in monitored acute care unit admissions during this period. Our secondary objectives were to identify predictors of monitored acute care unit admission, describe patient outcomes following monitored acute care unit admission, and explore the association between patient age and mortality at day 7 and 1 year postadmission.

Methods

Study Design

This single-center retrospective cohort study used routinely collected data from patients who visited the ED between 2009 and 2019. The study was approved by the institutional ethics committee of Geneva, Switzerland, and patient consent was waived based on Article 34 of the Swiss law on human research.

Study Setting and Population

The study was conducted at Geneva University Hospitals (HUG), the largest medical complex in Switzerland. The HUG is a tertiary care institution with 2 adult EDs located at different sites. The main ED, which operates 24/7, handles approximately 80,000 visits annually and includes a stretcher bay and a walk-in sector, providing comprehensive emergency care. The second ED, opened in 2016, is situated at a geriatric hospital and handles approximately 5000 visits annually. It operates only during daytime hours and is dedicated to patients aged older than 75 years with low-severity conditions. Although this geriatric ED is equipped with a full radiological division, it has limited access to specialty consultants, sometimes requiring the transfer of patients who need urgent specialized consultation or surgical procedures. Hospitals within the HUG network share the same computer system, and patients' medical data are accessible in each hospital.

HUG provides specialized care across all medical fields, serving as a reference center for regional hospitals and often receiving patients from smaller centers. The main location offers 24-hour cardiac catheterization services and is the only trauma and stroke center in the region, providing immediate care for patients requiring intravenous thrombolysis, interventional radiology, or the management of life-threatening injuries.

The HUG's critical care infrastructure includes a medical-surgical intensive care unit and several specialized intermediate care units. These intermediate care units serve as an intermediary level of care between the intensive care unit and regular wards and include a cardiac intermediate care unit, a neurological and neurosurgical intermediate care unit (with a stroke unit), a surgical intermediate care unit, and 2 medical intermediate care units (one located at the main hospital and one at the geriatric hospital).

The HUG is funded through a mixed financing model, with approximately half of the budget provided by the State of Geneva and the remainder generated from hospital revenues (including patient care and insurance reimbursements) as well as dedicated research funding from national grants and private foundations. Health care financing in Switzerland is based on mandatory basic health insurance (LAMal), which ensures equal access to care for all residents and is funded through individual premiums, complemented by regulated cost-sharing mechanisms (eg, deductibles and co-payments) and optional supplementary insurance.

Participants

We included all adult patients (aged ≥18 y) who visited either of HUG's EDs between January 1, 2009, and December 31,

XSL•FO
RenderX

2019. Patients were excluded if they (1) were aged younger than 65 years, (2) were transferred from another hospital or ED outside of HUG, or (3) refused the use of their data for research purposes. The age cutoff of 65 years was selected, as it is a widely recognized threshold for defining "older patients" and aligns with the retirement age in Switzerland.

Variables and Data Sources

The primary outcome was monitored acute care unit admission, defined as an admission to an operating room (including surgical rooms, interventional radiology, or a cardiac catheterization laboratory), an intermediate care unit, or the intensive care unit. This outcome was based on patient trajectories within the institution. The secondary outcomes included hospital length of stay, mortality at day 7 after monitored acute care unit admission, and mortality at 1 year. Owing to the lack of linkage between the electronic health record and the national death registry, data were missing for patients who did not die in the hospital and were not readmitted beyond 1 year after the initial admission.

Patient variables extracted included age, sex, marital status (married or in a relationship vs single, divorced, or widowed), primary language, emergency triage level (based on the Swiss Emergency Triage Scale [11], ranging from 1 [highest priority] to 4 [lowest priority]), triage motive category, nursing home residence, and means of transportation to the ED (ambulance or nonambulance).

Data were electronically extracted from the hospital's data warehouse by a dedicated team, which contains all information routinely gathered for clinical use in the patient's electronic health record. These structured data can be retrieved for quality assessment or research purposes after approval from the research ethics board. Each care episode is assigned a unique number, facilitating linkage between different database subsets. There was no linkage with other databases.

Study Size

The sample size was determined based on the study's objective to describe trends over a decade rather than on power calculations for statistical significance. With an expected inclusion of more than 150,000 patients, the sample size was considered sufficient to perform a multivariable logistic regression analysis without the risk of overfitting, based on established methodological recommendations regarding events per variable. Previous studies have shown that logistic regression models are unlikely to be overfitted when at least 10 events per predictor variable are available and that model performance remains robust even with fewer events per variable in large datasets with stable estimates. Given the large number of observations and outcome events in this study, the risk of overfitting was therefore considered minimal [12-14].

Statistical Analysis

Data cleaning and statistical analysis were conducted using the integrated statistical software Stata/SE (version 18; StataCorp LLC).

First, the data were cleaned and standardized. Variables were reformatted to facilitate analysis. Outcomes were constructed based on patients' trajectories through the institution. To evaluate trends in monitored acute care unit admissions over time, we graphically computed the number of ED visits for patients aged 65 years or older and the number and type of monitored acute care unit admissions by year. The absolute and relative changes over time were reported. A linear regression model was used to evaluate trends over time and determine the statistical significance of the observed changes.

To explore predictors of monitored acute care unit admission, baseline variables were summarized using descriptive statistics for the overall population. Mean and SD (or median and IQR) were used for continuous variables, whereas frequency and proportion were used for categorical variables. Patients admitted to a monitored acute care unit were compared to those not admitted to a monitored acute care unit using standardized mean differences (SMDs). We used SMDs rather than *P* values, as these comparisons are descriptive and not inferential. An absolute SMD greater than 0.2 was considered to reflect a potentially meaningful imbalance. The association between age and monitored acute care unit admission was graphically represented using restricted cubic splines, a flexible regression method, with the number of knots determined by the Akaike information criterion and their placement based on recommended quantiles [14]. Independent predictors of monitored acute care unit admission were identified using a multivariable logistic regression model, with monitored acute care unit admission as the dependent variable. Potential predictors were selected based on previous literature or clinical relevance and included sex, age, marital status, nursing home residence, primary language, arrival by ambulance, emergency triage level, and triage motives. Relationships modeled with restricted cubic splines were presented as odds ratios (ORs), comparing older patients (75th percentile of age) to younger patients (25th percentile of age) [14]. A sensitivity analysis using age as a categorical variable was also performed.

To explore patient outcomes following monitored acute care unit admission, outcomes were summarized using descriptive statistics.

Finally, to explore the association between patient age and mortality at day 7 and 1 year following monitored acute care unit admission, restricted cubic splines were used. Subsequently, multivariable models similar to those used previously were applied.

Missing data frequency was reported for each variable. For the initial logistic regression model, a complete case analysis was performed. A subsequent analysis included variables with higher levels of missingness without imputation to assess changes in the coefficient estimates.

Ethical Considerations

The study was conducted at HUG in accordance with the principles of Good Clinical Practice (the Declaration of Helsinki, 2002). This study was approved on July 11, 2022, by the institutional ethics committee of Geneva, Switzerland (project ID 2022 - 00987). Patient consent was waived by this committee. No compensation was offered to participants Data extraction generated deidentified data used for the analysis.

XSL•FO
RenderX

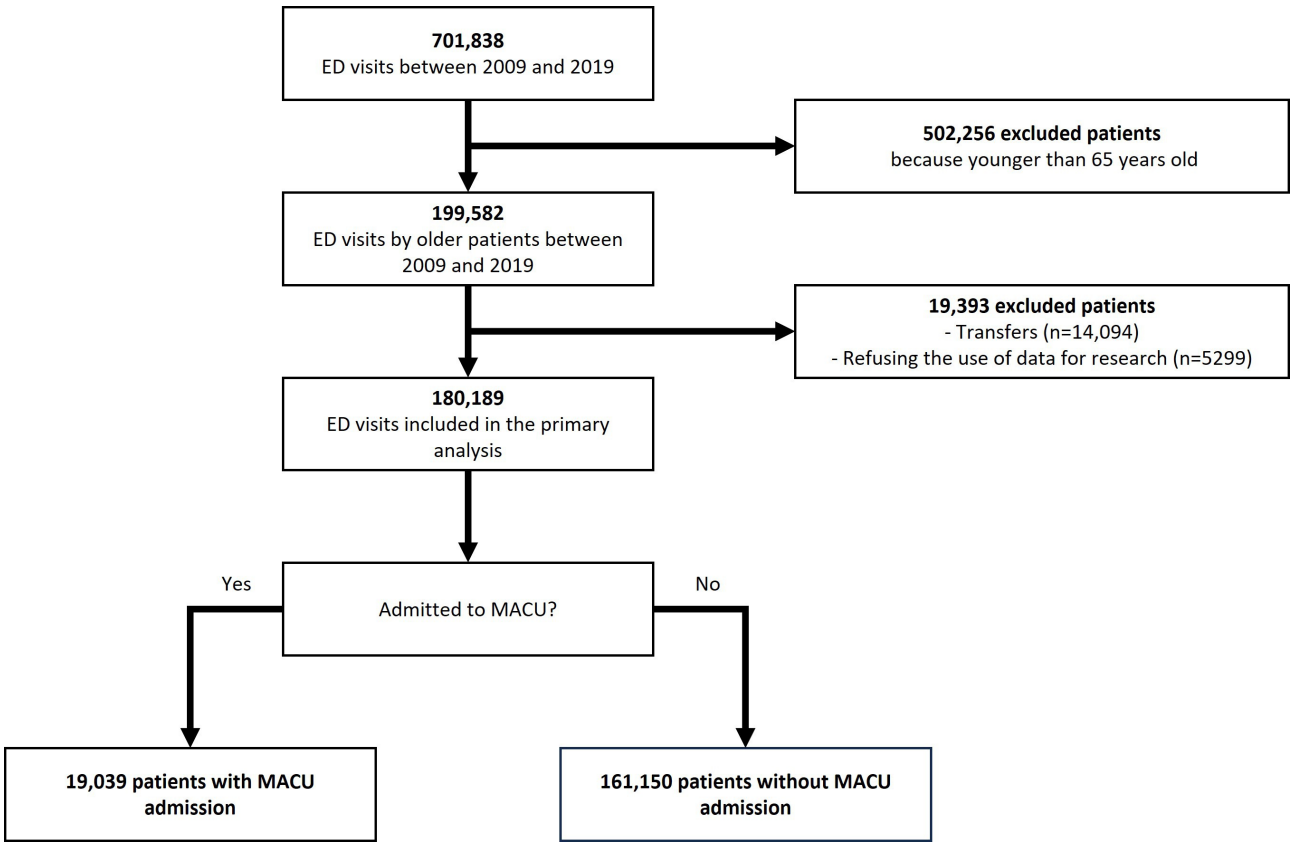
These data were subsequently anonymized for online publication.

Results

Inclusion of Patients and Characteristics

Between 2009 and 2019, a total of 701,838 ED visits were recorded at our institution (Multimedia Appendix 1). Of these, 180,189 (25.7%) patients met our inclusion criteria (Figure 1). The median age was 79 (IQR 72-86) years. Patient characteristics and the proportion of missing data are presented in Table 1.

Figure 1. Flowchart of our retrospective cohort study conducted at the Geneva University Hospitals (HUG) over a 10-y period (2009 - 2019). All adult patients (aged ≥18 y) who visited the HUG's emergency departments (EDs) were candidates for inclusion. Patients were excluded if they (1) were aged <65 y, (2) were transferred from another hospital or ED outside of HUG, or (3) refused the use of their data for research purposes. Of the 701,838 patients in our dataset, 180,189 met our inclusion criteria, representing 90.3% of all patients aged ≥65 y admitted to the ED during the study period. Of these patients, 19,039 (10.6%) were admitted to a monitored acute care unit (MACU).



XSL•FO
RenderX

Table . Baseline characteristics.

	Total (N=180,189)	With MACU[a] admission (n=19,039)	Without MACU admission (n=161,150)	SMD[b]
Age (y), median (IQR)	79 (72-86)	78 (72-84)	79 (72-86)	0.120
65 - 74, n (%)	61,511 (34.1)	6890 (36.1)	54,621 (33.9)	0.153
75 - 84, n (%)	66,603 (37.0)	7668 (40.3)	58,935 (36.6)	—[c]
85 - 94, n (%)	46,890 (26.0)	4190 (22.0)	42,700 (26.5)	—
≥95, n (%)	5185 (2.9)	291 (1.5)	4894 (3.0)	—
Sex, n (%)				0.150
Male	82,877 (46.0)	10,027 (52.7)	72,850 (45.2)	
Female	97,312 (54.0)	9012 (47.3)	88,300 (54.8)	
Marital status, n (%)				0.152
Married or in a relationship	79,499 (44.1)	9661 (50.7)	69,838 (43.3)	
Single, divorced, or widowed	100,435 (55.7)	9297 (48.8)	91,138 (56.6)	
Missing	255 (0.1)	81 (0.4)	174 (0.1)	
Nursing home resident, n (%)				0.058
No	164,454 (91.3)	17,552 (92.2)	146,902 (91.2)	
Yes	11,967 (6.6)	1027 (5.4)	10,940 (6.8)	
Missing	3768 (2.1)	460 (2.4)	3308 (2.1)	
Primary language, n (%)				0.123
French	56,788 (31.5)	6408 (34.0)	50,308 (31.2)	
Italian	5958 (3.3)	555 (2.9)	5403 (3.4)	
Spanish	3290 (1.8)	254 (1.3)	3036 (1.9)	
German	2069 (1.1)	233 (1.2)	1836 (1.1)	
English	1552 (0.9)	151 (0.8)	1401 (0.9)	
Portuguese	1355 (0.8)	111 (0.6)	1244 (0.8)	
Others	3821 (2.1)	326 (1.7)	3495 (2.2)	
Missing	105,536 (58.5)	10,929 (57.4)	64,427 (58.6)	
Arrival by ambulance, n (%)				0.478
No	66,881 (37.1)	3967 (20.8)	62,914 (39.0)	
Yes	91,829 (51.0)	13,572 (71.3)	78,257 (48.6)	
Missing	21,479 (12.0)	1500 (7.9)	19,979 (12.4)	
Triage scale, n (%)				1.152
1: vital emergency	20,921 (11.6)	8290 (43.5)	12,631 (7.8)	
2: urgent	65,305 (36.2)	7886 (41.4)	57,419 (35.6)	
3: Mildly urgent	89,840 (49.9)	2786 (14.6)	87,054 (54.0)	
4: nonurgent	3788 (2.1)	31 (0.2)	3757 (2.3)	
Triage motive category, n (%)				0.759
Cardiology-pneumology	48,348 (26.8)	7550 (39.7)	40,798 (25.3)	
Neurology-psychiatry	29,972 (16.6)	6318 (33.2)	23,654 (14.7)	
Traumatology	30,202 (16.8)	1598 (8.4)	28,604 (17.8)	

XSL•FO
RenderX

	Total (N=180,189)	With MACU[a] admission (n=19,039)	Without MACU admission (n=161,150)	SMD[b]
Digestive-OB/GYN[d]	17,731 (9.8)	1392 (7.3)	16,339 (10.1)	
Urology-nephrology	8285 (4.6)	263 (1.4)	8022 (5.0)	
Rheumatology	6793 (3.8)	105 (0.6)	6688 (4.2)	
Infectious disease	5813 (3.2)	435 (2.3)	5378 (3.3)	
Dermatology	4580 (2.5)	134 (0.7)	4446 (2.8)	
ENT[e]	3883 (2.2)	100 (0.5)	3783 (2.3)	
Others	24,263 (13.5)	1107 (5.8)	23,156 (14.4)	
Missing	319 (0.2)	37 (0.2)	282 (0.2)	

[a]MACU: monitored acute care unit.

[b]SMD: standardized mean difference.

[c]Not applicable.

[d]OB/GYN: obstetrics and gynecology.

[e]ENT: ear, nose, and throat.

Monitored Acute Care Unit Admissions and Evolution Over Time

Over the study period, 887 (0.5%) patients died in the ED, 97,238 (54.0%) patients were discharged home, 63,025 (35.0%) patients were admitted to a ward, and 19,039 (10.6%) patients were admitted to a monitored acute care unit. Of these, 4499 (23.6%) patients were admitted to the intensive care unit, 10,835 (56.9%) patients were admitted to an intermediate care unit, and 3705 (19.5%) patients were admitted directly to an operating room (Table 2).

Table . Patients' disposition from the emergency department by age categories.

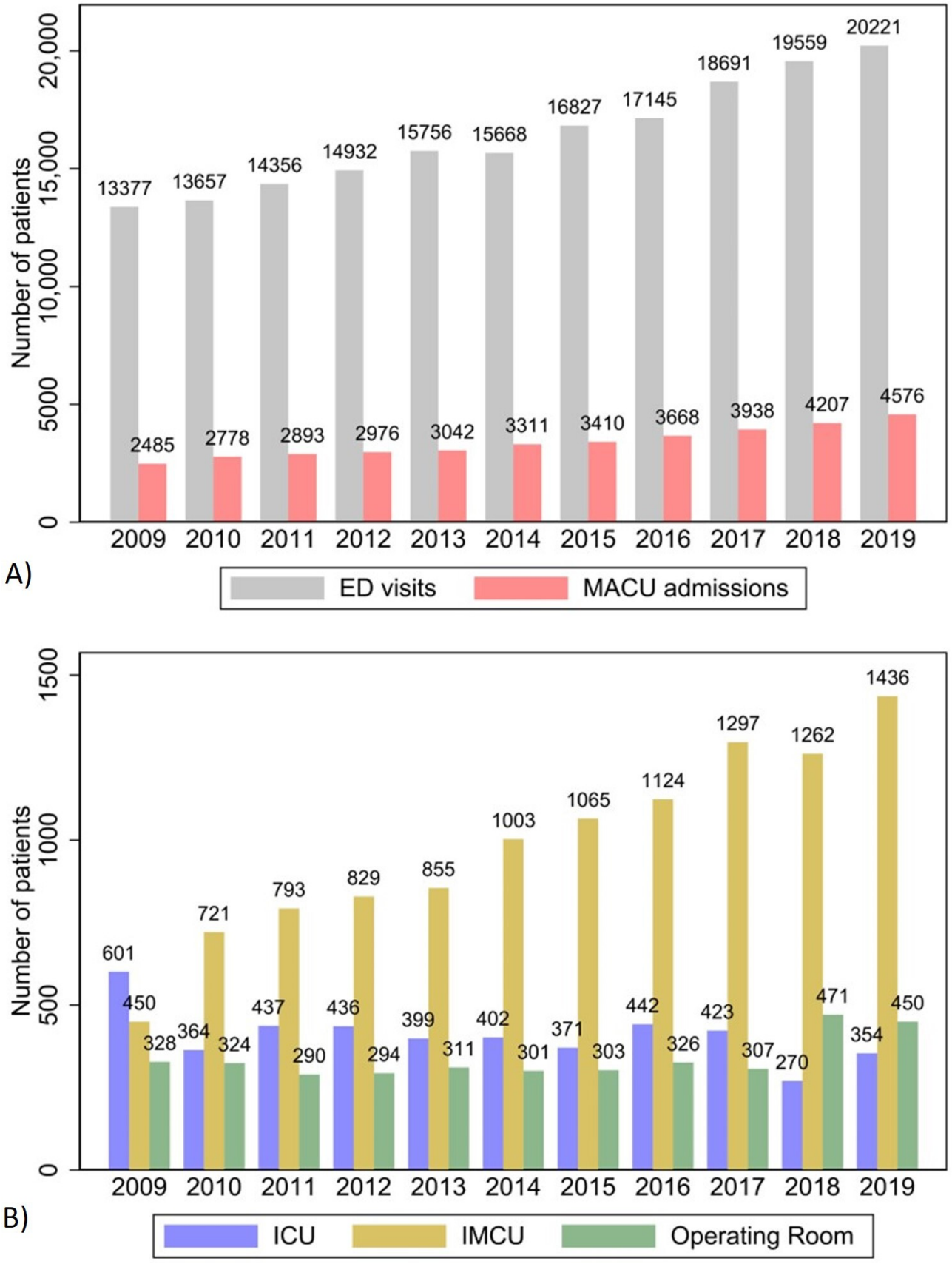
	Total (N=180,189), n (%)	65 - 74 y (n=61,511), n (%)	75 - 84 y (n=66,603), n (%)	85 - 94 y (n=46,890), n (%)	≥95 y (n=5185), n (%)
MACU[a]	19,039 (10.6)	6890 (11.2)	7668 (11.5)	4190 (8.9)	291 (5.6)
Intensive care unit	4499 (23.6)	1958 (28.4)	1917 (25.0)	611 (14.6)	13 (4.5)
Intermediate care unit	10,835 (56.9)	3508 (50.9)	4339 (56.6)	2788 (66.5)	200 (68.7)
Operating room	3705 (19.5)	1424 (20.7)	1412 (18.4)	791 (18.9)	78 (26.8)
Ward	63,025 (35.0)	17,917 (29.1)	23,999 (36.0)	18,953 (40.4)	2156 (41.6)
Discharged	97,238 (54.0)	36,534 (59.4)	34,608 (52.0)	23,411 (49.9)	2685 (51.8)
Died in the emergency department	887 (0.5)	170 (0.3)	328 (0.5)	336 (0.7)	53 (1.0)

[a]MACU: monitored acute care unit.

Over the 10-year study period, monitored acute care unit admissions increased by 62.4% in absolute terms (from 1379 in 2009 to 2240 in 2019) and by 0.8% in relative terms (from 10.3% in 2009 to 11.1% in 2019, *P*=.02; Figure 2A). This increase was mainly driven by a rise in intermediate care unit admissions, from 450 of 1379 (32.6%) in 2009 to 1436 of 2240 (64.1%) in 2019. In contrast, operating room admissions remained stable, from 328 of 1379 (23.8%) to 450 of 2240 (20.1%). Intensive care unit admissions decreased, from 601 of 1379 (43.6%) to 354 of 2240 (15.8%; Figure 2B). Notably, while the proportion of patients who died in the ED remained stable, the proportion of patients admitted to hospital wards increased over the 10-year period, from 4026 of 13,377 (30%) in 2009 to 7996 of 19,990 (40%) in 2019.

XSL•FO
RenderX

Figure 2. (A) Trends in emergency department (ED) visits and monitored acute care unit (MACU) admissions for patients aged ≥65 years and (B) distribution of MACU admissions to different types of MACUs over time. Over the 10-year study period (2009 - 2019), ED visits at the Geneva University Hospitals from patients aged ≥65 years increased by 56.1%, from 14,705 to 22,955. MACU admissions increased by 62.4%, from 1379 to 2240. The increase in MACU admissions primarily involved intermediate care units (IMCUs), from 450 of 1379 (32.6%) in 2009 to 1436 of 2240 (64.1%) in 2019. ICU: intensive care unit.

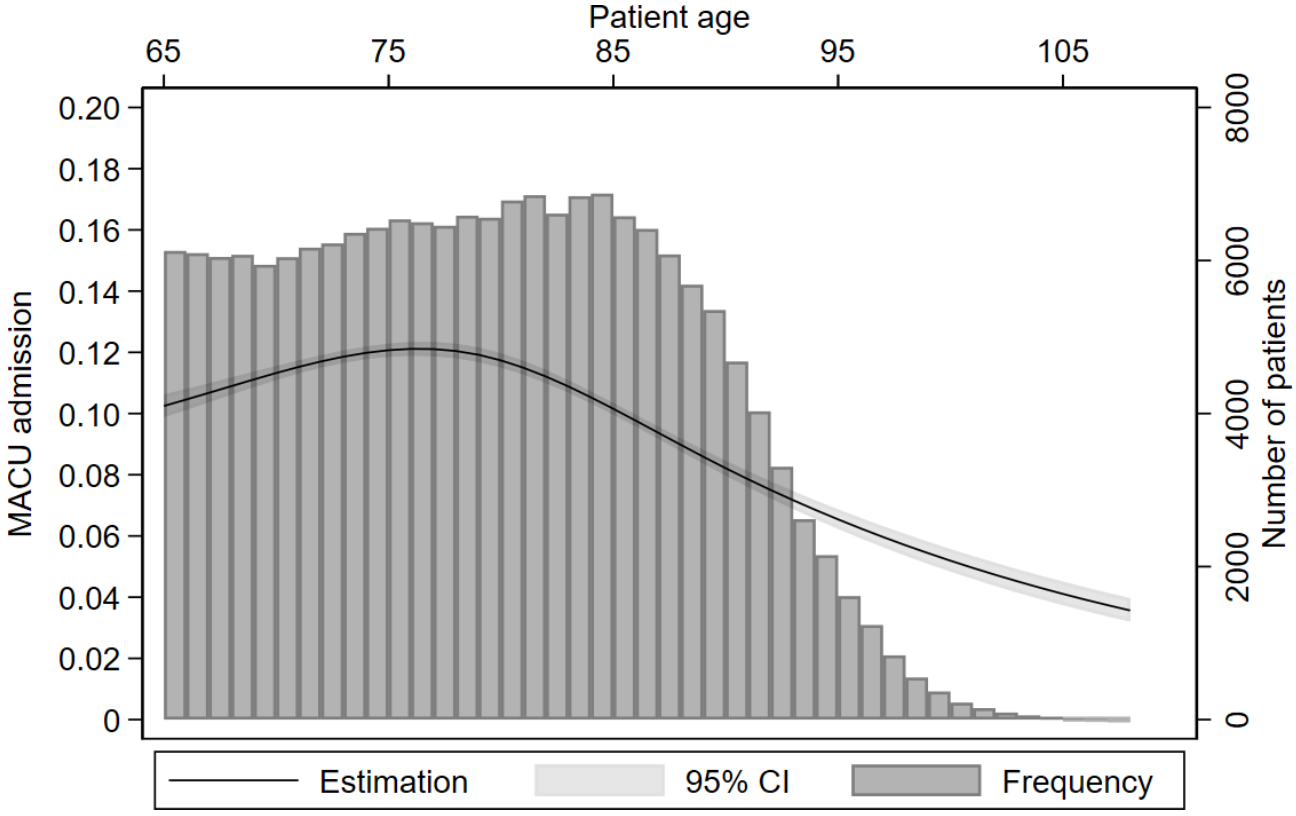


XSL•FO
RenderX

Predictors for Monitored Acute Care Unit Admission

Patients admitted to monitored acute care units were more frequently male, younger, and more often in a relationship (Table 1). Figure 3 shows the unadjusted association between age and monitored acute care unit admission among older patients. Patients admitted to a monitored acute care unit more frequently arrived by ambulance. They also had a higher triage level, with only 2817 (14.8%) of monitored acute care unit admissions having a lower triage level (3 or 4). In our multivariable analysis, younger age, male sex, marital status, absence of nursing home residence, arrival by ambulance, and higher triage level were all identified as independent predictors of monitored acute care unit admission (Table 3). A sensitivity analysis using age as a categorical variable did not significantly change these findings (Multimedia Appendix 2)

Figure 3. Association between age and monitored acute care unit (MACU) admissions among older patients. Older patients are less likely to be admitted to a MACU, with an inflection around 80 to 85 years old. In our multivariable analysis, younger age was identified as an independent predictor of MACU admission (75th vs 25th percentile, adjusted odds ratio 0.73, 95% CI 0.69-0.76).



XSL•FO
RenderX

Table . Predictors for monitored acute care unit admission.

	Multivariable model, adjusted OR[a] (95% CI)
Age (y)	
75th versus 25th percentile	0.73 (0.69-0.76)
Sex	
Female	Ref.[b]
Male	1.21 (1.14-1.28)
Marital status	
Married or in a relationship	Ref.
Single, divorced, or widowed	0.88 (0.83-0.93)
Nursing home resident	
No	Ref.
Yes	0.71 (0.63-0.80)
Primary language	
French	Ref.
Italian	0.83 (0.74-0.92)
Spanish	0.76 (0.66-0.89)
German	0.97 (0.83-1.14)
English	0.94 (0.78-1.15)
Portuguese	0.81 (0.65-1.02)
Others	0.78 (0.68-0.89)
Arrival by ambulance	
No	Ref.
Yes	1.87 (1.75-1.99)
Triage scale	
1: vital emergency	39.95 (25.10-63.58)
2: urgent	9.14 (5.76-14.51)
3: mild urgent	2.54 (1.60-4.03)
4: nonurgent	Ref.
Triage motive category	
Cardiology-pneumology	Ref.
Neurology-psychiatry	1.36 (1.28-1.45)
Traumatology	0.61 (0.55-0.67)
Digestive-OB/GYN[c]	1.28 (1.15-1.43)
Urology-nephrology	0.69 (0.56-0.84)
Rheumatology	0.39 (0.28-0.55)
Infectious disease	0.69 (0.58-0.83)
Dermatology	0.57 (0.43-0.75)
ENT[d]	0.35 (0.25-0.49)
Others	1.26 (1.12-1.42)

[a]OR: odds ratio.

[b]Ref.: reference.

[c]OB/GYN: obstetrics and gynecology.

[d]ENT: ear, nose, and throat.

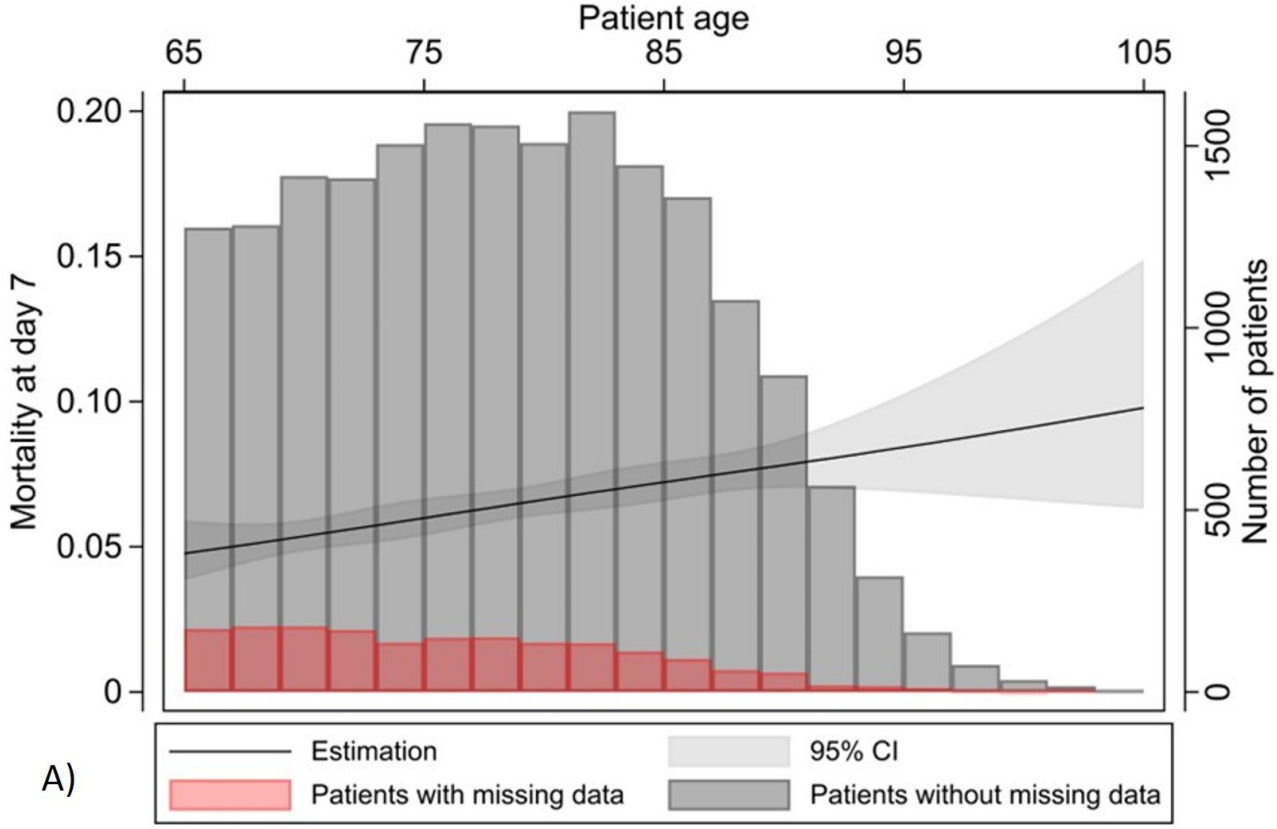
XSL•FO
RenderX

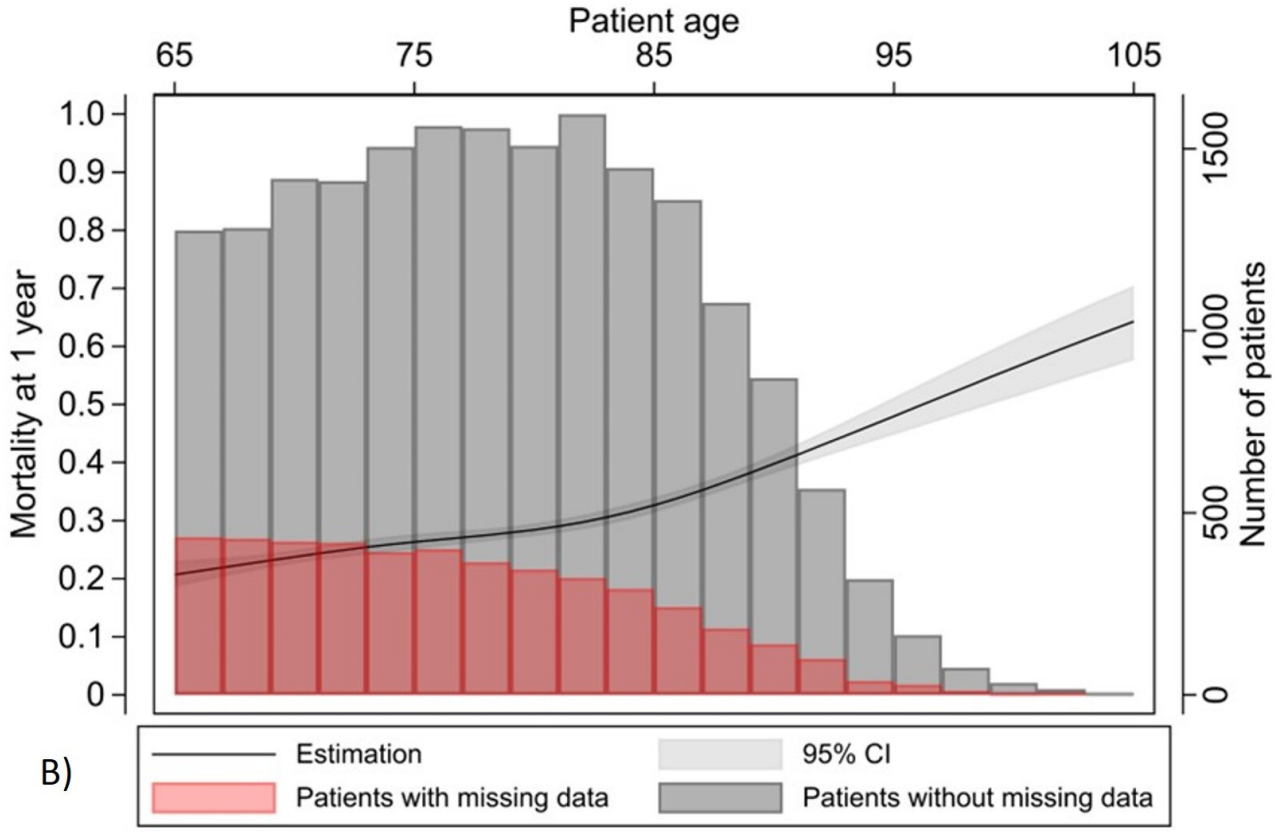
Outcomes

Seven days following monitored acute care unit admission, 1105 (5.8%) patients had died, 6184 (32.5%) patients had been discharged or transferred to long-term care, 10,139 (53.3%) patients were hospitalized in a ward, and 1611 (8.5%) were still in a monitored acute care unit (Table 4). Patients initially admitted to the intensive care unit had the highest mortality (584 patients, 13.0%) and the lowest discharge rate. Only 135 (1.2%) patients admitted to an intermediate care unit required transfer to a higher level of care. The average hospital length of stay ranged from 11.2 to 13.8 days, depending on the admitting unit. One-year mortality following monitored acute care unit admission was 22.3% (4251 patients), with 23.9% (4553 patients) of data missing for the cohort. Compared to patients admitted to an intermediate care unit or the operating room, patients initially admitted to the intensive care unit had worse outcomes (Table 4). After adjustment for other variables, there was a significant association between age and mortality at day 7 (adjusted OR 1.55, 95% CI 1.14-2.10) and at 1 year (adjusted OR 1.28, 95% CI 1.08-1.51), as illustrated in Figure 4.

XSL•FO
RenderX

Figure 4. Association between age and (A) mortality at day 7 and (B) mortality at 1 year. For patients aged ≥65 years or older with monitored acute care unit admission, there was a significant association between age and mortality at day 7 (adjusted odds ratio 1.55, 95% CI 1.14-2.10) and at 1 year (adjusted odds ratio 1.28, 95% CI 1.08-1.51). Seven-day mortality was 5.8% (1105 patients), whereas 1-year mortality was 22.3% (4251 patients), with 23.9% (4553 patients) of data missing for the cohort (19,039 patients).





XSL•FO
RenderX

Table . Outcomes of patients with monitored acute care unit admission.

	All (N=19,039)	Intensive care unit (n=4499)	Intermediate care unit (n=10,835)	Operating room (n=3705)
Status at day 7, n (%)				
Deceased	1105 (5.8)	584 (13.0)	334 (3.1)	187 (5.0)
Intensive care unit	931 (4.9)	560 (12.4)	135 (1.2)	236 (6.4)
Intermediate care unit	680 (3.6)	226 (5.0)	383 (3.5)	71 (1.9)
Ward	10,139 (53.3)	2203 (49)	6173 (57)	1763 (47.6)
Discharged or transferred	6184 (32.5)	926 (20.6)	3810 (35.2)	1448 (39.1)
Hospital length of stay (d), median (IQR)	9.0 (4.9-15.4)	10.5 (5.2-18.0)	8.9 (5.0-14.2)	8.0 (3.9-15.7)
Status at 1 year, n (%)				
Alive	10,230 (53.7)	2164 (48.1)	6096 (56.3)	1970 (53.2)
Deceased	4251 (22.3)	1355 (30.1)	2133 (19.7)	763 (20.6)
Unknown	4553 (23.9)	980 (21.8)	2606 (24.1)	972 (26.2)

Discussion

Principal Findings

This retrospective cohort study, spanning from 2009 to 2019, revealed a significant increase in monitored acute care unit admissions among older patients. Key predictors of monitored acute care unit admission, such as age, sex, and arrival by ambulance, were identified. While short-term outcomes were relatively good with a 7-day mortality rate of approximately 5%, long-term outcomes were poorer, with more than 20% mortality at 1 year.

Over the study period, monitored acute care unit admissions increased by 62%. This increase was not uniform and can be primarily attributed to the growing number of older patients presenting to the ED with a more modest rise in the proportion of these patients being admitted to a monitored acute care unit, from 10.3% (1379 patients) in 2009 to 11.1% (2240 patients) in 2019. Notably, this overall increase in admissions was primarily driven by a substantial rise in intermediate care unit admissions. Although changes in unit capacity over time could not be assessed, the increase in intermediate care unit admissions and the decrease in intensive care unit admissions are likely to be related. This trend may reflect a redistribution of patients and the development of new competencies within the intermediate care units. Such units can also be seen as a more tailored approach to managing older patients, providing an intermediate level of care that is perhaps more suitable for the needs of this population. Additionally, these units benefit from economic advantages, supporting their use as a cost-effective strategy to prioritize health care resource allocation.

Intensive care unit admission rates among older patients have been extensively studied, with the available literature reporting considerable variability. Some studies have found annual increases of up to 5.6% in patients over 80 years old [15], whereas others have found no significant change [16]. In contrast, data on intermediate care unit admissions are more limited, likely due to the wide variety of formats of such units (including intermediate care units, step-down units, and high dependency units), making it challenging to standardize the data [17-19]. The observed rise in intermediate care unit admissions in our study is closely linked to the overall increase in ED visits by older patients, which has directly led to higher demand for monitored care. Beyond this, several other factors may have contributed, including the growing prevalence of chronic diseases among older patients that require a level of care beyond general ward care but less than intensive care unit, along with advances in medical technologies, interventions, protocols, and admission criteria that allow for safer and more effective management of these patients in intermediate care units [20-24].

As anticipated, older patients were less likely to be admitted to a monitored acute care unit, with a notable decrease in admissions and an increase in mortality observed around the age of 85 years. This age threshold may be due to higher comorbidity rates and the presence of advanced care directives in this patient population [25,26]. Age is known to be associated with both a decrease in intensive care unit referral by emergency physicians and an increase in admission denial by intensive care physicians [27]. It might make sense, as intensive care unit admission for patients aged older than 80 years has not been shown to affect 2-year survival rates [28]. However, critical care societies recommend that decisions regarding monitored acute care unit admissions should be based on illness severity, comorbidities, and baseline functional status rather than age alone [29]. It is crucial for physicians to engage in discussions with patients and their next of kin about the potential implications of a monitored acute care unit admission, although the final decision rests with the receiving physician [30].

Our study also identified other predictors of a monitored acute care unit admission. Arrival by ambulance, which is up to 4.6 times more frequent for older patients [31] and serves as an indicator of higher clinical severity, was a strong predictor of monitored acute care unit admission. Nursing home residents were less likely to be admitted to a monitored acute care unit, likely due to higher levels of comorbidities and frailty levels,

XSL•FO
RenderX

often associated with a diminished quality of life and cognitive impairment [32-34]. Neurology-psychiatry–related triage motives were the strongest predictors of monitored acute care unit admission, likely due to conditions such as strokes, which typically require a monitored acute care unit–level care. Interestingly, the French language, male sex, and marital status were also associated with monitored acute care unit admission. Language barriers are well documented as contributing to worse health outcomes, including longer hospital length of stay, higher readmission rates, and limited access to health care systems, regardless of socioeconomic status [35,36]. While female patients were predominant in the ED, they were 21% less likely to be admitted to a monitored acute care unit compared to male patients, a finding consistent with existing literature [37-39]. This disparity may be influenced by more comorbidities in men, and the tendency for female patients to set medical limitations and have advanced directives, especially when divorced or widowed [40]. However, sociocultural factors and implicit or explicit biases may also contribute to this disparity. Unmarried patients often presented with more severe illness at admission [41], possibly due to delayed medical intervention, whereas patients with a partner are more likely to receive high-level care to meet family expectations. Physicians must be aware of such disparities and strive to mitigate them in their practice.

Short-term outcomes for patients admitted to a monitored acute care unit are quite encouraging, with one-third of patients discharged and more than half transferred to the ward by day 7. However, long-term outcomes were less favorable, particularly for patients admitted to the intensive care unit, who had a 1-year mortality rate exceeding 30%. While intensive care unit mortality in older patients is well documented [42], outcomes for intermediate care unit patients have been less studied, especially among ED patients. Torres et al [43] reported no significant difference between demographics regarding in-hospital mortality after intermediate care unit admission and a 34% mortality after 2 years for patients aged older than 65 years, significantly higher than the 10% for younger patients. In a secondary analysis of the ICE-CUB2 trial focusing on ED patients, Thietart et al [44] reported 6-month mortality rates of 44% and 31% for intensive care unit and intermediate care unit patients, respectively, which were higher than those reported in other studies, including this one, which the severity of the patients can explain, as they did not include uncomplicated acute coronary syndromes and strokes. Finally, D'Andrea et al [45] found a 43% 1-year mortality rate in patients aged older than 75 years admitted to a geriatric intermediate care unit, with a significant association between age and 1-year mortality.

Clinical and Scientific Implications

This study highlights the need to adapt health care structures to meet the growing demands of an aging population. Developing specialized geriatric acute care units or intensive care units, along with training physicians specifically to address the needs of older patients, could help raise awareness of these needs while addressing potential biases related to gender, language, and marital status. Further research should explore patient and family satisfaction and the alignment between care provided and patient wishes. Previous studies have shown that quality of life in older intensive care unit survivors is initially worse [46] but tends to improve within a year, with many older patients willing to undergo intensive care unit admission again if necessary [47,48]. Further studies could focus on developing clinical guidelines or tools specifically tailored to older patients to aid in the shared decision-making process regarding monitored acute care unit admission. These guidelines should incorporate measures such as the Functional Independence Measure or the Clinical Frailty Scale, which have been shown to correlate with resource use and outcomes [45,49,50].

Strengths and Limitations

Some strengths and limitations need to be acknowledged. The study analyzes a large cohort over a decade, providing a robust dataset for analysis. The innovative and comprehensive approach of considering all monitored acute care unit admissions offers a broad view of high-cost care, which has not been extensively studied. Additionally, we used a rigorous statistical plan, including restricted cubic spline models to account for the nonlinear relationship between age and outcomes, to enhance the robustness of our results. The primary limitations of this study are related to the design. As a retrospective cohort, the study is prone to bias and missing data, notably the variation in monitored acute care unit beds per unit over the years, which may influence patient disposition based on unit saturation. The second major limitation is the lack of information on comorbidities, place of residence (rural vs urban), and polypharmacy. These variables are either not available as structured data in our electronic health record or are at high risk of bias. To preserve high data quality, we chose not to report them. The authors acknowledge this as a limitation of the study, with a risk of residual confounding. Results involving mortality should be considered cautiously, as mortality data were missing for 1 in 4 patients. Our study is monocentric; however, its findings are likely generalizable to other university hospitals in Switzerland and Europe, given the similar monitored acute care unit admission criteria. The exclusion of private hospitals with minimal monitored acute care unit capacity is a minor limitation, as their patients tend to be younger than 65 years. The change in the slope of the association between age and monitored acute care unit admission, as well as age and 1-year mortality, may be driven by only a small number of patients in the oldest age group, potentially distorting the true association. Finally, this study did not cover the periods of the COVID-19 pandemic, which could limit the generalizability of the results, as monitored acute care units were particularly strained during this period.

Conclusions

This study highlights the sharp increase in monitored acute care unit admissions among older patients, reflecting the growing demand for high levels of care in this population. The rise was primarily driven by a substantial increase in intermediate care unit admissions, whereas intensive care unit admissions remained stable. These findings emphasize the need for hospitals to adapt health care infrastructure, clinical procedures, and resource allocation to meet the evolving needs of an aging population. Future work should focus on the importance of developing specialized geriatric acute care units, refining clinical guidelines, and engaging in shared decision-making to ensure equitable and effective care for all older patients.

XSL•FO
RenderX

Acknowledgments

Language and syntax were reviewed and corrected using artificial intelligence–based tools to improve clarity and consistency of the manuscript. The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing, adapting and adjusting emotional tone, and reformatting. The GAI tool used was ChatGPT. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

There was no specific funding for this study.

Data Availability

The datasets generated or analyzed during this study are available with this manuscript (Multimedia Appendix 3).

Authors' Contributions

Conceptualization: LvD, OG, CAF. Data curation: CAF. Formal analysis: CAF. Funding acquisition: TD. Investigation: CAF. Methodology: LvD, SvD, CAF. Project administration: CAF. Supervision: CAF. Visualization: LvD, CAF. Writing – original draft : LvD, SvD, CAF. Writing – review & editing : LvD, SvD, FR, AR, CM, SC, XR, TD, OG, CAF.

Conflicts of Interest

SvD received speaking honoraria from Löwenstein Medical. All other authors declare no competing interests.

Multimedia Appendix 1
Proportion of ED visits by older patients. Between 2009 and 2019, a total of 701,838 ED visits were recorded, with a global increase of 34.1% (56,944-76,368) in all ED visits. This increase was more pronounced for patients aged ≥65 years (14,705-22,955, +56.1%) than for patients aged <65 years (42,239 to 53,413, +26.5%; *P*<.001). ED: emergency department.
[PNG File, 40 KB - i-jmr_v15i1e80629_app1.png]

Multimedia Appendix 2
Potential predictors for MACU admission with age in categories. MACU: monitored acute care unit.
[DOCX File, 108 KB - i-jmr_v15i1e80629_app2.docx]

Multimedia Appendix 3
Dataset for publication. To ensure patient confidentiality, the dataset prepared for publication excludes the patient's first language, exact admission date, and identifying information, such as patient ID.
[ZIP File, 2865 KB - i-jmr_v15i1e80629_app3.zip]

References

1. Ukkonen M, Jämsen E, Zeitlin R, Pauniaho SL. Emergency department visits in older patients: a population-based survey. BMC Emerg Med 2019 Feb 27;19(1):20. [doi: 10.1186/s12873-019-0236-3] [Medline: 30813898]
2. Wittenberg R, Sharpin L, McCormick B, Hurst J. The ageing society and emergency hospital admissions. Health Policy 2017 Aug;121(8):923-928. [doi: 10.1016/j.healthpol.2017.05.007] [Medline: 28619464]
3. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul 7;380(9836):37-43. [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
4. Albert M, McCaig LF, Ashman JJ. Emergency department visits by persons aged 65 and over: United States, 2009-2010. : Centers for Disease Control and Prevention, National Center for Health Statistics; 2013 Oct URL: https://www.cdc.gov/nchs/data/databriefs/db130.pdf [accessed 2026-01-17]
5. Liston P, Conyngham G, Brady M, Byrne PJ, Gilligan P. Growing old in the emergency department. Ir Med J 2017 Aug 12;110(7):621. [Medline: 29169003]
6. Latham LP, Ackroyd-Stolarz S. Emergency department utilization by older adults: a descriptive study. Can Geriatr J 2014 Dec;17(4):118-125. [doi: 10.5770/cgj.17.108] [Medline: 25452824]
7. Wunsch H, Guerra C, Barnato AE, Angus DC, Li G, Linde-Zwirble WT. Three-year outcomes for Medicare beneficiaries who survive intensive care. JAMA 2010 Mar 3;303(9):849-856. [doi: 10.1001/jama.2010.216] [Medline: 20197531]

XSL•FO
RenderX

8. Fuchs L, Chronaki CE, Park S, et al. ICU admission characteristics and mortality rates among elderly and very elderly patients. Intensive Care Med 2012 Oct;38(10):1654-1661. [doi: 10.1007/s00134-012-2629-6] [Medline: 22797350]
9. Somme D, Maillet JM, Gisselbrecht M, Novara A, Ract C, Fagon JY. Critically ill old and the oldest-old patients in intensive care: short- and long-term outcomes. Intensive Care Med 2003 Dec;29(12):2137-2143. [doi: 10.1007/s00134-003-1929-2] [Medline: 14614546]
10. Dowdy DW, Eid MP, Sedrakyan A, et al. Quality of life in adult survivors of critical illness: a systematic review of the literature. Intensive Care Med 2005 May;31(5):611-620. [doi: 10.1007/s00134-005-2592-6] [Medline: 15803303]
11. Rutschmann OT, Hugli OW, Marti C, et al. Reliability of the revised Swiss emergency triage scale: a computer simulation study. Eur J Emerg Med 2018 Aug;25(4):264-269. [doi: 10.1097/MEJ.0000000000000449] [Medline: 28099182]
12. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996 Dec;49(12):1373-1379. [doi: 10.1016/s0895-4356(96)00236-3] [Medline: 8970487]
13. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007 Mar 15;165(6):710-718. [doi: 10.1093/aje/kwk052] [Medline: 17182981]
14. Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis: Springer International Publishing; 2015. [doi: 10.1007/978-3-319-19425-7]
15. Bagshaw SM, Webb SAR, Delaney A, et al. Very old patients admitted to intensive care in Australia and New Zealand: a multi-centre cohort analysis. Crit Care 2009;13(2):R45. [doi: 10.1186/cc7768] [Medline: 19335921]
16. Rai S, Brace C, Ross P, et al. Characteristics and outcomes of very elderly patients admitted to intensive care: a retrospective multicenter cohort analysis. Crit Care Med 2023 Oct 1;51(10):1328-1338. [doi: 10.1097/CCM.0000000000005943] [Medline: 37219961]
17. Plate JDJ, Leenen LPH, Houwert M, Hietbrink F. Utilisation of intermediate care units: a systematic review. Crit Care Res Pract 2017;2017:8038460. [doi: 10.1155/2017/8038460] [Medline: 28775898]
18. López-Jardón P, Martínez-Fernández MC, García-Fernández R, Martín-Vázquez C, Verdeal-Dacal R. Utility of intermediate care units: a systematic review study. Healthcare (Basel) 2024 Jan 24;12(3):296. [doi: 10.3390/healthcare12030296] [Medline: 38338181]
19. Prin M, Wunsch H. The role of stepdown beds in hospital care. Am J Respir Crit Care Med 2014 Dec 1;190(11):1210-1216. [doi: 10.1164/rccm.201406-1117PP] [Medline: 25163008]
20. Chou YS, Lin HY, Weng YM, et al. Step-down units are cost-effective alternatives to coronary care units with non-inferior outcomes in the management of ST-elevation myocardial infarction patients after successful primary percutaneous coronary intervention. Intern Emerg Med 2020 Jan;15(1):59-66. [doi: 10.1007/s11739-019-02037-z] [Medline: 30706252]
21. Meaudre E, Nguyen C, Contargyris C, et al. Management of septic shock in intermediate care unit. Anaesth Crit Care Pain Med 2018 Apr;37(2):121-127. [doi: 10.1016/j.accpm.2017.07.004] [Medline: 28790011]
22. Plate JDJ, Peelen LM, Leenen LPH, Hietbrink F. The intermediate care unit as a cost-reducing critical care facility in tertiary referral hospitals: a single-centre observational study. BMJ Open 2019 Jun 4;9(6):e026359. [doi: 10.1136/bmjopen-2018-026359] [Medline: 31167865]
23. Ranhoff AH, Rozzini R, Sabatini T, et al. Subintensive care unit for the elderly: a new model of care for critically ill frail elderly medical patients. Intern Emerg Med 2006;1(3):197-203. [doi: 10.1007/BF02934737] [Medline: 17120465]
24. Ohbe H, Matsui H, Yasunaga H. Intensive care unit versus high-dependency care unit for patients with acute heart failure: a nationwide propensity score-matched cohort study. J Intensive Care 2021 Dec 20;9(1):78. [doi: 10.1186/s40560-021-00592-2] [Medline: 34930470]
25. Gamertsfelder EM, Seaman JB, Tate J, Buddadhumaruk P, Happ MB. Prevalence of advance directives among older adults admitted to intensive care units and requiring mechanical ventilation. J Gerontol Nurs 2016 Apr;42(4):34-41. [doi: 10.3928/00989134-20151124-02] [Medline: 26651862]
26. Foley C, Bloomer M, Hutchinson AM. Factors that influence intensive care admission decisions for older people: a systematic review. Aust Crit Care 2023 Mar;36(2):274-284. [doi: 10.1016/j.aucc.2021.12.006] [Medline: 35144889]
27. Guidet B, Boumendil A, Garrousteorgeas M, Pateron D. Admission en réanimation du sujet âgé à partir du service des urgences. État des lieux [Article in French]. Réanimation 2008 Dec;17(8):790-801. [doi: 10.1016/j.reaurg.2008.09.004]
28. Abuhasira R, Anstey M, Novack V, Bose S, Talmor D, Fuchs L. Intensive care unit capacity and mortality in older adults: a three nations retrospective observational cohort study. Ann Intensive Care 2022 Mar 4;12(1):20. [doi: 10.1186/s13613-022-00994-x] [Medline: 35244803]
29. Nates JL, Nunnally M, Kleinpell R, et al. ICU admission, discharge, and triage guidelines: a framework to enhance clinical operations, development of institutional policies, and further research. Crit Care Med 2016 Aug;44(8):1553-1602. [doi: 10.1097/CCM.0000000000001856] [Medline: 27428118]
30. Blanch L, Abillama FF, Amin P, et al. Triage decisions for ICU admission: report from the Task Force of the World Federation of Societies of Intensive and Critical Care Medicine. J Crit Care 2016 Dec;36:301-305. [doi: 10.1016/j.jcrc.2016.06.014] [Medline: 27387663]
31. Baum SA, Rubenstein LZ. Old people in the emergency room: age-related differences in emergency department use and care. J Am Geriatr Soc 1987 May;35(5):398-404. [doi: 10.1111/j.1532-5415.1987.tb04660.x] [Medline: 3571788]

XSL•FO
RenderX

32. Manis DR, Katz P, Lane NE, et al. Clinical comorbidities and transitions between care settings among residents of assisted living facilities: a repeated cross-sectional study. J Am Med Dir Assoc 2023 Sep;24(9):1356-1360. [doi: 10.1016/j.jamda.2023.06.017] [Medline: 37507099]
33. Moore KL, Boscardin WJ, Steinman MA, Schwartz JB. Patterns of chronic co-morbid medical conditions in older residents of U.S. nursing homes: differences between the sexes and across the agespan. J Nutr Health Aging 2014 Apr;18(4):429-436. [doi: 10.1007/s12603-014-0001-y] [Medline: 24676326]
34. Toot S, Swinson T, Devine M, Challis D, Orrell M. Causes of nursing home placement for older people with dementia: a systematic review and meta-analysis. Int Psychogeriatr 2017 Feb;29(2):195-208. [doi: 10.1017/S1041610216001654] [Medline: 27806743]
35. Karliner LS, Kim SE, Meltzer DO, Auerbach AD. Influence of language barriers on outcomes of hospital care for general medicine inpatients. J Hosp Med 2010;5(5):276-282. [doi: 10.1002/jhm.658] [Medline: 20533573]
36. Lindholm M, Hargraves JL, Ferguson WJ, Reed G. Professional language interpretation and inpatient length of stay and readmission rates. J Gen Intern Med 2012 Oct;27(10):1294-1299. [doi: 10.1007/s11606-012-2041-5] [Medline: 22528618]
37. Lat TI, McGraw MK, White HD. Gender differences in critical illness and critical care research. Clin Chest Med 2021 Sep;42(3):543-555. [doi: 10.1016/j.ccm.2021.04.012] [Medline: 34353458]
38. Merdji H, Long MT, Ostermann M, et al. Sex and gender differences in intensive care medicine. Intensive Care Med 2023 Oct;49(10):1155-1167. [doi: 10.1007/s00134-023-07194-6] [Medline: 37676504]
39. Modra L, Pilcher D, Bailey M, Bellomo R. Sex differences in intensive care unit admissions in Australia and New Zealand. Crit Care Resusc 2021 Mar;23(1):86-93. [doi: 10.51893/2021.1.OA8] [Medline: 38075190]
40. McPherson K, Carlos WG, Emmett TW, Slaven JE, Torke AM. Limitation of life-sustaining care in the critically ill: a systematic review of the literature. J Hosp Med 2019 May;14(5):303-310. [doi: 10.12788/jhm.3137] [Medline: 30794145]
41. Gordon HS, Rosenthal GE. Impact of marital status on outcomes in hospitalized patients. Evidence from an academic medical center. Arch Intern Med 1995;155(22):2465-2471. [Medline: 7503606]
42. Vallet H, Schwarz GL, Flaatten H, de Lange DW, Guidet B, Dechartres A. Mortality of older patients admitted to an ICU: a systematic review. Crit Care Med 2021 Feb 1;49(2):324-334. [doi: 10.1097/CCM.0000000000004772] [Medline: 33332816]
43. Torres OH, Francia E, Longobardi V, Gich I, Benito S, Ruiz D. Short- and long-term outcomes of older patients in intermediate care units. Intensive Care Med 2006 Jul;32(7):1052-1059. [doi: 10.1007/s00134-006-0170-1] [Medline: 16791668]
44. Thietart S, Boumendil A, Pateron D, Guidet B, Vallet H, ICE-CUB2 Study Network. Impact on 6-month outcomes of hospital trajectory in critically ill older patients: analysis of the ICE-CUB2 clinical trial. Ann Intensive Care 2022 Jul 11;12(1):65. [doi: 10.1186/s13613-022-01042-4] [Medline: 35819563]
45. D'Andrea A, Le Peillet D, Fassier T, et al. Functional Independence Measure score is associated with mortality in critically ill elderly patients admitted to an intermediate care unit. BMC Geriatr 2020 Sep 9;20(1):334. [doi: 10.1186/s12877-020-01729-y] [Medline: 32907534]
46. Ariyo K, Canestrini S, David AS, Ruck Keene A, Wolfrum S, Owen G. Quality of life in elderly ICU survivors before the COVID-19 pandemic: a systematic review and meta-analysis of cohort studies. BMJ Open 2021 Oct 11;11(10):e045086. [doi: 10.1136/bmjopen-2020-045086] [Medline: 34635510]
47. Tabah A, Philippart F, Timsit JF, et al. Quality of life in patients aged 80 or over after ICU discharge. Crit Care 2010;14(1):R2. [doi: 10.1186/cc8231] [Medline: 20064197]
48. Israelsson-Skogsberg Å, Eriksson T, Lindberg E. A scoping review of older patients' health-related quality of life, recovery and well-being after intensive care. Nurs Open 2023 Sep;10(9):5900-5919. [doi: 10.1002/nop2.1873] [Medline: 37306357]
49. Fernando SM, McIsaac DI, Perry JJ, et al. Frailty and associated outcomes and resource utilization among older ICU patients with suspected infection. Crit Care Med 2019 Aug;47(8):e669-e676. [doi: 10.1097/CCM.0000000000003831] [Medline: 31135504]
50. Fehlmann CA, Nickel CH, Cino E, Al-Najjar Z, Langlois N, Eagles D. Frailty assessment in emergency medicine using the Clinical Frailty Scale: a scoping review. Intern Emerg Med 2022 Nov;17(8):2407-2418. [doi: 10.1007/s11739-022-03042-5] [Medline: 35864373]

Abbreviations

ED: emergency department
HUG: Hôpitaux Universitaires de Genève/Geneva University Hospitals
OR: odds ratio
SMD: standardized mean difference

XSL•FO
RenderX

Edited by M Balcarras, N Cahill; submitted 22.Jul.2025; peer-reviewed by C Thompson; revised version received 27.Dec.2025; accepted 29.Dec.2025; published 02.Feb.2026.

Please cite as:
van Delden L, von Düring S, Rouyer F, Rossel A, Marti C, Cereghetti S, Roux X, Desmettre T, Grosgurin O, Fehlmann CA
Trends, Predictors, and Outcomes of Monitored Acute Care Unit Admissions in Older Adults: 10-Year Retrospective Analysis
Interact J Med Res 2026;15:e80629
URL: https://www.i-jmr.org/2026/1/e80629
doi:10.2196/80629

© Ludovic van Delden, Stephan von Düring, Frederic Rouyer, Anne Rossel, Christophe Marti, Sara Cereghetti, Xavier Roux, Thibaut Desmettre, Olivier Grosgurin, Christophe A Fehlmann. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 2.Feb.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Sleep Disturbance and Its Association With Purchasing Behavior of COVID-19 Medicine Among the Public After the Adjustment of Zero-COVID Policy in China: Results From a Web-Based Survey Study

Yu Huang[1], MSc; Xiang Zhao[1], MSc; Lei Wang[1], MSc; Qiaohong Lv[1], MSc; Qingqing Wu[1], MSc; Shuiyang Xu[1], MPH; Xuehai Zhang[1], MSc; Shuxian Wu[2], MSc

[1]Department of Health Education, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

[2]Central Office, Zhejiang Provincial Center for Disease Control and Prevention, 630 Xincheng Road, Binjiang District, Hangzhou, Zhejiang, China

Corresponding Author:
Shuxian Wu, MSc
Central Office, Zhejiang Provincial Center for Disease Control and Prevention, 630 Xincheng Road, Binjiang District, Hangzhou, Zhejiang, China

Abstract

Background: In December 2022, in light of the weakened pathogenicity of the new variants and other scientific considerations, China optimized its zero-COVID policy. As the situation evolved, the virus spread more widely across the country.

Objective: This study aims to explore the public's sleep status and its association with purchasing behavior of COVID-19 medicine after the adjustment of zero-COVID policy in China.

Methods: A cross-sectional, internet-based survey among residents aged 18 - 69 years was conducted in Zhejiang province, China, from December 16 to 30, 2022, to collect data on sociodemographic characteristics, COVID-19 drug purchasing behavior, sleep disturbance levels, etc. Chi-square tests, univariate analyses, and multivariate analyses were used to explore the associations among these factors.

Results: Out of 38,480 participants, 20,803 (54.1%) reported sleep disruption after China's COVID-19 response policy adjustment. The degree of impact varied, with 10,964 (52.70%) reporting "slight," 3105 (14.93%) "moderate," 3493 (16.79%) "significant," and 3241 (15.58%) "very significant" impacts. Only 20.90% (782/3742) of those who deemed purchasing unnecessary had sleep disruptions, compared to 45.19% (6214/13,752) of those who acquired medications and 65.79% (13,807/20,986) of those who tried but failed to obtain them. Sleep disturbance levels were significantly associated with sociodemographic factors like age, education levels, occupation, marital status, and presence of family members diagnosed with COVID-19 (P<.05). By age, sleep disturbance proportions differed notably: 36.32% (409/1126) for those under 20 years, 54.81% (19,714/35,970) for the 20 to 60 age group, and 49.13% (680/1384) for individuals over 60 years. For education level, the proportions were 57.44% (517/900, primary school), 54.34% (3928 /7229, junior high school), 54.27% (3808/7017, senior high school), 53.99% (11,974/22,180, junior college/undergraduate), and 49.91% (576/1154, master's degree), showing a clear downward trend as education level increased. By occupation, farmers had the highest rate (855/1447, 59.09%), followed by business/service industry workers and stay-at-home/unemployed individuals (13,925/24,750, 56.26%) and government staff (4161/7712, 53.95%), while 1242 out of 3049 (40.73%) health workers and 620 out of 1522 (40.74%) students had lower rates. Married participants had a 55.21% (17,143/31,053) sleep disturbance rate, and those with COVID-positive family members had the highest rate (2023/2873, 70.41%). Multivariate logistic regression, adjusting for these sociodemographic factors, showed that compared to those who thought purchasing COVID-19 medications was unnecessary, those who acquired medications were 3.11 times (adjusted odds ratio 3.11, 95% CI 2.85 - 3.39) more likely, and those who tried but couldn't get medications were 7.11 times (adjusted odds ratio 7.11, 95% CI 6.53 - 7.74) more likely to experience sleep disturbance.

Conclusions: The adjustment of China's zero-COVID policy affected the sleep health of the public, which was closely linked to drug-purchasing status, especially among the older people, those with lower education levels, and those with family members diagnosed with COVID-19. It highlights the need to develop and deploy interventions aimed at promoting better sleep health in times of crisis.

(Interact J Med Res 2026;15:e79903) doi:10.2196/79903

KEYWORDS

adjustment of zero-COVID policy; China; COVID-19; drug purchasing; drug shortage; sleep disturbance

XSL•FO
RenderX

Introduction

In December 2019, an unprecedented pneumonia outbreak of unknown origin emerged in Wuhan, Hubei province, China [1,2]. Subsequently identified as caused by the novel coronavirus (COVID-19), this highly contagious pathogen rapidly spread beyond China's borders within just 30 days [3-6]. Characterized by potentially severe symptoms such as acute respiratory failure and sepsis [7], COVID-19 posed a significant threat to global public health, prompting the World Health Organization (WHO) to declare it a public health emergency of international concern on January 30, 2020 [8].

To contain the virus, China implemented the dynamic zero-COVID policy, which included strict measures like national lockdowns, school closures, travel restrictions, and the suspension of numerous economic activities [9-13]. However, as the virulence of new virus variants decreased and COVID-19 vaccines became more prevalent, China adjusted its policy on December 7, 2022. The mandatory large-scale nucleic acid testing requirements, the centralized isolation for asymptomatic cases and mild symptomatic patients, and the mobility restrictions were removed [14,15]. This suddenly loosened policy led to a change in the epidemic situation, with a wider spread of the virus across the country.

The COVID-19 pandemic has had a profound impact on people's lives, particularly on their mental health and sleep quality [16-20]. A large-scale global study (combining 250 studies with nearly 500,000 participants from 49 countries) found that about 40% of people had sleep disturbances during the COVID-19 pandemic (95% CI 37.56%-43.48%). Among specific populations, the prevalence of sleep problems varied, with 52.39% (95% CI 41.69%-62.88%) among COVID-19 patients, 45.96% (95% CI 36.90%-55.30%) among children and adolescents, 42.47% (95% CI 37.95%-47.12%) among health care workers, 41.50% (95% CI 32.98%-50.56%) among special populations with health care needs, 41.16% (28.76%-54.79%) among university students, and 36.73% (32.32%-41.38%) among the general population [21]. Research in China shows that the prevalence of insomnia during the early and late stages of the pandemic was 37.0% (95% CI 34.1% - 39.9%) and 41.8% (95% CI 33.6% - 50.0%) respectively. Notably, health care workers, COVID-19 patients, those with chronic diseases, and people with mental disorders are more likely to experience insomnia than the general population [22].

Sleep is a fundamental physiological process that is essential for maintaining physical and mental health [19,23]. It plays a crucial role in various bodily functions, including homeostasis, immune system regulation, cognitive function, neural plasticity, memory consolidation, energy maintenance, macromolecular biosynthesis, and metabolism regulation [24,25]. Conversely, sleep deprivation can lead to a range of adverse health outcomes, such as impaired physical and mental performance, increased risk of depression, stroke, chronic inflammation, cancers, and weakened immune defense, which may be particularly relevant during the COVID-19 pandemic [26,27].

During the pandemic, multiple factors influenced people's sleep quality. The spread of news and rumors through social media, television, and daily interactions about COVID-19's severity and fatality numbers fueled anxiety and fear. Additionally, lack of in-person social connections, changes in eating patterns, less outdoor activities which lead to limited exposure to sunlight, and consequently extended exposure to blue and bright light from screens may lead to disturbed circadian rhythms [21,28-32]. Above all, the suddenly loosened policy and the increased risk of infection may lead some individuals to panic-purchase medical resources—especially COVID-19 medicines (including cough medicine, nasal congestion/snot relievers, fever reducers, pain relievers, and other symptom-alleviating pharmaceuticals for COVID-19)—to ensure timely administration when needed [33,34]. The purchasing behavior towards COVID-19 medicines reflected people's perception of the disease and their self-protection strategies. However, it may also leave those in need unable to obtain them. This behavior was not only driven by the actual need for medications but also by psychological factors, such as the desire for a sense of security in the face of an uncertain health situation. Understanding the relationship between sleep status and these factors after the policy adjustment is of great significance for public health.

However, there is a lack of in-depth research on how the adjustment of China's zero-COVID policy affects the public's sleep status and its connection with the purchasing behavior of COVID-19 medicine. This study aims to fill this gap by conducting a cross-sectional, internet-based survey to assess the public's sleep status and explore its association with the purchasing behavior of COVID-19 medicine after the policy adjustment. The findings of this research can provide valuable insights for public health interventions and future pandemic-related policies.

Methods

Questionnaire

The questionnaire was formulated by the Zhejiang Provincial Center for Disease Control and Prevention based on pandemic-specific context, literature review, and expert consultation. To ensure its reliability, a pilot study was first conducted. Following the feedback from the pilot study and through two rounds of collaboration based on the Delphi method, the questionnaire was revised. During this process, subject matter experts evaluated its face validity, and a literature review was carried out to assess its content validity.

The questionnaire collected sociodemographic data: purchasing behavior of COVID-19 medications; specific types of COVID-19 drugs that respondents had purchased; channels through which they bought these medications; and degrees of their sleep disturbances. Sociodemographic data includes age, gender, education level, occupation, marital status, and diagnosis of COVID-19 among family members. The specific questions regarding the purchasing behavior of COVID-19 medications and the degrees of their sleep disturbances were as follows: (1) "Have you purchased COVID-19 drugs?" (2) "Which of the following COVID-19 drugs have you purchased?" (3) "Through

XSL•FO
RenderX

what channel did you successfully purchase these medicines?" (4) "In the past week, has the adjustment of COVID-19 policies affected your sleep?" No item randomization or adaptive questioning was used, and the 18-item questionnaire was distributed across 3 screens. Participants could review and revise answers via a "Back" button on each screen before submission. IP address tracking was used to prevent multiple entries from the same individual, and no two entries from the same IP address were allowed within 72 hours.

To maintain the quality of the survey data, a quality control question was incorporated into the questionnaire to identify survey respondents who did not answer the questions seriously. The question was, "This is a quality control question, please choose B." All the survey subjects who chose other options were not included in the final data analysis. In addition, server-side completeness checks were implemented upon submission, incomplete responses were rejected, and participants were prompted to fill in missing fields.

The questionnaire was developed and administered entirely in simplified Chinese, consistent with the native language of the target population—see Supplemental File for Chinese (in Multimedia Appendix 1) and translated English version (in Multimedia Appendix 2) of the questionnaire.

Recruitment

The survey was carried out between December 16 and 30, 2022, among residents aged 18-69 years in Zhejiang province, China. Initially, an online open survey questionnaire was dispatched to the 11 municipal Centers for Disease Control and Prevention. These centers then actively promoted and conducted the online investigation by sharing survey links via local official WeChat public accounts, neighborhood WeChat groups, and other regional lifestyle apps. Visitors to the survey link could choose to exit at any time, and no access to other content was restricted to survey completion. The questionnaire was hosted on Wenjuan, a professional web-based platform specialized in questionnaires, tests, assessments, and voting. This platform is dedicated to offering users powerful and user-friendly services such as online questionnaire design, data collection, customized reports, and analysis of survey results.

Measurements Control

Levels of Sleep Disturbance

The Likert scale was used to assess the levels of sleep disturbance related to the COVID-19 epidemic reported by participants. In our research, we used a 5-point Likert scale which consists of 5 answer options: (1) not at all, (2) slight, (3) moderate, (4) significant, and (5) very significant.

Purchasing Behavior of COVID-19 Medicine

Purchasing behavior of COVID-19 medicine was classified as follows: (1) participants who felt it unnecessary to purchase COVID-19 medicine, (2) participants who had tried, but were unable to access any COVID-19 medicine, and (3) participants who had successfully purchased COVID-19 medicine.

Participants were classified as having COVID-19 medication purchasing behavior during the early phase following China's zero-COVID policy adjustment if they had acquired the following drugs at that time: (1) cough medicine, (2) nasal congestion/snot relievers, (3) fever reducers, (4) pain relievers such as ibuprofen (Advil, Motrin IB, and others) or acetaminophen (Tylenol and others), and (5) other pharmaceuticals capable of alleviating COVID-19 infection symptoms.

Family Members Infected With COVID-19

"Family members infected with COVID-19" was defined as "individuals who reside in the same household as the participant and had been diagnosed with COVID-19 (via nucleic acid or antigen testing) at any time prior to the survey."

Statistical Analysis

Data were exported from the Wenjuan website to Excel (Microsoft), and were analyzed using the Statistical Analysis System (SAS), version SAS Viya Long-Term Support 2024.03 (SAS Institute Inc., Cary, NC, USA). Standard descriptive statistics were used for continuous and categorical variables to describe the characteristics of participants in this study. The χ^2 test was used to explore the association between the purchasing behavior of COVID-19 medicines and the levels of sleep disturbance. Univariate logistic regression analysis was conducted to identify the sociodemographic factors associated with sleep disturbance. We used multivariate logistic regression to identify the association between sleep disturbance and purchasing behavior of COVID-19 medicines (categorical variable; reference group was "deemed unnecessary to purchase;" comparison groups was "successfully purchased" and "tried but failed to purchase"). The initial candidate variables included in the model were: age, gender, education level, occupation, marital status, and diagnosis of COVID-19 among family members. A backward stepwise selection approach was used, with a significance level of $P<.10$ for variable retention. Odds ratios (OR) with 95% CI were used to express measures of the associations. P values less than .05 were considered to represent significance (two-sided).

Ethical Considerations

The Zhejiang Provincial Center for Disease Control and Prevention Ethics Board approved the study protocol (2022-027-01). Conducted under strict anonymity, the survey safeguarded the privacy of all participants. No personally identifiable information was collected at any stage of the survey, the dataset was stored by specified staff, and all researchers involved signed a data confidentiality agreement. Informed consent was obtained from all participants prior to their inclusion in the study. By voluntarily clicking an electronic consent box, participants provided their informed consent, thereby indicating their willingness to participate in this research project. As compensation, participants could collect free health education materials at their local Center for Disease Control and Prevention.

XSL•FO
RenderX

Results

Respondents' Characteristics

A total of 40,130 residents took part in the survey, and after implementing quality control procedures, 1650 samples were excluded, with 38,480 samples ultimately included in the study. The sample consisted of 10,431 (27.11%) male participants and 28,049 (72.89%) female participants, with 1129 (2.93%) participants aged 20 years or below, 35,970 (93.48%) participants aged between 20 and 60, and 1384 (3.60%) participants over 60. Regarding occupations, 3049 (7.92%) were health workers, 7712 (20.04%) were government staff, 1522 (3.96%) were students , 1447 (3.76%) were farmers, and 24,750 (64.32%) were business/service industry workers and stay-at-home/unemployed individuals. In terms of education level, 22,180 (57.64%) participants had an education of junior college or undergraduate. Marital status showed that 31,053 (80.70%) people were married and 7427 (19.30%) people were unmarried, while 2873 (7.47%) people resided in households with confirmed COVID-19 cases. The detailed demographic characteristics of the participants are presented in Table 1.

Table . Sociodemographic characteristics of participants completing the online survey in the early stage after the lifting of zero-COVID policy in China (n=38,480).

Characteristic	Count, n (%)
Age (y)[a]	
<20	1126 (2.93)
20 - 60	35,970 (93.48)
>60	1384 (3.60)
Gender	
Male	10,431 (27.11)
Female	28,049 (72.89)
Occupation	
Health workers	3049 (7.92)
Government staff	7712 (20.04)
Students	1522 (3.96)
Farmers	1447 (3.76)
Business/service industry workers and stay-at-home/unemployed individuals	24,750 (64.32)
Education level	
Primary school	900 (2.34)
Junior high school	7229 (18.79)
Senior high school	7017 (18.24)
Junior college/undergraduate	22,180 (57.64)
Postgraduate	1154 (3)
Family members infected with COVID-19	
Yes	2873 (7.47)
No	35,607 (92.53)
Marital status	
Married	31,053 (80.70)
Others (single/divorced/widowed/separated, etc)	7427 (19.30)

[a]In the preliminary analysis, participants were divided into 10-year age groups (<20, 20 - 29, 30 - 39, 40 - 49, 50 - 59, 60 - 69 y). Given similar purchasing behavior in 20 - 29 to 50 - 59 years groups, these 4 groups were merged into 20 - 60 years for concise result presentation, with this data-driven rationale added herein.

The Impact of China's Zero-COVID Policy Adjustment on the Sleep Status of the Public of Zhejiang, China

When the Likert scale was used to assess the degree of impact on participants' sleep, the findings revealed that out of 38,480 participants, 17,677 (45.94%) indicated that the adjustment of China's COVID-19 response policies had no effect on their sleep. In contrast, 20,803 (54.06%) reported experiencing varying levels of sleep disruption. Specifically, 10,964 (52.70%) described the impact as "slight," 3105 (14.93%) as "moderate," 3493 (16.79%) as "significant," and 3241 (15.58%) as "very significant."

XSL•FO
RenderX

Association Between Sleep Status and Purchasing Behavior of COVID-19 Medicine Among the Public in the Early Stage After the Adjustment of China's Zero-COVID Policy

As shown in Table 2, among individuals who had tried, but were unable to access any COVID-19 medicine, 34.21% (7179/20,986) reported no impact on their sleep by the lifting of China's zero-COVID policy in the past week. In contrast, 54.81% (7538/13,752) of those who successfully acquired COVID-19 medications indicated no sleep disturbances. Notably, the proportion of individuals who deemed it unnecessary to buy COVID-19 medications and experienced no sleep disruption was as high as 79.10% (2960/3742).

Table . Impact of purchasing behavior of the COVID-19 medicine on the sleeping status among the participants of an online survey in the early stage after the lifting of zero-COVID policy in China.

Purchasing of the COVID-19 medicine	How has your sleep been affected by the lifting of China's zero-COVID policy in the past week[a]? n (%)					Total
	Very significant	Significant	Moderate	Slight	Not at all	
Participants who felt it unnecessary to purchase COVID-19 medicine	60 (1.60)	62 (1.66)	99 (2.65)	561 (14.99)	2960 (79.10)	3742
Participants who purchased COVID-19 medicine	629 (4.57)	931 (6.77)	917 (6.67)	3737 (27.17)	7538 (54.81)	13,752
Participants who had tried, but were unable to access any COVID-19 medicine	2552 (12.16)	2500 (11.91)	2089 (9.95)	6666 (31.76)	7179 (34.21)	20,986

[a]χ^2=3639.2; *P*<.001.

Among the 3742 respondents who considered purchasing COVID-19 medications unnecessary, the numbers of those experiencing slight, moderate, significant, and very significant impacts on sleep were 561 (14.99%), 99 (2.65%), 62 (1.66%), and 60 (1.60%), respectively. Among the 13,752 participants who had successfully obtained COVID-19 medications, the corresponding figures were 3737 (27.17%), 917 (6.67%), 931 (6.77%), and 629 (4.57%). For the 20,986 individuals who attempted to purchase but failed to obtain COVID-19 medications, the numbers were 6666 (31.76%), 2089 (9.95%), 2500 (11.91%), and 2552 (12.16%). A χ^2 test for comparing multiple sample rates revealed that, overall, there were significant differences among the population rates (*P*<.001).

Factors Associated With Sleep Disturbance Among the Public in the Early Stage After the Adjustment of China's Zero-COVID Policy

As shown in Table 3, there were statistically significant differences in sleep disturbance after the adjustment of China's zero-COVID policy among participants by age, education level, occupation, marital status, and diagnosis of COVID-19 among family members .

XSL•FO
RenderX

Table . Sociodemographic factors associated with sleep disturbance among the public in the early stage after the lifting of China's zero-COVID policy.

Variables	Sleep status affected by the lifting of China's zero-COVID policy, n (%)		OR[a] (95%CI)
	Yes	No	
Sex			
Male	5716 (54.80)	4715 (45.20)	1.04 (0.99-1.09)
Female	15,087 (53.79)	12,962 (46.21)	Reference
Age (years)			
<20	409 (36.32)	717 (63.68)	Reference
20 - 60	19,714 (54.81)	16,256 (45.19)	2.13 (1.88-2.41)[b]
>60	680 (49.13)	704 (50.87)	1.69 (1.44-1.99)[b]
Education level			
Primary school	517 (57.44)	383 (42.56)	1.36 (1.14-1.61)[b]
Junior high school	3928 (54.34)	3301 (45.66)	1.19 (1.05-1.35)[c]
Senior high school	3808 (54.27)	3209 (45.73)	1.18 (1.04-1.35)[c]
Junior college/undergraduate	11,974 (53.99)	10,206 (46.01)	1.17 (1.03-1.31)[c]
Postgraduate	576 (49.91)	578 (50.09)	Reference
Occupation			
Health workers	1242 (40.73)	1807 (59.27)	Reference
Government staff	4161 (53.095)	3551 (46.05)	1.71 (1.57-1.86)[b]
Students	620 (40.74)	902 (59.26)	1.00 (0.88-1.13)
Business or service industry/stay-at-home/unemployed, etc	13,925 (56.26)	10,825 (43.74)	1.87 (1.73-2.02)[b]
Farmers	855 (59.09)	592 (40.91)	2.10 (1.85-2.39)[b]
Marital status			
Married	17,143 (55.21)	13,910 (44.79)	1.27 (1.21-1.33)[b]
Others (divorced/widowed/separated, etc)	3660 (49.28)	3767 (50.72)	Reference
Has anyone in your family been diagnosed with COVID-19?			
Yes	2023 (70.41)	850 (29.59)	2.13 (1.96-2.32)[b]
No	18,780 (52.74)	16,827 (47.26)	Reference

[a]OR: odds ratio.

[b]$P<.001$.

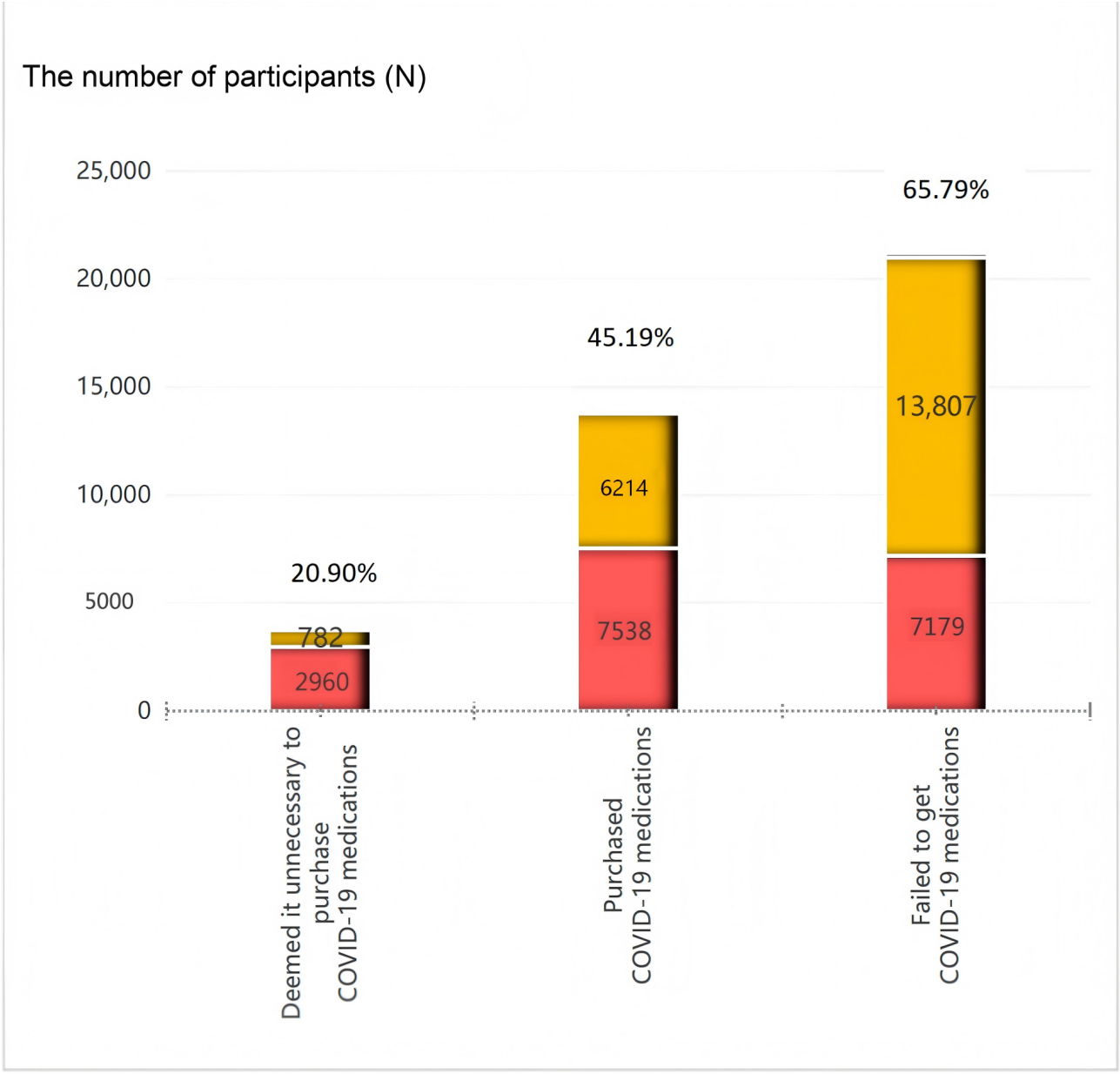
[c]$P<.01$.

Age emerged as a notable predictor: participants aged 20 - 60 years had the highest sleep disturbance rate (19,714/35,970, 54.80%), followed by those over 60 years (680/1384, 49.13%), while 409 out of 1126 participants under 20 years reported sleep disturbance (409/1126, 36.32%). Education level showed a clear inverse relationship with sleep disturbance: the rate was highest among primary school graduates (517/900, 57.44%) and gradually decreased with higher education, reaching 49.91% (576/1154) for master's degree holders. Occupation also played a role. Farmers had the highest rate (855/1447, 59.09%), followed by business/service industry workers and stay-at-home/unemployed individuals (13,925/24,750, 56.26%), government staff (4161/7712, 53.95%), while 1242 out of 3049 health workers (1242/3049, 40.74%) and 620 out of 1522 students (620/1522, 40.74%) had lower rates. Marital status and family COVID-19 diagnosis further contributed to sleep disturbance: married participants had a higher (17,143/31,053, 55.21%) sleep disturbance rate than unmarried/separated individuals (OR 1.27, 95% CI 1.21 - 1.33). Most notably, participants with family members diagnosed with COVID-19 had the highest sleep disturbance rate (2023/2873, 70.41%) across all groups, with odds of disturbance 2.13 times higher than those without infected family members (OR 2.13, 95% CI 1.96 - 2.32) (detailed values in Table 3).

XSL•FO
RenderX

As shown in Figure 1, among individuals who had tried, but were unable to access any COVID-19 medicine, 65.79% (13,807/20,986) reported experiencing sleep disturbance. In contrast, 45.19% (6214/13,752) of those who successfully acquired COVID-19 medications indicated sleep disturbances. Notably, the proportion of individuals who deemed it unnecessary to buy COVID-19 medications and experienced sleep disruption was a mere 20.9% (782/3742).

Figure 1. Stacked bar chart showing individuals affected by sleep disturbance stratified by COVID-19 medication access status among the public in the early stage after the lifting of China's zero-COVID policy. The yellow color represents those who reported experiencing sleep disturbance, and the red color represents those unaffected.



By using multivariate logistic regression analysis and after adjusting for variables including gender, age, education level, occupation, marital status, and presence of family members diagnosed with COVID-19, it was discovered that there was a statistically significant correlation between purchasing behavior of COVID-19 medicine and the sleep status of the participants. Table 4 shows that, compared to those who deemed it unnecessary to buy COVID-19 medications, those who successfully acquired COVID-19 medications were 3.11 times more likely to encounter sleep disturbance (adjusted odds ratio [aOR] 3.11, 95% CI 2.85 - 3.39). Moreover, those who had tried, but were unable to access any COVID-19 medicine were 7.11 times (aOR 7.11, 95% CI 6.53 - 7.74) more likely to experience sleep disturbance.

XSL•FO
RenderX

Table . Multivariate logistic analysis on the association between the sleep status and the purchasing of COVID-19 medicines among the public in the early stage after the lifting of zero-COVID policy in China.

Purchasing of COVID-19 medicine	Participants, n	Sleep status affected by COVID-19 epidemic, n (%)	aOR[a] (95% CI)	*P* value
Participants who felt it unnecessary to purchase COVID-19 medicine	3742	782 (20.90)	Reference	—[b]
Participants who purchased COVID-19 medicine	13,752	6214 (45.19)	3.11 (2.85-3.39)	<.001
Participants who had tried, but were unable to access any COVID-19 medicine	20,986	13,807 (65.79)	7.11 (6.53-7.74)	<.001

[a]aOR: adjusted odds ratio.

[b]Not applicable.

Discussion

Principal Findings

This study explores the sleep status of the public in Zhejiang province, China, following the adjustment of the zero-COVID policy and explores its relationship with the purchasing behavior towards COVID-19 medicine. The results offer valuable insights into the post-policy-adjustment situation. Given the dearth of similar studies, this research will serve as a foundation for future studies that highlight such crucial issues in the context of future public health emergencies.

After the zero-COVID policy adjustment, news and rumors about infection severity and fatalities—spread via social media, television, and daily interactions—fueled public anxiety and fear. Consequently, people act on their primal instincts and are prone to engage in panic buying to ease this psychological reaction [35]. The results of our study revealed that 35.74% (13,752/38,480) of the participants had successfully purchased COVID-19 medicines, 54.54% (20,986/38,480) had attempted to make a purchase but were unable to obtain any of the drugs, and 9.72% (3742/38,480) felt it was unnecessary to stockpile COVID-19 drugs. Psychologically, those who tried to purchase COVID-19 medicines may be driven by a fear of drug shortages, price increases, the surge in infections, and the availability of medicines and home delivery services during the pandemic [34,36]. Therefore, whether they successfully purchase medications directly affects their psychological state. Those who manage to buy medications will experience a certain degree of anxiety relief, while those who fail to do so may become more anxious [37-40].

Our study revealed that 54.06% (20,803/38,480) of the participants experienced sleep disturbances following the adjustment of the zero-COVID policy. It can be explained as a psychological burden during stressful circumstances. Pandemic-related sleep disturbance was also confirmed by many other studies worldwide. A study carried out in Beijing, China, found that, during the second wave of COVID-19, the overall prevalence of sleep disturbance was 50.8% [41]. Madeleine reported a 33.5% sleep disturbances during the COVID-19 shutdown phase in Germany [42]. It is also documented that there is a negative impact on the quality of sleep among the Spanish population during the lockdown period due to COVID-19 [43]. Studies found that COVID-19 had both negative and positive impacts on adolescent sleep. A study in Canada found that the shutdown of schools due to the COVID-19 pandemic might lead to a 2-hour shift in the sleep of typically developing adolescents, longer sleep duration, improved sleep quality, and less daytime sleepiness compared to those experienced under the regular school-time schedule [44]. Parents of 8th grade students from local schools across Ohio, Kentucky, and Virginia (USA) reported adolescents had more difficulties initiating and maintaining sleep during COVID-19 than before COVID-19, with clinically elevated rates increasing from 24% to 36%. Both bedtimes and wake times shifted later during COVID-19, and adolescents reported more delayed sleep/wake behaviors. Adolescents also reported less daytime sleepiness and longer school night sleep duration during COVID-19 [45].

Our study found that the degree of sleep disturbance varied, with 52.70% (10,964/20,803) reporting “slight,” 14.93% (3105/20,803) “moderate,” 16.79% (3493/20,803) “significant,” and 15.58% (3241/20,803) “very significant.” It suggests a complex and diverse range of effects on individuals, which may be related to various factors. Our study found that sociodemographic factors played a crucial role in sleep disturbances. Older people were more likely to experience sleep disturbances compared to the younger. A systematic review found that studies involving older participants had a higher pooled prevalence of insomnia symptoms during the COVID-19 pandemic [22]. Older individuals may be more sensitive to policy change impacts, due to preexisting health conditions, limited medical access, and greater reliance on stable daily routines—all of which increase their fear of infection [46]. Even though women were found to be more prone to sleep disturbance than men [47-49], the perceived impact of COVID-19 on sleep was not different between sexes in our study, which was also confirmed by research from the United States by Jeremy A Bigalke [50]. Literature reviews also revealed that sex had no bearing on the estimated prevalence of sleep disturbance [21]. Education level was inversely related to sleep disturbances; as the education level increased, the likelihood of experiencing sleep disturbances decreased. This could be because individuals with higher education levels are generally better at accessing and understanding information and medical resources, and thus better able to cope with the stress caused by the pandemic.

XSL•FO
RenderX

Studies found that people with higher socioeconomic status are more likely to attain medical resources during the COVID-19 pandemic [40,51]. While health care workers were documented as the most sleep-vulnerable group early in the COVID-19 pandemic [52,53], our study found the opposite in post-policy-adjustment in China. Different occupations face varying degrees of stress during the pandemic. For example, government staff may be burdened with epidemic-prevention work responsibilities, while business/service industry workers may worry about economic losses, and farmers may be concerned about the impact of the pandemic on agricultural production. However, at the very beginning after the adjustment of zero-COVID policy in China, the anxiety of getting infected without drugs in hand might relate to poor sleep quality. Our previous research found that 56.54% of health workers had successfully purchased COVID-19 medicines, which was the highest among all the occupation groups and might explain the less experience of sleep disturbance in this study [40]. Marital status and family members' COVID-19 diagnosis also had an impact. Married individuals had a higher probability of sleep disturbances, perhaps due to increased family responsibilities during the pandemic. Participants with family members diagnosed with COVID-19 were more prone to sleep disturbances, with a proportion as high as 70.41% (2023/2873), indicating that the direct impact of the disease on the family can cause significant psychological pressure, which was also confirmed by a study carried out in Bangladesh. Participants having relatives or acquaintances infected with COVID-19 were more likely to experience poor sleep quality (68.8% vs 49.0%) [47].

In addition to sociodemographic factors, our study found a strong association between the purchasing behavior of COVID-19 medicine and sleep status. Those who deemed purchasing unnecessary had the lowest proportion of sleep disruptions, while those who tried but failed to obtain medications had the highest. After adjusting for sociodemographic factors, it demonstrated that compared to those who thought purchasing was unnecessary, those who acquired medications were 3.11 times more likely, and those who tried but couldn't get medications were 7.11 times more likely to experience sleep disturbance. This difference highlights the role of medication availability and the psychological stress related to it [35]. For those who manage to buy the medicine, it may bring a sense of security and relieve some anxiety, as they feel prepared for potential infections. However, their higher sleep disturbance rate—compared to those who saw no need to purchase COVID-19 medicines—suggests lingering concerns about the disease and future drug shortages still impact their mental state. For those who fail to purchase it, the situation can exacerbate their anxiety. The fear of not being able to access necessary medications during illness and the uncertainty about their health can heighten stress levels, making them more likely to experience sleep disturbance. Our previous research revealed an inverse correlation between anxiety levels and the behavior of purchasing COVID-19 medications which can further explain the association between purchasing behavior of COVID-19 medicine and sleep status. Compared to those with severe anxiety, those with moderate anxiety were 1.76 times more likely to have purchased COVID-19 medicine (aOR 1.76, 95% CI 1.64 - 1.89); those with mild anxiety were 2.11 times (aOR 2.11, 95% CI 1.98 - 2.24) more likely to have purchased COVID-19 medicine; those with no anxiety were 2.48 times (aOR 2.48, 95% CI 2.31 - 2.67) more likely to have purchased COVID-19 medicine [40]. Individuals who were unable to obtain COVID-19 medications may have experienced heightened anxiety, which in turn could have contributed to increased instances of sleep disturbance. The very strong aOR of 7.11 for those unable to access medications and their likelihood of reporting sleep disturbances is notable. However, this causal inference requires further validation to account for unmeasured confounders (eg, baseline anxiety levels, preexisting sleep disorders).

Limitations

This study has certain limitations. Firstly, to quickly explore the association between sleep disturbance and COVID-19 medicine purchasing behavior, the questionnaire was kept concise, omitting potential confounders (eg, COVID-19 vaccination status, comorbidities). "Participants who had tried, but were unable to access any COVID-19 medicine" was not clearly defined, which may lead to misclassification of participants' purchasing behavior. Sleep disturbance was measured via a single Likert-scale item—practical for large surveys but lacking details on sleep frequency/duration/severity, limiting study depth. These should be explored in future studies. Secondly, the data collection relied on self-reported information, which is subject to recall bias and may distort the study results. Thirdly, among the participants in the online survey, the participation of women and those with a higher education level is significantly higher than that of men and those with a low education level. Thus, applying the survey results directly to the general population could lead to inaccuracies.

Conclusions

The findings reflect 54.1% of participants experiencing varying levels of sleep disturbance after the zero-COVID policy adjustment in Zhejiang, China. The older people, those with lower education levels, those with family members diagnosed with COVID-19, and those who had tried, but were unable to access any COVID-19 medicine were mostly affected. Public health agencies should design tailored communication campaigns for different demographic groups, especially for those high-risk groups, to minimize COVID-19-related fear and anxiety. Additionally, optimizing the supply and distribution of COVID-19 medications might help alleviate the stress of individuals who are unable to access them. A robust, transparent pharmaceutical supply chain monitoring system and dedicated channels (hotlines/online platforms) for reporting shortages would enable timely supply adjustments, ensuring high-risk groups access necessary drugs. Beyond China, our findings hold global relevance to observations in other countries where postpandemic policy shifts linked to anxiety and sleep disruptions [42,43].

XSL•FO
RenderX

Acknowledgments

The authors are grateful to the Wenjuan platform that helped carry out the online survey and staff from the municipal Center for Disease Control and Prevention involved in this study.

Data Availability

The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical restrictions.

Authors' Contributions

Conceptualization: YH, X Zhao, SW
Data curation: SW, SX
Formal analysis: YH, SW, X Zhao
Investigation: YH, LW, QL, QW, SX
Supervision: SW, X Zhang
Writing—original draft: YH, X Zhao, SW
Writing—review and editing: YH, X Zhao, LW, QL, SW, QW, SX, X Zhang
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire: Chinese version.
[DOC File, 31 KB - i-jmr_v15i1e79903_app1.doc]

Multimedia Appendix 2
Questionnaire: translated English version.
[DOC File, 41 KB - i-jmr_v15i1e79903_app2.doc]

References

1. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - the latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 2020 Feb;91:264-266. [doi: 10.1016/j.ijid.2020.01.009] [Medline: 31953166]
2. Lu H, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle. J Med Virol 2020 Apr;92(4):401-402. [doi: 10.1002/jmv.25678] [Medline: 31950516]
3. Spina S, Marrazzo F, Migliari M, Stucchi R, Sforza A, Fumagalli R. The response of Milan's emergency medical system to the COVID-19 outbreak in Italy. Lancet 2020 Mar 14;395(10227):e49-e50. [doi: 10.1016/S0140-6736(20)30493-1] [Medline: 32119824]
4. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. Zhonghua Liu Xing Bing Xue Za Zhi 2020 Feb 10;41(2):145-151. [doi: 10.3760/cma.j.issn.0254-6450.2020.02.003] [Medline: 32064853]
5. Lim J, Jeon S, Shin HY, et al. Case of the index patient who caused tertiary transmission of coronavirus disease 2019 in Korea: the Application of Lopinavir/Ritonavir for the Treatment of COVID-19 pneumonia monitored by quantitative RT-PCR. J Korean Med Sci 2020 Feb 17;35(6):e79. [doi: 10.3346/jkms.2020.35.e79] [Medline: 32056407]
6. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. JAMA 2020 Aug 25;324(8):782-793. [doi: 10.1001/jama.2020.12839] [Medline: 32648899]
7. Jernigan DB, CDC COVID-19 Response Team. Update: public health response to the coronavirus disease 2019 outbreak - United States, February 24, 2020. MMWR Morb Mortal Wkly Rep 2020 Feb 28;69(8):216-219. [doi: 10.15585/mmwr.mm6908e1] [Medline: 32106216]
8. Eurosurveillance editorial team. Note from the editors: World Health Organization declares novel coronavirus (2019-nCoV) sixth public health emergency of international concern. Euro Surveill 2020 Feb;25(5):200131e. [doi: 10.2807/1560-7917.ES.2020.25.5.200131e] [Medline: 32019636]
9. Tang S, Wang X, Tang B, et al. Threshold conditions for curbing COVID-19 with a dynamic zero-case policy derived from 101 outbreaks in China. BMC Public Health 2023 Jun 6;23(1):1084. [doi: 10.1186/s12889-023-16009-8] [Medline: 37280554]

XSL•FO
RenderX

10. Li H, Zhang H. Cost-effectiveness analysis of COVID-19 screening strategy under China's dynamic zero-case policy. Front Public Health 2023;11:1099116. [doi: 10.3389/fpubh.2023.1099116] [Medline: 37228729]
11. China's response to COVID-19: a chance for collaboration. Lancet 2021 Apr 10;397(10282):1325. [doi: 10.1016/S0140-6736(21)00823-0] [Medline: 33838746]
12. Huang Y, Wu Q, Wang P, et al. Measures undertaken in China to avoid COVID-19 infection: internet-based, cross-sectional survey study. J Med Internet Res 2020 May 12;22(5):e18718. [doi: 10.2196/18718] [Medline: 32396516]
13. Wei Y, Ye Z, Cui M, Wei X. COVID-19 prevention and control in China: grid governance. J Public Health (Oxf) 2021 Apr 12;43(1):76-81. [doi: 10.1093/pubmed/fdaa175]
14. Wilson O, Flahault A. China's U-turn in its COVID-19 policy. Anaesth Crit Care Pain Med 2023 Apr;42(2):101197. [doi: 10.1016/j.accpm.2023.101197] [Medline: 36646356]
15. Notice on further optimizing and implementing prevention and control measures for COVID-19 [Web page in Chinese]. National Health Commission of the People's Republic of China. 2022. URL: https://www.nhc.gov.cn/xcs/zhengcwj/202211/c95de5f0be9b4e49b99517385aa662a0.shtml [accessed 2025-12-12]
16. Huang Y, Zhao N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. Psychiatry Res 2020 Jun;288:112954. [doi: 10.1016/j.psychres.2020.112954] [Medline: 32325383]
17. Partinen M, Bjorvatn B, Holzinger B, et al. Sleep and circadian problems during the coronavirus disease 2019 (COVID-19) pandemic: the International COVID-19 Sleep Study (ICOSS). J Sleep Res 2021 Feb;30(1):e13206. [doi: 10.1111/jsr.13206] [Medline: 33179820]
18. Dongol E, Shaker K, Abbas A, et al. Sleep quality, stress level and COVID-19 in university students; the forgotten dimension. Sleep Sci 2022;15(Spec 2):347-354. [doi: 10.5935/1984-0063.20210011] [Medline: 35371400]
19. Jahrami H, BaHammam AS, Bragazzi NL, Saif Z, Faris M, Vitiello MV. Sleep problems during the COVID-19 pandemic by population: a systematic review and meta-analysis. J Clin Sleep Med 2021 Feb 1;17(2):299-313. [doi: 10.5664/jcsm.8930] [Medline: 33108269]
20. Magklara K, Kyriakopoulos M. The impact of the COVID-19 pandemic on children and young people. Psychiatriki 2023 Dec 29;34(4):265-268. [doi: 10.22365/jpsych.2023.024] [Medline: 37982248]
21. Jahrami HA, Alhaj OA, Humood AM, et al. Sleep disturbances during the COVID-19 pandemic: a systematic review, meta-analysis, and meta-regression. Sleep Med Rev 2022 Apr;62:101591. [doi: 10.1016/j.smrv.2022.101591] [Medline: 35131664]
22. Li Y, Chen B, Hong Z, et al. Insomnia symptoms during the early and late stages of the COVID-19 pandemic in China: a systematic review and meta-analysis. Sleep Med 2022 Mar;91:262-272. [doi: 10.1016/j.sleep.2021.09.014] [Medline: 34732293]
23. Kumar N, Gupta R. Disrupted sleep during a pandemic. Sleep Med Clin 2022 Mar;17(1):41-52. [doi: 10.1016/j.jsmc.2021.10.006] [Medline: 35216760]
24. Krueger JM, Frank MG, Wisor JP, Roy S. Sleep function: toward elucidating an enigma. Sleep Med Rev 2016 Aug;28:46-54. [doi: 10.1016/j.smrv.2015.08.005] [Medline: 26447948]
25. Vyazovskiy VV. Sleep, recovery, and metaregulation: explaining the benefits of sleep. Nat Sci Sleep 2015;7:171-184. [doi: 10.2147/NSS.S54036] [Medline: 26719733]
26. Richter K, Kellner S, Hillemacher T, Golubnitschaja O. Sleep quality and COVID-19 outcomes: the evidence-based lessons in the framework of predictive, preventive and personalised (3P) medicine. EPMA J 2021 Jun;12(2):221-241. [doi: 10.1007/s13167-021-00245-2] [Medline: 34122671]
27. Liew SC, Aung T. Sleep deprivation and its association with diseases- a review. Sleep Med 2021 Jan;77:192-204. [doi: 10.1016/j.sleep.2020.07.048] [Medline: 32951993]
28. Cox RC, Olatunji BO. Sleep in a pandemic: Implications of COVID-19 for sleep through the lens of the 3P model of insomnia. Am Psychol 2021 Oct;76(7):1159-1171. [doi: 10.1037/amp0000850] [Medline: 34138586]
29. Papagiouvanni I, Kotoulas SC, Vettas C, Sourla E, Pataka A. Sleep during the COVID-19 pandemic. Curr Psychiatry Rep 2022 Nov;24(11):635-643. [doi: 10.1007/s11920-022-01371-y] [Medline: 36192579]
30. Ara T, Rahman MM, Hossain MA, Ahmed A. Identifying the associated risk factors of sleep disturbance during the COVID-19 lockdown in Bangladesh: a web-based survey. Front Psychiatry 2020;11:580268. [doi: 10.3389/fpsyt.2020.580268] [Medline: 33093839]
31. Das S, Khan R, Banerjee S, Ray S, Ray S. Alterations in circadian rhythms, sleep, and physical activity in COVID-19: mechanisms, interventions, and lessons for the future. Mol Neurobiol 2024 Dec;61(12):10115-10137. [doi: 10.1007/s12035-024-04178-5] [Medline: 38702566]
32. Lissak G. Adverse physiological and psychological effects of screen time on children and adolescents: literature review and case study. Environ Res 2018 Jul;164:149-157. [doi: 10.1016/j.envres.2018.01.015] [Medline: 29499467]
33. Al Zoubi S, Gharaibeh L, Jaber HM, Al-Zoubi Z. Household drug stockpiling and panic buying of drugs during the COVID-19 pandemic: a study from Jordan. Front Pharmacol 2021;12:813405. [doi: 10.3389/fphar.2021.813405] [Medline: 35002744]

XSL•FO
RenderX

34. Miah MS, Mamun MR, Saif-Ur-Rahman KM, Rabby AA, Zakaria AFM. A qualitative exploration of purchasing, stockpiling, and use of drugs during the COVID-19 pandemic in an urban city of Bangladesh. Public Health Pract (Oxf) 2024 Jun;7:100477. [doi: 10.1016/j.puhip.2024.100477] [Medline: 38379753]
35. Arafat SMY, Kar SK, Marthoenis M, Sharma P, Hoque Apu E, Kabir R. Psychological underpinning of panic buying during pandemic (COVID-19). Psychiatry Res 2020 Jul;289:113061. [doi: 10.1016/j.psychres.2020.113061] [Medline: 32413711]
36. Jifar WW, Geneti GB, Dinssa SD. The impact of COVID-19 on pharmaceutical shortages and supply disruptions for non-communicable diseases among public hospitals of South West, Oromia, Ethiopia. J Multidiscip Healthc 2022;15:1933-1943. [doi: 10.2147/JMDH.S377319] [Medline: 36072279]
37. Sari E, Kağan G, Karakuş B, Özdemir Ö. Dataset on social and psychological effects of COVID-19 pandemic in Turkey. Sci Data 2022 Jul 23;9(1):441. [doi: 10.1038/s41597-022-01563-4] [Medline: 35871225]
38. Taylor S. Understanding and managing pandemic-related panic buying. J Anxiety Disord 2021 Mar;78:102364. [doi: 10.1016/j.janxdis.2021.102364] [Medline: 33517219]
39. Goodwin R, Wiwattanapantuwong J, Tuicomepee A, Suttiwan P, Watakakosol R, Ben-Ezra M. Anxiety, perceived control and pandemic behaviour in Thailand during COVID-19: results from a national survey. J Psychiatr Res 2021 Mar;135:212-217. [doi: 10.1016/j.jpsychires.2021.01.025] [Medline: 33497875]
40. Huang Y, Xu S, Zhao X, et al. Drug stockpiling behavior and its impact on anxiety among the general public in the early stage after the lifting of China's zero-COVID policy: results from a web-based survey. Front Pharmacol 2025;16:1524068. [doi: 10.3389/fphar.2025.1524068] [Medline: 40421214]
41. Liu Y, Wang X, Sun P, et al. Sleep disturbance and anxiety symptom among public during the second wave of COVID-19 in Beijing: a web-based cross-sectional survey. J Affect Disord 2022 Feb;298(Pt A):80-85. [doi: 10.1016/j.jad.2021.10.068] [Medline: 34728284]
42. Hetkamp M, Schweda A, Bäuerle A, et al. Sleep disturbances, fear, and generalized anxiety during the COVID-19 shut down phase in Germany: relation to infection rates, deaths, and German stock index DAX. Sleep Med 2020 Nov;75:350-353. [doi: 10.1016/j.sleep.2020.08.033] [Medline: 32950879]
43. Maestro-Gonzalez A, Sánchez-Zaballos M, Mosteiro-Díaz MP, Zuazua-Rico D. Quality of sleep among social media users during the lockdown period due to COVID-19 in Spain. Sleep Med 2021 Apr;80:210-215. [doi: 10.1016/j.sleep.2021.01.050] [Medline: 33607556]
44. Gruber R, Saha S, Somerville G, Boursier J, Wise MS. The impact of COVID-19 related school shutdown on sleep in adolescents: a natural experiment. Sleep Med 2020 Dec;76:33-35. [doi: 10.1016/j.sleep.2020.09.015] [Medline: 33070000]
45. Becker SP, Dvorsky MR, Breaux R, Cusick CN, Taylor KP, Langberg JM. Prospective examination of adolescent sleep patterns and behaviors before and during COVID-19. Sleep 2021 Aug 13;44(8):zsab054. [doi: 10.1093/sleep/zsab054] [Medline: 33631014]
46. Du M, Liu M, Wang Y, Qin C, Liu J. Global burden of sleep disturbances among older adults and the disparities by geographical regions and pandemic periods. SSM Popul Health 2024 Mar;25:101588. [doi: 10.1016/j.ssmph.2023.101588] [Medline: 38225953]
47. Islam MS, Rahman ME, Zubayer AA, et al. Investigating poor sleep quality and associated factors during the COVID-19 pandemic: a population-based survey in Bangladesh. Front Public Health 2021;9:724520. [doi: 10.3389/fpubh.2021.724520] [Medline: 34881215]
48. Mandelkorn U, Genzer S, Choshen-Hillel S, et al. Escalation of sleep disturbances amid the COVID-19 pandemic: a cross-sectional international study. J Clin Sleep Med 2021 Jan 1;17(1):45-53. [doi: 10.5664/jcsm.8800] [Medline: 32900428]
49. Casagrande M, Favieri F, Tambelli R, Forte G. The enemy who sealed the world: effects quarantine due to the COVID-19 on sleep quality, anxiety, and psychological distress in the Italian population. Sleep Med 2020 Nov;75:12-20. [doi: 10.1016/j.sleep.2020.05.011] [Medline: 32853913]
50. Bigalke JA, Greenlund IM, Carter JR. Sex differences in self-report anxiety and sleep quality during COVID-19 stay-at-home orders. Biol Sex Differ 2020 Oct 13;11(1):56. [doi: 10.1186/s13293-020-00333-4] [Medline: 33050941]
51. Elek P, Bíró A, Fadgyas-Freyler P. Income gradient of pharmaceutical panic buying at the outbreak of the COVID-19 pandemic. Health Econ 2021 Sep;30(9):2312-2320. [doi: 10.1002/hec.4378] [Medline: 34218496]
52. Wu K, Wei X. Analysis of psychological and sleep status and exercise rehabilitation of front-line clinical staff in the fight against COVID-19 in China. Med Sci Monit Basic Res 2020 May 11;26:e924085. [doi: 10.12659/MSMBR.924085] [Medline: 32389999]
53. Zhang C, Yang L, Liu S, et al. Survey of insomnia and related social psychological factors among medical staff involved in the 2019 novel coronavirus disease outbreak. Front Psychiatry 2020;11:306. [doi: 10.3389/fpsyt.2020.00306] [Medline: 32346373]

Abbreviations

aOR: adjusted odds ratio
OR: odds ratio
SAS: Statistical Analysis System

XSL•FO
RenderX

Edited by N Cahill; submitted 30.Jun.2025; peer-reviewed by MC Yang, OC Balogun; accepted 28.Oct.2025; published 06.Jan.2026.

Please cite as:
Huang Y, Zhao X, Wang L, Lv Q, Wu Q, Xu S, Zhang X, Wu S
Sleep Disturbance and Its Association With Purchasing Behavior of COVID-19 Medicine Among the Public After the Adjustment of Zero-COVID Policy in China: Results From a Web-Based Survey Study
Interact J Med Res 2026;15:e79903
URL: https://www.i-jmr.org/2026/1/e79903
doi:10.2196/79903

© Yu Huang, Xiang Zhao, Lei Wang, Qiaohong Lv, Qingqing Wu, Shuiyang Xu, Xuehai Zhang, Shuxian Wu. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 6.Jan.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

An Eye-Opening Approach: Cancer of Unknown Primary Source With Choroidal Metastasis Case Report

Erin Kaser[1], DO; Vedin Barve[2], BS; Alexander Blaschke[3], DPM; Constance Fry[4], MD; Corey Waldman[5], MD; Sarah Hackman[6], MD; William Kelly[1], MD

[1]Department of Internal Medicine, The University of Texas Health Science Center at San Antonio, 7979 Wurzbach Road, San Antonio, TX, United States

[2]Long School of Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[3]Division of Podiatry, Department of Orthopaedics, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[4]Department of Opthalmology, University of Tennessee Health Science Center, Memphis, TN, United States

[5]Department of Opthalmology, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[6]The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

Corresponding Author:
William Kelly, MD
Department of Internal Medicine, The University of Texas Health Science Center at San Antonio, 7979 Wurzbach Road, San Antonio, TX, United States

Abstract

Choroidal metastases (CM) represent a rare but clinically significant manifestation of systemic malignancy, most frequently from lung cancer. The choroid's vascular anatomy allows hematogenous tumor seeding. Although CM may be the first clinical sign of an underlying malignancy, evidence guiding its management in the modern immunotherapy era remains limited, as most published cases predate the widespread use of immune checkpoint-inhibitors. We describe a 33-year-old male patient presenting with ocular pain and visual disturbance, who was found to have an amelanotic choroidal lesion. Systemic workup revealed small pulmonary nodules and an iliac crest lesion. Sequential biopsies suggested that this was metastatic adenocarcinoma of unknown primary origin, but most likely of lung origin, without actionable mutations or PD-L1 (programmed death-ligand 1) expression. Management required multidisciplinary coordination and included carboplatin, paclitaxel, and pembrolizumab, followed by radiation to the orbit, iliac crest, and mediastinal sites of disease. Unfortunately, he experienced progression while on maintenance immunotherapy with new rib and brain lesions, for which he underwent treatment with platinum, pemetrexed, and bevacizumab with additional radiotherapy. Despite loss of vision in the affected eye, he achieved durable disease control and remains free of radiographic recurrent disease>4 years after diagnosis. This case illustrates that multimodality salvage strategies—integrating systemic therapy with aggressive local radiation—can provide unexpectedly prolonged survival even after immunotherapy failure. Importantly, current guidelines offer minimal direction on managing CM in this context, and prior case reports do not reflect present-day treatment realities. The key message for clinicians is that CM should not automatically be approached with palliative intent; carefully selected patients may benefit from an oligometastatic strategy that actively targets limited metastatic sites to prolong survival. Our findings underscore the need for ophthalmology, radiation oncology, and medical oncology collaboration when vision-threatening or occult metastatic lesions arise. For readers, the takeaway is that choroidal metastasis—particularly in the era of immunotherapy—warrants individualized, multidisciplinary evaluation rather than default palliation. Our case demonstrates that coordinated multimodality management can achieve long-term disease control, highlighting a treatment paradigm worth considering for selected patients and calling for updated guidelines that reflect modern therapeutic capabilities.

(Interact J Med Res 2026;15:e77895) doi:10.2196/77895

KEYWORDS

choroidal; metastasis; non-small cell lung cancer; cancer of unknown primary origin; multi-modal therapy

Introduction

Rare presentations of cancer can include metastases to the orbit through hematogenous dissemination. The uveal tract is a highly vascular complex made from the choroid, ciliary body, and iris. Importantly, the eye contains the blood-retinal barrier, analogous to the blood-brain barrier. However, the choroid layer lies outside retinal pigment epithelial tight junctions and is therefore exposed to unimpeded drug delivery from the vasculature [1]. One of the most common sources of choroidal metastases is primary lung cancer, with 88% of lung metastases to the eye involving the choroid, with an associated overall survival at one-year of 54% [2]. Consequently, choroidal metastases (CM) of unknown origin are often treated similarly to lung cancer as

XSL•FO
RenderX

the clinical, imaging, and histopathologic features of choroidal lesions frequently resemble those seen in lung cancer patients [3,4]. The prognosis for carcinoma of unknown primary (CUP) with CM is generally poor, with median overall survival ranging from 3 to 12 months, as patients tend to present with advanced and disseminated disease [5,6].

Case Presentation

Patient Information

The patient was a 33-year-old male at the time of initial presentation with no significant past medical history. He presented with right ophthalmalgia, blurry vision, and sinus congestion. Clinical course was as follows (Table 1).

Table . Timeline of key events in the described patient presentation

Day	Description of key events.
Day 0-2	Initial presentation: chief complaint of right ophthalmalgia and blurry vision. Ophthalmology observed an amelanotic right choroidal lesion. MRI[a] Orbit showed enhancing plaque of the posterior-lateral wall of right globe. CT[b] chest showed RML and RUL nodules
Day 44-113	Pathological workup: right intraophthalmic choroidal biopsy with scant atypical cells. PET (Positron Emission Tomography) revealed non-avid RML[c] and RUL[d] nodules and mildly avid L iliac crest. R choroidal FNA[e] read as metastatic carcinoma, compatible with pulmonary adenocarcinoma. L iliac biopsy with further suggestion of lung origin (CK7, TTF1+)
Day 125-313	Initial chemoimmunotherapy: started on carboplatin, paclitaxel and pembrolizumab for 4 cycles followed by pembrolizumab maintenance. Completed stereotactic radiation to R orbit
Day 293-324	Disease progression: clinical and radiographic progression at ribs and lungs as well as intracranial site
Day 313-743	Oligometastatic salvage therapy: given radiation to chest wall and left frontal operculum. Started on carboplatin, pemetrexed, and bevacizumab for 4 cycles followed by pemetrexed and bevacizumab maintenance. Bevacizumab was continued as maintenance for 14 cycles
Day 743-Present	Surveillance: no clinical or radiographic evidence of disease progression as of day 1563.

[a]MRI: magnetic resonance imaging.

[b]CT: computed tomography.

[c]RML: right middle lobe.

[d]RUL: right upper lobe.

[e]FNA: fine-needle aspiration.

Ethical Considerations

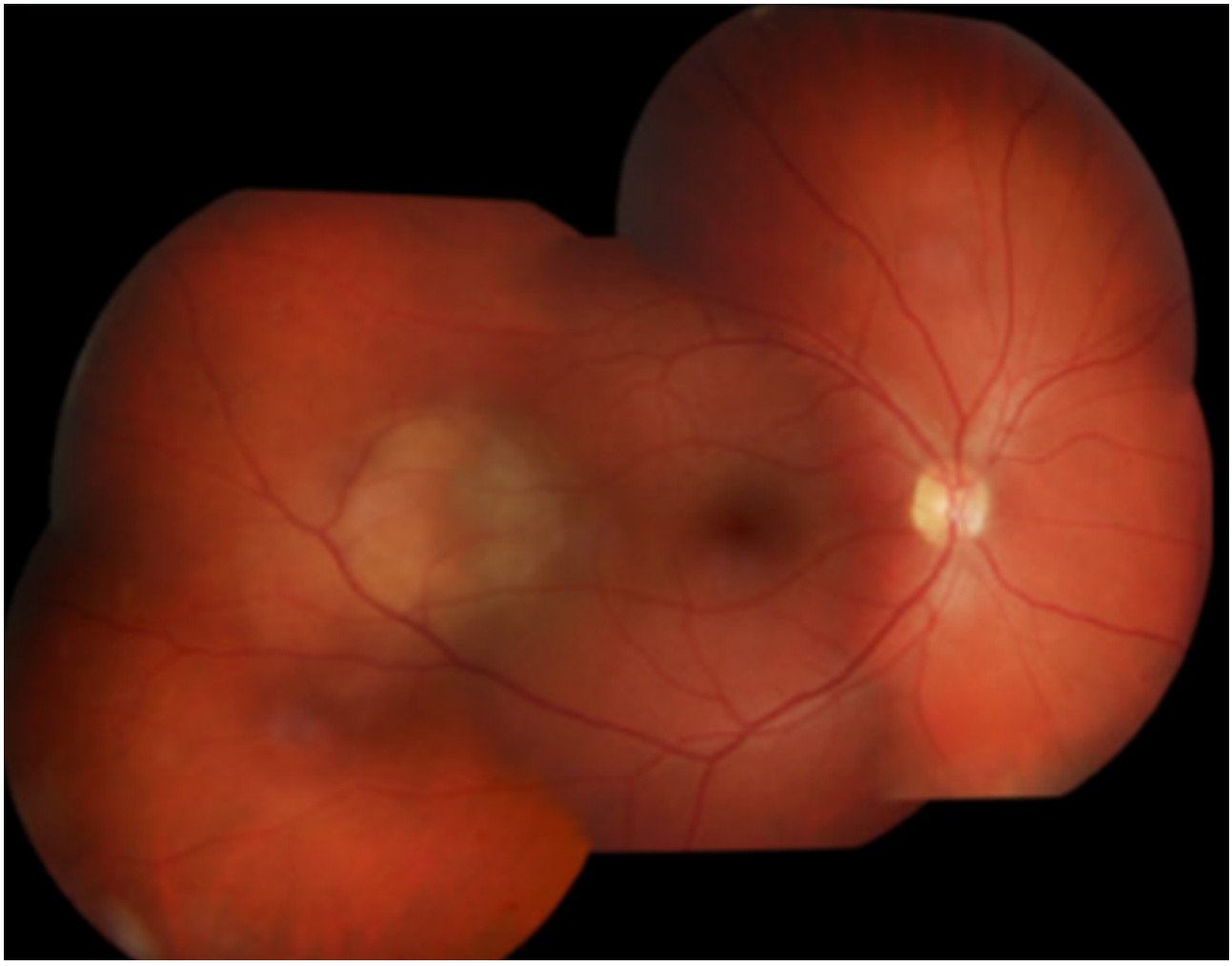
The patient involved in this case report has given his informed consent authorizing the use and disclosure of his health information.

Clinical Findings

Evaluation by ophthalmology revealed an amelanotic right choroidal lesion (1.04 × 1.15×.31 cm, Figure 1). The patient had 20/20 - 1 vision bilaterally with full visual fields.

XSL•FO
RenderX

Figure 1. Ophthalmologic evaluation showed an amelanotic right choroidal lesion.



Diagnostic Assessment

MRI Orbit on Day 0 (Figure 2A) confirmed an enhancing plaque along the posterior-lateral wall of his right globe. A CT Chest (Day 2) demonstrated a right middle lobe (RML) (0.5cm) and right upper lobe (RUL) nodule (0.2cm). The patient underwent intraophthalmic choroidal biopsy (Day 44). Pathology read as scant atypical cells, with few cells showing 'neuroendocrine-like nuclear features.' A DOTATATE PET (Positron Emission Tomography) on Day 64 was performed, which showed mild avidity in the right globe mass (SUV 2.6, Krenning 1) (Figure 3A), a nonavid RML nodule (0.5cm, below PET resolution) and RUL nodule (0.2cm), and a mildly avid left iliac crest lesion (1cm) (Figure 3B, C). The patient subsequently underwent a right choroidal fine-needle aspiration (FNA) on Day 69, with pathology read as metastatic carcinoma, compatible with pulmonary adenocarcinoma. Notably, the histology did not resemble a retinal pigment epithelium tumor and was positive for CK7 and TTF1 but negative for PD-L1, thyroglobulin, PAX8, CD20, GATA3, CD3, and Melan-A. CK20 was only weakly focally positive. Synaptophysin highlighted pieces of retina but not the tumor. Molecular testing showed no actionable mutation in EGFR, KRAS, BRAF, or HER2. Unfortunately, testing for ALK, RET, ROS1, NTRK1/3 and MET was inconclusive due to insufficient sample. A left iliac biopsy (Day 113, Figure 4) and mediastinal lymph node (4R) FNA biopsy (Day 114) were also read as carcinoma showing CK7 and TTF1 positivity. An insufficient sample of tumor was procured for sequencing; however, immunohistochemical (IHC) analysis was negative for EGFR, ALK, ROS, and BRAF driver mutations. No PD-L1 expression (0%) was observed. The patient was discussed in a multidisciplinary tumor board with the ultimate determination that this was metastatic adenocarcinoma of unknown primary, but most likely of pulmonary origin and to treat as non-small cell lung cancer (NSCLC).

XSL•FO
RenderX

Figure 2. (A) MRI orbit (Day 0) showed enhancing plaque along the posterior-lateral wall of his right globe. (B) MRI orbit (Day 1,123) showed new right optic nerve enhancement.

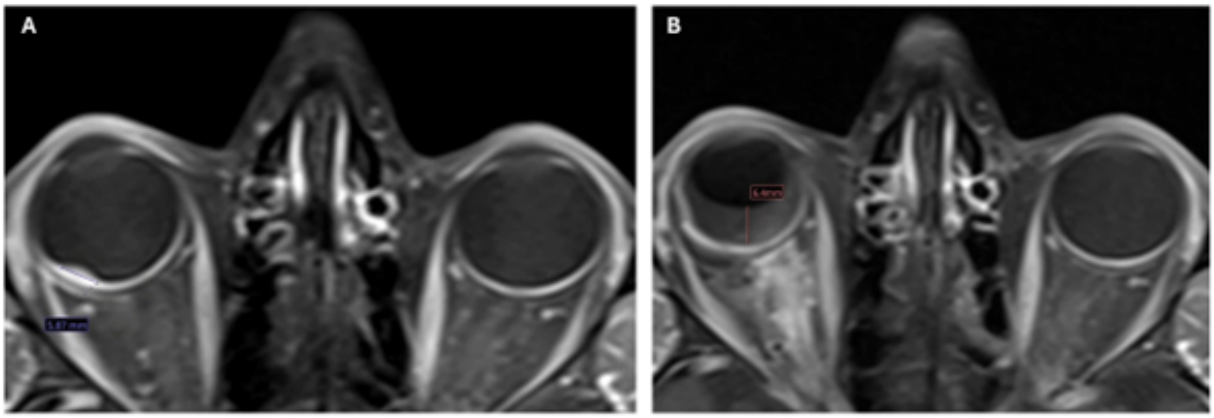
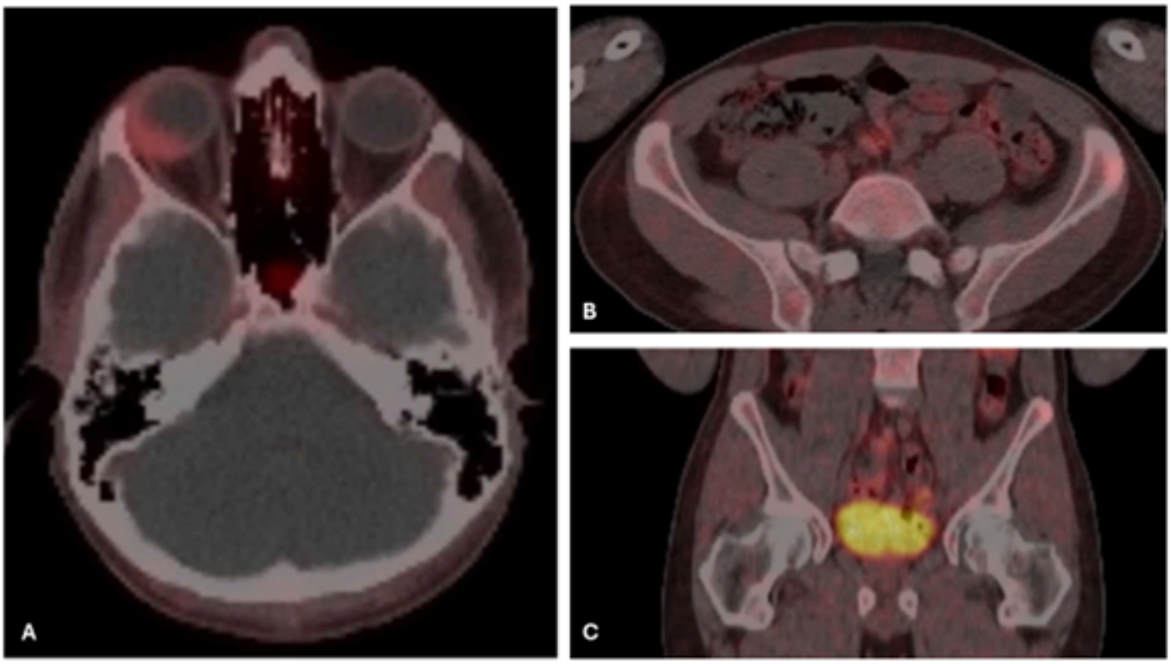
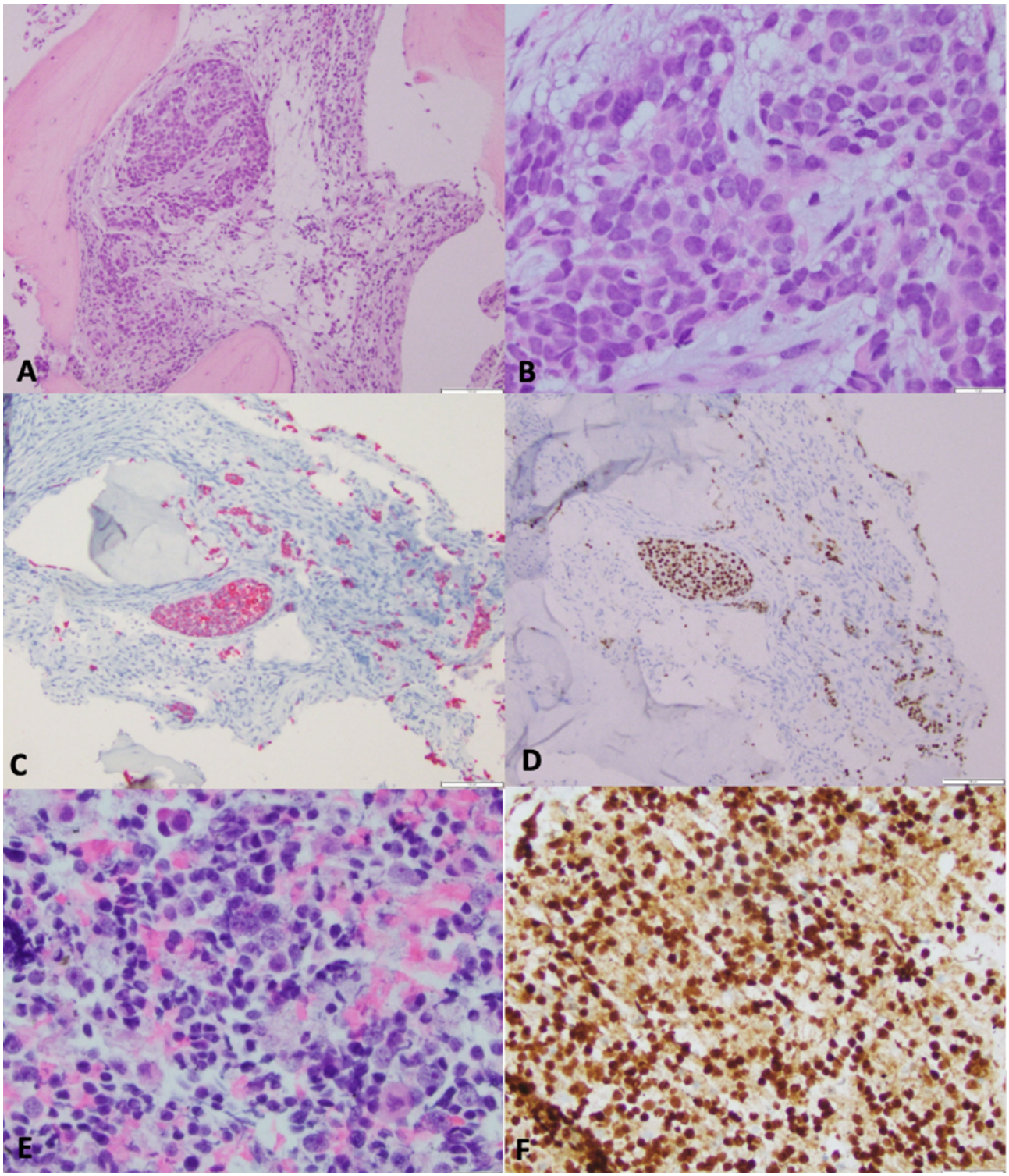


Figure 3. DOTATATE PET Imaging on d 64 demonstrating sites of metastatic involvement(A) Axial DOTATATE PET of the orbits showing mild radiotracer avidity within the right choroidal lesion (SUV 2.6; Krenning score 1), corresponding to the enhancing plaque previously identified on MRI; (B) Axial PET of the pelvis showing a mildly avid lytic lesion in the left iliac crest (approximately 1 cm), consistent with metastatic involvement; (C) Coronal PET confirming the left iliac crest lesion with focal uptake.



XSL•FO
RenderX

Figure 4. Iliac crest biopsy (Day 113) and pulmonary lymph node biopsy (Day 114). (A, B) Iliac crest bone lesion, H&E, 10 x and 40 x: tumor cells with bone; (C) Iliac crest bone lesion, Cytokeratin AE1/3 immunostaining, 10 x: cytokeratin stain confirming the tumor cells are carcinoma; (D) Iliac crest bone lesion, TTF-1 immunostaining, 10 x: TTF-1 stain supports lung origin; (E) Pulmonary lymph node: H&E, 40 x; tumor cells are morphologically comptabile with those in bone; (F) pulmonary lymph node: TTF-1 immunostaining, 20 x: TTF-1 stain supports lung origin. TTF: thyroid transcription factor.



Therapeutic Intervention

Following tumor board discussion, treatment was initiated with carboplatin (AUC 6), paclitaxel (200mg/m^2), and pembrolizumab (200mg) every 21 days (C1D1 on Day 125) for 4 cycles, with pembrolizumab continued as maintenance after cycle 4. Alternative therapies considered included a palliative approach with chemotherapy alone, but age, functional status, limited extent of disease, and the patient's wishes to pursue an aggressive therapeutic approach led to the chosen treatment protocol. Paclitaxel was chosen over pemetrexed as the taxane used because it was deemed to cover more histologies in the event that the cancer was not actually of bronchopulmonary origin. Patient completed stereotactic radiosurgery (SRS) (Day

XSL•FO
RenderX

173) with 36 Gray (Gy) in 4 fractions (fx) to the right orbit lesion. Following cycle 4, he was placed on pembrolizumab maintenance. PET (Day 211) showed an interval increase in L iliac crest lesion and increased conspicuity of R lung base grouped nodularity, but otherwise stable R paratracheal (0.7cm) and mediastinal (0.4cm) lymph nodes and a decreased RML (0.3cm) pulmonary nodule. He also completed radiation to the left iliac (Day 252, 50Gy in 5 fx), RML (Day 252, 50Gy in 5 fx), and mediastinal nodes (Day 273, 40Gy in 10 fx). However, he began having rib pain and PET/CT (Day 293) showed a new lesion in the left 4th rib and RLL, as well as an increase in additional right pulmonary nodules, concerning for progression. Concurrently, MRI Brain (Day 324) had also shown a new left frontal gyrus lesion. Given his age and continued limited disease involvement, an oligometastatic approach was used, and the patient completed radiation to the left chest wall (Day 313, 50Gy in 5 fx) and left frontal operculum (Day 337, 1 fx). Pembrolizumab was stopped (last dose on Day 315) due to radiographic progression, and he was started (C1D1 on Day 344) on carboplatin (AUC 5), pemetrexed (500mg/m^2), and bevacizumab (15mg/kg) every 21 days for 4 cycles, with pemetrexed and bevacizumab continued as maintenance after cycle 4. Pemetrexed and bevacizumab were held after 14 cycles (last dose on Day 743) in preparation for cataract removal of the right eye, and he was subsequently placed on surveillance with MRI Brain and PET imaging every 3 months.

Follow-Up and Outcomes

He developed eye pain, and MRI orbit (Day 1123, Figure 2B) showed right optic nerve enhancement, which was new from previous imaging and felt secondary to prior radiotherapy. He received one dose of intraocular bevacizumab injection, which completely resolved the pain. The patient had developed blindness in his right eye after initial treatment, but remained with 20/20 vision on the left. He maintained his employment and continues with regular follow-up and surveillance imaging every 3 months. His scans as of Day 1563 show no evidence of recurrence.

Discussion

Given CUP with CM tends to present with late, advanced, and disseminated disease, the general treatment protocol is palliative, focusing on symptom control and vision preservation. For most patients, external beam radiotherapy is the standard local treatment for symptomatic choroidal metastases, with systemic therapy including platinum- or taxane-based chemotherapy used for widespread disease [7]. Systemic therapy for CM of CUP often mirrors lung cancer regimens because empirical chemotherapy and targeted therapies have demonstrated efficacy in controlling both systemic and ocular disease in lung cancer patients [8-10]. For metastatic non-small cell lung cancer, platinum-based chemotherapy with PD-1 blockade is the standard first-line treatment; however, data regarding CM in the age of immunotherapy are limited. Recently, Li et al presented a case of a 65-year-old male patient with NSCLC and CM who demonstrated ultrasonic regression, improved visual acuity, and symptomatic relief after treatment with pemetrexed and capecitabine plus pembrolizumab [11]. A second case report by Matsuyama et al describes treatment with nivolumab plus ipilimumab plus chemotherapy, which resulted in symptomatic improvement and no relapse at 18 months follow-up [12]. These isolated cases underscore both the rarity of CM and the need for evidence-guided management strategies.

Bevacizumab, an extracellular VEGF-A inhibitor, has strong anti-angiogenic properties and has been used to treat CM from pulmonary, breast, and colorectal metastases. For instance, Riess JW et al described a case series of 3 patients with NSCLC and CM who were treated with a backbone of platinum-based chemotherapy plus bevacizumab [13]. The first patient received SRS to the symptomatic CM with improved vision and 12 months of progression-free survival. The second patient was found to have CM after right eye vision loss. This vision normalized following whole-brain radiotherapy that included the choroidal regions. Thereafter, the patient received 14 cycles of chemotherapy followed by erlotinib for a discovered EGFR mutation and progression. The third patient tolerated 16 months of chemotherapy followed by a recurrence of CM, which was treated with an ALK inhibitor for a found translocation. Furthermore, George et al reported on a 42-year-old female patient diagnosed with NSCLC with CM who achieved complete response of her 1.2cm CM following 3 cycles of carboplatin, gemcitabine, and bevacizumab, and was still alive with reported normal vision following her 7th maintenance dose of bevacizumab [14].

In the case of our patient, timely referral to ophthalmology led to recognition of a uveal lesion within three weeks of symptom onset. PET imaging can be helpful in defining dominant lesions and disease burden but histologic evaluation is necessary to guide diagnosis. In situations such as ours, where the primary site cannot be readily identified from imaging, judicious use of immunohistochemistry and molecular profiling is called for. In the era of immunotherapeutics, the median survival of newly diagnosed NSCLC without targetable mutation ranges from 14 to 22 months [15,16]. At the time of writing, our patient has survived 4 years with no radiographic or clinical signs of recurrence. While we must be careful not to read too much into an isolated case, it is the authors' belief that an aggressive oligometastatic approach with multimodality salvage options may be best for all NSCLC patients or patients with suspected primary lung malignancy who initially present with choroid disease and limited disease elsewhere.

As presented, this case illustrates an innovative management strategy for choroidal metastasis in the era of immunotherapy. While prior reports describe short-term responses to chemotherapy, anti-VEGF therapy, or immune checkpoint inhibitors, none have documented long-term survival following an oligometastatic approach that was applied both at diagnosis and after systemic progression. Our patient's survival of more than 4 years exceeds historical expectations for metastatic NSCLC without targetable alterations or PD-1 expression and highlights the potential benefits of aggressive local therapy even at the cost of vision.

The substantial collaboration required to execute this treatment approach is equally important. Coordination between imaging,

XSL•FO
RenderX

multiple ophthalmologic biopsies, and subspecialty pathologic review was essential for diagnosis. Treatment required communication between medical and radiation oncology to prioritize systemic control while delivering targeted radiation to multiple metastatic sites. This approach reflects the real-world practice and serves as a model for clinicians managing similarly rare and complex presentations.

Conclusion

CM remains an uncommon presentation of systemic malignancy, and contemporary management is challenged by the absence of evidence-based guidelines that reflect the current era of immunotherapy. Our case demonstrates that even after progression on checkpoint inhibitors, carefully selected patients may achieve prolonged survival through a coordinated multimodality approach that integrates systemic therapy with site-directed radiation. This experience suggests that CM should not automatically be considered a purely palliative entity, particularly when disease burden is low and lesions are anatomically accessible for local therapy.

More broadly, this report emphasizes the importance of early ophthalmologic evaluation for unexplained visual symptoms, the value of thorough systemic staging when CM is suspected, and the critical role of multidisciplinary collaboration involving medical oncology, radiation oncology, and ophthalmology. As prior published cases predate modern immunotherapy, this case highlights an unmet need for updated clinical guidance on managing CM in patients who have progressed on or failed to respond to immune checkpoint inhibitors.

Acknowledgments

No Generative AI was used throughout the generation of this manuscript.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Authors' Contributions

EK and WK were involved in the preparation and writing of the initial draft. VB reviewed and edited the initial draft and prepared the manuscript for submission. WK provided overall supervision of the development of the manuscript. AB, CF, CW, and SH were involved in providing histology and biopsy images as well as reviewing the original manuscript for revisions.

Conflicts of Interest

None declared.

References

1. Yang X, Yu XW, Zhang DD, Fan ZG. Blood-retinal barrier as a converging pivot in understanding the initiation and development of retinal diseases. Chin Med J (Engl) 2020 Nov 5;133(21):2586-2594. [doi: 10.1097/CM9.0000000000001015] [Medline: 32852382]
2. Kim RS, Nguyen MK, Card KR, Shields CL. Utility of systemic brigatinib therapy in tumour recurrence of choroid metastasis from non-small cell lung carcinoma. Can J Ophthalmol 2024 Aug;59(4):e397-e399. [doi: 10.1016/j.jcjo.2023.04.013] [Medline: 37236265]
3. Shields CL, Kalafatis NE, Gad M, et al. Metastatic tumours to the eye. Review of metastasis to the iris, ciliary body, choroid, retina, optic disc, vitreous, and/or lens capsule. Eye (Lond) 2023 Apr;37(5):809-814. [doi: 10.1038/s41433-022-02015-4]
4. Blasi MA, Maceroni M, Caputo CG, et al. Clinical and ultrasonographic features of choroidal metastases based on primary cancer site: Long-term experience in a single center. PLoS One 2021;16(3):e0249210. [doi: 10.1371/journal.pone.0249210] [Medline: 33765045]
5. Jardel P, Sauerwein W, Olivier T, et al. Management of choroidal metastases. Cancer Treat Rev 2014 Dec;40(10):1119-1128. [doi: 10.1016/j.ctrv.2014.09.006] [Medline: 25451606]
6. Mathis T, Jardel P, Loria O, et al. New concepts in the diagnosis and management of choroidal metastases. Prog Retin Eye Res 2019 Jan;68:144-176. [doi: 10.1016/j.preteyeres.2018.09.003] [Medline: 30240895]
7. Raghav K. Cancer of unknown primary site. N Engl J Med 2025 May 29;392(20):2035-2047. [doi: 10.1056/NEJMcp2402691] [Medline: 40435465]
8. Yang CJ, Tsai YM, Tsai MJ, Chang HL, Huang MS. The effect of chemotherapy with cisplatin and pemetrexed for choroidal metastasis of non-squamous cell carcinoma. Cancer Chemother Pharmacol 2014 Jan;73(1):199-205. [doi: 10.1007/s00280-013-2341-4] [Medline: 24202667]
9. Chee ASH, Mak ACY, Kam KW, et al. Efficacy of epidermal growth factor receptor tyrosine kinase inhibitor in treating choroidal metastasis from non-small-cell lung cancer: a 10-year review. Retina 2025 Feb 1;45(2):207-214. [doi: 10.1097/IAE.0000000000004314] [Medline: 39504451]

XSL•FO
RenderX

10. Maller B, Salvatori S, Tanvetyanon T. Outcomes of intraocular metastasis from lung cancer in the era of targeted therapy: a systematic review and pooled analysis. Clin Lung Cancer 2022 Dec;23(8):e519-e525. [doi: 10.1016/j.cllc.2022.07.018] [Medline: 36030188]
11. Li X, Liang Y, Wang J, Hu Z, Yuan ZL, Xie P. Pembrolizumab for choroidal metastases from non-small cell lung cancer: a case report and literature review. Hum Vaccin Immunother 2021 Aug 3;17(8):2626-2629. [doi: 10.1080/21645515.2021.1885282] [Medline: 33769211]
12. Matsuyama T, Oniwa M, Tsuruzono K, et al. Improving visual acuity with nivolumab plus ipilimumab plus two cycles of chemotherapy following a diagnosis of lung adenocarcinoma with choroidal metastasis: a case report and literature review. Respirol Case Rep 2024 Jan;12(1):e01262. [doi: 10.1002/rcr2.1262] [Medline: 38045825]
13. Riess JW, Nagpal S, Das M, Neal JW, Kim JW, Wakelee HA. A case series of NSCLC patients with different molecular characteristics and choroidal metastases: improvement in vision with treatment including pemetrexed and bevacizumab. J Thorac Oncol 2013 Feb;8(2):e17-e18. [doi: 10.1097/JTO.0b013e31827690da] [Medline: 23328555]
14. George B, Wirostko WJ, Connor TB, Choong NW. Complete and durable response of choroid metastasis from non-small cell lung cancer with systemic bevacizumab and chemotherapy. J Thorac Oncol 2009 May;4(5):661-662. [doi: 10.1097/JTO.0b013e31819c9a73] [Medline: 19395911]
15. Garassino MC, Gadgeel S, Speranza G, et al. Pembrolizumab plus pemetrexed and platinum in nonsquamous non-small-cell lung cancer: 5-year outcomes from the phase 3 KEYNOTE-189 study. J Clin Oncol 2023 Apr 10;41(11):1992-1998. [doi: 10.1200/JCO.22.01989] [Medline: 36809080]
16. Johnson ML, Cho BC, Luft A, et al. Durvalumab with or without tremelimumab in combination with chemotherapy as first-line therapy for metastatic non-small-cell lung cancer: the phase III POSEIDON study. J Clin Oncol 2023 Feb 20;41(6):1213-1227. [doi: 10.1200/JCO.22.00975] [Medline: 36327426]

Abbreviations

ALK: anaplastic lymphoma kinase
AUC: area under the curve
BRAF: B-rapidly accelerated fibrosarcoma (BRAF) gene
CK7: Cytokeratin 7
CM: choroidal metastases
CT: computed tomography
CUP: cancer of unknown primary
CxDy: Cycle x Day y
DOTATATE PET: DOTA-Tyr³-octreotate Positron Emission Tomography
EGFR: epidermal growth factor receptor
FNA: fine-needle aspiration
Fx: fraction
Gy: gray (unit of radiation dose)
IHC: immunohistochemistry
MRI: magnetic resonance imaging
NSCLC: non-small cell lung cancer
PAX8: paired box gene 8
PD-L1: programmed death-ligand 1
RLL: right lower lobe
RML: right middle lobe
ROS1: ROS proto-oncogene 1 receptor tyrosine kinase
RUL: right upper lobe
SRS: stereotactic radiation
TTF1: thyroid transcription factor-1
VEGF-A: vascular endothelial growth factor – A

Edited by A Schwartz, M Balcarras; submitted 22.May.2025; peer-reviewed by T Matsuyama; accepted 02.Jan.2026; published 05.Feb.2026.

Please cite as:
Kaser E, Barve V, Blaschke A, Fry C, Waldman C, Hackman S, Kelly W
An Eye-Opening Approach: Cancer of Unknown Primary Source With Choroidal Metastasis Case Report
Interact J Med Res 2026;15:e77895
URL: https://www.i-jmr.org/2026/1/e77895
doi:10.2196/77895

XSL•FO
RenderX

© Erin Kaser, Vedin Barve, Alexander Blaschke, Constance Fry, Corey Waldman, Sarah Hackman, William Kelly. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 5.Feb.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 the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX