Review

Effectiveness of Virtual Reality–Based Early Rehabilitation Strategies on Pain, Sleep, Anxiety, Balance, Cognition, and Limb Motor Function in Adult Intensive Care Unit Patients: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Fei Wu[1,2], MSN; Yunting Wu[1,2], MSN; Yana Xing[1,3], MSN; Weixin Cai[1*], MSN; Ran Zhang[1,3*], MSN

[1]Nursing Department, Beijing Tiantan Hospital, Capital Medical University, Beijing, Beijing, China

[2]School of Nursing, Capital Medical University, Beijing, Beijing, China

[3]Department of Critical Care Medicine, Beijing Tiantan Hospital, Capital Medical University, Beijing, Beijing, China

*these authors contributed equally

Corresponding Author:

Ran Zhang, MSN
Nursing Department
Beijing Tiantan Hospital, Capital Medical University
No. 119, South 4th Ring West Road, Fengtai District, Beijing, China
Beijing, Beijing
China
Phone: 86 59978336
Email: 125166686@qq.com

Abstract

Background: Early rehabilitation is vital for functional recovery in critically ill patients. Virtual reality–based early rehabilitation intervention (VR-ERI) is an emerging strategy, but evidence on its feasibility, safety, and efficacy remains inconsistent and unsynthesized.

Objective: We synthesized evidence from randomized controlled trials (RCTs) on the feasibility and safety of VR-ERI in adult critically ill patients and evaluated its effects on functional outcomes during intensive care unit (ICU) stay and at short-term follow-up (≤3 months post ICU).

Methods: Following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines, we searched 10 databases (eg, PubMed, Web of Science, Cochrane Library, and Embase) from inception to October 5, 2025, for Chinese and English publications. We included RCTs comparing VR-ERI with control measures initiated early (within 72 h of ICU admission) in adults. FW and YW independently screened studies and extracted data; YW and YX independently conducted the revised Cochrane Risk of Bias tool assessment. Data were synthesized narratively or via meta-analysis in R Studio using a random-effects model (Hartung-Knapp-Sidik-Jonkman adjustment). Effects were expressed as standardized mean differences (SMD; with Hedges g correction) with 95% CIs and 95% prediction intervals (95% PIs). Subgroup and sensitivity analyses were performed to explore heterogeneity and assess robustness. The Grading of Recommendations Assessment, Development, and Evaluation framework was used to assess the quality of evidence.

Results: Sixteen RCTs (published 2020-2025) involving 1356 patients were included. Bias assessment found 1 study at low risk, 5 with some concerns, and 10 at high risk. Meta-analysis suggested potential trends for VR-ERI in improving ICU anxiety (SMD −0.86, 95% CI −1.85 to 0.13, 95% PI −3.75 to 2.03; very low certainty) and subjective sleep quality (SMD 3.36, 95% CI 0.77-5.94; very low certainty), with a more pronounced effect in the Richards-Campbell Sleep Questionnaire–assessed subgroup (SMD 5.12, 95% CI 0.54-9.71). At follow-up, VR-ERI showed trends toward improved balance (Berg Balance Scale: SMD 0.97, 95% CI 0.74-1.20, 95% PI 0.37-1.58; moderate certainty), limb motor function (Fugl-Meyer: SMD 1.40, 95% CI −0.23 to 3.02; low certainty), and cognitive function (SMD 0.78, 95% CI 0.16-1.39; low certainty). No significant differences were found for objective sleep measures or ICU pain (low to very low certainty). No serious adverse events were reported; only a few studies mentioned mild reactions, such as dizziness, nausea, and fatigue.

Conclusion: This review indicates VR-ERI's potential to improve anxiety, subjective sleep, balance, cognition, and motor function in early ICU rehabilitation, while its effects on pain and objective sleep remain unclear and safety protocols need refinement. Given the high risk of bias, substantial heterogeneity, and imprecision, the overall certainty of evidence is low. Thus, VR-ERI may serve as a nonpharmacological adjunct, but its clinical translation requires consideration of cost and patient suitability, and more rigorous research is needed.

Trial Registration: PROSPERO CRD420251114439; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251114439

Introduction

Background

Patients in the intensive care unit (ICU) often experience a series of physical, cognitive, and psychological impairments after discharge due to factors such as physiological dysfunction, surgical trauma, pain, and reduced social interaction [1]. In 2010, the Society of Critical Care Medicine first introduced the term "post-intensive care syndrome (PICS)" to define this phenomenon [2,3]. The manifestations commonly include weakness, pain, anxiety, sleep disturbances, and cognitive dysfunction, which can persist for months or even years. Studies indicate that 47.9% of patients still exhibit cognitive impairment 3-6 months after transfer from the ICU [4]; 68% (355/523) report persistent pain after discharge, with pain intensity remaining at a moderate level even 1 year later [5]. Additionally, 55% of patients experience sleep disturbances within the first month after discharge [6]. Furthermore, 12.3% (98,530/799,645) of critically ill patients are diagnosed with depression or anxiety after discharge, and these patients demonstrate higher long-term mortality rates [7]. PICS severely compromises patients' quality of life, reduces functional independence, and imposes a significant care burden on families and society [8].

To reduce long-term impairments after critical illness, the Society of Critical Care Medicine recommends that health care professionals assess patients and coordinate care strategies for injury management beginning at ICU admission [9]. As clinical treatment advances, early functional recovery has become a key focus for preventing PICS [10]. Early rehabilitation is generally initiated within 72 hours of ICU admission, following achievement of hemodynamic and respiratory stability (typically within 24-48 h) [11, 12]. It commonly includes early physical activity [13,14], neurocognitive stimulation [15,16], psychological intervention [17], and multimodal programs [18]. Existing meta-analyses confirm that routine early rehabilitation started within 72 hours may improve physical and cognitive function, help prevent PICS without increasing adverse events [19], and reduce the incidence of ICU-acquired weakness [20]. However, no differences have been found in cognitive-related delirium-free days or Hospital Anxiety and Depression Scale scores. Another meta-analysis on nurse-led early rehabilitation showed it significantly shortens ICU length of stay, although effects on functional outcomes, such as mobility, muscle strength, and mortality, were not significant

[21]. These inconsistencies may relate to factors including the time-intensive nature of current measures, insufficient standardization in implementation, and low patient participation and adherence.

Virtual reality (VR), a term first introduced by computer scientist Jaron Lanier in 1989, refers to a simulation technology providing users with an immersive environmental experience [22]. VR technology integrates computer, electronic information, and simulation technologies to enable users to interact with the VR system and experience an immersive, tactile, visual, and multisensory environment via VR helmets, all-in-one devices, and other gadgets [23]. The system is known for its novelty, interactivity, and immersive qualities [24]. VR technology–based rehabilitation techniques have been widely adopted in clinical rehabilitation. There is evidence suggesting that these techniques can effectively improve cognitive function in patients with stroke [25], reduce pain levels in patients with chronic low back pain and neck pain [26,27], and alleviate anxiety and depression [28,29].

Virtual reality–based early rehabilitation intervention (VR-ERI) is an emerging ICU recovery strategy that is gaining attention and has shown preliminary application results. Using VR technology, it provides ICU patients with specific virtual environments, such as natural landscapes [30], meditation music [31], or virtual tasks (such as memory games [32]), and cognitive exercises like "catching coconuts" and "avoiding obstacles" [33]. This aims to enhance rehabilitation motivation, participation, and physiological and psychological function. Its interactive, immersive, and feedback-driven design makes VR-ERI a promising tool for personalized early rehabilitation in critically ill patients. However, current research findings are inconsistent. Some studies report that VR-ERI effectively reduces anxiety [32], while others show no significant difference [34,35]. Such discrepancies may affect the certainty and consistency of clinical decision-making, potentially hindering the technology's wider adoption in ICU early rehabilitation.

To our knowledge, researchers [36] have explored the interventional effects of VR therapy on patients' PICS through a meta-analysis, but its included population did not specifically focus on early rehabilitation strategies during the ICU period. This analysis pooled data from patients receiving VR therapy both within the ICU and during follow-up after discharge. Due to differences between these 2 populations in aspects such as intervention timing and

environment, combining them in the analysis may introduce clinical heterogeneity, thereby weakening the credibility of the results. Another meta-analysis indicated that compared with traditional rehabilitation, VR-based early motor rehabilitation can improve balance, functional walking, and upper and lower limb motor function in critically ill patients [37]. However, this review only assessed motor function and did not comprehensively examine the effects on other early rehabilitation outcome indicators.

Objectives

This systematic review and meta-analysis aim to (1) characterize the types and content of VR-ERI activities and analyze the feasibility and safety of VR technology implementation in critically ill patients; (2) systematically evaluate the effectiveness of VR-ERI on physiological, psychological, and cognitive outcomes during ICU treatment and at short-term follow-up (within 3 months post-ICU discharge); and (3) provide recommendations for future research.

Methods

Study Protocol and Registration

The review protocol was registered in PROSPERO. This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [38] to ensure transparency (Checklists 1-3).

Deviations From the PROSPERO Protocol

During implementation, the search language was limited to Chinese and English to ensure assessment accuracy, and the search cutoff date was extended to October 5, 2025, to incorporate the latest evidence. Regarding statistical analysis, due to heterogeneity among the included studies, all analyses adopted the more conservative random-effects model. To enhance statistical robustness, the analysis was switched to the R language (version 4.4.1; R Core Team), and the Hartung-Knapp-Sidik-Jonkman method was applied to correct random-effects outcomes. This method provides more conservative CIs when the number of studies is small, reducing the risk of false-positive results in the presence of high heterogeneity. Apart from the above adjustments, all other procedures remained consistent with the original protocol.

Data Sources and Search Strategy

Two independent researchers (FW and YW) conducted literature searches across 6 English-language and 4 Chinese-language databases. MEDLINE and Embase were searched simultaneously via the Ovid platform. The other English databases were searched on their respective platforms—PubMed (via National Center for Biotechnology Information), Web of Science (via Clarivate Analytics), the Cochrane Library, and PsycINFO (via ProQuest). The Chinese databases searched were China National Knowledge Infrastructure (CNKI), WanFang Data, VIP Database (VIP),

and the SinoMed Service System (SinoMed). The search strategy combined Medical Subject Headings with free-text terms using Boolean operators. The search was performed in 2 stages. An initial search covered records from the inception of each database until May 11, 2025, and to incorporate the most recent evidence, the full search strategy was rerun on October 5, 2025, extending the cutoff date accordingly. All searches were limited to Chinese and English publications to align with the team's language capabilities and ensure feasibility.

All search strategies were independently designed based on the research question and were peer-reviewed and refined by experts in evidence-based medicine and critical care medicine in accordance with the PRISMA-S (Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension) checklist [39]. No published search filters were used, nor were search strategies from other systematic reviews adapted or reused. To cover the literature as comprehensively as possible, we also used the snowballing method to trace references and citations and attempted to obtain full-text articles by contacting librarians or the original authors. Furthermore, as this study focused on published efficacy evidence, we did not specifically search clinical trial registries or gray literature (including conference proceedings). The complete search strategies and procedures are detailed in Multimedia Appendix 1.

Study Selection

The rationale for formulating inclusion and exclusion criteria was based on the Population, Intervention, Comparison, Outcome, and Study Design (PICOS) framework, (1) population: adults aged ≥18 years receiving treatment during their ICU stay; (2) intervention: VR-ERI; (3) control: no treatment, usual care measures, or non–VR-based interventions; (4) outcomes: studies reporting at least 1 physical, psychological, or cognitive function outcome (such as cognition, pain, delirium, anxiety, depression, and sleep quality) or feasibility and safety outcomes during ICU treatment or after ICU transfer; and (5) study design: randomized controlled trials (RCTs) published in Chinese or English. The exclusion criteria were (1) studies that did not report original outcome measures and for which relevant outcome data could not be obtained after contacting the authors; (2) duplicate publications; and (3) case reports, letters, conference abstracts, and review articles.

Screening

In total, 2 independent researchers (FW and YW) conducted the literature screening and data extraction, followed by cross-checking. Any uncertainties were resolved by WC. First, duplicate records were automatically identified and removed using the built-in function of the EndNote X9 (Clarivate) reference management software, based on fields such as title, author, publication year, and journal name. Subsequently, the 2 researchers (FW and YW) independently examined the titles, authors, and other key information of the remaining records to further exclude any duplicates not detected by the software. Next, titles and abstracts were reviewed to exclude obviously irrelevant literature. Finally,

full texts were retrieved and assessed against the eligibility criteria to determine the final set of included studies.

Data Extraction

A total of 2 independent researchers (FW and YW) performed data extraction using a uniformly developed standardized form without using automated tools. The data that were systematically extracted were (1) general information, including first author, country, and publication year; (2) participant characteristics, including disease type, sample size, and baseline demographic and clinical characteristics; and (3) intervention characteristics, including specific type of VR, content of the VR-based intervention, content of the control group intervention, intervention duration, intervention frequency, outcome measures, measurement tools, time points of data collection and raw data (such as mean, SD, median, and IQR), for outcomes reported at baseline, post intervention, or follow-up for both groups.

To manage multiple data points objectively and consistently, we predefined selection criteria following the Cochrane Handbook [40]. The core principle was to prioritize data that best aligned with the review's primary objective, had the highest validity, or greatest clinical representativeness. For multiple measures, we selected the tool that best matched the outcome definition or had the highest validity; for multiple intervention groups, we included the primary intervention group, the most clinically representative, or the largest in sample size; for multiple control groups, we chose the "most representative usual care" group. Where data were missing, unclear, or incomplete, we contacted corresponding authors by email for clarification or missing data. If no response was received after 2 attempts, we used recommended formulas to estimate means and SDs from available data (eg, based on sample size, median, and range) [41].

Risk-of-Bias Assessment

The second author (YW) and the third author (YX) independently assessed the risk of bias of the included studies using the revised Cochrane Risk of Bias tool (RoB 2) for randomized trials [42]. The RoB 2 tool assesses five domains for randomized trials, which are (1) bias arising from the randomization process, (2) bias due to deviations from the intended interventions, (3) bias due to missing outcome data, (4) bias in the measurement of the outcome, and (5) bias in the selection of the reported result. Each domain is judged against a series of signaling questions. The overall risk of bias for a study is categorized as "low," "some concerns," or "high." A study is judged to have a low overall risk of bias if all domains are rated as "low." An overall rating of "some concerns" is given if at least 1 domain is rated as having "some concerns" and no domain is rated as "high." An overall "high" risk of bias is judged if any single domain is rated as "high" or if multiple domains are rated as having "some concerns."

Data Synthesis and Analysis

To ensure transparency in the data synthesis process, we followed a systematic decision-making procedure. First, all studies had to meet the predefined PICOS criteria and provide the raw data for the relevant outcome measures. Second, we established a rule that each study would contribute only a single data point for a given outcome. If a study contained multiple measurements or multiple group comparisons, a single dataset was selected for pooling according to the prespecified criteria, aiming to minimize heterogeneity as much as possible.

All analyses were performed using the R language (version 4.4.1) within the R Studio integrated development environment (version 2023.12). The metafor package was used for effect size pooling and forest plot generation [43], providing visual representations of each study's effect size with its 95% CI, along with the pooled overall effect size. All included outcome measures were continuous variables. We conducted a meta-analysis using change scores (from baseline to postintervention or follow-up) in means and SDs for both groups. For studies reporting only pre- and post-intervention means and SDs, the mean and SD of change scores were calculated as $\text{Mean}_{change} = \text{Mean}_{post} - \text{Mean}_{pre}$ and $\text{SD}_{change} = \sqrt{\text{SD}_{pre}^2 + \text{SD}_{post}^2 - 2 \times r \times \text{SD}_{pre} \times \text{SD}_{post}}$, where r represents the correlation coefficient between pre- and postintervention values. When correlations were not reported in the original publications, they were calculated according to the Cochrane Handbook for Systematic Reviews of Interventions. To harmonize different assessment tools and ensure consistent direction of results, we reversed the scoring for subjective sleep quality assessed by the Pittsburgh Sleep Quality Index and cognitive scores assessed by the Perceived Deficits Questionnaire for Depression, following Cochrane Handbook guidelines [40]. Consequently, higher scores in the pooled results indicate better sleep quality and cognitive function in patients. Additionally, we categorized outcome measures by assessment time points, specifically divided into (1) during ICU treatment and (2) short-term follow-up (2 wk to 3 mo after ICU discharge).

Given the variability in assessment tools and the presence of clinical and methodological heterogeneity among the included studies, the standardized mean difference (SMD) was adopted as the effect size measure for meta-analysis outcomes. The adjusted Hedges g was used to estimate the effect sizes [44,45]. Absolute SMD values of 0.2, 0.5, and 0.8 were interpreted as small, medium, and large effect sizes, respectively [46]. A random-effects model was applied to derive more conservative effect estimates. To enhance the robustness of statistical inference, particularly in reducing the risk of false-positive results when the number of studies was limited, the Hartung-Knapp-Sidik-Jonkman adjustment was used during model fitting to estimate the pooled effect size and its 95% CI [47].

To quantify between-study heterogeneity and enhance the generalizability of the results in real-world settings, we used a comprehensive assessment strategy. First, absolute heterogeneity was quantified using τ^2 (between-study variance) and its square root τ (between-study SD) [48]. τ^2 provides an absolute measure of the variation between studies, while τ, being on the same scale as the effect size, facilitates clinical interpretation. Second, the Cochran Q test was used for

the hypothesis testing of effect size homogeneity, supplemented by the I^2 statistic to quantify the relative extent of heterogeneity [49], with thresholds of 25%, 50%, and 75% used to define low, moderate, and high levels, respectively [50]. Finally, for outcome measures involving at least 3 studies, 95% prediction intervals (PIs) were calculated [51] to estimate the potential range of the true effect size in future similar interventions, thereby providing more clinically informative guidance for practice.

Subgroup analyses were performed to explore potential sources of heterogeneity. A leave-one-out sensitivity analysis was conducted to assess the robustness of the pooled results. For outcome measures with fewer than 2 included studies, descriptive analyses were performed. Additionally, to assess the potential for small-study effects (such as publication bias), we performed several checks for the anxiety outcome, which had the largest number of included studies (n=8). These included a visual inspection of funnel plot symmetry, the Egger test (regression test), and the Begg test (rank correlation test). Furthermore, the trim-and-fill method was used to estimate the potential number of missing studies and to compute an adjusted effect size, thereby evaluating the potential impact of small-study effects on the pooled result. The threshold for statistical significance was set at $P \le .05$.

Evidence Quality

To evaluate the reliability of the evidence, we used the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework [52] to rate the certainty of evidence for each outcome. In addition, 2 independent raters (FW and YW) assessed the certainty of evidence using the official GRADE pro GDT software (Evidence Prime Inc), with any disagreements resolved by the corresponding author (RZ). The evidence was rated as very low, low, moderate, or high. Footnotes were provided to explain the reasons for any downgrading of the evidence, which included considerations of risk of bias, inconsistency, indirectness, imprecision, and publication bias.

Ethical Considerations

This study used literature data and did not require approval from an ethics review committee or patient consent.

Results

Search Results

A total of 3737 studies were identified from 10 databases. After removing 243 duplicate articles, the titles and abstracts of the studies were initially screened, and 58 full-text manuscripts relevant to this study were identified. Following full-text reading and quality assessment, 16 RCTs [16,17,30,32-35,53-61] were ultimately included for analysis. The specific search and inclusion process is illustrated in Figure 1, and the excluded articles after rescreening, along with the reasons, are detailed in Multimedia Appendix 2.

Figure 1. Flowchart of literature screening. CNKI: China National Knowledge Infrastructure; SinoMed: SinoMed Service System; VIP: VIP Database.
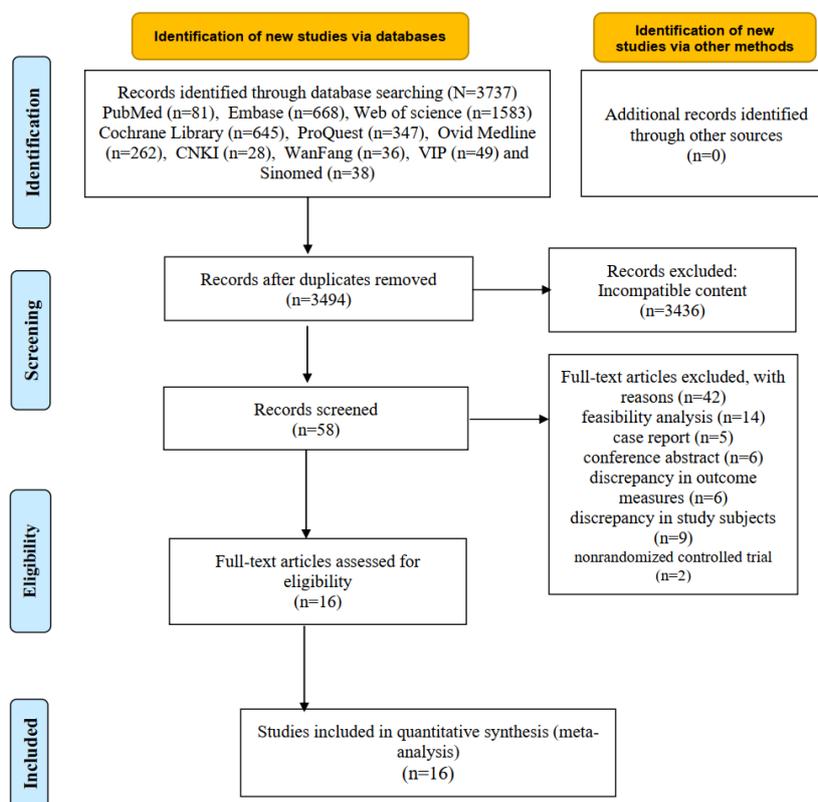
Characteristics of the Studies

Multimedia Appendix 3 summarizes the basic characteristics of the included studies. The 16 studies (published between 2020 and 2025) comprised a total of 1356 participants, with sample sizes ranging from 42 to 180. The studies originated from Brazil [53], Korea [30,54], France [17,35], Spain [16], Turkey [34], and China [32,33,55-61]. Patient conditions included cardiovascular diseases [30,35,53,55,60], postcraniotomy [33], postliver transplantation [61], postlaparoscopic sleeve gastrectomy [34], conventional oncologic surgery [32], severe respiratory diseases [58], and postmultiple fracture surgery [56], among others. The mean age of patients was generally between 60 and 70 years. Interventions were initiated within 24 hours after ICU admission in 3 studies [35,53,59], within 6 hours in 2 studies [32,34], and on the day of ICU admission once the patient's condition was stabilized in the remaining 11 studies [16,17,30,33,54-58,60, 61].

One study [16] used nonimmersive VR technology, while 15 studies [17,30,32-35,53-61] used fully immersive VR. The intervention directions included (1) early cognitive rehabilitation (training reaction and executive function through games or tasks) [16,33,55,57], (2) early motor rehabilitation (engaging in virtual activities, such as fruit cutting or cycling combined with motion capture) [53,56,58,59,61], (3) early delirium prevention (providing emotional support and cognitive training for multidimensional prevention) [32], (4) VR relaxation therapy (conducting physical and mental relaxation in natural scenes) [17,34,35], (5) VR meditation (improving sleep and mood by combining mindfulness) [30, 54], and (6) VR preoperative visit (structured introduction to the ICU environment, treatment process, and rehabilitation) [60]. The intervention protocols were categorized into (1) single-session short-term intervention (15-30 min at key clinical time points) [17,30,34,35,60], (2) high-frequency short-term intervention during ICU stay (approximately 20 min per session, 1-3 times daily, lasting 3-7 d) [16,32,54, 55,57-59], (3) medium- to long-term structured intervention (30-40 min per session, 5 d per wk, lasting 2 wk to 3 mo)

[33,56,61], and (4) flexible individualized protocol (adjusted based on patient tolerance) [53]. The studies involved a total of 24 outcome measures, such as cognitive function, delirium, anxiety, depression, pain, sleep, and quality of life, although the assessment tools and observation windows varied across studies.

Risk of Bias

The overall risk of bias for the studies included in this review was high, with summary results presented in Figures 2 and 3, and detailed assessment results available in Multimedia Appendix 4. Furthermore, 1 study [58] was rated as "low," 5 studies [17,32,34,54,55] raised "some concerns," and 10 [16,30,33,35,53,56,57,59-61] were rated as "high." In domain 1 (randomization process), 7 studies [33-35,53,56, 60,61] did not describe allocation concealment, and 1 study [32] described neither the method for generating the random sequence nor allocation concealment; all were rated as "some concerns." For domain 2 (deviations from intended interventions), 4 studies [16,32,55,57] were rated as "some concerns" due to lack of blinding of patients and implementers; 1 study [59] was rated as "high" because it lacked blinding and did not adhere to the intention-to-treat analysis principle. In domain 3 (missing outcome data), 2 studies [16,59] were rated as "high" due to a high rate of loss to follow-up and because the reasons for missing data were potentially related to health status. For domain 4 (measurement of the outcome), 11 studies [16,17,30,32,33,35,53,56,57,60,61] had issues with the lack of blinding of outcome assessors. Among these, 9 studies [16,30,33,35,53,56,57,60,61] were rated as "high" because they measured subjective patient-reported outcomes (eg, pain and anxiety), and 2 studies [17,32] were rated as "some concerns" as their outcome measures were relatively objective. In domain 5 (selection of the reported result), transparency was generally insufficient. Moreover, 4 studies [16,17,35,55] mentioned trial registration but lacked verifiable prespecified analysis descriptions for the reported results, while the remaining 11 studies [30,32-34,53,54,56, 57,59-61] did not mention a preregistered protocol; consequently, all were rated as "some concerns."

Figure 2. Risk of bias graph for proportion of judgements per item in included studies.

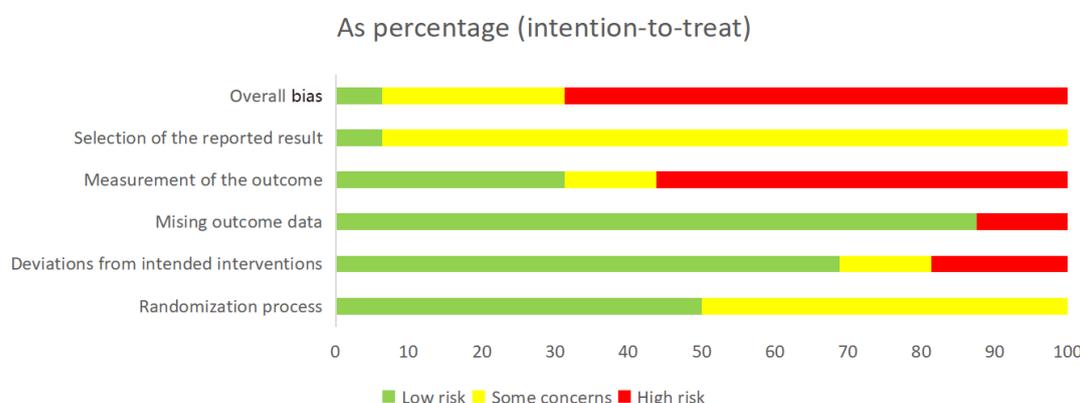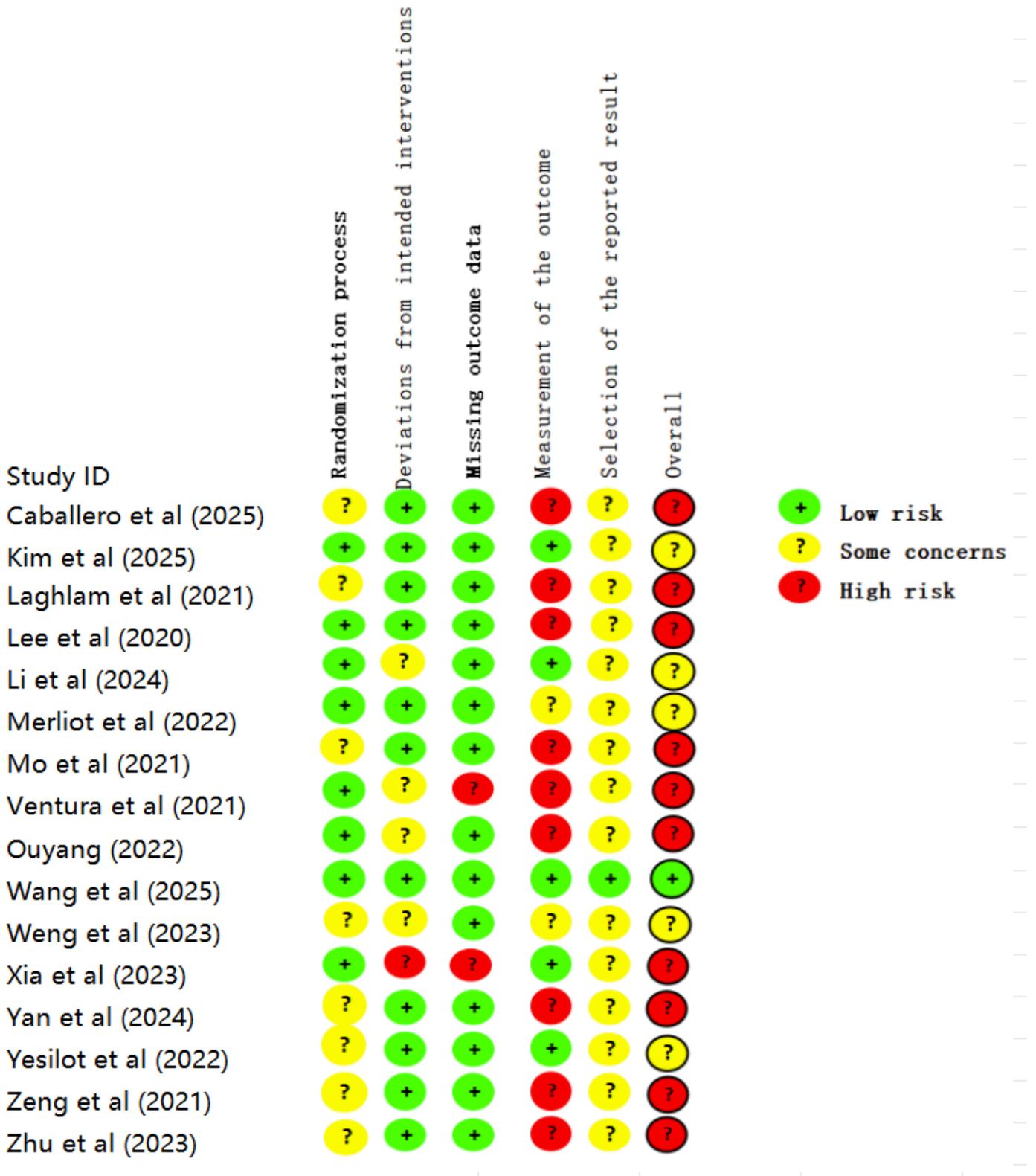As percentage (intention-to-treat)

Figure 3. Risk of bias summary Items assessed in each included study [16,17,30,32-35,53-61].

Clinical Safety and Feasibility

In total, 8 studies [16,17,35,53,55,57-59] reported on safety, 5 studies [16,55,57-59] confirmed the absence of serious adverse events but documented related reseactions including dizziness [16,35,57], nausea [35,57], fatigue [16,17], anxiety [16,17], confusion [16,55], visual disturbance [17], headache or discomfort from wearing the headset [55], claustrophobia,

agitated withdrawal, and dyspnea due to device displacement [17]. Furthermore, 1 study reported a significant drop in blood pressure during intervention in a patient [59], another reported a case of excessive drowsiness [16], and 1 noted that vital signs in all patients fluctuated within the normal range (±10%) during the intervention [58]. Regarding feasibility, 1 study reported a 78.3% patient recommendation rate and a 90% "good/excellent" experience rating [53]. Meanwhile, 2

studies reported that patients in the VR group demonstrated significantly better outcomes in terms of daily total exercise time, average daily out-of-bed activity time [58], and exercise adherence (with a high adherence rate of 71.74% vs 54.35% at 3 months) [56].

Results of the Meta-Analysis

Among the included trials, due to the limited number of studies, we conducted meta-analyses on outcome indicators with at least 2 studies. The specific analysis results are as follows.

Anxiety

A total of 8 studies (n=801) included anxiety as an outcome measure. According to the predefined data pooling criteria, for the study by Li et al [55], we only pooled the data assessed during ICU treatment using the State-Trait Anxiety Inventory scale. For the study by Merliot-Gailhoustet et al [17], we only pooled the data from the group receiving the DEEPSEN VR system (Saint-Didier-au-Mont-d'Or, France). All studies were analyzed using change-from-baseline values for the pooling.

The meta-analysis revealed that, in terms of the average effect, VR-ERI did not show a statistically significant effect on reducing patient anxiety (SMD –0.86, 95% CI –1.85 to 0.13). Given that higher scores on all assessment tools indicated greater anxiety, the negative point estimate suggests that VR-ERI may reduce anxiety. However, the 95% PI was wide and included the null value (−3.75 to 2.03), indicating considerable uncertainty about the intervention's effect in future settings, ranging from a substantial benefit to potential harm. Therefore, these findings require validation in further studies.

Extremely high heterogeneity was detected among the studies (I^2=95.4%, τ=1.1480, τ^2=1.3179; Cochran Q=150.82, df=7; P<.01; Figure 4). Exploratory subgroup analyses were performed based on VR intervention content, timing, patient population, and type of anxiety assessment tool, but no potential source of heterogeneity was identified. Sensitivity analysis (leave-one-out method) indicated that the pooled effect size was sensitive to the removal of individual studies, with a maximum change of 0.3478, exceeding the prespecified threshold of 0.1, suggesting that the results may not be robust. The study by Weng et al [32] had the most substantial influence on the overall effect estimate. After excluding this study, the pooled effect size remained statistically nonsignificant (SMD –0.51, 95% CI –1.17 to 0.15; 95% PI –2.31 to 1.28), and heterogeneity remained high (I^2=92.5%, τ=0.6806, τ^2=0.4632; Cochran Q=132.89, df=6; P<.01), further supporting the instability of the findings.

Figure 4. Forest plot of the meta-analysis for anxiety outcomes [16,17,32,34,35,55,57,60].

Study	Experimental N	Mean (SD)	Control N	Mean (SD)	Hedge g	SMD (95% CI)	Weight
Weng et al (2023)	42	-46.3 (5.70)	42	-27.16 (5.61)		-3.35(-4.03 to -2.68)	12%
Li et al (2024)	70	-2.82 (2.71)	78	1.38 (2.29)		-1.67(-2.05 to -1.30)	12.7%
Ouyang (2022)	23	-3 (3.70)	24	4 (6.37)		-1.31(-1.95 to -0.68)	12.1%
Yesilot et al (2022)	55	-1.53 (1.15)	55	-1.06 (1.67)		-0.33(-0.70 to 0.05)	12.7%
Yan et al (2024)	39	-3 (6.59)	41	-1 (7.30)		-0.28(-0.73 to 0.16)	12.6%
Merliot et al (2022)	56	-1.3 (3.39)	54	-0.5 (3.39)		-0.23(-0.61 to 0.14)	12.7%
Ventura et al (2021)	21	3 (4.56)	21	4 (4.16)		-0.22(-0.83 to 0.38)	12.2%
Laghlam et al (2021)	90	-2 (2.67)	90	-3 (2.67)		0.37(0.08 to 0.67)	12.9%
Random effects model	**396**		**405**			**-0.86(-1.85 to 0.13)**	**100%**
Prediction interval						**[-3.75; 2.03]**	

Knapp-Hartung adjustment applied

Heterogeneity: I^2 = 95.4%, τ = 1.1480, χ^2_7 = 150.82 (P < .01)
Test for overall effect: t_7 = -2.06 (P = .08)

Favors experimental   Favors control

According to the GRADE framework, the certainty of evidence for anxiety was rated as "very low." The downgrading was based on (1) serious risk of bias in the included studies (4 studies [16,35,57,60] with high risk), (2) very high heterogeneity among studies, (3) imprecision of the pooled effect estimate (the 95% CI spanned both the line of null effect and the threshold for clinically meaningful benefit), and (4) potential publication bias. Although the pooled effect size suggested a potential moderate-to-large effect, an upgrade was not considered after weighing the above limitations (detailed assessment is available in Multimedia Appendix 5).

Subjective Sleep Quality (Scale Assessment)

In total, 6 studies [30,32,54,55,57,60] (n=500) assessed subjective sleep quality during ICU stay. In accordance with predefined pooling criteria, we extracted data from the second intervention day for Kim and Kang [54] and the Richards-Campbell Sleep Questionnaire (RCSQ)–measured sleep data only for Ouyang [57]. While all studies reported comparable baseline characteristics, specific data were provided by only 3 [32,55,60]. To maximize data utility, the meta-analysis was performed using postintervention outcomes.

Meta-analysis revealed considerable heterogeneity (I^2=97.5%, τ=2.4143, τ^2=5.8288; Cochran Q=198.44, df=5; P<.01), likely due to differences in populations, interventions,

and assessments. Sensitivity analysis (leave-one-out method) confirmed the robustness of the results, as the maximum change in effect size after excluding any single study was 0.0347 (below the prespecified 0.1 threshold), with no specific source of heterogeneity identified. The pooled results (Figure 5) indicated a significant positive effect of VR-ERI on improving patients' subjective sleep quality compared

with usual care (SMD 3.36, 95% CI 0.77-5.94). For consistent interpretation, Pittsburgh Sleep Quality Index scores were reversed so that higher scores uniformly indicated better sleep quality across all scales; thus, the positive SMD reflects an enhancing effect of VR-ERI, although more studies are needed to confirm the precision of this estimate.

Figure 5. Forest plot of the subgroup analysis for subjective sleep quality [30,32,54,55,57,60].
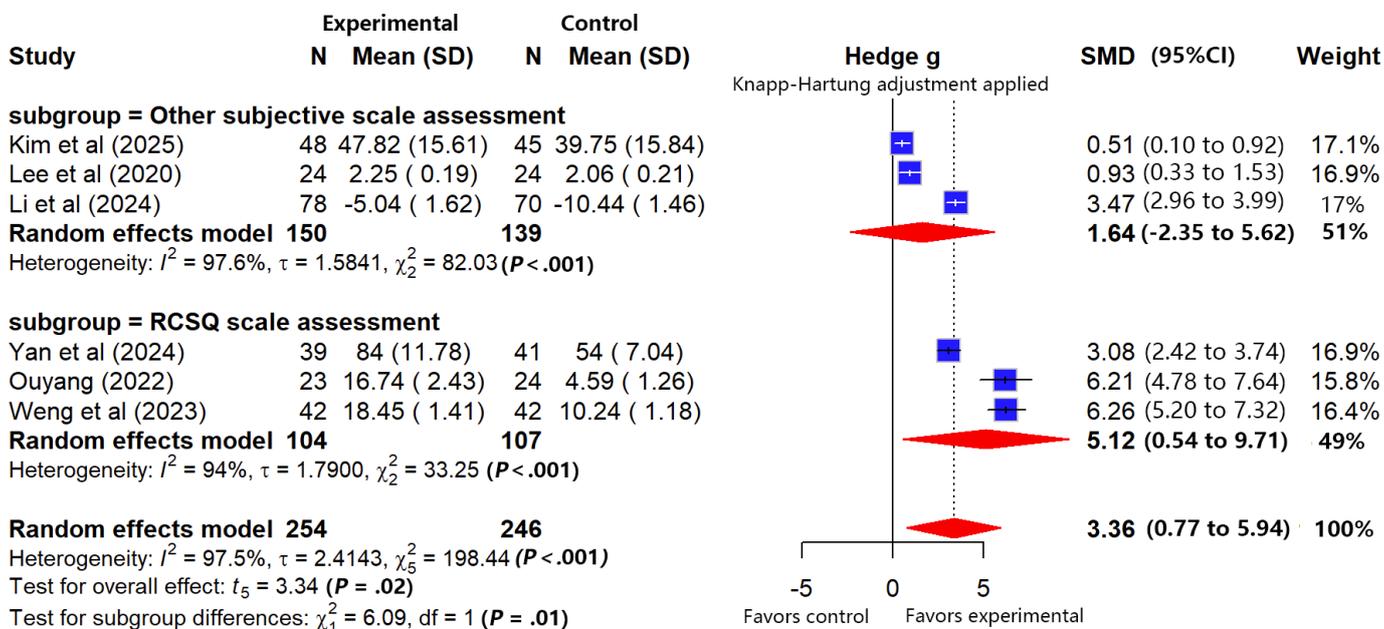
Subgroup analysis was conducted after unifying scale directions to assess how the inherent characteristics of different instruments might influence the pooled result. The analysis showed that VR-ERI did not significantly improve sleep when assessed with generic scales (SMD 1.64, 95% CI –2.35 to 5.61). In contrast, a positive effect was observed with the ICU-specific RCSQ (SMD 5.12, 95% CI 0.54 to 9.71). The between-subgroup difference was statistically significant (χ^2_1=6.1; *P*=.014), indicating that the type of assessment tool is a significant effect modifier, with the RCSQ potentially being more sensitive to VR-ERI–related sleep improvements in the ICU setting.

According to the GRADE approach, the certainty of evidence for subjective sleep quality during ICU treatment was rated as "very low." The reasons for downgrading included (1) a serious risk of bias in the included studies (3 studies [30,57,60] with high risk), (2) very high heterogeneity among studies, and (3) extreme imprecision of the pooled effect estimate (the 95% CI was extremely wide, spanning the range from moderate to very large effects). Considering all the downgrading factors above, an upgrade was not applied despite the large pooled effect size (for detailed assessment, refer to Multimedia Appendix 5).

Objective Sleep Quality (Smart Wristband Monitoring)

In total, 2 studies [30,54] (n=141) monitored objective sleep quality using smart wristbands. Key metrics included total

sleep time (TST), defined as the actual duration of sleep from onset to awakening; wake after sleep onset (WASO), representing total wake time during the sleep period as an indicator of sleep fragmentation; and sleep efficiency, calculated as the percentage of TST relative to total time in bed, with higher values indicating better sleep quality. A pooled analysis was performed using change-from-baseline values, with specific results as follows.

TST

The meta-analysis indicated no heterogeneity among the studies (I^2=0%, τ=0, τ^2=0; Cochran Q=0, df=1; *P*=.95). Sensitivity analysis (leave-one-out method) indicated that the results were robust. The maximum change in the effect size after excluding any single study was 0.0375, which is below the prespecified threshold of 0.1. The meta-analysis (Figure 6) showed that, regarding the average effect, VR-ERI did not lead to a statistically significant improvement in patients' TST compared with usual care (SMD 0.08, 95% CI –0.07 to 0.22). The positive SMD suggests that the direction of the effect may favor VR-ERI, indicating an increase in TST. According to the GRADE approach, the certainty of evidence for objective TST during ICU treatment was rated as "moderate." The rating was downgraded due to a high risk of bias in 1 study [30]. No downgrading was applied for other domains, as there was no heterogeneity among studies and the pooled effect estimate was relatively precise (for detailed assessment, refer to Multimedia Appendix 5).

Figure 6. Forest plot of the pooled analysis for objective sleep quality [30,54].
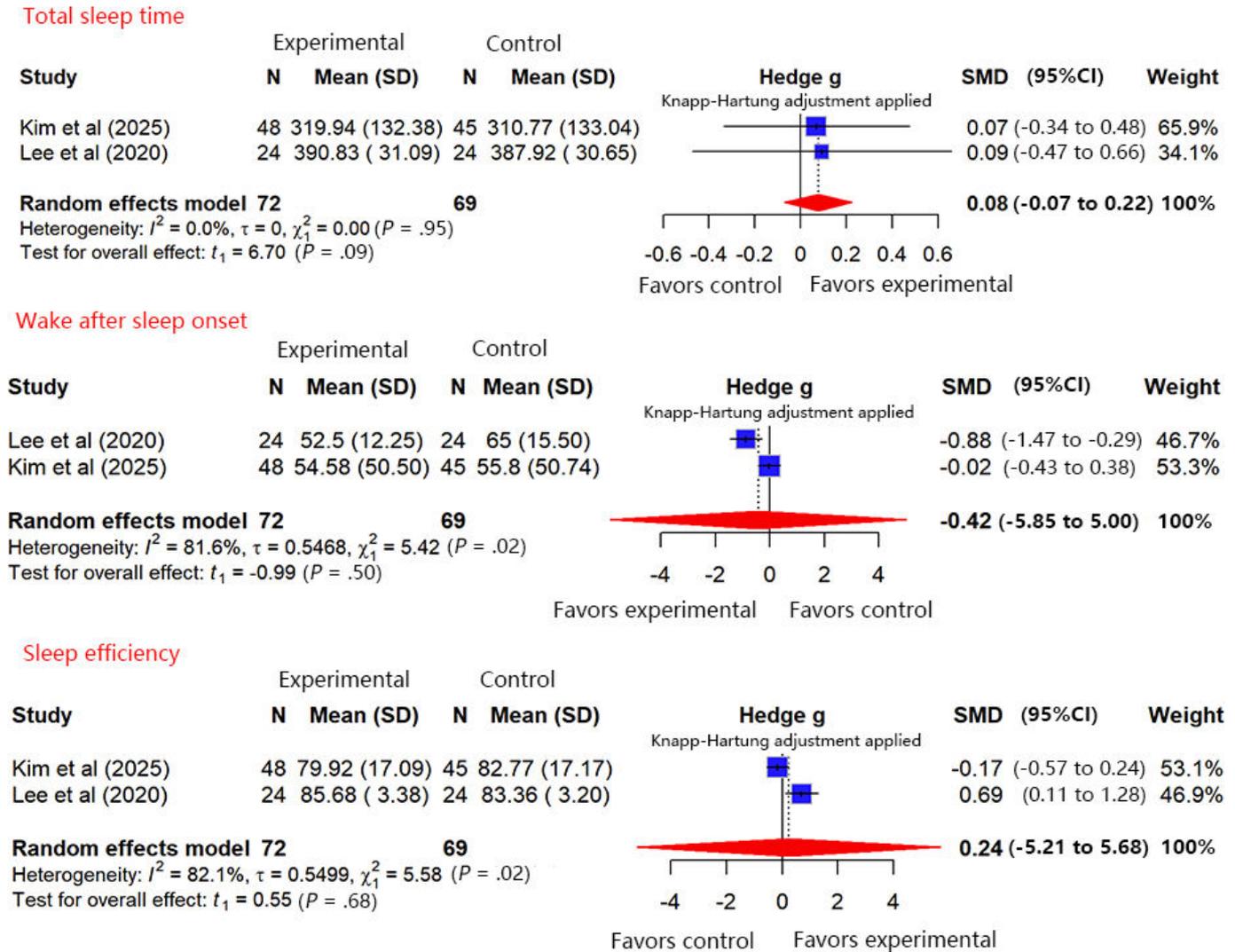
WASO

The meta-analysis indicated substantial heterogeneity among the studies (I^2=81.6%, τ=0.5468, τ^2=0.2990; Cochran Q=5.42, df=1; P=.02). Sensitivity analysis showed that the results were robust (the maximum change in the effect size after excluding any single study was 0.0397, below the threshold of 0.1). The meta-analysis (Figure 6) showed that, regarding the average effect, VR-ERI did not lead to a statistically significant reduction in WASO compared with usual care (SMD −0.42, 95% CI −5.85 to 5.00). The negative SMD indicates that the direction of the effect points toward a reduction in WASO, signifying an improvement in sleep maintenance. According to the GRADE approach, the certainty of evidence for objective WASO during ICU treatment was rated as "very low." The rating was downgraded due to (1) a very limited number of included studies with risk of bias, (2) high heterogeneity among studies, and (3) extreme imprecision of the pooled effect estimate (the 95% CI was exceptionally wide; for detailed assessment, refer to Multimedia Appendix 5).

Sleep Efficiency

The meta-analysis indicated substantial heterogeneity among the studies (I^2=82.1%, τ=0.5499, τ^2=0.3024; Cochran Q=5.58, df=1; P=.02). Sensitivity analysis indicated robust results (the maximum change in the effect size after excluding any single study was 0.0432, below the threshold of 0.1). Regarding the average effect, the meta-analysis (Figure 6) showed that VR-ERI did not lead to a statistically significant improvement in sleep efficiency compared with usual care (SMD 0.24, 95% CI −5.21 to 5.68). The positive SMD indicates that the direction of the effect points toward an improvement in sleep efficiency. According to the GRADE approach, the certainty of evidence for objective sleep efficiency during ICU treatment was rated as "very low." The rating was downgraded due to (1) risk of bias in the included studies, (2) substantial heterogeneity among studies, and (3) extreme imprecision of the pooled effect estimate (refer to Multimedia Appendix 5 for detailed assessment).

Pain

In total, 3 studies included pain level during ICU treatment as an outcome measure. According to the predefined data pooling criteria, for the study by Merliot-Gailhoustet

et al [17], we only included the pain data from patients who received the DEEPSEN VR system. The pain outcome measure was analyzed using change-from-baseline values for the pooled analysis.

Initial meta-analysis revealed extremely high heterogeneity among the studies (I^2=95.1%, τ=0.7550, τ^2=0.5700; Cochran Q=41.10, df=2; P<.01). Regarding the average effect, VR-ERI did not show a statistically significant difference in reducing patients' ICU-related pain compared with usual care (SMD –0.13, 95% CI –2.06 to 1.79). Sensitivity analysis (leave-one-out method) indicated that the results may not be robust. The maximum change in the effect size after excluding a single study was 0.4062, which exceeds the prespecified threshold of 0.1. The study with the largest influence was by Laghlam et al [35]. Upon verification,

this study was a noninferiority trial comparing VR therapy with another analgesic intervention. Pooling its control group with standard usual care controls could confound the pain outcome. Therefore, the study by Laghlam et al [35] was excluded, and the pooled result from the remaining 2 studies [17,34] (n=220) was used as the primary evidence for assessing the effect of VR-ERI on pain during ICU treatment. The meta-analysis of the 2 remaining studies showed low heterogeneity (I^2=23.7%, τ=0.1086, τ^2=0.0118; Cochran Q=1.31, df=1; P=.25; Figure 7). Regarding the average effect, VR-ERI did not demonstrate a statistically significant effect on reducing pain during ICU treatment compared with usual care (SMD –0.58, 95% CI –2.58 to 1.43). Subgroup analysis was not performed due to the limited number of studies.

Figure 7. Forest plot for pain outcomes during the intensive care unit stay [17,34].
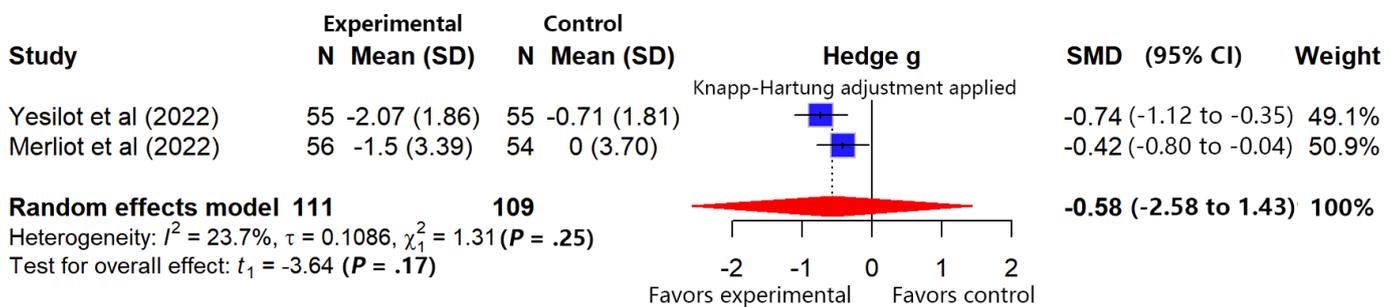
According to the GRADE criteria, the certainty of evidence for the outcome of pain during ICU treatment was rated as "low." The reasons for downgrading include (1) a high risk of bias in the included primary studies and (2) serious imprecision, as the 95% CI of the pooled effect is extremely wide and includes the line of null effect. As the heterogeneity among the remaining studies was low after the exclusion of 1 study based on clinical heterogeneity, no downgrading was applied for inconsistency (refer to Multimedia Appendix 5 for detailed assessment).

Balance

In total, 3 studies [33,56,61] (n=228) assessed patients' balance function using the Berg Balance Scale at short-term follow-up. We pooled change-from-baseline values for

analysis. No heterogeneity was observed among the included studies (I^2=0%, τ=0, τ^2=0; Cochran Q=0.29, df=2; P=.86). The meta-analysis showed that, in terms of the average effect, VR-ERI improved patients' balance function during short-term follow-up after discharge compared with usual care (SMD 0.97, 95% CI 0.74 to 1.20). Furthermore, the 95% PI (0.37-1.58) lies entirely to the right of the line of null effect, suggesting that the positive effect of VR-ERI on balance function is likely to remain at a clinically meaningful level in most future, similar studies (Figure 8). Sensitivity analysis (leave-one-out method) indicated that the results were robust. The maximum change in the effect size after excluding any single study was 0.0461, which is below the prespecified threshold of 0.1. Subgroup analysis was not performed due to the limited number of studies.

Figure 8. Forest plot for balance outcomes [33,56,61].
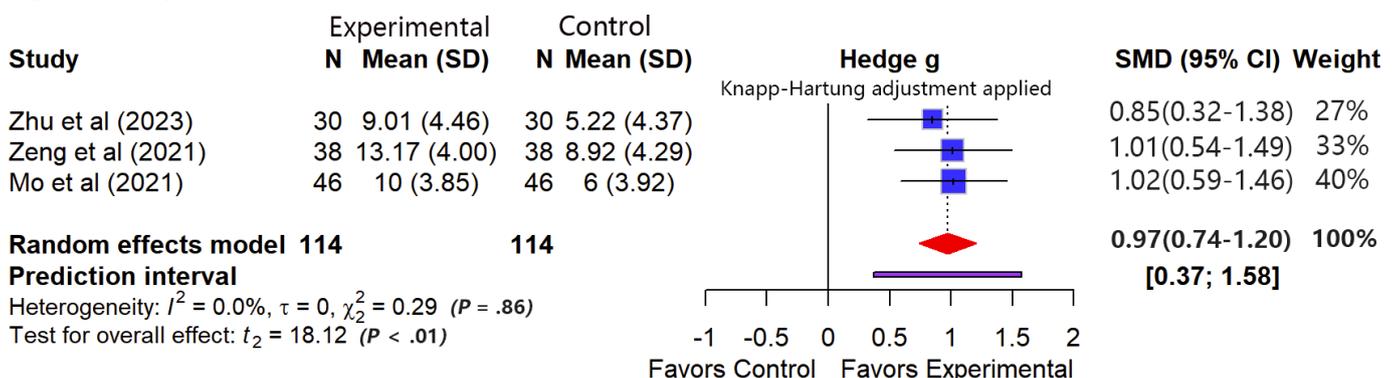
According to the GRADE approach, the certainty of evidence for balance function at short-term follow-up was rated as "moderate." The rating was downgraded because all included primary studies were assessed as having a high risk of bias. No downgrading was applied for inconsistency or imprecision, as no heterogeneity was observed among studies, and the pooled effect estimate was precise. Although the pooled effect size was large and clinically meaningful, an upgrade was not performed. This decision was made to maintain a cautious interpretation, considering that the methodological limitations commonly present in the existing studies could affect the validity of the effect estimate (refer to Multimedia Appendix 5 for the detailed evaluation process).

Cognitive Function

A total of 3 studies [33,55,57] (n=271) were included that assessed cognitive function at short-term follow-up. According to the predefined data pooling criteria, for the study by Zeng [33], we only included the results obtained using the Montreal Cognitive Assessment (MoCA) scale. All analyses were based on change-from-baseline values.

The meta-analysis (Figure 9) indicated mild heterogeneity among the studies (I^2=20.6%, τ=0.1498, τ^2=0.0224; Cochran Q=2.52, df=2; P=.28). Sensitivity analysis (leave-one-out method) showed that the results are robust, as the maximum change in the effect size after excluding any single study was 0.0218, which is below the threshold of 0.1. The pooled results indicated that, on average, VR-ERI showed a statistically significant trend toward improving patients' cognitive function compared with usual care (SMD 0.78, 95% CI 0.16 to 1.39). However, the 95% PI (–0.16 to 1.71) crossed the line of null effect, suggesting that the beneficial effect of VR-ERI on cognitive function in most future, similar studies is unstable—it could be beneficial or show no difference and therefore requires further exploration in future research. Subgroup analysis was not performed due to the limited number of included studies.

Figure 9. Forest plot of the meta-analysis for cognitive function outcomes [33,55,57].
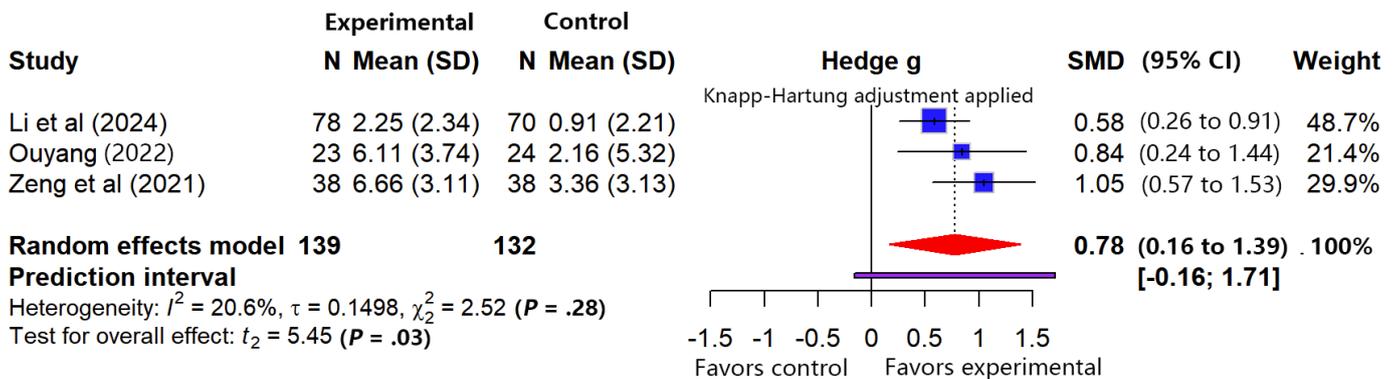
Study	Experimental		Control		Hedge g	SMD	(95% CI)	Weight
	N	Mean (SD)	N	Mean (SD)	Knapp-Hartung adjustment applied			
Li et al (2024)	78	2.25 (2.34)	70	0.91 (2.21)		0.58	(0.26 to 0.91)	48.7%
Ouyang (2022)	23	6.11 (3.74)	24	2.16 (5.32)		0.84	(0.24 to 1.44)	21.4%
Zeng et al (2021)	38	6.66 (3.11)	38	3.36 (3.13)		1.05	(0.57 to 1.53)	29.9%
Random effects model	**139**		**132**			**0.78**	**(0.16 to 1.39)**	**100%**
Prediction interval							**[-0.16; 1.71]**	

Heterogeneity: I^2 = 20.6%, τ = 0.1498, χ^2_2 = 2.52 (**P** = .28)
Test for overall effect: t_2 = 5.45 (**P** = .03)

According to the GRADE framework, the quality of evidence was rated as "low." This rating was based on (1) all 3 included studies [33,55,57] were rated as having a high risk of bias, and (2) imprecision of the pooled effect estimate. Although the pooled effect size suggests a moderate effect, the methodological flaws in the studies significantly reduce our confidence in this estimate. Specific rating details are provided in Multimedia Appendix 5.

Limb Motor Function

In total, 2 studies [33,56] (n=168) assessed the improvement of patients' limb motor function during short-term follow-up using the Fugl-Meyer Assessment. As the study by Zeng [33] reported upper limb and lower limb function scores separately, we applied the weighted mean method to combine the function scores from different body parts within the same study into an overall limb motor function score. All analyses were based on change-from-baseline values.

No heterogeneity was observed between the studies (I^2=0%, τ=0, τ^2=0; Cochran Q=0.55, df=1; P=.46). The meta-analysis (Figure 10) indicated that, based on the average effect, VR-ERI showed a trend toward improving patients' limb motor function at short-term follow-up after discharge compared to usual care (SMD 1.40, 95% CI –0.23 to 3.02), although the difference was not statistically significant and requires further verification. According to the GRADE framework, the quality of evidence was rated as "low." This rating was based on (1) both included studies had a high risk of bias, and (2) the pooled effect estimate was extremely imprecise. Although the pooled effect size suggests a potentially large effect, we did not upgrade the rating due to the very limited evidence base and the low quality of the studies (Multimedia Appendix 5).

Figure 10. Forest plot for limb motor function [33,56].

	Experimental		Control				
Study	N	Mean (SD)	N	Mean (SD)	Hedge g	SMD (95% CI)	Weight
					Knapp-Hartung adjustment applied		
Zeng et al (2021)	38	8.76 (3.17)	38	4.66 (3.27)		1.26 (0.77 to 1.75)	47%
Mo et al (2021)	46	6 (2.96)	46	2 (2.22)		1.52 (1.05 to 1.98)	53%
Random effects model	**84**		**84**			**1.40 (−0.23 to 3.02)**	**100%**

Heterogeneity: $I^2 = 0.0\%$, $\tau = 0$, $\chi_1^2 = 0.55$ ($P = .46$)
Test for overall effect: $t_1 = 10.93$ ($P = .06$)

Favors control   Favors experimental
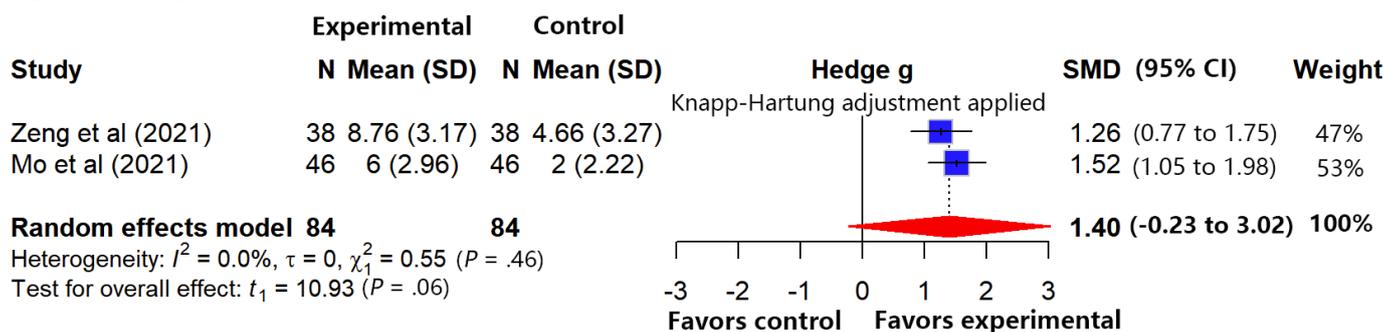
Descriptive Integration Results

Overview

We performed descriptive analysis on outcomes with fewer than 2 included studies and those unsuitable for meta-analysis due to differences in assessment methods or data collection time points. Given the limited number of studies, further validation of the intervention effects on these outcomes is currently lacking, and more high-quality RCTs are needed in the future to evaluate them.

Motor Function Outcomes (Muscle Strength, Grip Strength, Patient Engagement, and Functional Outcomes)

VR-ERI demonstrated multifaceted benefits on motor function, including improved muscle strength. In total, 2 studies supported this finding; Zhu et al [61] reported significantly higher Manual Muscle Testing scores in the VR-ERI group versus controls ($t_{58}=2.33$, $P=.02$), while Wang et al [58] observed stronger muscle strength via the Medical Research Council scale ($F_{1,66}=7.9$, $P=.006$), with a maintained intervention effect over time ($F_{7,66}=50.1$, $P<.001$). Due to insufficient raw data, a meta-analysis was not conducted, and further studies are needed to confirm these results. In contrast, findings regarding grip strength improvement were inconsistent. Both Xia et al [59] and Wang et al [58], using a dynamometer, reported no statistically significant between-group difference in absolute grip strength post intervention. Notably, Wang et al [58] observed significant improvement over time in both groups and a significant group-time interaction, suggesting VR-ERI may influence the recovery trajectory. However, its advantage over conventional rehabilitation requires further validation.

Furthermore, VR-ERI also enhanced patient engagement. Both Xia et al [59] and Wang et al [58] reported that the intervention group had significantly longer daily exercise duration than the control group. Mo et al [56] further supported this, finding a higher proportion of patients with high exercise compliance in the VR-ERI group (33/46, 71.74%) versus the control group (25/46, 54.35%) via a self-developed questionnaire. This may be attributed to the inherent enjoyment and real-time feedback of VR-based training, although limited raw data necessitates further research. Regarding functional outcomes, VR-ERI also demonstrated positive effects. Mo et al [56] observed a higher proportion of patients attaining the highest walking ability level (Functional Ambulation Category grade 5) in the intervention group (22/46, 47.93%) versus the control group (14/46, 30.43%). Meanwhile, Xia et al [59] reported a significant reduction in ICU-acquired weakness incidence with VR-ERI (intervention: 13.5% vs control: 35.5%). Future high-quality studies are needed to confirm these effects.

Psychological Function Outcomes (Depression, Posttraumatic Stress Disorder, Somatic Symptoms, and Perceived Well-Being)

The effects of VR-ERI on psychological function were inconsistent. Navarra-Ventura et al [16] found no significant between-group differences in depression or posttraumatic stress disorder at 1 and 12 months post-ICU discharge. In contrast, Ouyang [57] reported a significant reduction in depression during ICU treatment. However, due to differing assessment timepoints between the 2 studies, meta-analysis was not performed, and these divergent conclusions require further validation. Li et al [55] observed better improvement in somatic symptoms and higher perceived well-being in the intervention group at 1 week during ICU stay, but these differences were not maintained 1 month after discharge. Due to differing assessment timepoints, meta-analysis was not performed, and the long-term psychological effects of VR-ERI require further investigation.

Cognitive Function Outcomes (Delirium)

Kim and Kang [54] used the Confusion Assessment Method (CAM) for ICU scale to evaluate the effect of VR-ERI on delirium incidence and found no significant difference between the 2 groups (12.2% vs 12.8%, $\chi_1^2=0.01$; $P=.94$). In contrast, Weng et al [32] used the CAM scale to assess delirium severity and reported a statistically significant difference in CAM scores between groups (mean 2.25, SD 0.44 vs mean 6.37, SD 1.05, $t_{82}=23.453$; $P<.05$). Future high-quality studies with standardized assessment criteria are recommended to validate the intervention effects.

ICU Length of Stay and Recovery Quality

Regarding ICU length of stay, both Xia et al [59] and Wang et al [58] reported significantly shorter stays for the VR-ERI group versus controls, suggesting faster patient turnover. However, a meta-analysis was not feasible due to incomplete raw data. For quality of recovery, Zhu et al [61] (using the Postoperative Quality Recovery Scale) found significantly

better recovery rates in the VR-ERI group for physiological function and activity domains. Li et al [55] noted improved mental health scores (12-Item Short Form Health Survey Mental Component Score) at 1 week post intervention, though physical health scores (12-Item Short Form Health Survey Physical Component Score) showed no difference, and no between-group differences remained significant at the 3-month follow-up.

Small-Study Effects

Although the number of included studies for all outcomes in this review did not meet the recommended threshold for conventional statistical tests of publication bias (≥10 studies), we still performed an exploratory analysis on the anxiety outcome, which had the largest number of included studies (n=8). This was done to preliminarily assess the potential impact of bias on the robustness of the conclusions.

The funnel plot (Figure 11) appeared approximately symmetrical by visual inspection. Statistical test results showed that the Egger linear regression test did not indicate significant small-study effects (t_6=−1.94; P=.10). However, the Begg rank correlation test suggested the possible presence of publication bias (z=−2.47, P=.01). To quantify the potential impact of missing studies on the pooled result, a trim-and-fill analysis was conducted. The analysis estimated that 1 potentially missing study needed to be imputed. The adjusted pooled effect size changed from the original SMD −0.862 (95% CI −1.853 to 0.129) to SMD −0.495 (95% CI −1.719 to 0.729), a change of 0.367. The decrease in the absolute value suggests that, if bias exists, its influence may have led to an overestimation of the intervention effect.

Figure 11. Funnel plot for anxiety outcomes[16,17,32,34,35,55,57,60].
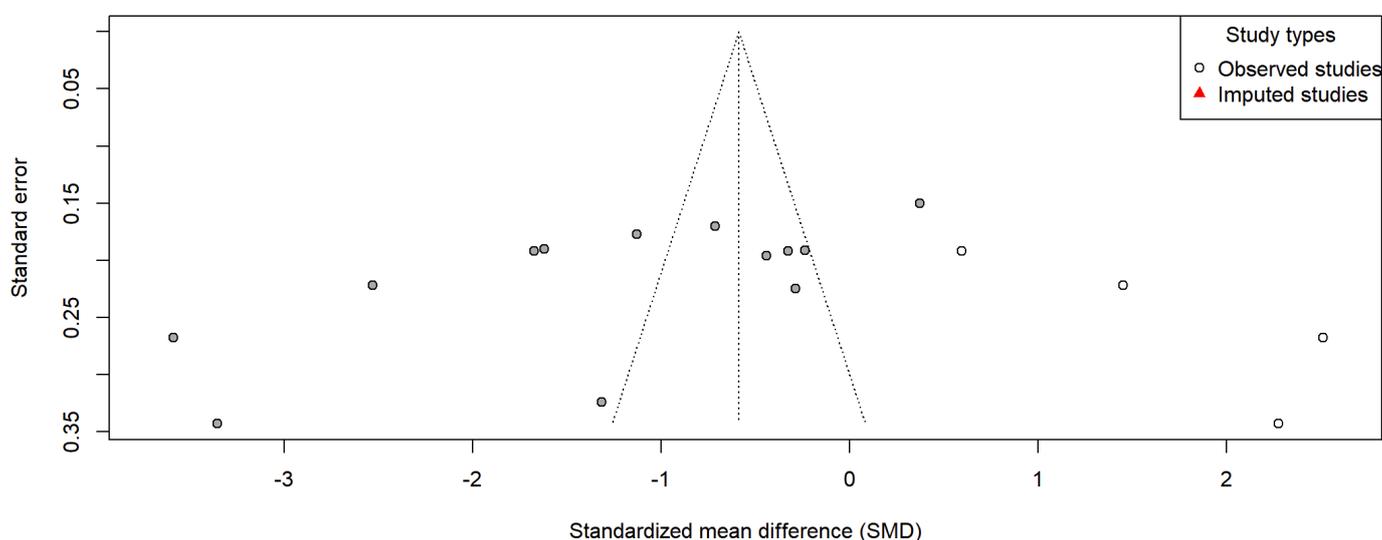
We fully acknowledge that with only 8 included studies [16, 17,32,34,35,55,57,60] and substantial heterogeneity present, the statistical power of these tests is limited. Furthermore, heterogeneity itself can interfere with the interpretation of funnel plot asymmetry. Therefore, the above results should be considered exploratory. In summary, the current analysis, while suggesting that publication bias cannot be entirely ruled out, does not provide evidence from the Egger test supporting significant small-study effects. Moreover, the CI for the effect size adjusted by the trim-and-fill method remained wide and crossed the line of null effect. This indicates that, even if potential bias exists, the current conclusion regarding the anxiety outcome is stable. However, it serves as a reminder to interpret the magnitude of this effect size with caution in the future.

Discussion

Principal Results

This review systematically assessed the feasibility, safety, and rehabilitation outcomes of VR-ERI in critically ill patients across 16 RCTs [16,17,30,32-35,53-61] (n=1356). VR-ERI was well-accepted and associated with high exercise adherence. Vital signs remained within acceptable ranges during use, and no serious adverse events were reported; however, mild dizziness, nausea, and fatigue occurred in some studies, underscoring the need for standardized safety monitoring. Regarding efficacy, VR-ERI improved balance during short-term follow-up, supported by low heterogeneity, narrow PIs, and moderate-certainty evidence, indicating robust and potentially generalizable benefit. VR-ERI also shows potential to improve in-ICU anxiety and subjective sleep quality, as well as short-term cognitive and motor function, particularly sleep, when assessed with the RCSQ. However, it did not significantly improve objective sleep

parameters or in-ICU pain. For most outcomes, the certainty of evidence was low or very low due to implementation bias (eg, lack of blinding), study inconsistency, and imprecision in pooled estimates. Thus, results should be interpreted cautiously, and current evidence remains insufficient for definitive conclusions. Some outcomes could not be pooled due to limited studies. Further high-quality trials are needed to confirm these findings and clarify the clinical value of VR-ERI in critical care.

VR-ERI may reduce anxiety in critically ill patients, a finding consistent with recent reviews in both critical care [36] and elective surgery [62], suggesting VR could broadly mitigate hospital-related anxiety. Stress recovery theory [63] offers a possible explanation, proposing that exposure to natural landscapes can activate the parasympathetic nervous system and promote emotional relaxation. Research confirms that biophilic elements aid stress recovery [64]. Most VR-ERI scenarios in the included studies incorporated natural views (eg, mountains and rivers), which may help block typical ICU stressors, such as constant light, noise, and isolation. However, the pooled effect in this analysis was not statistically significant, and the 95% PI crossed zero, indicating instability in the estimate. The included studies also had a high risk of bias, with substantial inconsistency and imprecision, yielding very low certainty of evidence. Thus, future rigorous RCTs with adequate sample sizes and transparent reporting are needed to clarify VR's effect on anxiety in this population.

VR-ERI shows potential in improving ICU patients' subjective sleep quality, likely due to its intervention design —2 protocols [30,54] used VR meditation and hypnosis modules to promote relaxation and distraction, enhancing perceived sleep quality [30,54]. However, objective parameters measured with smart wristbands (eg, TST and sleep efficiency) did not improve concurrently, suggesting VR may modulate sleep perception rather than alter sleep physiology. A recent network meta-analysis [65] found that simple interventions like eye masks and earplugs offer more consistent sleep benefits in critically ill patients, while the effect of VR meditation remains unestablished, underscoring VR's uncertain role in sleep intervention. Subgroup analysis indicated positive effects only when sleep was assessed with the RCSQ, not with other subjective scales. This highlights how tool selection influences outcomes, possibly because the RCSQ focuses more on immediate sleep depth and is designed specifically for critically ill patients. Future research should use standardized, population-specific sleep assessment tools to accurately capture intervention effects. Current evidence is limited by low certainty. Future studies should develop standardized VR sleep protocols, clarify optimal dosage, and incorporate objective measures, such as polysomnography or validated wearable monitoring, to systematically evaluate VR's impact on sleep architecture, thereby building a more reliable evidence base for practice.

We also observed that VR-ERI can improve balance (measured by Berg Balance Scale) and limb motor function (measured by Fugl-Meyer Assessment) in critically ill patients during short-term follow-up. This finding is consistent with the review by He et al [37] on VR for early motor intervention in this population. By providing an immersive environment, VR compensates for the sensory and interaction deficits resulting from prolonged immobilization and isolation. It enhances sensorimotor pathways and balance-related neuroplasticity through visual-vestibular-proprioceptive integration [66]. Furthermore, VR-ERI engages both cognitive and motor functions simultaneously [67] and promotes cortical activation via feedback-driven learning and task-specific adaptation [68,69]. The engaging and gamified design of VR-ERI may also improve patient adherence to exercise, which could explain its advantage over traditional motor rehabilitation. Improvement in balance is supported by moderate-quality evidence and relatively narrow PIs, suggesting that standardized VR-ERI motor activities could be implemented early in critical care rehabilitation, with attention to personalized parameters and long-term effect tracking. However, evidence for limb motor improvement remains low in certainty and requires further rigorous validation.

However, VR-ERI did not demonstrate a statistically significant effect on reducing pain in critically ill patients during ICU treatment, a result that may be associated with several factors. The pain experience of critically ill patients is particularly complex, as frequent threats to psychological and physical integrity make their pain difficult to manage effectively [70]. Neuroimaging studies have shown that receiving VR intervention during painful stimuli can significantly reduce excitability in brain regions associated with pain processing, with a higher level of VR immersion correlating with stronger inhibitory effects [71]. Concurrently, immersive VR devices can divert attention away from nociceptive stimuli by isolating patients' audiovisual connection to the external environment [72]. Nevertheless, in this study, the reported pain scores were relatively low in both groups, which might be related to insufficient immersion of the VR experience, inadequate intervention dosage, or the widespread use of analgesic medications. This warrants further investigation in future studies.

VR-ERI shows a potential trend for improving cognitive function in short-term follow-up, although the 95% PI crosses the null effect line, indicating unstable results. This may stem from several factors. First, the pathophysiology of post-ICU cognitive impairment is complex (eg, neuroinflammation and blood-brain barrier disruption) [73]; for instance, underlying neural damage in conditions like posterior cortical atrophy may limit response to conventional interventions [74], highlighting the need for more targeted programs. Second, widely used screening tools, such as the Mini-Mental State Examination and MoCA, may lack sensitivity to detect subtle, clinically meaningful improvements in critically ill patients. The Mini-Mental State Examination, although brief and simple, has a ceiling effect and inadequately assesses executive function and complex attention [75]. The MoCA is more sensitive to mild impairment but remains a broad screening tool that may miss specific post-ICU executive deficits (eg, inhibitory control and cognitive flexibility) [76]. Therefore, future studies should develop VR modules tailored

to ICU-related cognitive impairment and consider longer intervention duration. A multilevel assessment strategy is recommended, combining screening scales with computerized neuropsychological tests (eg, flanker tasks) for objective, quantitative data [77]. Integrating VR with other rehabilitation methods into a multidimensional intervention may also help achieve meaningful cognitive gains. Given the low certainty of current evidence, these findings should be interpreted with caution.

This systematic review is the first to focus on VR intervention studies conducted during the early ICU rehabilitation phase. By systematically searching ten Chinese and English databases and implementing independent dual-reviewer screening and data extraction, it strictly adhered to the PRISMA 2020 guidelines, ensuring a transparent process and reproducible results. The review comprehensively used the Cochrane RoB 2.0 tool for risk of bias assessment and the GRADE framework for evidence grading, supplemented by sensitivity analyses to verify the robustness of the findings. It summarizes and analyzes the safety and clinical feasibility of VR intervention in critically ill patients, while systematically elucidating the efficacy profile of VR on relevant outcomes during both the ICU stay and at short-term follow-up. This provides an important reference for advancing the translational application of VR technology in this patient population.

Limitations and Future Recommendations

Several limitations should be acknowledged. First, the literature search was restricted to Chinese and English publications and excluded trial registries and gray literature, potentially introducing omission and language bias. Second, the limited number of studies and small total sample size reduce statistical power, which may affect the robustness and generalizability of the pooled results. Third, substantial clinical and methodological heterogeneity existed across studies, including variations in VR immersion levels, protocol design, dosage, and outcome measures, contributing to high statistical heterogeneity. Inadequate reporting of precise dosage parameters also limited dose-response analyses. Fourth, most included studies had a high risk of bias due to issues such as unclear allocation concealment, lack of assessor blinding, and inherent challenges in blinding participants and personnel. Consequently, the certainty of evidence for most outcomes was low or very low, necessitating cautious interpretation. Finally, some outcomes were supported by only a few studies, resulting in meta-analyses with limited sample sizes and potentially unstable effect estimates. In summary, further well-designed, large-scale studies are required to validate these findings.

At the clinical level, VR-ERI can supplement traditional early rehabilitation, but widespread adoption faces systematic challenges. First, cost remains a key barrier. Although commercial VR prices are falling, expenses may still burden some institutions or patients, with VR for early pain management costing approximately €47.48 (approximately US $51.30) per patient [78]. Policymakers should

thus consider including VR in insurance reimbursement or supporting it via dedicated funding. Second, patient vulnerability limits generalizability [79]. Those who are older adults, or have vestibular dysfunction, severe cognitive impairment, or complex life-support needs may be unable to tolerate VR. Health care institutions should form multidisciplinary teams to create standardized screening protocols—focusing on vestibular and cognitive function—and implement a safety system covering pre-, intra-, and postsession phases. Implementation should be phased and individualized—starting at low intensity and adjusted gradually under therapist supervision, while training ward staff and patients in device use. Finally, before broader promotion, safety and feasibility must be validated in heterogeneous critically ill populations [14].

At the policy level, relevant departments should establish a comprehensive support framework by developing evidence-based clinical guidelines for VR rehabilitation that specify indications, contraindications, and protocols, and implementing unified device certification and quality control systems to ensure safe and standardized application. Future research should adopt a phased strategy—priorities include conducting large-scale, rigorous pragmatic trials using blinded outcome assessment, diverse control groups, objective monitoring, and long-term follow-up to minimize bias; creating personalized interventions for specific patient groups, defining a core outcome set, and using objective, standardized assessment tools; and enhancing implementation science to identify optimal rollout pathways across different resource settings—beginning with pilot projects at regional medical centers before scaling to primary care. Concurrently, interdisciplinary collaboration among engineering, rehabilitation medicine, and critical care nursing should be strengthened to jointly advance technology innovation and support the development of a more integrated early rehabilitation system for critically ill patients.

Conclusions

This review is the first to examine VR rehabilitation effectiveness specifically during early intervention in the ICU. Unlike previous reviews, we focused strictly on the in-ICU early rehabilitation period to establish an evidence base for initiating such care and to minimize the clinical heterogeneity from combining interventions across different care settings. Evidence was synthesized across physiological, psychological, and cognitive domains. Results indicate that VR-ERI shows considerable potential for improving anxiety and subjective sleep quality during ICU treatment, as well as balance, limb motor function, and cognitive function in short-term follow-up. However, effects on pain and objective sleep quality in the ICU remain unclear. Given the high risk of bias in included studies, substantial heterogeneity, and imprecise pooled effect estimates, the evidence certainty for most outcomes is low, and conclusions should be interpreted cautiously. Additionally, establishing a comprehensive safety monitoring system is essential for future clinical translation. This review offers a new perspective on early ICU rehabilitation, suggesting VR technology could serve as a nonpharmacological adjunct in this phase. However, the current evidence is insufficient to support widespread implementation. Clinical

translation requires careful consideration of cost and patient suitability, and long-term benefits need validation through large-scale and rigorous research.

Data Availability

Relevant data are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed search strategy.
[DOCX File (Microsoft Word File), 49 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Excluded literature and reasons.
[DOCX File (Microsoft Word File), 70 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Characteristics of included studies.
[DOCX File (Microsoft Word File), 34 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Dual Revised Cochrane Risk-of-Bias tool for randomized trials assessment process with interrater consistency.
[XLS File (Microsoft Excel File), 1604 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Grading of Recommendations Assessment, Development, and Evaluation assessment results.
[DOCX File (Microsoft Word File), 34 KB-Multimedia Appendix 5]

Checklist 1

PRISMA checklist for abstracts.
[PDF File (Adobe File), 120 KB-Checklist 1]

Checklist 2

PRISMA-S checklist.
[PDF File (Adobe File), 67 KB-Checklist 2]

Checklist 3

PRISMA checklist.
[PDF File (Adobe File), 322 KB-Checklist 3]

References

1. Gao S, Liang X, Pan Z, Zhang X, Zhang L. Effect size estimates of risk factors for post-intensive care syndrome: a systematic review and meta-analysis. Intensive Crit Care Nurs. Apr 2025;87:103888. [doi: 10.1016/j.iccn.2024.103888] [Medline: 39561481]
2. Fresenko LE, Rivera ZC, Parry SM, Mayer KP. Post-intensive care syndrome: physical impairments and function. Crit Care Clin. Jan 2025;41(1):1-20. [doi: 10.1016/j.ccc.2024.08.004] [Medline: 39547718]

3. Needham DM, Davidson J, Cohen H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. Crit Care Med. Feb 2012;40(2):502-509. [doi: 10.1097/CCM.0b013e318232da75] [Medline: 21946660]

4. Ho MH, Lee YW, Wang L. Estimated prevalence of post-intensive care cognitive impairment at short-term and long-term follow-ups: a proportional meta-analysis of observational studies. Ann Intensive Care. Jan 10, 2025;15(1):3. [doi: 10.1186/s13613-025-01429-z] [Medline: 39792310]

5. Valsø Å, Rustøen T, Småstuen MC, et al. Occurrence and characteristics of pain after ICU discharge: a longitudinal study. Nurs Crit Care. Sep 2022;27(5):718-727. [doi: 10.1111/nicc.12701] [Medline: 34382725]

6. Fukamachi Y, Sakuramoto H, Sato T, et al. Prevalence, risk factors, and intervention of long-term sleep disturbance after intensive care unit discharge: a scoping review. Cureus. Apr 2025;17(4):e83011. [doi: 10.7759/cureus.83011] [Medline: 40421347]

7. Yoo KH, Lee J, Oh J, et al. Depression or anxiety and long-term mortality among adult survivors of intensive care unit: a population-based cohort study. Crit Care. May 6, 2025;29(1):179. [doi: 10.1186/s13054-025-05381-z] [Medline: 40329374]

8. Ahn S, LaNoue M, Su H, et al. Post-intensive care syndrome and caregiver burden: a post hoc analysis of a randomized clinical trial. JAMA Netw Open. Apr 1, 2025;8(4):e253443. [doi: 10.1001/jamanetworkopen.2025.3443] [Medline: 40198074]

9. Mikkelsen ME, Still M, Anderson BJ, et al. Society of Critical Care Medicine's International Consensus Conference on Prediction and Identification of Long-Term Impairments After Critical Illness. Crit Care Med. Nov 2020;48(11):1670-1679. [doi: 10.1097/CCM.0000000000004586] [Medline: 32947467]

10. Dayton K, Lindroth H, Engel HJ, et al. Creating a culture of an awake and walking intensive care unit: in-hospital strategies to mitigate post-intensive care syndrome. Crit Care Clin. Jan 2025;41(1):121-140. [doi: 10.1016/j.ccc.2024.08.002] [Medline: 39547720]

11. Tipping CJ, Harrold M, Holland A, Romero L, Nisbet T, Hodgson CL. The effects of active mobilisation and rehabilitation in ICU on mortality and function: a systematic review. Intensive Care Med. Feb 2017;43(2):171-183. [doi: 10.1007/s00134-016-4612-0] [Medline: 27864615]

12. Ding N, Zhang Z, Zhang C, et al. What is the optimum time for initiation of early mobilization in mechanically ventilated patients? A network meta-analysis. PLoS ONE. 2019;14(10):e0223151. [doi: 10.1371/journal.pone.0223151] [Medline: 31589642]

13. Haghedooren E, Haghedooren R, Langer D, Gosselink R. Feasibility and safety of interactive virtual reality upper limb rehabilitation in patients with prolonged critical illness. Aust Crit Care. Nov 2024;37(6):949-956. [doi: 10.1016/j.aucc.2024.06.004] [Medline: 39054204]

14. de Vries M, Beumeler LFE, van der Meulen J, Bethlehem C, den Otter R, Boerma EC. The feasibility of virtual reality therapy for upper extremity mobilization during and after intensive care unit admission. PeerJ. 2025;13:e18461. [doi: 10.7717/peerj.18461] [Medline: 39802193]

15. Godoy-González M, López-Aguilar J, Fernández-Gonzalo S, et al. Efficacy and safety of a non-immersive virtual reality-based neuropsychological intervention for cognitive stimulation and relaxation in patients with critical illness: study protocol of a randomized clinical trial (RGS-ICU). BMC Psychiatry. Dec 18, 2024;24(1):917. [doi: 10.1186/s12888-024-06360-4] [Medline: 39696098]

16. Navarra-Ventura G, Gomà G, de Haro C, et al. Virtual reality-based early neurocognitive stimulation in critically ill patients: a pilot randomized clinical trial. J Pers Med. Nov 29, 2021;11(12):1260. [doi: 10.3390/jpm11121260] [Medline: 34945732]

17. Merliot-Gailhoustet L, Raimbert C, Garnier O, et al. Discomfort improvement for critically ill patients using electronic relaxation devices: results of the cross-over randomized controlled trial E-CHOISIR (Electronic-CHOIce of a System for Intensive care Relaxation). Crit Care. Sep 3, 2022;26(1):263. [doi: 10.1186/s13054-022-04136-4] [Medline: 36057612]

18. Alvarado Sánchez JI, Castillo Morales LM, Cardenas Bolivar YR, et al. Efficacy and safety of the early implementation of a multimodal rehabilitation program in mechanically ventilated patients: a randomized clinical trial protocol. PLoS ONE. 2025;20(5):e0324335. [doi: 10.1371/journal.pone.0324335] [Medline: 40388519]

19. Matsuoka A, Yoshihiro S, Shida H, et al. Effects of mobilization within 72 h of ICU admission in critically ill patients: an updated systematic review and meta-analysis of randomized controlled trials. J Clin Med. Sep 11, 2023;12(18):5888. [doi: 10.3390/jcm12185888] [Medline: 37762829]

20. Fuke R, Hifumi T, Kondo Y, et al. Early rehabilitation to prevent postintensive care syndrome in patients with critical illness: a systematic review and meta-analysis. BMJ Open. May 5, 2018;8(5):e019998. [doi: 10.1136/bmjopen-2017-019998] [Medline: 29730622]

21. Xu J, Wang S, Zhang Q, Yao Y, Yu J. Effectiveness of nurse-led early mobility protocols on the outcomes of critical care patients: a systematic review and meta-analysis. Nurs Open. May 2025;12(5):e70206. [doi: 10.1002/nop2.70206] [Medline: 40371936]

22. Lanier J. Virtual Reality. CITIC Publishing Group Corporation Ltd; 2018. ISBN: 978-7-5086-8634-9

23. Abrams Z. Enhancing mental health care with VR. IEEE Pulse. 2022;13(5):16-20. [doi: 10.1109/MPULS.2022.3208824] [Medline: 36301813]

24. Bauer ACM, Andringa G. The potential of immersive virtual reality for cognitive training in elderly. Gerontology. 2020;66(6):614-623. [doi: 10.1159/000509830] [Medline: 32906122]

25. Zhang Q, Fu Y, Lu Y, et al. Impact of virtual reality-based therapies on cognition and mental health of stroke patients: systematic review and meta-analysis. J Med Internet Res. Nov 17, 2021;23(11):e31007. [doi: 10.2196/31007] [Medline: 34787571]

26. Li R, Li Y, Kong Y, et al. Virtual reality-based training in chronic low back pain: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res. Feb 26, 2024;26:e45406. [doi: 10.2196/45406] [Medline: 38407948]

27. Guo Q, Zhang L, Gui C, et al. Virtual reality intervention for patients with neck pain: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res. Apr 3, 2023;25:e38256. [doi: 10.2196/38256] [Medline: 37010891]

28. Schröder D, Wrona KJ, Müller F, Heinemann S, Fischer F, Dockweiler C. Impact of virtual reality applications in the treatment of anxiety disorders: a systematic review and meta-analysis of randomized-controlled trials. J Behav Ther Exp Psychiatry. Dec 2023;81:101893. [doi: 10.1016/j.jbtep.2023.101893] [Medline: 37453405]

29. 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. May 2021;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]

30. Lee SY, Kang J. Effect of virtual reality meditation on sleep quality of intensive care unit patients: a randomised controlled trial. Intensive Crit Care Nurs. Aug 2020;59:102849. [doi: 10.1016/j.iccn.2020.102849] [Medline: 32241625]

31. Ong TL, Ruppert MM, Akbar M, et al. Improving the intensive care patient experience with virtual reality-a feasibility study. Crit Care Explor. Jun 2020;2(6):e0122. [doi: 10.1097/CCE.0000000000000122] [Medline: 32695991]

32. Weng D, Cui M, Jiang J. Application of 5g enabled virtual reality technology in the intervention of delirium in intensive care unit patients. Chinese Journal of Integrated Traditional and Western Medicine in Intensive and Critical Care. 2023;30(6):693-696. [doi: 10.3969/j.issn.1008-9691.2023.06.011]

33. Zeng Y. Application effects of cognitive behavioral intervention combined with virtual reality training in postoperative patients with severe traumatic brain injury. Henan Medical Research. 2021;30(28). [doi: 10.3969/j.issn.1004-437X.2021.28.058]

34. Yesilot SB, Yeşilkuş R, Beyaz F. Use of virtual reality for reducing pain and anxiety after laparoscopic sleeve gastrectomy: a randomized controlled trial. Pain Manag Nurs. Dec 2022;23(6):826-831. [doi: 10.1016/j.pmn.2022.07.001] [Medline: 35934661]

35. Laghlam D, Naudin C, Coroyer L, et al. Virtual reality vs. Kalinox® for management of pain in intensive care unit after cardiac surgery: a randomized study. Ann Intensive Care. May 13, 2021;11(1):74. [doi: 10.1186/s13613-021-00866-w] [Medline: 33983498]

36. Liu Q, Zhang Y, Chen X, Ding B, Guo X, Gai Y. Efficacy of virtual reality in alleviating post-ICU syndrome symptoms: a systematic review and meta-analysis. Nurs Crit Care. Mar 2025;30(2):e70004. [doi: 10.1111/nicc.70004] [Medline: 40070039]

37. He Y, Yang Q, Dai X, et al. Effects of virtual reality technology on early mobility in critically ill adult patients: a systematic review and meta-analysis. Front Neurol. 2024;15:1469079. [doi: 10.3389/fneur.2024.1469079] [Medline: 39975851]

38. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. Mar 29, 2021;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

39. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. J Med Libr Assoc. Apr 1, 2021;109(2):174-200. [doi: 10.5195/jmla.2021.962] [Medline: 34285662]

40. Higgins JPT, Thomas J. Chandler J, editor. Cochrane Handbook for Systematic Reviews of Interventions Version 65. Cochrane; 2024. URL: https://www.cochrane.org/authors/handbooks-and-manuals/handbook [Accessed 2025-12-04]

41. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005;5(1):13. [doi: 10.1186/1471-2288-5-13]

42. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. Aug 28, 2019;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]

43. Viechtbauer W. Conducting meta-analyses in r with the metafor package. J Stat Softw. 2010;36(3). [doi: 10.18637/jss.v036.i03]

44. Murad MH, Wang Z, Chu H, Lin L. When continuous outcomes are measured using different scales: guide for meta-analysis and interpretation. BMJ. Jan 22, 2019;364:k4817. [doi: 10.1136/bmj.k4817] [Medline: 30670455]

45. Lin L, Aloe AM. Evaluation of various estimators for standardized mean difference in meta-analysis. Stat Med. Jan 30, 2021;40(2):403-426. [doi: 10.1002/sim.8781] [Medline: 33180373]

46. Braitman LE. Statistical power analysis in medical research. Ann Intern Med. Aug 1983;99(2):269-271. [doi: 10.7326/0003-4819-99-2-269] [Medline: 6349459]

47. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. Feb 18, 2014;14:25. [doi: 10.1186/1471-2288-14-25] [Medline: 24548571]

48. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. Sep 1986;7(3):177-188. [doi: 10.1016/0197-2456(86)90046-2] [Medline: 3802833]

49. Bowden J, Tierney JF, Copas AJ, Burdett S. Quantifying, displaying and accounting for heterogeneity in the meta-analysis of RCTs using standard and generalised Q statistics. BMC Med Res Methodol. Apr 7, 2011;11:41. [doi: 10.1186/1471-2288-11-41] [Medline: 21473747]

50. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. Sep 6, 2003;327(7414):557-560. [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]

51. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. Jul 12, 2016;6(7):e010247. [doi: 10.1136/bmjopen-2015-010247] [Medline: 27406637]

52. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. Apr 26, 2008;336(7650):924-926. [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]

53. Caballero LG, Fraga IB, Tremea CEM, et al. Experience of heart failure patients in mobilization with virtual reality: mixed methods study. Rev Esc Enferm USP. 2025;59:e20250114. [doi: 10.1590/1980-220X-REEUSP-2025-0114en] [Medline: 40955916]

54. Kim S, Kang J. Effects of virtual reality meditation on sleep and delirium in ICU patients: a randomized controlled trial. Comput Inform Nurs. Apr 2, 2025;43(8):e01307. [doi: 10.1097/CIN.0000000000001307] [Medline: 40194914]

55. Li YY, Peng J, Ping YY, et al. Virtual reality-based cognitive-behavioural therapy for the treatment of anxiety in patients with acute myocardial infarction: a randomised clinical trial. Gen Psychiatr. 2024;37(2):e101434. [doi: 10.1136/gpsych-2023-101434] [Medline: 38645380]

56. Mo Y, Wang H, Ye P. Application of strengthening rehabilitation nursing multidisciplinary combined with virtual reality training in severe patients with multiple fractures. China Medical Herald. 2021;18(14):173-176. [doi: 10.20047/j.issn1673-7210.2021.14.042]

57. Ouyang Y. Study on Influencing Factors of Cognitive Dysfunction and Intervention of Virtual Reality Nature Therapy in Critical Patients. Fujian Medical University; 2022. [doi: 10.27020/d.cnki.gfjyu.2022.001308]

58. Wang J, Shi C, Jia Y, Xiao Q. Effectiveness of virtual reality assisted active limb movement exercises for patients in the respiratory intensive care unit: a randomized pilot study. J Rehabil Med. Jun 3, 2025;57:jrm28399. [doi: 10.2340/jrm.v57.28399] [Medline: 40462528]

59. Xia Y, Pan W, Zhao Y, Chen Z, Cai S, Tang Y, et al. Application of virtual reality technology in early activities of ICU patients. Chinese Journal of Nursing. 2023;58(9):1050-1055.

60. Yani Y, Wenyi T, Zhenhong F, Wenyu L, Mengyao W, Ruhui C. Improvement of sleep quality and anxiety among cardiac surgery patients by preoperative visits based on virtual reality technology. Journal of Nursing Science. 2024;39(22):112-115. [doi: 10.3870/j.issn.1001-4152.2024.22.112]

61. Zhu H, Jie H, Wang X, et al. Application of virtual reality technology in early rehabilitation training of patients after liver transplantation. Nurs Res. 2023;37(20):3790-3792. URL: https://doaj.org/article/7afabfa3a7a14f3aa6030b21d1f1ee1b [Accessed 2026-02-22] [Medline: 37916845]

62. Li H, Chiu PL, Efendi D, Huang H, Ko KY, Wong CL. Effects of virtual reality-based interventions on preoperative anxiety in patients undergoing elective surgery with anesthesia: systematic review and meta-analysis. J Med Internet Res. Apr 30, 2025;27:e55291. [doi: 10.2196/55291] [Medline: 40305092]

63. Brosschot JF, Pieper S, Thayer JF. Expanding stress theory: prolonged activation and perseverative cognition. Psychoneuroendocrinology. Nov 2005;30(10):1043-1049. [doi: 10.1016/j.psyneuen.2005.04.008] [Medline: 15939546]

64. Yin J, Yuan J, Arfaei N, Catalano PJ, Allen JG, Spengler JD. Effects of biophilic indoor environment on stress and anxiety recovery: a between-subjects experiment in virtual reality. Environ Int. Mar 2020;136:105427. [doi: 10.1016/j.envint.2019.105427] [Medline: 31881421]

65. Huang D, Li Y, Ye J, Liu C, Shen D, Lv Y. Different nursing interventions on sleep quality among critically ill patients: a systematic review and network meta-analysis. Medicine (Baltimore). Dec 29, 2023;102(52):e36298. [doi: 10.1097/MD.0000000000036298] [Medline: 38206715]

66. Wankhede NL, Koppula S, Ballal S, et al. Virtual reality modulating dynamics of neuroplasticity: innovations in neuro-motor rehabilitation. Neuroscience. Feb 6, 2025;566:97-111. [doi: 10.1016/j.neuroscience.2024.12.040] [Medline: 39722287]

67. Morone G, Spitoni GF, De Bartolo D, et al. Rehabilitative devices for a top-down approach. Expert Rev Med Devices. Mar 2019;16(3):187-195. [doi: 10.1080/17434440.2019.1574567] [Medline: 30677307]

68. Cervera MA, Soekadar SR, Ushiba J, et al. Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. Ann Clin Transl Neurol. May 2018;5(5):651-663. [doi: 10.1002/acn3.544] [Medline: 29761128]

69. Lavi R, Pau M, Dreyer-Alster S, Drori T, Cocco E, Kalron A. Real-time cortical activity during virtual reality practice in people with multiple sclerosis: a pilot fNIRS study. J Neuroeng Rehabil. Jul 4, 2025;22(1):148. [doi: 10.1186/s12984-025-01681-1] [Medline: 40615921]

70. Leininger S. The challenges of managing pain in the intensive care unit. Crit Care Nurs Q. 2022;45(4):325-331. [doi: 10.1097/CNQ.0000000000000424] [Medline: 35980794]

71. Hoffman HG, Richards TL, Van Oostrom T, et al. The analgesic effects of opioids and immersive virtual reality distraction: evidence from subjective and functional brain imaging assessments. Anesth Analg. Dec 2007;105(6):1776-1783. [doi: 10.1213/01.ane.0000270205.45146.db] [Medline: 18042882]

72. Matamala-Gomez M, Donegan T, Świdrak J. VR for pain relief. Curr Top Behav Neurosci. 2023;65:309-336. [doi: 10.1007/7854_2022_402] [Medline: 36592274]

73. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. N Engl J Med. Oct 3, 2013;369(14):1306-1316. [doi: 10.1056/NEJMoa1301372] [Medline: 24088092]

74. Suárez-González A, Bier N, Sauvageau H, Pelak VS, Ahmed S. Cognitive rehabilitation in posterior cortical atrophy. Pract Neurol. Jan 16, 2025;25(1):40-44. [doi: 10.1136/pn-2024-004259] [Medline: 39515839]

75. Pfoh ER, Chan KS, Dinglas VD, et al. Cognitive screening among acute respiratory failure survivors: a cross-sectional evaluation of the Mini-Mental State Examination. Crit Care. May 5, 2015;19(1):220. [doi: 10.1186/s13054-015-0934-5] [Medline: 25939482]

76. Fu C, Jin X, Chen B, et al. Comparison of the Mini-Mental State Examination and Montreal Cognitive Assessment executive subtests in detecting post-stroke cognitive impairment. Geriatr Gerontol Int. Dec 2017;17(12):2329-2335. [doi: 10.1111/ggi.13069] [Medline: 28675607]

77. Lin LT, Lin HA, Bai CH, Lin SF. Flanker task assessment of inhibitory control in ICU survivors: preliminary insights into cognitive impairment. Behav Brain Res. Oct 18, 2025;495:115791. [doi: 10.1016/j.bbr.2025.115791] [Medline: 40907667]

78. Groenveld TD, van Dorst P, Ten Broek RPG, et al. Early health economic analysis of virtual reality therapy for pain management after surgery. Int J Surg. Jun 1, 2025;111(6):3798-3809. [doi: 10.1097/JS9.0000000000002385] [Medline: 40387696]

79. Li M, Race M, Huang F, Escalon MX. Role of virtual reality to promote mobilization in the critical care setting: a narrative review. Am J Phys Med Rehabil. May 1, 2025;104(5):487-494. [doi: 10.1097/PHM.0000000000002648] [Medline: 39642293]

Abbreviations

CAM: Confusion Assessment Method
CNKI: China National Knowledge Internet
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
ICU: intensive care unit
MoCA: Montreal Cognitive Assessment
PI: prediction interval
PICOS: Population, Intervention, Comparison, Outcome, and Study Design
PICS: post-intensive care syndrome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension
RCSQ: Richards-Campbell Sleep Questionnaire
RCT: randomized controlled trials
RoB 2: revised Cochrane Risk of Bias tool
SinoMed: SinoMed Service System
SMD: standardized mean difference
TST: total sleep time

VIP: VIP Database
VR: virtual reality
VR-ERI: virtual reality–based early rehabilitation intervention
WASO: wake after sleep onset