Extended Reality Interventions for Chronic Pain

March 2024

U.S. Department of Veterans Affairs
Veterans Health Administration
Health Systems Research

Recommended citation: Goldsmith E, Anthony M, Landsteiner A, et al. Extended Reality Interventions for Chronic Pain: A Systematic Review. Washington, DC: Evidence Synthesis Program, Health Systems Research, Office of Research and Development, Department of Veterans Affairs. VA ESP Project #09-009; 2024.

AUTHORS

Author roles, affiliations, and contributions (using the CRediT taxonomy) are listed below.

Author	Role and Affiliation	Report Contribution
Elizabeth Goldsmith, MD, PhD	Project Lead Staff Physician, Core Investigator, Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System Assistant Professor of Medicine, University of Minnesota Twin Cities Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing, Supervision, Project administration
Maylen Anthony, MPH	Project Coordinator Research Associate, Minneapolis Evidence Synthesis Program (ESP) Center Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing, Project administration
Adrienne Landsteiner, PhD, MPH	Senior Scientist, Minneapolis ESP Center Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing
Kristen Ullman, MPH	Program Manager, Minneapolis ESP Center Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing
Caleb Kalinowski, MA	Research Associate, Minneapolis ESP Center Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing
Nick Zerzan, MPH	Research Associate, Minneapolis ESP Center Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing
David Ewart, MD	Staff Physician, Minneapolis VA Health Care System Assistant Professor of Medicine, University of Minnesota Twin Cities Minneapolis, MN	Conceptualization, Methodology, Writing – review & editing
Tonya Rich, PhD	Research Occupational Therapist, Minneapolis VA Health Care System Assistant Professor of Rehabilitation Medicine, University of Minnesota Minneapolis, MN	Conceptualization, Methodology, Writing – review & editing
Wendy Miller, MD	Staff Physician, Hennepin Healthcare Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Writing – review & editing
Collin Calvert, PhD	HSR&D Post-Doctoral Fellow, Minneapolis VA Health Care System Minneapolis, MN	Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – review & editing

Author	Role and Affiliation	Report Contribution
David Thomas, MD, MS	Resident Physician in Internal Medicine, University of Minnesota Twin Cities Minneapolis, MN	Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – review & editing
Timothy Wilt, MD, MPH	Co-Director, Minneapolis ESP Center Staff Physician, Core Investigator, Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System Professor of Medicine and Public Health, University of Minnesota Twin Cities Minneapolis, MN	Conceptualization, Methodology, Writing – review & editing, Supervision
Wei (Denise) Duan-Porter, MD, PhD	Co-Director, Minneapolis ESP Center Staff Physician, Core Investigator, Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System Assistant Professor of Medicine, University of Minnesota Twin Cities Minneapolis, MN	Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing, Supervision, Project administration

PREFACE

The VA Evidence Synthesis Program (ESP) was established in 2007 to conduct timely, rigorous, and independent systematic reviews to support VA clinicians, program leadership, and policymakers improve the health of Veterans. ESP reviews have been used to develop evidence-informed clinical policies, practice guidelines, and performance measures; to guide implementation of programs and services that improve Veterans' health and wellbeing; and to set the direction of research to close important evidence gaps. Four ESP Centers are located across the US. Centers are led by recognized experts in evidence synthesis, often with roles as practicing VA clinicians. The Coordinating Center, located in Portland, Oregon, manages program operations, ensures methodological consistency and quality of products, engages with stakeholders, and addresses urgent evidence synthesis needs.

Nominations of review topics are solicited several times each year and submitted via the ESP website. Topics are selected based on the availability of relevant evidence and the likelihood that a review on the topic would be feasible and have broad utility across the VA system. If selected, topics are refined with input from Operational Partners (below), ESP staff, and additional subject matter experts. Draft ESP reviews undergo external peer review to ensure they are methodologically sound, unbiased, and include all important evidence on the topic. Peer reviewers must disclose any relevant financial or non-financial conflicts of interest. In seeking broad expertise and perspectives during review development, conflicting viewpoints are common and often result in productive scientific discourse that improves the relevance and rigor of the review. The ESP works to balance divergent views and to manage or mitigate potential conflicts of interest.

ACKNOWLEDGMENTS

The authors are grateful to external peer reviewers and the following individuals for their contributions to this project:

Operational Partners

Operational partners are system-level stakeholders who help ensure relevance of the review topic to the VA, contribute to the development of and approve final project scope and timeframe for completion, provide feedback on the draft report, and provide consultation on strategies for dissemination of the report to the field and relevant groups.

Anne Lord Bailey, PharmD, BCPS
Executive Director, VA Strategic Initiatives Lab
VA Immersive Lead, VA Office of Healthcare Innovation and Learning

Friedhelm Sandbrink, MD
Executive Director
National Pain Management, Opioid Safety and PDMP (PMOP)

Jennifer L Murphy, PhD
Director of Pain Management
National Pain Management, Opioid Safety and PDMP (PMOP)

David Atkins, MD, MPH
Former Director

VA Health Services Research and Development (HSR&D)

Jessica Cleveland, MSOR
AAAS Science & Technology Policy Fellow
VA Health Services Research and Development (HSR&D)

Technical Expert Panel

To ensure robust, scientifically relevant work, the technical expert panel (TEP) guides topic refinement; provides input on key questions and eligibility criteria, advising on substantive issues or possibly overlooked areas of research; assures VA relevance; and provides feedback on work in progress. TEP members included:

Thiru Annaswamy, MD, MA
Professor and Chair, Physical Medicine & Rehabilitation
Penn State College of Medicine

Jeremy Bailenson, PhD
Professor of Communication, Education, and Program in Symbolic Systems
Director, Virtual Human Interaction Lab
Stanford University

Albert 'Skip' Rizzo, PhD
Research Professor, Schools of Gerontology and Medicine
Director for Medical Virtual Reality
University of Southern California - Davis

Disclosures

This report was prepared by the Evidence Synthesis Program Center located at the **Minneapolis VA Health Care System,** directed by Timothy J. Wilt, MD, MPH and Wei Duan-Porter, MD, PhD and funded by the Department of Veterans Affairs, Veterans Health Administration, Health Systems Research.

The findings and conclusions in this document are those of the author(s) who are responsible for its contents and do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. The final research questions, methodology, and/or conclusions may not necessarily represent the views of contributing operational and content experts. No investigators have affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

Main Report

Evidence Synthesis Program

TABLE OF CONTENTS

ABBREVIATIONS TABLE

Abbreviation	Definition
2D	2-dimensional
3D-TV	3-dimensional television
6MWT	6-minute walk test
ACL	Anterior cruciate ligament
ACR	American College of Rheumatology
ADIM	Abdominal drawing-in maneuver
AE	Adverse event
AR	Augmented reality
ARAT	Action Research Arm Test
ASQOL	Ankylosing Spondylitis Quality of Life
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BPI	Brief Pain Inventory
BSME	Biering-Sorensen test of Statis Muscular Endurance
CAP	Concerns About Pain
CAVE	Cave Automatic Virtual Environment
CBT	Cognitive behavioral therapy
CDC	Center for Disease Control
CI	Confidence interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CLBP	Chronic low back pain
Cm	Centimeter
COE	Certainty of Evidence
CPM	Continuous passive motion
CRPS	Complex Regional Pain Syndrome
CTMT	Conventional thermal magnetic therapy
Diff Δ	Difference in change scores
DTxP	Digital therapeutics, virtual reality, psychological intervention for pain
DVPRS	Defense and Veterans Pain Rating Scale
ECOG	Eastern Cooperative Oncology Group
EuroQoL-5D-5L	European Quality of Life-5 questionnaire
EuroQoL-5D-VAS	European Quality of Life – Visual Analog Scale
FABQ	Fear Avoidance Belief Questionnaire
FACIT	Functional Assessment of Chronic Illness Therapy
FIM	Functional Independence Measure
FIQ	Fibromyalgia Impact Questionnaire
GPE	Global Perceived Effect scale
GRADE	Grading of Recommendations, Assessment, Development, and Evaluation

Abbreviation	Definition
GRADEpro GDT	GRADEpro Guideline Development Tool
HAQ	Health Assessment Questionnaire
Hr	Hour
HSM	Horse simulator machine
HSR	VA Health Systems Research
IKT	Isokinetic training
IQR	Interquartile range
IREX	Interactive Rehabilitation and Exercise
KOOS	Knee Injury and Osteoarthritis Outcome Score
KQ	Key questions
KT	Kinematic training
LBP	Low back pain
M	Meters
MAS	Modified Ashworth Scale for Grading Spasticity
MCE	Motor control exercise
MI	Motricity Index
MIDAS	Migraine Disability Assessment Test
MME	Morphine milligram equivalence
MODI	Modified Oswestry Disability Index
MSK	Musculoskeletal
MVR	Mediated virtual reality
N/A	Not applicable
NC	Not calculable
NDI	Neck Disability Index
No.	Number
NR	Not reported
NRS	Numeric Rating Scale
NSAIDS	Non-steroidal anti-inflammatory drugs
OA	Osteoarthritis
ODI	Oswestry Disability Index
OKS	Oxford Knee Score
oMEDD	Oral morphine equivalent daily dose
OP	Operational Partners
PASS-20	Pain Anxiety Symptom Scale
PCS	Pain Catastrophizing Scale
PDI	Pain Disability Index
PGIC	Pain Global Impression of Change scale
PI	Prediction interval
PMOP	VA National Pain Management, Opioid Safety, and Prescription Drug Monitoring Program

Abbreviation	Definition
PRO	Patient-reported outcomes
PROMIS	Patient-Reported Outcomes Measurement Information System
PROSPERO	International Prospective Register of Systematic Reviews
PSFS	Patient Specific Function Scale
PT	Physical therapy
QLQ-C30	The European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30
RA	Rheumatoid arthritis
RAPT	Risk Assessment Prediction Tool
RCT	Randomized controlled trial
RFIQ	Revised Fibromyalgia Impact Questionnaire
RMDQ	Roland-Morris Disability Questionnaire
RoB	Risk of bias
ROBINS-I	Cochrane Risk Of Bias In Non-randomised Studies – of Interventions
ROM	Range of motion
SD	Standard deviation
SF-12	12-item Short Form health survey
SF-36	36-item Short Form health survey
SF-36 MCS	36-item Short Form health survey mental component score
SF-36 PCS	36-item Short Form health survey physical component score
SMD	Standardized mean difference
SMT	Sensorimotor training
SPADI	Shoulder Pain and Disability Index
SPPB	Short Performance Physical Battery
TENS	Transcutaneous electrical nerve stimulation
TEP	Technical Expert Panel
THA	Total hip arthroplasty
TKA	Total knee arthroplasty
TNF	Tumor necrosis factor
TSK	Tampa Scale of Kinesiophobia
TUG	Timed Up and Go test
UK	United Kingdom
US	United States
VA	Veteran's Affairs
VAS	Visual Analog Scale
VERA	Virtual Exercise Rehabilitation Assistant
VHA	Veteran's Health Administration
VR	Virtual reality
WHOQOL-BREF	World Health Organization Quality of Life Brief Version
WOMAC	Western Ontario and McMaster Universities Arthritis Index

Abbreviation	Definition
XR	Extended reality
Yr	Years

BACKGROUND

Chronic pain is prevalent among United States (US) Veterans. In 2021, the age-adjusted prevalence of chronic pain among US adults was 27.5% for Veterans and 19.2% for non-Veterans, with 11.6% of Veterans and 6.5% of non-Veterans reporting that their pain resulted in frequent functional limitations in daily life or work activities.[1] Non-drug therapies (*eg*, exercise and psychological interventions) are first-line treatments for common pain conditions due to their evident benefits and low risks, particularly as compared to opioids and invasive procedures.[2-5] Accordingly, Veterans Health Administration (VHA) initiatives have emphasized use of non-drug therapies to improve chronic pain management and decrease inappropriate opioid prescribing. The VHA Directive on Pain Management (#2009-053)[6] also mandated national implementation of a multimodal stepped care model that would enable timely access to evidence-based non-drug therapies, among other features. Effects of non-drug therapies may vary across common pain conditions and patient characteristics. For example, while targeted exercise is generally effective for chronic low back pain,[7] integration of psychotherapy techniques with exercise may be needed for maximal benefit in those with high fear of movement (*ie*, kinesiophobia).[8] Outcomes of non-drug therapies may also differ depending on whether they are used to treat established chronic pain (*eg*, chronic low back pain) or to prevent the development of persistent or chronic symptoms from acute pain (*eg,* post-surgical pain management and rehabilitation). Additionally, non-drug treatments often require long-term adherence to yield maximum benefit, making patient engagement a key factor in effectiveness over time.

Extended reality (XR) is a spectrum of digital technologies and software that serve as modalities for delivering different types of interventions for pain, such as teaching pain self-management psychological skills or guiding engagement in physical activity.[9,10] A user's level of *immersion* in the XR digital environment is used to define categories within the XR spectrum.[10] Virtual reality (VR) has the highest level of immersion in an interactive, fully digital environment: for example, relaxation training in a digitally created peaceful forest that is separate from the user's real-world physical environment. Augmented reality (AR) provides partially immersive user experiences by augmenting depictions of real-world physical environments with digital elements or translating user activities into the digital world. For example, an AR intervention may involve visualizing a simulated digital leg on real-time video of the user's physical body to facilitate rehabilitation after amputation. Like VR, AR can also incorporate sensors to monitor participant movements and translate these into actions in digitally created worlds, as in gaming systems such as Microsoft Kinect and Nintendo Wii. With the increased availability of low-cost XR devices and widespread popularity of these technologies, interest in a broad range of XR clinical applications has also grown.

Although XR interventions have been implemented in clinical settings for a range of health conditions, including acute and chronic pain treatment,[11] the ways in which XR contributes to therapeutic change remain under investigation and likely vary by type of intervention (*eg*, building mindfulness skills vs performing targeted exercises). The use of XR to decrease acute pain through distraction is well documented (*eg*, during dental treatment or other procedures that cause pain).[12] For treatment of chronic pain and rehabilitation, distraction may enable participation in recommended exercises, leading to better engagement and adherence. Some have also proposed that XR technology can specifically facilitate embodiment (the perception of one's body as comprising both digital and real-world elements), a potentially important mechanism for addressing conditions like phantom limb pain.[13] However, it remains unclear whether XR interventions can improve outcomes across different common chronic pain conditions, and if so, how they should be integrated into a comprehensive plan

for pain treatment. As the field continues to expand and diverse XR interventions are developed, understanding the current evidence for effects of XR pain therapies is critically needed to guide future implementation efforts and further research to address knowledge gaps.

The VHA's XR Network is a nationwide resource hub for dissemination and pilot testing of XR technologies across VHA facilities; current pilots of XR interventions for pain are occurring in post-operative care, Community Living Centers, and various outpatient settings.[14] To inform future research on XR interventions for pain and implementation of XR treatments at VA facilities, the VHA XR Network, in collaboration with VA Health Systems Research (HSR) and VA National Pain Management, Opioid Safety, and Prescription Drug Monitoring Program (PMOP) Office, requested this evidence review on the benefits and harms of XR interventions for chronic pain.

In this systematic review, we synthesized evidence on benefits and harms of XR interventions for treatment of chronic pain or to prevent development of chronic pain (if treating acute pain). We present findings by pain condition, beginning with chronic low back pain and chronic neck pain, followed by other conditions. Within each condition, we provide results separately for VR and AR interventions, and then types of XR interventions (and comparators) within these categories. We conducted meta-analyses where feasible and provide qualitative summaries otherwise.

METHODS

TOPIC DEVELOPMENT

Collaboratively with representatives from VHA XR Network, VA PMOP, VA HSR, and our Technical Expert Panel, we defined the scope, formulated key questions (KQs), and determined eligibility criteria. We included a wide variety of technologies and types of XR interventions that may be used to treat various pain conditions.

REGISTRATION AND REVIEW

A preregistered protocol for this review can be found on the PROSPERO international prospective register of systematic reviews (CRD42023439903). A draft version of this report was reviewed by external peer reviewers; their comments and author responses are in Appendix L.

KEY QUESTIONS AND ELIGIBILITY CRITERIA

The following key questions were the focus of this review: For adults, what are the benefits and harms of XR interventions for 1) treating chronic pain and 2) preventing the development of chronic pain? Study eligibility criteria are shown in the table below.

	Inclusion Criteria	Exclusion Criteria
Population	Key Question 1 (KQ1): Adults (≥18 years) with chronic pain, defined as a condition that has chronic pain as a major symptom (*eg*, arthritis, fibromyalgia, phantom limb pain, diabetic neuropathy) AND/OR assessed pain of ≥90 days duration Key Question 2 (KQ2): Adults (≥18 years) with any pain	Children (<18 years)
Intervention	Extended reality (*ie*, augmented, mixed or fully immersive virtual reality) treatments with a primary aim of treating or preventing chronic pain	XR tools for providers; telehealth or virtual care that does not involve XR (only videoconferencing, *etc*) KQ1 only: XR interventions solely during surgery or other medical procedures
Comparator	Any	—
Outcomes	Primary outcomes* Pain-related functioning or interference (PRO) Pain intensity or severity (PRO) Pain global change (PRO) Pain catastrophizing and kinesiophobia (PRO) Quality of life (PRO) Adverse events Secondary outcomes Opioid dose or use Physical performance Completion of medical treatments or adherence	—

	Inclusion Criteria	Exclusion Criteria
Setting	KQ1: Non-acute care settings (*eg*, outpatient clinic, home) KQ2: Any	KQ1 only: Hospitalizations or emergency room visits
Study Design	KQ1: RCTs with $N \geq 30$ and follow-up ≥ 1 day post-treatment KQ2: Observational cohorts and RCTs with $N \geq 30$ and follow-up ≥ 90 days	Commentary, case series, case reports, reviews; $N < 30$ KQ1 only: observational studies

Notes. * To be eligible, study had to address at least 1 primary outcome.
Abbreviations. PRO=patient-reported outcomes; RCT=randomized controlled trial; XR=extended reality.

SEARCHING AND SCREENING

We searched MEDLINE, Embase, CINHAL, PsycINFO, and Scopus databases from inception to May 2023, using key words and subject headings for virtual and augmented reality, exergaming, pain, and a variety of pain-related conditions (*eg*, neuralgia and fibromyalgia). See Appendix A for complete search strategies. We also hand-searched included studies of relevant systematic reviews identified via the database searches. We searched clinical trial registries (clinicaltrials.gov, Australian and New Zealand Trials Registry, European Union Trials Register, and www.isrctn.com) for recently completed and ongoing trials. For completed trials, we looked for publications associated with these trials using the protocol title, investigator names, and locations. Ongoing trials (without identified publications) are noted in Appendix C.

Duplicate search results were removed, and abstracts were screened using DistillerSR version 2.35 (Evidence Partners, Ottawa, Canada).[15] Exclusion of abstracts required agreement of 2 reviewers. Included abstracts underwent full-text review by 2 individuals, with eligibility decisions requiring consensus of both reviewers.

DATA ABSTRACTION AND RISK OF BIAS ASSESSMENT

Data abstraction was completed by 1 reviewer and verified by a second reviewer. Abstracted data included participant characteristics and inclusion/exclusion criteria, intervention characteristics (technology and devices used, content and goals of intervention), study design and settings, and findings for primary and secondary outcomes, as noted above. If findings were only reported in figures, we used PlotDigitizer (https://plotdigitizer.com/app) to derive numbers from figure images, per recommended practices.[16]

Using reported intervention characteristics, we classified interventions as VR or AR according to the proposed framework by Rauschnabel et al.[10] Additionally, we coded the type of XR intervention into 4 categories based on intervention content and goals:

1) Pain psychology and coping skills (referred to as "psychological skills" hereafter)
2) Guide and engage in physical activity
3) Embodiment only
4) Distraction only.

Risk of bias (RoB) assessments were conducted independently by 2 researchers, and discrepancies were resolved by consensus or with a third reviewer. RCTs were assessed with Cochrane Risk of Bias

2.0[17] and observational studies with the Cochrane Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I).[18] RoB ratings (including domain ratings) are provided in Appendix D.

SYNTHESIS

We first grouped studies by pain condition (*eg*, chronic back pain, knee pain) and then by intervention characteristics. We separately examined studies of VR and AR interventions and then by type of intervention (*eg*, pain psychology and coping skills, or guiding and engaging in physical activity). We conducted meta-analyses when there were ≥3 studies for a given pain condition that evaluated sufficiently similar interventions and reported the same outcome (*eg*, comparable measures of pain-related functioning or interference). Otherwise, we provided narrative syntheses of study characteristics and findings. For meta-analyses, we used random-effects models (with Hartung–Knapp-Sidik–Jonkman estimator) due to the anticipated heterogeneity in effects arising from variation in patient populations, clinical settings, and other study characteristics. We focused on between-group comparisons of the mean change in continuous outcomes (*ie*, difference in change scores [Diff Δ]), preferentially as standardized effect sizes (Diff Δ/standard deviation [SD] of change). To calculate Diff Δ, we subtracted the mean change in the comparator group from the mean change in the XR intervention group (Δ_{XR} - Δ_C); thus, for outcome measures where lower scores are better (*eg*, pain intensity or severity), a negative value for Diff Δ indicates that there were greater improvements in the XR intervention group. When SDs of change scores were not reported, we used imputation techniques to estimate SD using data from other sufficiently similar studies, when these were available. As described in the Cochrane Handbook for Systematic Reviews of Interventions,[16] we calculated the correlation coefficients for mean change and SD of change scores using data from studies that provided all data, and used these correlation coefficients to impute the SD of change scores from the reported mean change (or calculated mean change, using baseline and follow-up mean scores). We evaluated for statistical heterogeneity using visual inspection, τ^2, and 95% prediction intervals (PI). We planned to assess publication bias using funnel plots if there were ≥10 sufficiently similar studies (according to considerations described above). We used *meta* and *metafor* packages and R version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria)[19] to conduct meta-analyses and generate forest plots.

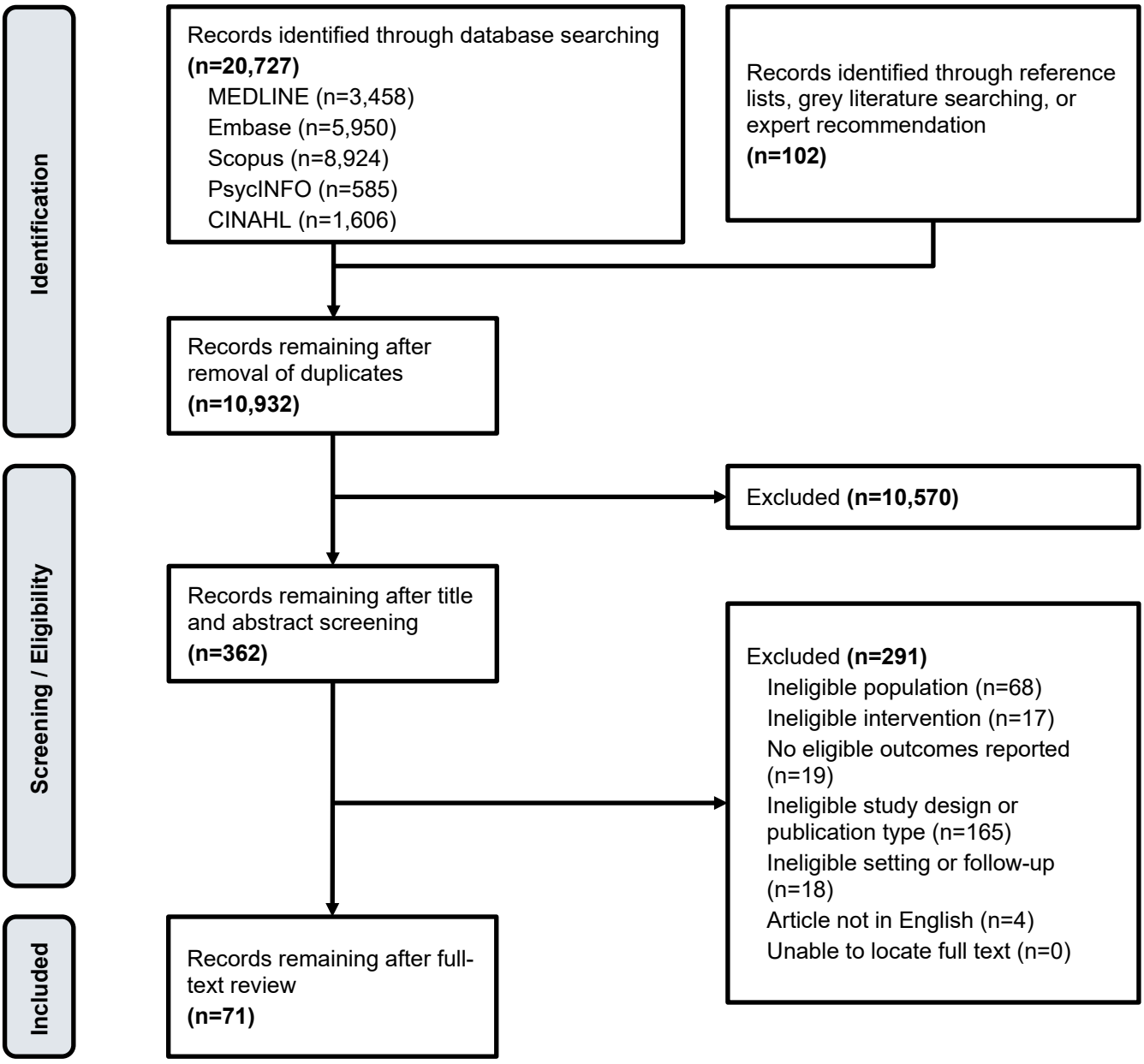
Certainty of Evidence

For certainty of evidence (COE) assessments, we prioritized 3 primary outcomes with input from Operational Partners and Technical Expert Panel members. Before analysis and synthesis of eligible study findings, we met with partners and the expert panel to discuss prioritization of outcomes for COE assessments and then conducted an online survey requesting feedback on the top 3 outcomes in importance (*ie*, indicate which primary outcome is first, second, or third, from among the 6 eligible primary outcomes). The top 3 prioritized outcomes were pain-related functioning or interference, pain intensity or severity, and adverse events. We rated COE for each prioritized outcome separately for types of VR and AR interventions for the following pain conditions: chronic low back pain, chronic neck pain, fibromyalgia, chronic knee pain, and post-surgical pain and rehabilitation (KQ2). We used Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) methodology to rate overall COE as high, moderate, low, or very low.[20,21] Briefly, for each prioritized outcome, we used GRADEpro Guideline Development Tool (GDT)[22] to systematically evaluate 5 domains: study limitations (risk of bias), imprecision (limitations in precision of effect estimates), inconsistency (in direction and magnitude of effects across studies), indirectness (applicability of the results), and other considerations (including publication bias).

RESULTS

LITERATURE FLOW DIAGRAM

The literature flow diagram summarizes the results of the study selection process. A full list of excluded studies is provided in Appendix B.



Abbreviations. CINAHL=Cumulative Index to Nursing and Allied Health Literature; ISRCTN=International Standard Randomised Controlled Trial Number; US=United States.

OVERVIEW OF INCLUDED STUDIES

We screened 10,933 unique citations and reviewed the full text for 362. We identified 71 eligible articles reporting 60 unique primary studies: 49 studies (58 articles) that addressed KQ1 and 11 studies (13 articles) for KQ2. Table 1 provides summary characteristics for all eligible studies. Most studies assessed pain-related functioning or interference ($k = 43$) and pain intensity or severity ($k = 53$). In contrast, few reported on adverse events ($k = 15$), and of these, most did not systematically assess adverse effects for all study arms. Half of studies were very small ($k = 30$ with total $n \leq 50$), and only 4 studies had total $n > 100$. Most studies were conducted outside of the US ($k = 54$), and most included young and middle-aged adult participants ($k = 46$). All studies were rated high or some concerns for RoB, with half being high RoB ($k = 30$). Detailed RoB ratings for all articles are provided in Appendix D. We also identified 47 ongoing or recently completed studies (Appendix C). Four of the ongoing studies have planned $n > 250$; 3 of these are currently recruiting (NCT04933474, NCT05005026, and NCT049067643), and the fourth has since published results during the drafting of this report (NCT05263037).[23]

Among KQ1 studies evaluating XR interventions for treatment of chronic pain, most addressed low back pain ($k = 22$), and fewer examined chronic neck pain ($k = 6$), fibromyalgia ($k = 5$), and chronic knee pain ($k = 5$). There were also a variety of other chronic pain conditions with only 1 eligible study (*eg*, headache or phantom limb pain). Most KQ1 studies evaluated AR interventions ($k = 32$), and physical activity was the most common type of XR intervention ($k = 36$; Table 1; Figure 1).

Eleven KQ2 studies examined XR interventions for the prevention of chronic pain, with 7 of these addressing post-surgical pain and rehabilitation. Nearly all post-surgical studies involved XR physical activity interventions ($k = 6$; Table 1; Figure 1). The remaining 4 studies included 2 RCTs on AR physical activity interventions for post-stroke pain and rehabilitation, 1 RCT on AR physical activity for flight-associated neck pain (both acute and chronic), and 1 cohort study on VR psychological skills intervention for pain from work-related injuries.

Below, we first describe findings for KQ1 studies, grouped by VR and AR interventions within each pain condition. We then briefly present characteristics and findings of single studies for a variety of KQ1 conditions, including those with mixed populations or general chronic pain (*eg*, chronic pain in ≥2 joints). Next, we describe results for post-surgical pain and rehabilitation (KQ2), grouped similarly by VR and AR interventions. Finally, we present the findings for the 4 KQ2 studies on a variety of conditions. Within sections on chronic low back pain, chronic neck pain, fibromyalgia, chronic knee pain, and post-surgical pain, we also provide COE ratings.

Table 1. Overview of Characteristics for Included Studies

Characteristics		Chronic Low Back Pain (*k* = 22)	Chronic Neck Pain (*k* = 6)	Fibromyalgia (*k* = 5)	Chronic Knee Pain (*k* = 5)	KQ1 Other Conditions* (*k* = 11)	Post-Surgical (*k* = 7)	KQ2 Other Conditions† (*k* = 4)
XR Technology	Virtual reality (VR)	6	5	-	1	5	2	2
	Augmented reality (AR)	16	1	5	4	6	5	2
Type of XR Intervention	Engage & guide in physical activity	16	6	4	5	5	6	3
	Pain psychology & coping skills	3	-	1	-	3	-	1
	Embodiment only	3	-	-	-	2	-	-
	Distraction only	-	-	-	-	1	1	-
Outcomes Reported	Pain-related functioning or interference	14	5	5	4	8	6	1
	Pain intensity or severity	21	6	3	5	9	6	3
	Adverse events	3	4	-	1	5	1	1
	Pain catastrophizing & kinesiophobia	11	3	-	-	2	-	-
	Pain global change	2	1	-	-	2	-	-
	Quality of life	4	1	4	2	4	3	2
	Physical performance	4	6	5	4	3	6	3
	Opioid dose/use	2	-	-	-	1	-	-
Sample Size	30-50	13	5	2	-	7	-	3
	51-100	8	1	3	5	4	4	1
	101-200	1	-	-	-	-	2	-
	>200	-	-	-	-	-	1	-
Country	USA	2	-	-	-	2	1	1
	Europe	4	2	3	2	5	3	2
	Middle East	5	2	1	2	-	1	1
	Asia	6	-	-	1	1	2	-
	Australia/New Zealand	2	2	-	-	2	-	-
	Others‡	3	-	1	-	1	-	-
Follow-Up Duration	7-30 days	7	1	-	-	3	-	-
	31-90 days	7	2	4	4	4	4	2
	>90 days	8	3	1	1	4	3	2

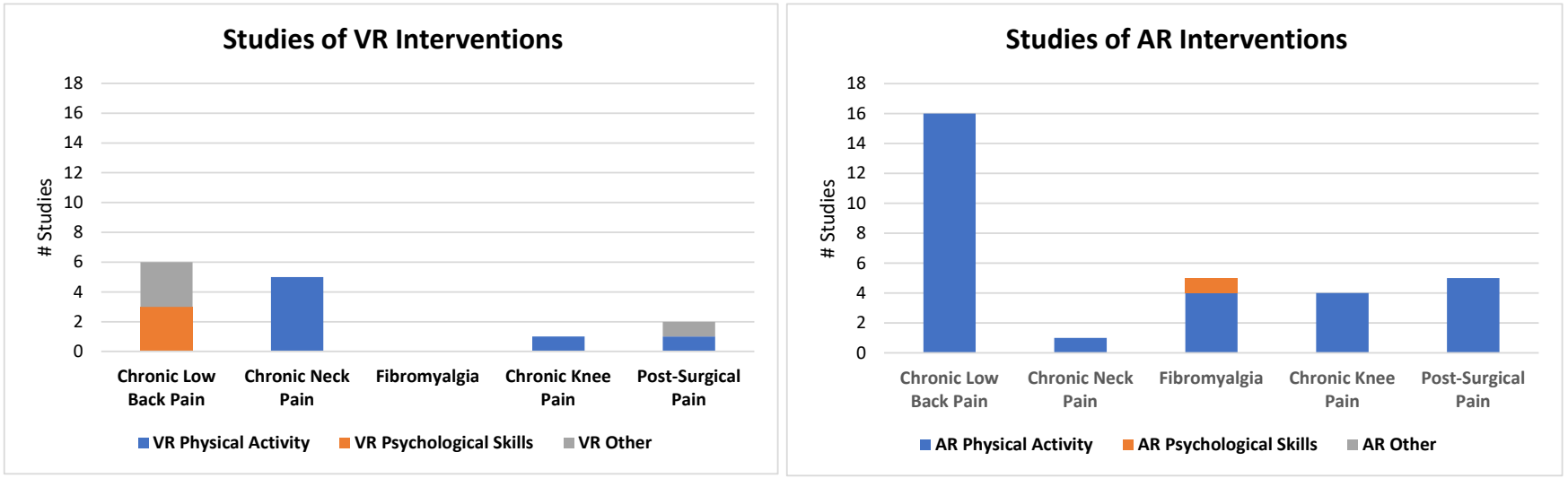
Mean/Median Age	<30	9	-	-	1	1	-	1
	30-64	11	6	4	4	7	1	1
	65+	1	-	-	-	1	5	2
	NR	1	-	1	-	2	1	-

* Includes 1 study each on the following conditions: headache/migraine, cancer-related neuropathic pain, musculoskeletal pain in ≥2 joints, pain related to metastatic breast cancer, complex regional pain syndrome, post-polio syndrome, rheumatoid arthritis, phantom limb pain, ankylosing spondylitis, frozen shoulder, & back pain + fibromyalgia

† Includes 2 studies on participants post-stroke, 1 study on flight-related neck pain (acute and chronic), & 1 on those work-related injury

‡ Includes 3 studies from Brazil and 2 studies from Nigeria

Figure 1. Studies Evaluating Various XR Intervention Types for Pain Conditions



CHRONIC LOW BACK PAIN (KQ1)

We identified 22 trials evaluating the use of VR (k = 6) or AR (k = 16) interventions for chronic low back pain. VR interventions were either primarily embodiment or psychological skills. In contrast, all AR interventions involved engagement and guidance in physical activity. There was variation in definitions of chronic low back pain, with the minimal duration ranging 2-6 months and not specified in some studies. Below, we first describe findings for VR interventions, and then results for AR studies. Detailed trial characteristics and findings are found in Appendix E.

VR Intervention Trials

Three studies evaluated VR embodiment,[24-26] and 3 examined VR psychological skills interventions for chronic low back pain.[27-33] Study characteristics and findings for VR interventions are summarized in Table 2. The largest study examined the efficacy of a VR psychological skills intervention (RelieVRx); it was conducted in the US and included 188 middle-aged participants (mean age 51 years), mostly women (67-78%). The remaining 5 studies were all small (total n = 30-46), included middle-aged women and men, and occurred outside the US. Below, we first describe results from studies evaluating VR embodiment, and then present findings for the VR psychological skills interventions.

VR Embodiment Interventions

The evidence is very uncertain on the effect of VR embodiment on pain-related functioning and pain intensity compared with conventional therapy (very low COE; Table 3), based on results from a single study. Yilmaz Yelvar, 2016[25] compared VR embodiment in addition to conventional therapy (n = 23) with conventional therapy alone (n = 23). The VR arm used iPods with video glasses for passive viewing of a virtual walking video clip. The study was rated as high RoB due to concerns about randomization, adherence to the intervention, measurement of outcomes, and the selection of reported results. Pain-related functioning was assessed with Oswestry Disability Index (ODI) at baseline and 2-week follow-up, with both groups improving at 2 weeks and little difference between groups (Diff Δ = 1.2; baseline mean ODI scores were 20.7 in VR group and 26.1 in the control group). Pain intensity was measured using a Visual Analog Scale (VAS) at baseline and 2 weeks, with the VR group having greater improvement at 2 weeks (Diff Δ = -2.8; baseline mean VAS scores were 6.0 in VR group and 5.6 in the control group). Similarly, there were greater improvements at 2 weeks in the VR group in quality of life, assessed with the Nottingham Health Profile (NHP), with Diff Δ of -12.5 (baseline mean NHP scores were 226 in VR group and 158 in the control group). VR group also had greater improvements in physical performance tests, including 6-minute walk and timed up and go (TUG). For example, the VR group had greater increases in distance for the 6-minute walk at 2 weeks (Diff Δ = 91.0 m, baseline mean 414.3 m in VR group and 401.1 m in the control group). Kinesiophobia was measured using the Tampa Scale of Kinesiophobia (TSK), but the authors did not report follow-up scores for the control group. Yilmaz Yelvar, 2016[25] did not report on adverse events.

Table 2. Summary of Findings for VR Interventions for Chronic Low Back Pain

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes			
				Pain-Related Functioning or Interference	**Pain Intensity or Severity**	**Pain Catastro-phizing & Kinesiophobia**	**Adverse Events & Other Eligible Outcomes**
VR Embodiment vs Physical Therapy							
Yilmaz Yelvar, 2016[25] High Turkey	LBP > 2 mo and no surgical treatments for disc herniation, spina bifida, or spinal stenosis; mean ages 46-53 yr, 46%-82% female	Viewing a virtual walking clip using iPod with video glasses, conventional therapy (hot pack, TENS, deep heat with ultrasound, and therapeutic exercises) *N* = 23 (22) Clinic; 2 wk	Conventional therapy (hot pack, TENS, deep heat with ultrasound, and therapeutic exercises) *N* = 23 (22) Clinic; 2 wk	**ODI** Baseline means (SD): Intervention: 20.7 (7.2) Comparator: 26.1 (11.0) Diff Δ (2 wk): 1.2*	**VAS** Baseline means (SD) Intervention: 6.0 (1.1) Comparator: 5.6 (2.4) Diff Δ (2 wk): -2.8*	**TSK** Baseline means (SD) Intervention: 43.7 (4.3) Comparator 1: 40.4 (5.6) Diff Δ (2 wk): NC	NR **Quality of life (2 wk)** Nottingham Health Profile **Physical performance (2 wk)** TUG 6MWT
VR Embodiment vs VR Other							
Harvie, 2022[24] Some concerns Australia	LBP ≥ 6 mo; mean ages 52-57 yr, 45%-50% female	Games using Oculus Rift S, participants embodied a boxer, superhero, and rock climber *N* = 20 (20) Clinic; 1 wk	Relaxing experiences using Oculus Rift S (*eg*, building a sandcastle, standing at beach) *N* = 10 (10) Clinic; 1 wk	NR	NRS Baseline means (CI)†: Intervention: 6.4 (5.8, 7.0) Comparator: 6.6 (4.9, 8.2) Diff Δ (1 wk): 0.5*	**Photograph Series of Daily Activities** Baseline (and follow-up) scores NR‡	NR Pain global change (1 wk) PGIC
Kammler-Sücker, 2023[26] Some concerns Germany	Back pain > 6 mo, excluding acute causes of back pain, neurological complications, or contraindication to exercise; mean ages 46-	Perform activities, movements guided by virtual doppelganger avatar in immersive environment (projections on 4 walls) *N* = 17 (17)	Perform activities, movements guided by a videotaped model, 2D projection on only 1 wall of same immersive environment	**NRS (limitations due to pain)** Baseline means (SD): Intervention: 2.7 (2.0) Comparator: 1.7 (1.8) Diff Δ (session 2-session 1, mean 14 days): -0.7*	**NRS (pain intensity)** Baseline means (SD): Intervention: 2.6 (1.1) Comparator: 1.9 (1.6) Diff D (session 2-session 1, mean 14 days): -0.4*	NR	NR NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning or Interference	Pain Intensity or Severity	Pain Catastro-phizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
	52 yr, 56%-71% female	Clinic; 4-117 days (3 sessions)	*N* = 16 (16) Clinic; 4-117 days (3 sessions)	Diff Δ (session 3-session 1, mean 27 days): -0.4*	Diff Δ (session 3-session 1, mean 27 days): -0.02*		
VR Psychological Skills							
Eccleston, 2022[27] Some concerns Finland	LBP ≥ 3 mo, with average pain intensity ≥ 4/10 over past week on NRS; mean ages 53-57 yr, 82-100% female	DTxP—using Oculus Quest headsets for program of 24 progressive modules informed by cognitive behavioral therapy *N* = 14 (14) Home; 6-8 wk	Using Oculus Quest headsets for seaside virtual environment *N* = 17 (17) Home; 6-8 wk Usual care *N* = 11 (11) NA; NA	**ODI** Baseline means (SD): DTxP: 36.0 (7.6) VR control: 37.2 (9.4) Usual care: 36.2 (7.6) Diff Δ * (DTxP-VR control): 8 wk: -8.5* Diff Δ * (DTxP-usual care): 8 wk: -3.8 **PROMIS pain interference** Baseline means (SD): DTxP: 64.5 (3.7) VR control: 63.1 (3.4) Usual care: 63.1 (2.5) Diff Δ * (DTxP-VR control): 8 wk: -5.0 Diff Δ * (DTxP-usual care): 8 wk:-3.3	**NRS** Baseline means (SD) DTxP: 6.0 (1.4) VR control: 6.1 (1.4) Usual care: 5.7 (1.6) Diff Δ * (DTxP-VR control): 8 wk: -0.6 Diff Δ * (DTxP-usual care): 8 wk: -0.6* **PROMIS pain intensity** Baseline means (SD) DTxP: 66.5 (4.1) VR control: 65.1 (5.4) Usual care: 63.0 (5.5) Diff Δ * (DTxP-VR control): 8 wk: -3.3 Diff Δ * (DTxP-usual care): 8 wk: -4.5	**TSK** Baseline means (SD) DTxP: 41.9 (4.4) VR control: 43.2 (6.0) Usual care: 42.5 (5.4) Diff Δ * (DTxP-VR control): 8-wk: -8.1 Diff Δ * (DTxP-usual care): 8 wk: -5.5	Severe AE: DTxP: 50% VR control: 29% Usual care: 36% Any treatment-related AE: DTxP: 25% VR control: 35% Usual care: 18% No serious AE in any group **Quality of life (8 wk)** EuroQoL-5D-5L **Pain Global** Change **(8 wk, 5 mo)** PGIC

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes			
				Pain-Related Functioning or Interference	**Pain Intensity or Severity**	**Pain Catastro-phizing & Kinesiophobia**	**Adverse Events & Other Eligible Outcomes**
Maddox, 2022[32]; Maddox, 2023[33]; Garcia, 2021[29]; Garcia,2022a[30]; Garcia, 2022b[31] Some concerns USA	LBP ≥ 6 mo, average pain intensity ≥4/10 for past month; mean age 51 yr, 67-78% female	RelieVRx—using Pico G2 4K headset for a progressive series of modules to provide pain neuroscience education, mindfulness exercises, and biofeedback *N* = 94 (89) Home; 8 wk	Rotation of 20 nature videos, displayed in 2D in Pico G2 4K headset *N* = 94 (90) Home; 8 wk	**DVPRS-overall interference** Baseline means (SD) Intervention: 4.8 (NR) Comparator: 5.1 (NR) Diff Δ *: 20 mo: -0.7 26 mo: -0.7	**DVPRS-pain intensity** Baseline means (SD) Intervention: 5.1 (1.2) Comparator: 5.2 (1.1) Diff Δ *: 20 mo: -1.0 26 mo: -0.5	NR	NR NR
Groenveld, 2023[28] Some concerns Netherlands	LBP ≥ 3 mo, average pain score ≥ 4/10 scale in the week preceding enrollment, no current treatment other than PT or medications, no invasive treatment in the past year; mean age 51-52 yr, 83% female	Using Oculus Go for pain education program consisting of 5 games based on psychotherapy principles (eg, acceptance and commitment therapy, mindfulness) *N* = 20 (20) Home; 4 wk	Wait-list control *N* = 21 (20) NA; NA	**BPI-Interference** Baseline means (SD): Intervention: 5.9 (1.7) Comparator: 6.3 (2.0) Diff Δ *: 4 wk: -0.3 4 mo: -0.2 ODI Baseline means (SD): Intervention: 40.1 (19.1) Comparator: 42.8 (18.1) Diff D*: 4 wk: -3.7 4 mo: -0.9	**VAS-worst pain** Baseline means (SD): Intervention: 6.1 (NR) Comparator: 7.0 (NR) Diff Δ *: 30 days: -0.6 **VAS-least pain** Baseline means (SD) Intervention: 3.7 (NR) Comparator: 4.1 (NR) Diff Δ *: 30 days: -0.4	**PCS** Baseline means (SD): Intervention:21.7 (12.2) Comparator 1: 24.7 (7.8) Diff Δ *: 4 wk: 0.7 4 mo: -0.2	Dizziness reported for VR group only: 15% Quality **of life (4 wk, 4 mo)** SF-12 physical SF-12 mental **Opioid use (1, 4 wk)** Using at least weekly

Notes. * Diff Δ calculated by review team.

† Data abstracted from graphs using Plotdigitizer.

‡ No scores were reported and assessment only completed by 16 of 30 (53%) participants.

Abbreviations. 2D=2 dimensional; 6MWT=6-minute walk test; AE=adverse event; BPI=Brief Pain Inventory; CI=confidence interval; Diff Δ= Difference between groups of mean change scores; DVPRS=Defense and Veterans Pain Rating Scale; LBP=lower back pain; mo=month; NC=not calculable; NR=not reported; NRS=Numeric Rating Scale; ODI=Oswestry Disability Index; PCS=Pain Catastrophizing Scale; PGIC=Pain Global Impression of Change scale; PROMIS=Patient-Reported Outcomes Measurement Information System; PT=physical therapy; SF-12=Short Form Health Survey-12; SD=standard deviation; TENS=transcutaneous electrical nerve stimulation; TSK=Tampa Scale of Kinesiophobia; TUG=timed up and go; VAS=Visual Analogue Scale; wk=week; yr=year.

Table 3. Certainty of Evidence: VR Embodiment versus Conventional Therapy for Chronic Low Back Pain

Outcome Outcome Measure(s)	Follow-Up # of Participants and Studies	Anticipated Absolute Effects on Mean Change (Δ)			Certainty	What Happens
		VR Embodiment	Conventional Therapy	Diff. Δ		
Pain-Related Functioning or Interference ODI	2 weeks *N* = 46 1 RCT[25]	-3.8	-5.1	**1.2**	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of VR-embodiment on pain-related functioning, compared with conventional therapy.
Pain Intensity or Severity VAS	2 weeks *N* = 46 1 RCT[25]	-3.5	-0.7	**-2.8**	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of VR-embodiment on pain intensity, compared with conventional therapy.

Notes. GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
Explanations:
a. Downgraded 2 levels for study limitations (study rated high for risk of bias).
b. Downgraded 1 level for indirectness (age in study not generalizable to back pain population).
Abbreviations. CI=confidence interval; Diff. Δ=difference in mean change scores; ODI=Oswestry Disability Index; RCT=randomized controlled trial; VAS=Visual Analog Scale; VR=virtual reality.

The evidence is also very uncertain on the effect of VR embodiment on pain-related functioning and pain intensity, compared with other VR experiences (very low COE; Table 3), based on results from 2 studies.[24,26] Neither study reported findings on adverse events. The first study, Kammler-Sücker, 2023,[26] compared the effect of using a virtual doppelganger avatar to guide participants in various activities in a fully immersive environment (n = 17) with a videotaped model performing the same activities and projected on 1 wall of the immersive environment (n = 16). Outcomes were assessed during each of 3 sessions for both groups (occurring over 4-117 days). This study was rated as some concerns for RoB due to concerns about randomization methods and measurement of the outcomes. Pain-related interference was assessed with a Numeric Rating Scale (NRS) 0-10 for how much participants were limited in movements by their pain (10 = "complete incapability"), with both groups having similar reductions comparing session 3 with session 1 (Diff Δ = -0.4, session 1 mean NRS scores were 2.7 in VR embodiment group and 1.7 in the VR other group). Similarly, there were no differences between groups in pain intensity, as assessed with NRS for pain during movement comparing session 3 with session 1 (Diff Δ = -0.02, session 1 mean NRS scores were 2.6 in VR embodiment group and 1.9 in the VR other group).

In the second study, Harvie, 2022[24] used Oculus Rift S to create visual experiences and activities consistent with a boxer, superhero, or rock climber (*ie*, embodiment, n = 20), and compared this with vacation experiences (*eg*, standing at a beach) also using Oculus Rift S (n = 10). This study was rated as some concerns for RoB due to concerns about randomization methods and measurement of the

outcomes. Harvie, 2022[24] did not evaluate pain-related functioning or adverse events. Average pain intensity over the past week was assessed using NRS, measured at baseline and 1-week follow-up; both VR groups had very slight improvements in pain intensity (Diff Δ = 0.5; baseline mean NRS scores were 6.4 in VR group and 6.6 in the control group).[24] At 1-week follow-up, Pain Global Impression of Change (PGIC) was also assessed, with the VR-embodiment arm having greater proportion reporting at least minimal improvement (37% vs 11% for VR control), but also slightly more with minimal worsening (5% vs 0% for VR control). Fear of movement was assessed using Photograph Series of Daily Activities, which participants rated perceived harmfulness from 0-100; however, these assessments were only completed by 53% of participants and no baseline or follow-up scores were reported.

Table 4. Certainty of Evidence: VR Embodiment versus VR Other for Chronic Low Back Pain

Outcome Outcome Measure(s)	Follow-Up # of Participants and Studies	Anticipated Absolute Effects on Mean Change (Δ)			Certainty	What Happens
		VR Embodiment	VR Other	Diff. Δ		
Pain-Related Functioning or Interference NRS	27 days *N* = 30 1 RCT[26]	-0.4	0.0	**-0.4**	⊕◯◯◯ Very low[a,b,c]	The evidence is very uncertain on the effect of VR embodiment on pain-related functioning compared to VR control.
Pain Intensity or Severity NRS	1 wk – 27 days *N* = 63 2 RCTs[24,26]	-0.1*	-0.1*	**-0.02***	⊕◯◯◯ Very low[a,b,c]	The evidence is very uncertain on the effect of VR embodiment on pain intensity compared to VR control.

Notes. * Values for mean change in VR embodiment, VR other, and Diff. Δ taken from Kammler-Sucker.[26]

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations:

a. Downgraded 1 level for study limitations (studies rated some concerns for risk of bias).

b. Downgraded 1 level for indirectness (age in study not generalizable to back pain population).

c. Downgraded 1 level for imprecision (SDs large relative to means).

Abbreviations. CI=confidence interval; Diff. Δ=difference in mean change scores; NRS=Numeric Rating Scale; RCT=randomized controlled trial; VR=virtual reality; wk=week.

VR Psychological Skills Intervention Trials

VR psychological skills, compared to VR control or usual care, may result in better pain-related functioning and greater reductions in pain intensity (low COE), but the evidence on adverse effects is very uncertain (very low COE; Table 5). Three studies (reported in 7 articles[27-33]) evaluated VR psychological skills, compared with either VR control or usual care. VR psychological skills interventions occurred over 4-8 weeks and participants were predominantly middle-aged women. The range of follow-up was 2-26 months. Maddox, 2022[29,31-34] evaluated a psychological skills program involving progressive modules on pain neuroscience education, mindfulness, and biofeedback

(RelieVRx; *n* = 94) against a VR control of 20 2D nature videos, shown in the same Pico G2 4K headsets (*n* = 94). Eccleston, 2022[27] compared a VR program involving 24 modules and 30 unique sessions (*n* = 14), with a VR control of a relaxing seaside in the same Oculus Quest headsets (*n* = 17) and usual care (*n* = 11). Finally, Groenveld, 2023[28] evaluated a VR psychological skills program of 5 games informed by psychotherapy principles (*n* = 20) against a wait-list control (*n* = 21). These studies were all rated some concerns for RoB, due to concerns regarding adherence to the intervention and/or outcome measurement. All studies assessed pain-related functioning and pain intensity, 2 studies examined pain catastrophizing or kinesiophobia, and 2 reported on adverse events.

Pain-Related Functioning or Interference

A variety of measures were used to assess pain-related functioning or interference, including the Defense and Veterans Pain Rating Scale (DVPRS), ODI, the Brief Pain Inventory (BPI) interference subscale, and the Patient Reported Outcomes Measurement Information System (PROMIS) pain interference measure. Overall, there was greater improvement in the VR psychological skills group, although some of the between-group differences were quite small (Diff Δ range = -0.2 to -8.5 on various scales). In the study with the longest follow-up, the Diff Δ at both 20 and 26 months was -0.7, as assessed with DVPRS-overall interference (baseline mean scores were 4.8 for VR psychological skills and 5.1 for VR control).[32,33]

Pain Intensity or Severity

Similarly, the VR psychological skills group had greater improvement in pain intensity, but also with a large range in Diff Δ (-0.4 to -4.5, on various scales). There appeared to be sustained effects of VR psychological skills, with greater reductions in pain intensity still at 20 and 26 months (Diff Δ of -1.0 and -0.5 at 20 and 26 months, using the DVPRS-pain intensity subscale; baseline mean scores were 5.1 for VR psychological skills and 5.2 for VR control).[32,33] Other measures of pain intensity included the NRS and PROMIS pain intensity,[27] and daily worst and least pain VAS.[28]

Adverse Events

Eccleston, 2022[27] reported a range of adverse events, finding for example that severe adverse events (defined as symptoms leading to inability to perform daily or work activities) were relatively common: 50% of participants experienced any severe events in the VR psychological skills group, 29% in VR control, and 36% in usual care. No serious adverse events (defined as any event leading to death or serious deterioration in health) were reported in any group.[27] Groenveld, 2023[28] only assessed for dizziness in the VR psychological skills group, reporting 3 participants (15%) had dizziness but none discontinued the intervention due to symptoms.

Pain Catastrophizing and Kinesiophobia

Two studies assessed pain catastrophizing or kinesiophobia, showing inconsistent results. Eccleston, 2022[27] used TSK to assess kinesiophobia and found that the VR psychological skills group had greater reductions, compared to VR control (Diff Δ -8.1) or usual care (Diff Δ -5.5); baseline mean scores were 41.9 for VR psychological skills, 43.2 for VR control, and 42.5 for usual care. In contrast, Groenveld, 2023[28] used the Pain Catastrophizing Scale (PCS) and found that this worsened for all groups, with slightly greater catastrophizing in the VR psychological skills group at 4 weeks (Diff Δ 0.7) and slightly worse scores in usual care group at 4 months (Diff Δ -0.2); baseline mean PCS scores were 21.7 for VR psychological skills and 24.7 for usual care group).

Other Outcomes

Eccleston, 2022[27] assessed the European Quality of Life 5-dimension, 5-level scale (EuroQoL-5D), but only reported domain scores (no index score given). Groenveld, 2023[28] also evaluated quality of life, using instead the 12-item Short Form Survey (SF-12). SF-12 physical component scores improved somewhat for both groups, with little difference between groups (Diff Δ -0.3 at 4 months, baseline mean scores were 34.9 for VR psychological skills and 32.9 for usual care), but SF-12 mental component score improved more for the usual care group (Diff Δ -2.4 at 4 months, baseline mean scores were 45.6 for VR psychological skills and 43.0 for usual care).[28]

Eccleston, 2022[27] assessed global improvement using PGIC, reporting mean scores at 8 weeks and 5 months (VR psychological skills and VR control only at the later time point). There were no differences between groups at 5 months (mean score 3.0 for both VR psychological skills and VR control), and somewhat worse for the VR psychological skills group at 8 weeks (mean score 2.7 for VR psychological skills, 3.8 for VR control, and 3.9 for usual care).[27]

Groenveld, 2023[28] evaluated the proportion of participants reporting use of opioids at least weekly, finding that the VR psychological skills group had fewer participants using opioids at 4 weeks compared with at 1 week (28% vs 47%), while no change was found in the usual care group (37% at both time points).

Table 5. Certainty of Evidence: VR Psychological Skills versus VR Control or Usual Care for Chronic Low Back Pain

Outcome Outcome Measure(s)	Follow-Up # of Participants and Studies	Anticipated Absolute Effects			Certainty	What Happens
		VR Psychological Skills	Comparator	Difference		
Pain-Related Functioning or Interference BPI; ODI; DVPRS; PROMIS	1–26 mo *N* = 271 3 RCTs[27,28,33]	-2.0*	-1.3*	**Diff Δ: -0.7***	⨁⨁◯◯ Low[a,b]	VR psychological skills may result in better pain-related functioning, compared to VR control or usual care.
Pain Intensity or Severity DVPRS; NRS; PROMIS; VAS;	1–26 mo *N* = 271 3 RCTs[27,28,33]	-1.2*	-0.7*	**Diff Δ: -0.5***	⨁⨁◯◯ Low[a,b]	VR psychological skills may result in decreased pain intensity, compared to VR control or usual care.
Adverse Events Severe Adverse Events	2–8 wk *N* = 83 2 RCTs[27,28]	50%†	32%†	**18% more†**	⨁◯◯◯ Very low[a,c]	The evidence is very uncertain about the effect of VR psychological skills on adverse events, compared to VR control or usual care.

Notes. * Mean change in DVPRS for each group from baseline to for 26 mo, as reported in Maddox, 2023[32] and Diff. Δ calculated by review team.

† Proportion with severe adverse event (symptoms leading to inability to perform daily or work activities) in each group at 8 weeks as reported in Eccleston, 2022[27] and differences between groups calculated by review team.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of

the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. Downgraded 1 level for study limitations (study rated some concerns for risk of bias).

b. Downgraded 1 level for indirectness (women overrepresented in study population, 67-100%).

c. Downgraded 2 levels for indirectness (women overrepresented; adverse event data only reported for one arm in 1 study).

Abbreviation. BPI=Brief Pain Inventory; CI=confidence interval; Diff. Δ=difference in mean change scores; DVPRS=Defense and Veterans Pain Rating Scale; mo=month; no.=number; ODI=Oswestry Disability Index; PROMIS=Patient-Reported Outcomes Measurement Information System; RCT=randomized controlled trial; VR=virtual reality; wk=week.

AR Interventions for Back Pain

Sixteen studies evaluated AR interventions for chronic low back pain and all involved guiding and engagement in physical activity.[35-50] Study characteristics and findings for AR interventions are summarized in Table 6. Below, we first present results from studies evaluating AR physical activity interventions versus active comparators. Then, we summarize findings from trials comparing AR physical activity with usual care.

AR Physical Activity versus Active Comparator

Twelve studies compared AR physical activity to non-AR physical activity [35-39,41-43,46,48,50] or new medications.[40] All of these studies were small (n = 30-90). Five AR studies used the Nintendo Wii or Microsoft Kinect,[36,39,40,48,50] and 7 used other devices for AR physical activity.[35,37,38,41,42,43 ,46] Four studies in this latter group were conducted by the same group (lead author Gopal Nambi) and evaluated the Prokin system (TecnoBody); these were all conducted in Saudi Arabia with young male participants.[35,37,41,42] Other studies were conducted in Korea (k = 2),[43,48] Nigeria (k = 2),[36,50] Brazil,[46] China,[39] Japan,[40] and Pakistan.[38] These studies included a range of young and middle-aged or older adults, and 1 included only women.[48] Most studies (k = 9) were rated high for RoB, due to concerns about randomization, deviation from the intended intervention, missing outcome data, and/or selective reporting bias. The remaining 4 studies were rated some concerns for RoB.[35,37,38,45] Six studies evaluated pain-related functioning and most (k = 11) assessed pain intensity. No study reported adverse events.

Pain-Related Functioning or Interference

The evidence is very uncertain on the effect of AR physical activity on pain-related functioning or interference, compared with non-AR physical activity (very low COE; Table 7). Six studies assessed pain-related functioning using the ODI, the Modified ODI (MODI), and/or the Roland-Morris Disability Questionnaire (RMDQ).[36,38,39,43,48,50] Due to lack of reported results on SD of mean change scores, we were unable to conduct meta-analyses. In general, outcomes were assessed post-intervention (2-8 weeks from baseline) and were inconsistent across studies. Some showed that AR physical activity had greater improvement in pain-related functioning (*eg*, Diff Δ -1.73[43] and -9.46[48]), while others found that that the non-AR physical activity group had greater improvement (*eg*, Diff Δ 3.68 and 1.23).[36,39]

Pain Intensity or Severity

The evidence is also very uncertain on the effect of AR physical activity on pain intensity or severity, compared with non-AR physical activity or medication (very low COE; Table 7). Eleven studies

assessed pain intensity using either VAS or NRS.[35,37-43,46,48,50] We included results from 10 studies that evaluated AR physical activity interventions lasting 4-8 weeks (Figure 2).[35,37,38,40-43,46,48,50] The pooled estimate indicated AR physical activity had greater reductions in pain intensity but the CI was quite large (standardized Diff Δ -0.7 [-1.2, -0.2]). Heterogeneity was also substantial. The only study not included in the meta-analysis was a 3-arm trial comparing AR physical activity with non-AR exercises or thermal magnetic therapy, all lasting only 2 weeks.[39]

Table 6. Summary of Findings for AR Interventions for Chronic Low Back Pain

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
AR Physical Activity vs Active Comparator							
Afzal, 2022[38] Some concerns Pakistan	History of chronic LBP, excluding congenital deformity, history of trauma, fracture of spine or lower extremity, systemic or neurologic disease, on steroids or pregnant; mean age 38 yr, 67% women	Sensor enabled games involving various movements (eg, trunk flexion; jumping, arm, leg, head movements) N = 45 (42) Clinic; 4 wk	Conventional physical therapy N = 45 (42) Clinic; 4 wk	**MODI** Baseline means (SD): Intervention—69.16 (9.13) Comparator—65.08 (8.94) Diff Δ (4 wk): -28.6*	**VAS** Baseline means (SD): Intervention—6.50 (1.24) Comparator —6.62 (1.04) Diff Δ (4 wk): -2.2*	NR	NR NR
Fatoye, 202236 High Nigeria	LBP ≥ 3 mo, as determined by McKenzie Institute Lumbar Spine Assessment Algorithm; mean ages 48-49 yr, % women NR	Games using Microsoft Kinect where participants headed virtual balls N = 29 (22) NR; 8 wk	Conventional physical therapy (McKenzie Protocol) N = 28 (24) NR; 8 wk	**ODI** Baseline means (SD) Intervention: 14.23 (9.41) Comparator: 21.12 (10.68) Diff Δ (8 wk): 3.68*	NR	NR	NR NR
Kim, 2014[48] High Republic of Korea	Chronic LBP; mean ages 44-50 yrs, 100% women	30-minutes of yoga using Nintendo Wii Fit N = 15 (15) Clinic; 4 wk	Trunk-stabilizing exercises N = 15 (15) Clinic; 4 wk	**ODI** Baseline means (SD): Intervention: 34.91 (6.19) Comparator: 36.18 (5.02) Diff Δ (4 wk): -9.46*	**VAS** Baseline means (SD) Intervention: 7.00 (0.89) Comparator: 6.95 (0.79) Diff Δ (4 wk): -2.41*	**FABQ** Baseline means (SD) Intervention: 65.46 (9.64) Comparator : 70.82 (4.58) Diff Δ (4 wk): -18.73*	NR NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
				RMDQ Baseline means (SD): Intervention: 18.64 (2.84) Comparator: 19.09 (2.91) Diff Δ (4 wk): -4.73*			
Kim, 2020[43] High Korea	LBP ≥ 3 mo and average pain intensity ≥4/11 on tNRS, excluding those with serious medical conditions or physical limitations; mean ages 26-29, 32%-58% women	Simulated horseback riding system, consisting of walking, slow trotting, and fast trotting N = 24 (16) Clinic; 4 wks	Stabilization exercises using Redcord suspension including supine pelvic lift, supine/prone bridging, and side-lying hip abduction N = 24 (15) Clinic; 4 wks	**ODI** Baseline means (SD): Intervention: 20.24 (7.69) Comparator: 21.77 (7.11) Diff Δ *: 8 wk: -1.73 6 mo: 0.58 **RMDQ** Baseline means (SD): Intervention: 7.00 (4.40) Comparator: 5.11 (2.74) Diff Δ *: 8 wk: -1.45 6 mo: -1.81	**NRS** Baseline means (SD): Intervention: 4.70 (1.04) Comparator: 4.73 (0.82) Diff Δ *: 8 wk: -0.34 6 mo: 0.23	**FABQ-physical** Baseline means (SD): Intervention: 15.35 (4.12) Comparator: 11.93 (5.62) Diff Δ *: 8 wk: -5.72 6 mo:-8.94 **FABQ-work** Baseline means (SD): Intervention: 17.11 (5.33) Comparator: 20.47 (7.89) Diff Δ *: 8 wk: 0.07 6 mo: 0.71	NR NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
Li, 2021[39] High China	LBP > 3 mo, excluding those with specific lumbar pathologies, obesity, or other medical conditions; mean ages 22 – 25 yrs, 64%-83% female	Fruit Ninja using Nintendo Wii or game using Microsoft Kinect Xbox where participants wave their hands without moving their trunk N = 11 (11) Home, clinic; 2 wk	Motor control exercises (MCE), ultrasound-guided abdominal maneuver N = 12 (12) Home, clinic; 2 wk Conventional thermal magnetic therapy (CTMT) N = 11 (11) Home, clinic; 2 wk	**ODI** Baseline means (SD) Intervention: 15.65 (6.39) MCE: 18.42 (9.36) CTMT:12.72 (4.84) Diff Δ * (intervention-MCE): 2 wk: 1.23 Diff Δ * (intervention-CTMT): 2 wk: 0.28	**VAS** Baseline means (SD) Intervention: 4.36 (1.36) MCE: 4.58 (1.83) CTMT: 3.64 (1.36) Diff Δ * (intervention-MCE): 2 wk: 1.25 Diff Δ * (intervention-CTMT): 2 wk: 0.21	NR	NR NR
Mbada, 2019[50] High Nigeria	Long-term mechanical LBP and directional preference for extension; mean ages 33-49 yrs, 49%-79% female	Microsoft Kinect game where patients head balls with feet stationary but moving their head and trunk N = 28 (22) NR; 8 wk	Physical therapy (McKenzie extension protocol in standing) N = 29 (24) NR; 8 wk	**ODI** Baseline means (SD): Intervention: 18.7 (NR) Comparator: 27.8 (NR) Diff Δ (8 wk): 4.5* **RMDQ** Baseline means (SD): Intervention: 6.8 (4.9) Comparator: 11.3 (4.6) Diff Δ (8 wk): 1.3*	**VAS** Baseline means (SD): Intervention: 4.1 (1.8) Comparator: 5.0 (1.9) Diff Δ (8 wk): 0.9*	**TSK** Baseline means (SD) Intervention: 16.5 (NR) Comparator: 29.9 (NR) Diff Δ (8 wk): 23.5* **FABQ-physical** Baseline means (SD) Intervention: 21.0 (NR) Comparator: 25.8 (NR) Diff Δ: NC **FABQ-work** Baseline means (SD) Intervention: 21.2 (NR)	NR **Quality of life (8 wk)** SF-12 mental and physical **Physical performance (8 wk)** Biering-Sorensen test of Statis Muscular Endurance (BSME)

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes			
				Pain-Related Functioning	**Pain Intensity/Severity**	**Pain Catastrophizing & Kinesiophobia**	**Adverse Events & Other Eligible Outcomes**
						Comparator: 25.7 (NR) Diff Δ (8 wk): 6.4*	
Monteiro-Junior, 2015[46] High Brazil	Chronic LBP, not participating in a systematic exercise program; mean age 68 yr, 100% women	8 exercises using Nintendo Wii and Wii Balance Board (*eg*, tightrope walk, ski slalom, balance bubble, lunge), and core and strength exercises. N = 17 (14) Clinic; 8 wks	Core and strength exercises N = 17 (16) Clinic; 8 wks	NR	**NRS** Baseline means (SD) Intervention: 6.5 (1.1) Comparator: 6.6 (1.2) Diff Δ (8 wk): 0.4*	NR	NR **Physical** performance **(8 wk)** Sit-to-stand test
Nambi, 2020a[37] Some concerns Saudi Arabia	Male university soccer players aged 18-25 with chronic LBP and pain intensity 4-8 on the VAS; mean age NR	Using the ProKin system, participants control the game by moving their trunk, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 15 (15) Home, clinic; 4 wk	Isokinetic dynamometer to perform extension and flexion exercises, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 15 (15) Home, clinic; 4 wk --- Conventional balance training, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 15 (15)	NR	VAS Baseline means (SD) Intervention: 7.1 (0.6) Isokinetic dynamometer: 7.3 (0.5) Balance training: 7.3 (0.6) Diff Δ * (intervention-isokinetic dynamometer): 4 wk: -0.7 8 wk: -1.3 6 mo: -0.9 Diff Δ * (intervention-balance training): 4 wk: -2.1 8 wk: -4.0	NR	NR **Physical performance (4 wk, 8 wk, 6 mo)** 40 m sprint 4x5 spring Submaximal shuttle running Countermovement jump Squat jump

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
			Home, clinic; 4 wk		6 mo: -3.2		
Nambi, 2020b[42] High Saudi Arabia	Male university students with chronic LBP ≥ 3 mo and pain intensity 4-8 on the VAS, excluding serious medical conditions or participation in other weight and balance training programs; mean age 23 yr	Using the ProKin system, participants control the game by moving their trunk, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 20 (19) Home; 4 wk	Isokinetic dynamometer to perform extension and flexion exercises, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 20 (20) Home; 4 wk Conventional balance training, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 20 (19) Home; 4 wk	NR	**VAS** Baseline means (SD): Intervention: 7.5 (0.4) Isokinetic dynamometer: -7.3 (0.3) Conventional training: -7.4 (0.4) Diff Δ * (intervention-isokinetic dynamometer): 4 wk: -0.5 6 mo: -0.5 Diff Δ * (intervention-conventional training.): 4 wk: -2.2 6 mo: -3.0	**TSK** Baseline means (SD): Intervention: 57.52 (4.8) Conventional training: 58.11 (4.5) Conventional training: 57.93 (4.3) Diff Δ * (intervention-isokinetic dynamometer): 4 wk: -0.5 6 mo: -0.5 Diff Δ * (intervention-conventional training.): 4 wk: -19.4 6 mo: -18.1	NR NR
Nambi, 2021[41] High Saudi Arabia	Male university American soccer players with chronic LBP ≥ 3 mo and pain intensity 4-8 on the VAS,	Using the ProKin system, participants control the game by moving their trunk, and home-based exercise	Isokinetic dynamometer to perform extension and flexion exercises, and home-based exercise protocol,	NR	**VAS** Baseline means (SD) Intervention: 7.8 (0.6)	**TSK** Baseline means (SD) Intervention: 56.45 (3.2)	NR NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
	excluding serious medical conditions or participation in other weight and balance training programs; mean ages 21-22 yr	protocol, hot pack therapy and ultrasound therapy N = 18 (18) Home; 4 wk	hot pack therapy and ultrasound therapy N = 18 (18) Home; 4 wk Conventional balance training, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 18 (18) Home; 4 wk		Isokinetic dynamometer: -7.5 (0.5) Conventional training: -7.6 (0.4) Diff Δ * (intervention-isokinetic dynamometer): 4-wk: -2.1 6-mo: -2.6 Diff Δ * (intervention-conventional training.): 4 wk: -2.2 6 mo: -2.9	Isokinetic dynamometer: 58.02 (3.8) Conventional training: 57.68 (4.1) Diff Δ * (intervention-isokinetic dynamometer): 4-wk: -8.7 6-mo: -8.0 Diff Δ * (intervention-conventional training.): 4 wk: -15.9 6 mo: -17.8	
Nambi, 2022[35] Some concerns Saudi Arabia	Male soccer players with chronic LBP ≥ 3 mo and pain score 4-8 on 10 cm VAS; mean age NR	Using the ProKin system, participants control the game by moving their trunk, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 20 (19) Clinic; 3 wk	Isokinetic dynamometer to perform extension and flexion exercises, and home-based exercise protocol, hot pack therapy and ultrasound therapy 20 (19) Clinic; 3 wk	NR	VAS Baseline means (SD) Intervention: 7.2 (0.4) IKT: 7.3 (0.3) In-person training: 7.2 (0.3) Diff Δ * (intervention-IKT): 4 wk: -0.6	NR	NR NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
			Conventional balance training, and home-based exercise protocol, hot pack therapy and ultrasound therapy N = 20 (20) Clinic; 3 wk		Diff Δ * (intervention-in-person training): 4 wk: -3.0		
Sato, 2021[40] High Japan	LBP ≥ 3 mo, had not responded to conservative treatment; mean age 49 yr, 55% female	Games on Nintendo Wii Fit where participants control the character by jogging and squatting N = 20 (20) NR; 8 wk	New medications in order: NSAIDS, tramadol, and duloxetine N = 20 (20) Clinic; 8 wk	NR	**VAS** Baseline means (SD) Intervention: RE (1.99) Comparator: 7.01 (0.93) Diff Δ * 8 wk: -2.21	**PCS** Baseline means (SD) Intervention: 43.50 (7.97) Comparator: 40.77 (RE) Diff Δ * 8 wk: -4.95 **TSK** Baseline means (SD) Intervention: 42.50 (5.94) Comparator: 38.92 (5.35) Diff Δ * 8 wk: -0.11	NR NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
AR Physical Activity vs Usual Care							
Oh, 2014[49] High Korea	LBP > 3 mo; mean ages 20-21 yr, 100% men	10 minutes on horse simulator machine daily (10-min HSM) N=10 (10) Clinic; 8 wk 20 minutes horse simulator machine daily (20-min HSM) N = 9 (9) Clinic; 8 wk 30 minutes horse simulator machine daily (30-min HSM) N = 9 (9) Clinic; 8 wk	Usual care N = 9 (9) Clinic; 8 wk	NR	**VAS** Baseline means (SD) 10-min HSM: 3.8 (0.5) 20-min HSM: 4.9 (0.5) 30-min HSM: 5.6 (0.7) Usual care: 3.1 (0.6) Diff Δ * (10-min HSM) 8 wk: -3.0 Diff Δ * (20-min HSM) 8 wk: -4.5 Diff Δ * (30-min HSM) 8 wk: -2.8	NR	NR NR
Thomas, 2016[45] Some concerns US	LBP category 1-3 (Classification System of the Quebec Task Force on Spinal Disorders), excluding recent LBP onset, low kinesiophobia, or on current treatment for LBP; mean ages 24-27 yr, 46% female	Dodgeball game displayed on 60-inch, high definition 3D-TV N = 27 (26) Clinic; 3 days	Usual care N = 26 (26) NR; 3 days	**RMDQ** Baseline means SD) Intervention: 4.8 (3.0) Comparator: 5.3 (3.9) Diff Δ * 4 days: NC	**VAS** Baseline means (SD) Intervention: 21.1 (10.3) Comparator: 25.2 (16.7) Diff Δ * 4 days: NC	**TSK** Baseline means (SD) Intervention: 38.9 (4.1) Comparator: 38.3 (4.6) Diff Δ * 4 days: NC	No adverse events in either group NR

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting/Duration	Comparator N Randomized (N Analyzed) Setting/Duration	Outcomes			
				Pain-Related Functioning	Pain Intensity/Severity	Pain Catastrophizing & Kinesiophobia	Adverse Events & Other Eligible Outcomes
Yoo, 2014[47] Some concerns Korea	Young men with LBP > 3 mo, excluding history of neurological, hypertension, cardiopulmonary diseases, chronic disease or spine surgery; mean ages 20-21 yrs	Exercises using a horse simulator machine consisting of a warm-up, workout, and cool-down N = 24 (24) Clinic; 8 wk	Usual care N = 23 (23) NA; 8 wk	NR	**VAS** Baseline means (SD) Intervention: 4.37 (2.13) Comparator: 1.50 (0.15) Diff Δ * 8 wk: -1.65	NR	NR NR
Zadro, 2019[44] Some concerns Australia	Mechanical LBP ≥3 mo, excluding serious pathology in the spine, cognitive limitations, high risk of falls; mean ages 68-69 yr, 52% women	Wii Fit U game consisting of flexibility, bodyweight, and aerobic exercises N = 30 (28) Home; 8 wk	Usual care N = 30 (30) Home; 8 wk	**RMDQ** Baseline means (SD) Intervention: 6.3 (4.8) Comparator: 7.4 (5.2) Diff Δ (8 wk): -0.4* **PSFS** Baseline means (SD) Intervention: 5.3 (1.4) Comparator: 4.3 (2.1) Diff Δ (8 wk): 0.7*	**NRS** Baseline means (SD) Intervention: 5.2 (1.6) Comparator: 4.8 (1.7) Diff Δ (8 wk): -1.0*	**TSK** Baseline means (SD) Intervention: 33.6 (6.1) Comparator: 34.7 (5.8) Diff Δ (8 wk): -2.5*	No adverse events in either group NR

Notes. * Diff Δ calculated by review team, unable to calculate standardized Diff Δ.

Abbreviations. 3D-TV=3-dimensional television; AEs=adverse event; CI=confidence interval; CLBP=chronic low back pain; CM=centimeter; CTMT=conventional thermal magnetic therapy; Diff Δ = difference in change scores; FABQ=Fear Avoidance Belief Questionnaire; HSM=horse simulator machine; IKT=isokinetic training; LBP=low back pain; mo=month; MODI=Modified Oswestry Disability Index; NC=not calculable; NR=not reported; NRS=Numeric Rating Scale; NSAIDS=non-steroidal anti-inflammatory drugs; ODI=Oswestry Disability Index; PCS=Pain Catastrophizing Scale; PSFS=Patient Specific Functional Scale; PT=physical therapy; RMDQ=Roland-Morris Disability Questionnaire; RE=reporting error; RoB=risk of bias; SF-12=12-item Short Form health survey; SD=standard deviation; TSK=Tampa Scale of Kinesiophobia; US=United States of America; VAS=Visual Analogue Scale; wk=week; yr=year.

Figure 2. AR Physical Activity versus Active Comparator: Pain Intensity 4-8 Weeks

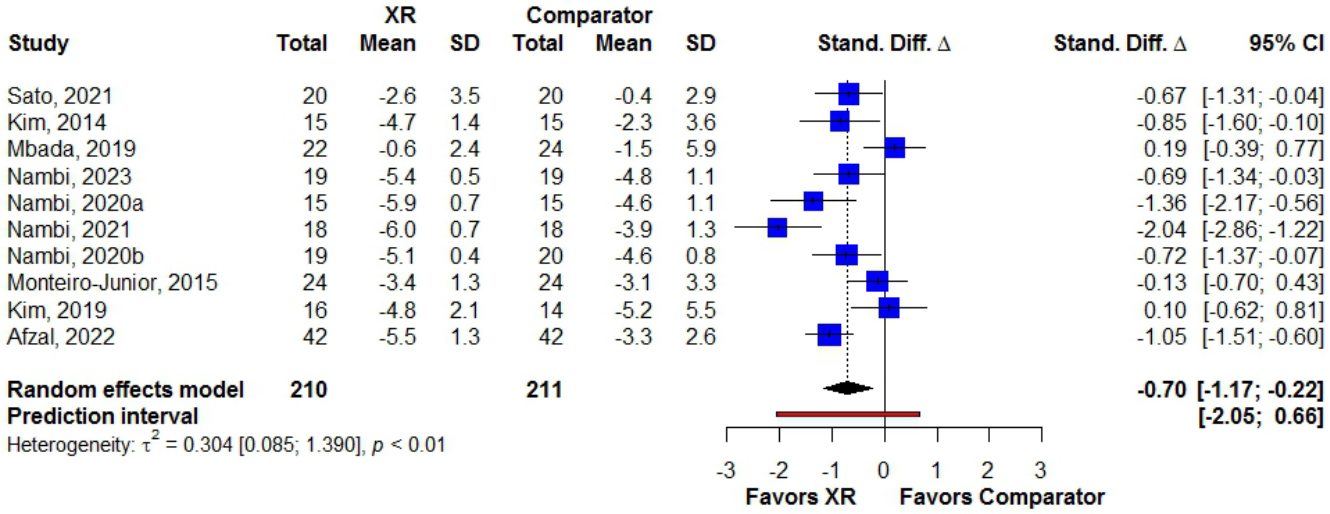


Table 7. Certainty of Evidence: AR Physical Activity vs. Active Comparators for Chronic Low Back Pain

Outcome Outcome Measure(s)	Follow-up No. of Participants (Studies)	Anticipated Absolute Effects (95% CI)			Certainty	What Happens
		AR Physical Activity	Active Comparator	Difference		
Pain-Related Functioning or Interference ODI; MODI; RMDQ	2 –8 wk N = 316 6 RCTs[36,38,39,43,48,50]	-53.1*	-24.5*	**Diff. Δ: -28.6***	⊕◯◯◯ Very low[a,b,c]	The evidence is very uncertain on the effect of AR physical activity on pain-related interference compared to non-AR physical activity.
Pain Intensity or Severity VAS; NRS	2–8 wk N = 552 11 RCTs[35,37-43,46,48,50]	-6.9†	-3.3†	**Stand. Diff. Δ: -0.7 (-1.2, -0.2)**	⊕◯◯◯ Very low[a,c,d]	The evidence is very uncertain on the effect of AR physical activity on pain intensity compared to an active comparator.

Notes. * Values for mean change in AR physical activity, active comparator and Diff. Δ taken from Afzal, 2022.[38]
† Mean change for active comparator taken from Afzal, 2022[38], mean change for AR physical activity calculated using pooled estimate for standardized Diff. Δ.

GRADE Working Group grades of evidence
High certainty**:** We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations:

a. Downgraded 2 levels for study limitations (studies rated high for risk of bias).

b. Downgraded 1 level for inconsistency (direction of effect varied across studies).

c. Downgraded 1 level for indirectness (age in study not generalizable to back pain population).

d. Downgraded 1 level for inconsistency (high heterogeneity).

Abbreviations. CI=confidence interval; MODI=Modifided Oswestry Disability Index; NRS=Numeric Rating Scale; ODI=Oswestry Disability Index; RCT=randomized controlled trial; RMDQ=Roland-Morris Disability Questionnaire; VAS=Visual Analog Scale; wk=week.

Pain Catastrophizing and Kinesiophobia

Six studies evaluated pain catastrophizing or kinesiophobia using the Pain Catastrophizing Scale (PCS), Tampa Scale of Kinesiophobia (TSK), or the Fear Avoidance Beliefs Questionnaire (FABQ).[40-43,48,50] We were unable to conduct meta-analyses as no study provided SD for mean change. Most studies found that the AR group showed greater reductions (Diff Δ -0.1 to -18.7),[40-42,48] while 1 study showed greater improvement in the comparator group (*eg*, TSK Diff Δ 23.5)[50] and another trial reported inconsistent results between FABQ-physical and FABQ-work subscales.[43]

AR Physical Activity versus Usual Care

AR physical activity may result in better pain-related functioning (low COE), but the evidence is very uncertain regarding its effects on pain intensity or severity and adverse events, compared with usual care (very low COE; Table 8). Four studies compared AR physical activity to usual care.[44,45,47,49] Two studies used simulated equine therapy (both conducted in Korea),[47,49] 1 used Nintendo Wii,[44] and 1 used a digital dodgeball game.[45] The latter 2 were conducted in Australia[44] and the US.[45] Three studies[44,45,47] were rated some concerns for RoB due to concerns about randomization, adherence, measurement of outcomes, and/or selected reporting bias. One was rated high RoB[49] due to similar concerns as the other study and additionally missing outcomes data.

Zadro, 2019[44] assessed pain-related functioning using the RMDQ and the Patient Specific Functional Scale (PSFS), finding greater improvement in the AR group at 8 weeks (*eg*, Diff Δ -0.4 on RMDQ). Three studies provided data on pain intensity, assessed using either VAS[47,49] or NRS,[44] and all showed greater reductions in the AR group at 8 weeks (Diff Δ -1.0 to -4.5). One study reported data on kinesiophobia, assessed with TSK, demonstrating greater improvement in the AR group at 8 weeks (Diff Δ -2.5).[44] The fourth study, Thomas, 2016,[45] did not report mean scores at follow-up after the 3-day intervention (digital dodgeball) but stated that there were no significant differences between groups on pain-related functioning or pain intensity. Two studies reported that no adverse events were detected in either group.[44,45]

Table 8. Certainty of Evidence: AR Physical Activity versus Usual Care for Low Back Pain

Outcome Outcome Measure(s)	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects (95% CI) AR Physical Activity	 Usual Care	 Difference	Certainty	What Happens
Pain-Related Functioning or Interference RMDQ; PSFS	8 wk N = 60 1 RCT[44]	-1.4*	-1.0*	**Diff. Δ: -0.4***	⊕⊕◯◯ Low[a,b]	AR physical activity may result in better pain-related functioning, compared to usual care.
Pain Intensity or Severity VAS; NRS	8 wk N = 144 3 RCTs[44,47,49]	-1.4*	-0.4*	**Diff. Δ: -1.0***	⊕◯◯◯ Very low[c,d]	The evidence is very uncertain on the effect of AR physical activity on pain intensity compared to usual care.
Adverse Events	4 days – 8 wk N = 113 2 RCTs[44,45]	0†	0†	0†	⊕◯◯◯ Very low[a,e,f]	The evidence is very uncertain on the effect of AR physical activity on adverse events compared to usual care.

Notes. * Values for mean change in AR physical activity, usual care, and Diff. Δ taken from Zadro, 2019[44] (RMDQ for pain-related functioning).

† In both studies, no events reported in either group.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations
a. Downgraded 1 level for study limitations (study rated some concerns for risk of bias).
b. Downgraded 1 level for indirectness (only older adults included).
c. Downgraded 2 levels for study limitations (studies rated some concerns or high for risk of bias).
d. Downgraded 1 level for indirectness (only older adults or young men included).
e. Downgraded 1 level for indirectness (only young or older adults included).
f. Downgraded 1 level for imprecision (sample size too small to detect adverse events).

Abbreviations. AR=augmented reality; CI=confidence interval; No.=number; NRS=Numeric Rating Scale; PSFS=Patient Specific Function Scale; RMDQ=Roland-Morris Disability Questionnaire; VAS=Visual Analog Scale.

CHRONIC NECK PAIN (KQ1)

We identified 6 trials evaluating the use of VR ($k = 5$)[51-55] or AR ($k = 1$)[56] interventions for chronic neck pain. All interventions involved engagement and guidance in physical activity for 3-6 weeks. Range for total duration of follow-up was 3 weeks to 4 months. Minimum duration of neck pain was specified as 3-6 months in all but a single study.[53] Study characteristics and findings are summarized in Table 6, and certainty of evidence for efficacy and harms are presented in Table 10 and Table 11. Detailed trial characteristics and findings are found in Appendix E. Below, we first describe findings for VR studies, and then the single AR study.

VR Interventions Trials

Five trials compared VR interventions with a variety of non-VR physical activity programs for chronic neck pain.[51-55] VR interventions and comparator programs all encouraged a range of movements at the neck (flexion, extension, rotation). All studies were small (total *n* range = 36-92) and involved young to middle-aged (mean or median ages 27-53 years) men and women. Studies were conducted in Australia ($k = 2$),[54,55] Europe ($k = 2$),[52,53] and Turkey.[51] Three studies were rated some concerns for RoB, primarily related to the potential bias in patient-reported measures when participants are unmasked to the assignment.[53-55] The remaining 2 studies were rated high RoB due to substantial drop-outs and/or individuals removed from the analyses for unclear reasons.[51,52]

Pain-Related Functioning or Interference, and Pain Intensity or Severity

All VR trials assessed pain-related functioning and pain intensity immediately post-intervention (3-6 weeks) and 3 studies also provided data at 3 months after the end of the intervention (3-4 months since baseline).[53,54,56] VR interventions may result in little to no difference in pain-related functioning at 3-6 weeks (low COE; pooled standardized Diff Δ -0.2 [-0.5, 0.2], Figure 3) and the evidence is very uncertain for effects at 3-4 months (very low COE; pooled standardized Diff Δ -0.1 [-0.8, 0.5], Figure 4). VR interventions may result in decreased pain intensity at 3-6 weeks (low COE; pooled standardized Diff Δ -0.5 [-0.8, -0.1], Figure 5), but lead to little to no difference in pain intensity at 3-4 months (low COE; pooled standardized Diff Δ -0.2 [-0.9, 0.5], Figure 6).

Table 9. Summary of Findings for Chronic Neck Pain

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Adverse Events & Other Eligible Outcomes
VR Intervention Trials						
Cetin, 2022[51] High Turkey	Neck pain ≥ 6 mo, baseline NDI score ≥10, excluding history of cervical spine surgery, rheumatologic, vestibular, neurological, or cardiopulmonary disease, receiving exercise or physical therapy in past 6 mo; mean ages 40-42 yrs, 56% female	Guided movements while watching relaxing videos using Oculus Go VR glasses N = 21 (17) Clinic; 6 wk	Motor control exercises N = 20 (17) Clinic; 6 wk	**ProFitMap-Neck** Baseline means (SD): Intervention—69.3 (11.3) Comparator—65.22 (13.9) Diff Δ (6 wk): -0.17 Stand Diff Δ (6 wk): -0.01	**VAS** Baseline means (SD): Intervention—5.77 (1.39) Comparator—5.98 (1.93) Diff Δ (6 wk): -1.25 Stand Diff Δ (6 wk): -0.62	"No adverse effects were observed in either group." **Quality of life (6 wk)** SF-36 – physical (domain scores only) **Physical performance (6 wk)** ROM measures (flexion, extension, right lateral flexion, left lateral flexion, right rotation, and left rotation)
Nusser, 2021[52] High Germany	Non-traumatic neck pain ≥3 mo, excluding history of cervical fracture/dislocation, operations in cervical spine, damage to inner ear, vertebrobasilar insufficiency, neurological disease, ROM <10° in cervical spine; mean ages 50-53 yrs, 68% women	Neck-specific sensorimotor training using VR headset + standard rehabilitation program N = 17 (17) Clinic; 3 wk	In-person general sensorimotor (SMC) training and standard rehabilitation program N = 18 (16) Clinic; 3 wk Standard rehabilitation program N = 20 (18) Clinic; 3 wk	**NDI** Baseline means (SD): VR—18.7 (5.2) SMC—21.5 (6.4) Control—18.2 (6.7) Diff Δ (3 wk): VR-SMC: 0.5* VR-control: -2.8*	**VAS** Baseline means (SD): VR—4.9 (2.1) Sensorimotor—4.4 (3.1) Control—4.2 (2.6) Diff Δ (3 wk): VR-SMC: -1.2* VR-control: -1.7*	"Besides the weight of the helmet, which some patients found unpleasant, no other negative side effects were reported regarding the VR device or in general." **Physical performance (3 wk)** ROM (flexion, extension, left rotation, right rotation)
Sarig Bahat, 2014[55] Some Concerns Australia	Neck pain >3 mo, NDI >10%, excluding vestibular pathology, cervical fracture or dislocation, neurologic/ cardiovascular/	Kinematic training using VR headset N = 16 (14)	Kinematic training using a head-mounted laser pointer N = 16 (12) Clinic, home; 4 mo	**NDI** Baseline means (SD): Intervention—20.4 (7.6) Control—20.2 (6.5)	**VAS** Baseline means (SD): Intervention—35.7 (17.7) Control—35.2 (16.7)	Motion sickness in 4 participants (25%) VR group only, no assessment of comparator group

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Adverse Events & Other Eligible Outcomes
	respiratory disorders affecting physical performance, history of traumatic head injury, pregnancy; mean age 41 yr, % women NR	Clinic, home; 4 mo		Diff Δ*: 4 wk: -2.1 4 mo: -3.5	Diff Δ*: 4 wk: -4.9 4 mo: -0.7	**Kinesiophobia (5 wk, 4 mo)** TSK **Pain global change (4 mo)** GPE on pain **Physical performance (4 mo)** Cervical ROM Head movement velocity & accuracy
Sarig Bahat, 2018[54] Some Concerns Australia	≥18 years of age, neck pain for >3 months, NDI score >12%, VAS score >20 mm in the past week	Kinematic training using VR headset N = 30 (25) Home; 4 mo	Kinematic training using a head-mounted laser pointer N = 30 (26) Home; 4 mo Waitlist control (re-randomized to VR and head-mounted laser training in phase 2) N = 30 (25) N/A; N/A	**NDI** Baseline means (SD): VR—32.9 (12.5) KT—32.2 (13.3) Control—24.7 (10.7) Diff Δ (4 wk)*: VR-KT: -4.0 VR-control: -8.0 Diff Δ (4 mo)†: VR-KT: -2.7	**VAS** Baseline means (SD): VR KT—47.8 (20.9) KT—52.5 (19.5) Control—45.8 (21.5) Diff Δ (4 wk)*: VR-KT: -0.2 VR-control: -10.4 Diff Δ (4 mo)†: VR-KT: -7.3	"…few cases of side effects from the VR use. 5 [drop-outs] were due to VR-associated sickness and headache." **Kinesiophobia (4 wk)** TSK **Physical performance (4 wk)** Velocity Number of velocity peaks, Time to peak velocity percentage, Accuracy error Cervical ROM
Tejera, 2020[53] Some Concerns	Non-specific chronic neck pain, excluding pregnancy, neck pain caused by cancer, infectious or	Progressive head & neck movements using VR Vox Play	In-person neck exercises N = 22 (22) Clinic; 4 mo	**NDI** Baseline means (SD): Intervention—13.7 (6.7)	**VAS** Baseline means (SD): Intervention—5.0 (1.9)	NR **Pain catastrophizing & Kinesiophobia** PCS (4 mo)

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Adverse Events & Other Eligible Outcomes
Spain	inflammatory disorders, fracture or trauma, positive neurological signs or symptoms, cervical osteoarthritis, spondylarthritis, vertigo, previous cervical surgery, headaches; mean ages 27-33 yrs, 52% female	glasses and smartphone N = 22 (22) Clinic; 4 mo		Comparator—14.1 (9.3) Diff Δ*: 4 wk: -0.2 4 mo: 0.5	Comparator—4.3 (1.4) Diff Δ*: 4 wk: -1.1* 4 mo: -0.3*	PASS-20 (4 mo) FABQ (4 mo) TSK (4 wk, 4 mo) **Physical performance (4 mo)** Flexion/extension Lateroflexion Rotation
AR Intervention Trials						
Rezaei, 2019[56] Some Concerns Iran	Non-traumatic neck pain for >3 mo, 10-14 on NDI, excluding cervical or thoracic trauma in past 6 mo, neurological signs or symptoms in upper extremities, nerve injury, cervical spine injury or pathology or surgery; mean ages 31-36 yrs, 43-52% female	Videogame promoting head movements using Head Mouse Extreme N = 22 (21) Clinic; 9 wk	Conventional proprioceptive training N = 22 (21) Clinic; 9 wk	**NDI** Baseline means (SD): Intervention—13.0 (1.3) Control—12.3 (1.4) Diff Δ*: 4 wk: -4.3 9wk: - 5.6	**VAS** Baseline means (SD): Intervention—47.1 (10.2) Control—39.0 (10.1) Diff Δ*: 4 wk: -17.0 9 wk: -18.8	NR **Physical performance** Dynamic balance Y-balance test

Notes. * Diff Δ calculated by review team, unable to calculate standardized Diff Δ.

† 4 months data only available including phase 2 participants, where wait-list control was re-randomized and new participants recruited, *N* = 18 more for VR group, N = 14 more for KT group.

Abbreviations. AR=augmented reality; CI=confidence interval; Diff Δ= difference between groups of mean change scores; FABQ=Fear Avoidance Belief Questionnaire; GPE=Global Perceived Effect scale; hr=hour; KT=kinematic training; mo=month; N/A=not applicable; NDI=Neck Disability Index; NR=not reported; PASS-20=Pain Anxiety Symptom Scale; PCS=Pain Catastrophizing Scale; RoB=risk of bias; ROM=range of motion; SD=standard deviation; SF-36=36-item Short Form health survey; TSK=Tampa Scale of Kinesiophobia; VAS=Visual Analog Scale; VR=virtual reality; wk=week.

Figure 3. VR Physical Activity versus Non-VR Physical Activity: Pain-Related Functioning 3-6 Weeks

Study	XR Total	XR Mean	XR SD	Comparator Total	Comparator Mean	Comparator SD	Stand. Diff. Δ	Stand. Diff. Δ	95% CI
Cetin, 2022	17	-13.0	9.9	17	-10.6	10.8		-0.22	[-0.90; 0.45]
Nusser, 2021	17	-7.3	9.1	18	-7.3	9.7		0.00	[-0.66; 0.66]
Tejara, 2020	22	-6.8	9.2	22	-6.6	10.8		-0.02	[-0.61; 0.57]
Sarig Bahat, 2018	48	-8.1	18.2	44	-4.5	16.8		-0.20	[-0.61; 0.21]
Sarig Bahat, 2014	16	-7.8	6.2	14	-5.6	7.0		-0.31	[-1.04; 0.41]
Random effects model	**120**			**115**				**-0.16**	**[-0.52; 0.21]**
Prediction interval									**[-0.57; 0.26]**

Heterogeneity: $\tau^2 = 0$ [0.000; 0.045], $p = 0.95$

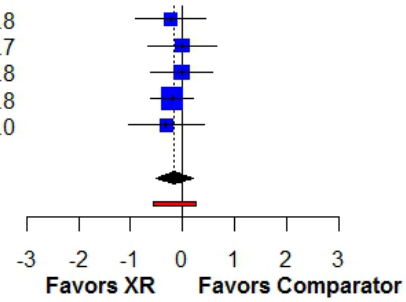


Figure 4. VR Physical Activity versus Non-VR Physical Activity: Pain-Related Functioning 3-4 Months

Study	XR Total	XR Mean	XR SD	Comparator Total	Comparator Mean	Comparator SD	Stand. Diff. Δ	Stand. Diff. Δ	95% CI
Tejara, 2020	22	-8.8	9.4	22	-8.3	10.4		-0.04	[-0.64; 0.55]
Sarig Bahat, 2018	48	-12.6	20.1	44	-9.8	16.4		-0.15	[-0.56; 0.26]
Sarig Bahat, 2014	14	-6.9	6.0	12	-3.4	14.9		-0.31	[-1.08; 0.47]
Random effects model	**84**			**78**				**-0.14**	**[-0.82; 0.53]**
Prediction interval									**[-2.15; 1.86]**

Heterogeneity: $\tau^2 = 0$ [0.000; 0.568], $p = 0.87$

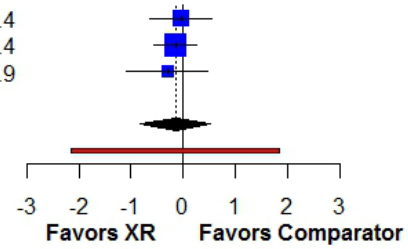


Figure 5. VR Physical Activity versus Non-VR Physical Activity: Pain Intensity 3-6 Weeks

Study	XR Total	XR Mean	XR SD	Comparator Total	Comparator Mean	Comparator SD	Stand. Diff. Δ	Stand. Diff. Δ	95% CI
Cetin, 2022	17	-3.7	1.8	17	-2.4	2.1		-0.61	[-1.30; 0.08]
Nusser, 2021	17	-2.7	2.4	18	-1.0	4.0		-0.50	[-1.18; 0.17]
Tejara, 2020	22	-2.3	2.7	22	-1.2	2.0		-0.47	[-1.07; 0.13]
Sarig Bahat, 2018	48	-20.7	32.9	44	-5.8	29.6		-0.47	[-0.89; -0.06]
Sarig Bahat, 2014	16	-12.5	19.5	14	-7.7	20.2		-0.24	[-0.96; 0.48]
Random effects model	**120**			**115**				**-0.47**	**[-0.83; -0.10]**
Prediction interval									**[-0.89; -0.04]**

Heterogeneity: $\tau^2 = 0$ [0.000; 0.021], $p = 0.97$

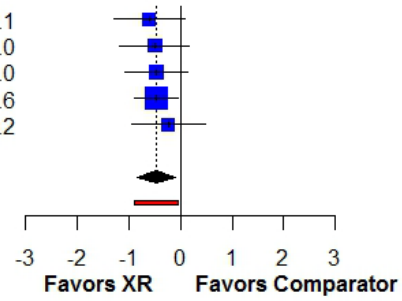
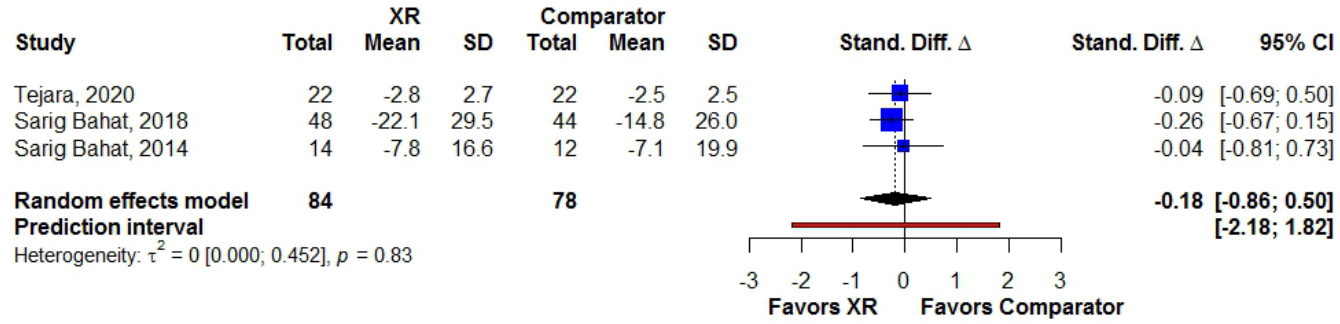


Figure 6. VR Physical Activity versus Non-VR Physical Activity: Pain Intensity 3-4 Months



Adverse Events

Although 4 trials reported some information on adverse events, we were unable to conduct a pooled analysis of this outcome due to inconsistencies in how these were assessed and reported. Overall, the evidence is very uncertain about harms due to VR interventions compared with non-VR physical activity (very low COE). One study reported that no adverse effects were observed in either group,[51] while 1 noted that the helmet weight was "unpleasant" for some participants without providing any counts.[52] The other 2 studies noted that VR-related motion sickness or headache contributed to drop-outs before and after randomization in the VR group, but it was unclear if the comparator groups were also assessed for potential adverse events.[54,55]

Table 10. Certainty of Evidence: VR Physical Activity for Chronic Neck Pain

Outcome Outcome Measure(s)	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects (95% CI) VR Physical Activity	 Non-VR Physical Activity	 Difference	Certainty	What Happens
Pain-Related Functioning or Interference NDI; ProFitMap-Neck	3-6 weeks N = 235 5 RCTs[51-55]	-8.6*	-6.6*	**Stand. Diff Δ: -0.2 (-0.5, 0.2)**	⊕⊕◯◯ Low[a,b]	VR physical activity may result in little to no difference in pain-related functioning compared to non-VR physical activity.
Pain-Related Functioning or Interference NDI	3-4 months N = 162 3 RCTs[53,54,56]	-9.3*	-8.3*	**Stand. Diff Δ: -0.1 (-0.8, -0.1)**	⊕◯◯◯ Very low[b,c]	The evidence is very uncertain on the effect of VR physical activity on pain-related functioning compared to non-VR physical activity.
Pain Intensity or Severity VAS	3-6 weeks N = 235 5 RCTs[51-55]	-2.4*	-1.2*	**Stand. Diff Δ: -0.5 (-0.8, -0.1)**	⊕⊕◯◯ Low[c]	VR physical activity may result in decreased pain intensity compared to non-VR physical activity.
Pain Intensity or Severity VAS	3-4 months N = 162 3 RCTs[53,54,56]	-3.1*	-2.6*	**Stand. Diff Δ: -0.2 (-0.9, 0.5)**	⊕⊕◯◯ Low[a,b]	VR physical activity may result in little to no difference in pain intensity compared to non-VR physical activity.

Outcome Outcome Measure(s)	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects (95% CI) VR Physical Activity	 Non-VR Physical Activity	 Difference	Certainty	What Happens
Adverse Events	4 weeks N = 159 4 RCTs[51,52,54,55]	-†	-†	**NC**†	⊕◯◯◯ Very Low[a,d]	The evidence is very uncertain on the effect of VR physical activity on adverse events compared to non-VR physical activity.

Notes. * Mean change values for non-VR physical activity taken from Tejera, 2022[53], mean change for VR physical activity calculated using pooled estimates for standardized Diff. Δ.

† 1 study reported no adverse events in either group; the others only assessed events in the VR group.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations:

a. Downgraded 1 level for study limitations (studies rated some concerns for risk of bias).

b. Downgraded 1 level for imprecision (PI crosses 0).

c. Downgraded 2 levels for study limitations (studies with some concerns and high risk of bias).

d. Downgraded 1 level for indirectness (authors only describe events in VR arm, not the control arm).

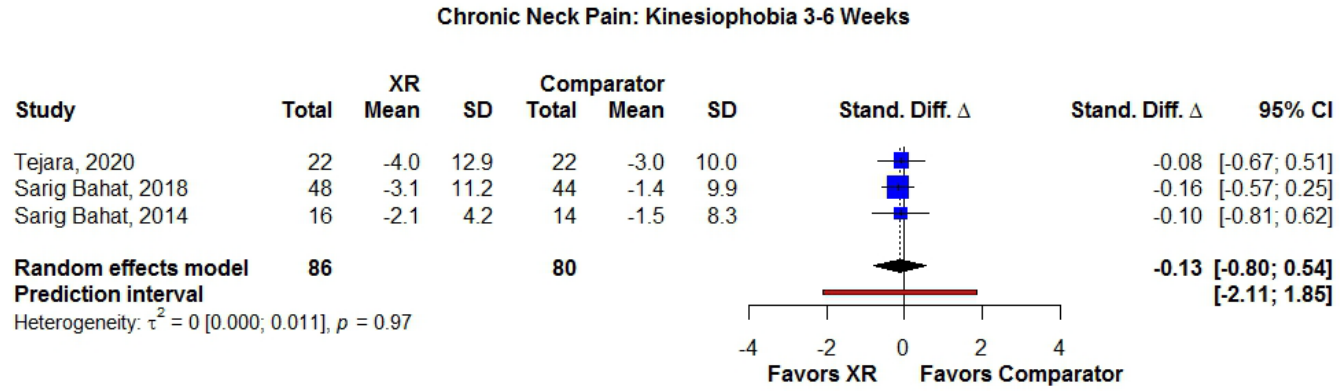
Abbreviations. CI=confidence interval; Diff Δ=between-group difference in mean change scores (Intervention-Comparator); NC=not calculable; NDI=Neck Disability Index; No.=number; RCT=randomized controlled trial; SD=standard deviation; VAS=Visual Analog Scale; VR=virtual reality.

Pain Catastrophizing and Kinesiophobia

Three studies evaluated kinesiophobia, all using TSK.[53-55] At 3-6 weeks, the pooled standardized Diff Δ was -0.1 (-0.8, 0.5), indicating little to no difference compared with the non-VR physical activity programs (Figure 7). Similarly, the pooled standardized Diff Δ at 3-4 months was -0.3 (-1.1, 0.6), indicating little to no difference (Figure 8).

Figure 7. VR Physical Activity versus Non-VR Physical Activity: Kinesiophobia 3-6 Weeks

A. Standardized Difference in Change (Stand. Diff Δ)



B. Difference in Change (Diff Δ) of Tampa Scale of Kinesiophobia (TSK) Scores

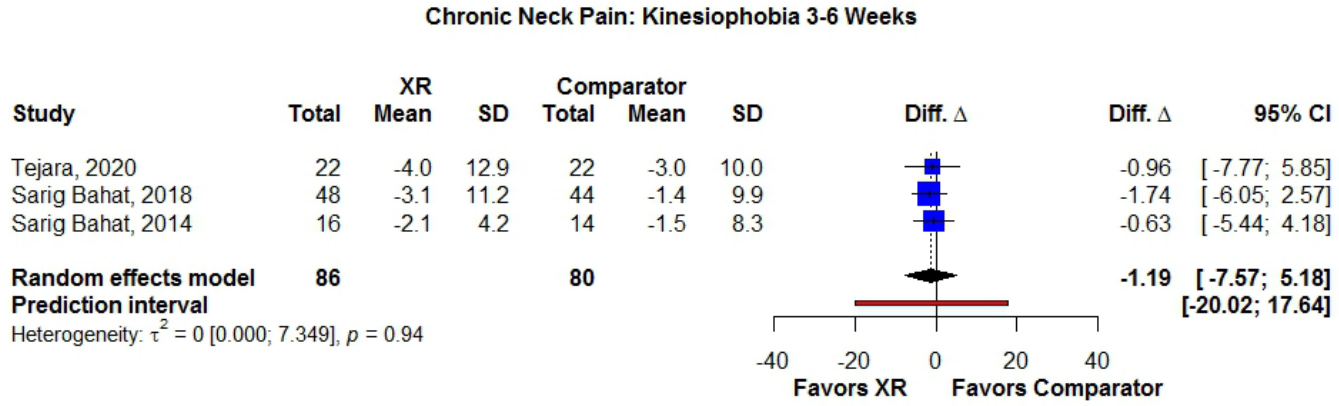
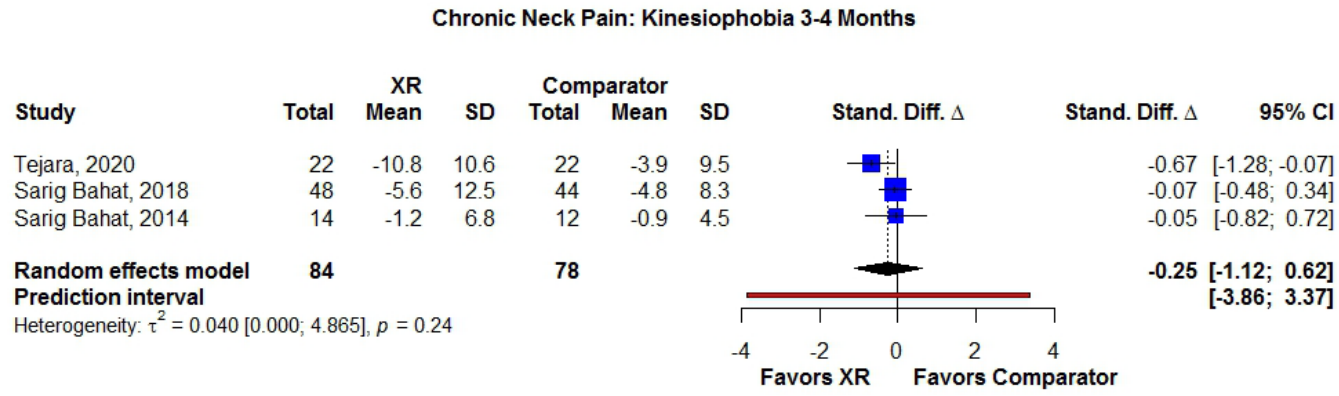
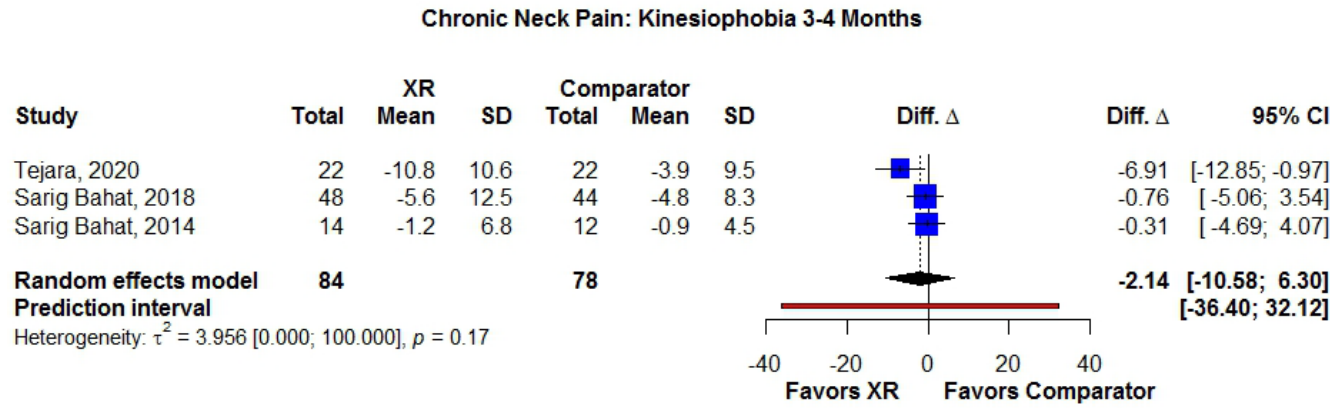


Figure 8. VR Physical Activity versus Non-VR Physical Activity: Kinesiophobia 3-4 Months

A. Standardized Difference in Change (Diff Δ)



B. Difference in Change (Diff Δ) of Tampa Scale of Kinesiophobia (TSK) Scores



Other Eligible Outcomes

Quality of life was evaluated by 1 study using the 36-item Short Form Health Survey (SF-36) but authors did not report the physical and mental component scores (PCS and MCS, respectively).[51] Scores for the individual domains are provided in the detailed results in Appendix tables. Global perceived effect (GPE) on change was only reported by 1 study,[55] while physical performance was assessed by all 5 studies using a variety of different measures, including ROM and velocity and accuracy in making certain movements.

AR Interventions Trials

One trial compared an AR physical activity intervention (n = 21) with conventional exercises (n = 21) for chronic neck pain.[56] Both physical activity programs lasted 4 weeks. This study was conducted in Iran and included young men and women (mean ages 31-36 years, 43-52% women). This study was rated high RoB due to concerns about allocation concealment, amount and treatment of missing data, and bias in self-reported measures when participants are unmasked to assignment.

Pain-related functioning was assessed using NDI at baseline, 4 weeks (immediately post-intervention), and 9 weeks (5 weeks after completion of the intervention). There was greater improvement in pain-related functioning in the AR group (Diff Δ -4.3 at 4 weeks, and -5.6 at 9 weeks). Pain intensity was assessed at the same time points using VAS 0-100 mm. The AR group also had greater reductions in pain intensity (Diff Δ -17.1 at 4 weeks, and -18.7 at 9 weeks). For physical performance, multiple measures from a dynamic Y-balance test were reported at baseline, 4 weeks, and 9 weeks. This study did not evaluate adverse events, quality of life, pain catastrophizing or kinesiophobia, global change in pain, or any of the secondary outcomes.

Table 11. Certainty of Evidence: AR Physical Activity for Chronic Neck Pain

Outcome Outcome Measure	Follow-up No. of Participants (Studies)	Anticipated Absolute Effects			Certainty	What Happens
		AR Physical Activity	Non-AR Physical Activity	Diff. Δ		
Pain-Related Functioning or Interference NDI	4-9 weeks N = 42 1 RCT[56]	-8.6	-3.1	**-5.6***	⊕⊕◯◯ Low[a]	AR physical activity may improve pain-related functioning compared to non-AR physical activity.
Pain Intensity or Severity VAS	4-9 weeks N = 42 1 RCT[56]	-37.5	-18.8	**-18.8***	⊕⊕◯◯ Low[a]	AR physical activity may reduce pain intensity compared to non-AR physical activity.

Notes. * Calculated by review team.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations:

a. Downgraded 2 levels for study limitations (study rated high risk of bias).

Abbreviations. AR=augmented reality; CI=confidence interval; Diff Δ=difference in change scores; NDI=Neck Disability Index; no.=number; RCT=randomized controlled trial; VAS=visual analogue scale.

FIBROMYALGIA (KQ1)

We identified 5 trials (reported in 8 articles) evaluating the use of XR interventions for fibromyalgia. All were AR interventions: 4 studies compared AR physical activity interventions using Microsoft Kinect or Nintendo Wii with either conventional exercises or usual care, [57-63] and 1 trial[64] evaluated AR-enhanced cognitive behavioral therapy (CBT). All trials included only female participants with a confirmed diagnosis of fibromyalgia (according to American College of Rheumatology criteria). All studies reported results for pain-related functioning,[57,58,60,64] 3 reported findings for pain intensity,[57,59,64] and no trials addressed adverse events. Study characteristics and findings are summarized in Table 12, and detailed trial characteristics and findings are found in Appendix G. Below, we first describe findings for AR physical activity trials, and then discuss the single trial of AR-enhanced CBT.

AR Physical Activity Interventions

Of 4 trials on AR physical activity interventions for fibromyalgia, 3 involved exercise programs over 7-8 weeks,[57,58,60,61] and the fourth lasted 6 months.[59,62,63] Studies were conducted in Spain ($k = 2$)[59-63], Brazil,[58] and Turkey.[57] All studies were small with total n range 35-83. Two studies were rated high for RoB,[57,58] and 2 as some concerns,[59-63] due to a variety of factors, including concerns about randomization, missing outcomes data, and bias in measurement of outcomes.

Pain-Related Functioning or Interference, and Pain Intensity or Severity

The evidence is very uncertain on the effects of AR physical activity on pain-related functioning and pain intensity, compared with conventional therapy or usual care (very low COE; Table 13). For the meta-analysis on pain-related functioning, we included data from the 3 trials evaluating AR physical activity interventions lasting 7-8 weeks, all of which used the Fibromyalgia Impact Questionnaire (FIQ).[57,58,60,61] One of these trials did not include the work-related items of FIQ.[57] The pooled estimate for standardized Diff Δ was -0.6 (-1.4, 0.2; Figure 9). The fourth trial used the revised version of FIQ (FIQR) to assess pain-related functioning at baseline and post-intervention (6 months), finding greater improvement in the AR group (Diff Δ -2.8).[63]

Both trials reporting pain intensity used the VAS, but on different scales. Polat, 2021[57] used 0-10 VAS and found greater reductions in the AR group during and immediately after the intervention (Diff Δ -0.6 at 4 weeks and -0.9 at 8 weeks). Villafaina, 2019[59] used 0-100 VAS and showed greater improvement in the AR group immediately post-intervention at 6 months (Diff Δ -11.1).

Figure 9. AR Physical Activity versus Control: Pain-Related Functioning 7-8 Weeks

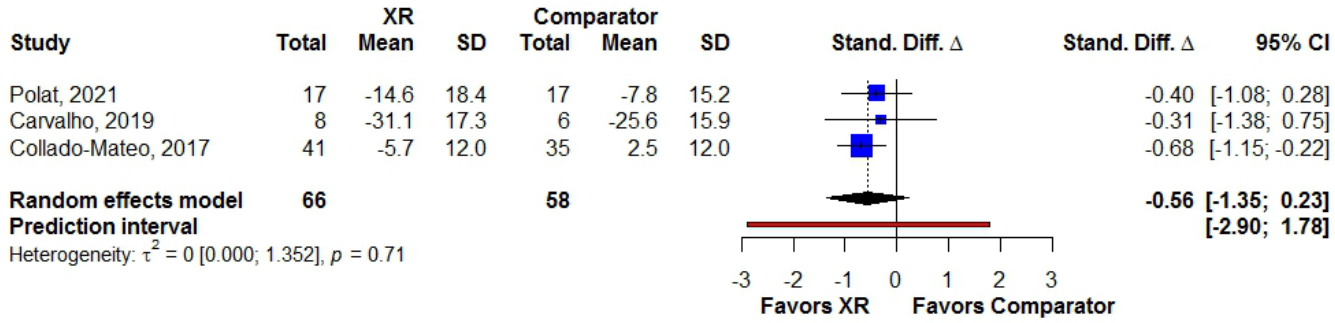


Table 12. Summary of Findings for Fibromyalgia

Author, Year **RoB** **Country**	**Key Participant Characteristics**	**Intervention** **N Randomized (N Analyzed)** **Setting; Duration**	**Comparator** **N Randomized (N Analyzed)** **Setting; Duration**	**Outcomes**		
				Pain-Related Functioning or Interference	**Pain Intensity or Severity**	**Adverse Events & Other Eligible Outcomes**
AR Physical Activity Interventions						
Collado-Mateo, 2017a[60]; Collado-Mateo, 2017b[61] Some concerns Spain	Women diagnosed with fibromyalgia by ACR criteria, excluding pregnancy, any changes in therapies in past 8 mo, and had contraindications to exercise; mean ages 52-53 yrs	Games using Microsoft Kinect that involved dance, walking, etc. N = 42 (41) Clinic; 8 wk	Usual care N=41 (35) Clinic; 8 wk	**FIQ** Baseline (SD): Intervention—50.6 (12.9) Comparator—49.2 (15.3) Diff Δ (8 wk): -8.3*	NR	NR **Quality of life (8 wk)** EuroQoL-5D **Physical performance (8 wk)** TUG Functional reach Balance tests
De Carvalho, 2019[58] High Brazil	Women diagnosed with fibromyalgia by ACR criteria, excluding cardiovascular, pulmonary, orthopedic, neurologic, or dermatologic conditions affecting strength and physical capabilities, and pregnancy; mean ages 48-56 yrs	Games using Nintendo Wii Fit that involve variety of lower limb, upper limb and trunk exercises N = 16 (11) Clinic; 7 wk	Chain muscle stretching N = 19 (10) Clinic; 7 wk	**FIQ** Baseline (SD): Intervention—64.5 (16.1) Comparator—72.0 (9.1) Diff Δ*: 4 wk: -14.3 7 wk: -5.6	NR	NR **Physical performance (4 wk, 7 wk)** Number of Steps up/down (25 cm height)
Polat, 2021[57] High Turkey	Women diagnosed with fibromyalgia by ACR criteria, ≥ 8 years of formal education and same medications for fibromyalgia ≥3 mo; mean ages 43-47 yrs	Beach volleyball games using Microsoft Kinect N = 20 (20) Home, clinic; 8 wk	Conventional strength training N = 20 (20) Home, clinic; 8 wk	**FIQ** (no work domain) Baseline (SD): Intervention—54.7 (13.3) Comparator—58.5 (9.2) Diff Δ*: 4 wk: -5.2 8 wk: -6.8	**VAS** Baseline (SD): Intervention—6.40 (1.4) Comparator—6.45 (1.3) Diff Δ*: 4 wk: -0.6 8 wk: -0.9	NR **Quality of life (4 wk, 8 wk)** EuroQoL-5D **Physical performance (4, 8 wk)** 6-minute walk

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Outcomes Pain Intensity or Severity	Outcomes Adverse Events & Other Eligible Outcomes
Villafaina, 2019[59]; Martin-Martinez, 2019[62]; Villafina, 2020[63] Some concerns Spain	Women diagnosed with fibromyalgia by ACR criteria; mean ages 53-54 yrs	Games using Microsoft Kinect that involved dance, walking, etc N = 28 (25) Clinic; 24 wk	Usual care N = 27 (25) Clinic; 24 wk	**FIQR** Baseline (SD): Intervention—52.6 (17.1) Comparator—55.0 (20.3) Diff Δ (24 wk): -2.8*	**VAS** Baseline (SD): Intervention—62.1 (19.3) Comparator—60.4 (19.3) Diff Δ (24 wk): -11.1*	NR **Quality of life (24 wk)** EuroQoL-5D-5L **Physical performance (24 wk)** 6-minute walk Sit-to-stand TUG 10 stairs (timed) Sit and reach Arm curl Back scratch
AR-Enhanced Cognitive Behavioral Therapy						
Garcia-Palacios, 2015[64] High Spain	Women 18-70 years old, diagnosed with fibromyalgia by ACR criteria, excluding severe mental disorders, mental retardation, substance abuse, physical disease "that could interfere with receiving…treatment," and not requesting disability; mean age NR	Cognitive behavioral therapy sessions with projected images of meadow or beach environments alongside environment sounds, music, or narratives. N = 31 (30) Clinic; 6 wk	Usual care N = 30 (29) Clinic; 6 wk	**FIQ** Baseline (SD): Intervention—61.6 (19.9) Comparator—60.6 (21.4) Diff Δ (6 wk): -15.7* **BPI-Interference** Baseline (SD): Intervention—32.2 (14.8) Comparator—32.3 (16.7) Diff Δ (6 wk): -5.7*	**BPI-Intensity** Baseline (SD): Intervention—23.6 (5.1) Comparator—22.4 (7.97) Diff Δ (6 wk): 0.7*	NR **Quality of life (8 wk)** Quality of Life Index

Notes. * Diff Δ calculated by review team, unable to calculate standardized Diff Δ.

Abbreviations. ACR=American College of Rheumatology; BPI=Brief Pain Inventory; Diff=difference in differences; EuroQoL-5D=European Quality of Life 5 dimensions; FIQ=Fibromyalgia Impact Questionnaire; FIQR=revised version of Fibromyalgia Impact Questionnaire; NR=not reported; RoB=risk of bias; SD=standard deviation; TUG=timed up and go; wk=week; yrs=years.

Table 13. Certainty of Evidence: AR Physical Activity for Fibromyalgia

Outcome Outcome Measure(s)	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects AR Physical Activity	 Comparator	 Difference	Certainty	What Happens
Pain-Related Functioning or Interference FIQ; FIQR	7-8 weeks N = 158 3 RCTs)[57,58,60]	-4.6*	2.5*	**Stand. Diff Δ: -0.6 (-1.4, 0.2)**	⊕◯◯◯ Very low[a,b,c]	The evidence is very uncertain about the effect of AR physical activity on pain-related functioning, compared with conventional therapy or usual care.
Pain Intensity or Severity VAS	2-4 months N = 95 2 RCTs[57,59]	-2.5†	-1.7†	**Diff Δ: -0.9†**	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain about the effect of AR physical activity on pain intensity compared with conventional therapy or usual care.

Notes. * Mean change for active comparator taken from Collado-Mateo, 2017a[60], mean change for AR physical activity calculated using pooled estimate for standardized Diff. Δ.

† Values for mean change in AR physical activity, active comparator, and Diff. Δ taken from Polat, 2021.[57]

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations:

a. Downgraded 2 levels for study limitations (studies rated high for risk of bias).

b. Downgraded 1 level for indirectness (all studies included only female participants).

c. Downgraded 1 level for imprecision (CI crosses 0).

Abbreviations. AR=augmented reality; CI=confidence interval; Diff Δ=difference in change scores; FIQ=Fibromyalgia Impact Questionnaire; FIQR=Revised version of Fibromyalgia Impact Questionnaire; No.=number; RCT=randomized controlled trial; VAS=Visual Analogue Scale.

Other Outcomes

Three AR physical activity studies reported findings on health-related quality of life, all using the European Quality of Life-5 (EuroQoL-5D).[57,59,60] All 3 found no substantial differences between groups, with both groups showing slight improvements or none (Diff Δ 0.1 across all studies and time points). These same 3 studies also provided results on a range of physical performance assessments, including 6-minute walk, TUG, and sit-to-stand. In general, the AR group demonstrated greater improvements, compared with either usual care or conventional exercises. For example, both Collado-Mateo, 2017b[61] and Martin-Martinez, 2019[62] showed greater reductions in time on the TUG (Diff Δ -0.7 s and -1.8 s, respectively). Similarly, the AR group had larger increases in distance on the 6-minute walk (Diff Δ 26.9 m[57] and 34.6 m[63]).

AR-Enhanced CBT Intervention

The evidence is also very uncertain on the effects of AR-enhanced CBT on pain-related functioning and pain intensity, compared with usual care (very low COE; Table 14). Garcia-Palacios, 2015[64] was

conducted in Spain, and compared AR-enhanced CBT for 6 weeks (n = 31) with usual care (n = 30).[64] The AR arm involved group CBT sessions that provided coaching on activity management while a landscape image was projected on a large screen and specific music, sounds, or narratives were playing. Pain-related functioning was assessed with FIQ and BPI-Interference. The AR group showed greater improvements on both measures at 6 weeks (Diff Δ −15.7 for FIQ and -5.7 on BPI-Interference). Pain intensity was measured with BPI-Intensity and although both groups had reductions at 6 weeks, the control group had a greater improvement (Diff Δ 0.7). Finally, quality of life was assessed with the Quality of Life Index, and the AR group showed greater improvement at 6 weeks (Diff Δ 1.4).

Table 14. Certainty of Evidence: AR-Enhanced CBT for Fibromyalgia

Outcome Outcome Measure	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects			Certainty	What Happens
		AR-CBT	Usual Care	Diff Δ		
Pain-Related Functioning or Interference FIQ; BPI-Interference	6 weeks N = 59 1 RCT[64]	-19.2	-3.6	-15.7	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain about the effect of AR-enhanced CBT on pain-related functioning, compared with usual care.
Pain Intensity or Severity BPI-Intensity	6 weeks N = 59 1 RCT[64]	-0.9	-1.7	0.7	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain about the effect of AR-enhanced CBT on pain-related functioning, compared with usual care.

Notes. GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. Downgraded 2 levels for study limitations (rated high for risk of bias).

b. Downgraded 1 level for indirectness (trials included only female participants.

Abbreviations. AR=augmented reality; BPI=Brief Pain Inventory; CBT=cognitive behavioral therapy; Diff Δ=difference in change scores; FIQ=Fibromyalgia Impact Questionnaire; No.=number; RCT=randomized controlled trial.

CHRONIC KNEE PAIN (KQ1)

We identified 5 trials evaluating XR interventions for chronic knee pain, all of which included only participants with knee osteoarthritis. All studies involved VR (k = 1) and AR (k = 4) interventions with physical activity, compared to conventional therapy. Summary characteristics and findings are summarized in Table 15, and detailed trial characteristics and results are found in Appendix H. Below, we first present findings from the single VR trial, and then describe results from the 4 AR studies.

Table 15. Summary of Findings for Chronic Knee Pain

Author, Year RoB Country	Key Participant Characteristics	XR Intervention N Randomized (N Analyzed) Setting; Duration	Comparator(s) N Randomized (N Analyzed) Setting; Duration	Outcomes		
				Pain-Related Functioning	**Pain Intensity/Severity**	**Adverse Events & Other Eligible Outcomes**
VR Intervention Trials						
Ozlu, 2023[65] High Turkey	Knee OA stages 2-3 (ACR criteria), MMSE≥ 22; mean ages 53-54 yr, 49-68% women	Games using Oculus headset that focus on lateral movement and trunk flexion to interact with virtual targets, in addition to ultrasound and TENS treatment N = 41 (35) Clinic; 3 wk	Conventional treatment with ultrasound and TENS N = 41 (38) Clinic; 3 wk	**WOMAC (total)** Baseline means (SD): Intervention—31.7 (6.8) Comparator—33.0 (7.9) Diff Δ*: 3 wk: -8.7 7 wk: -5.1	**VAS** Baseline means (SD): Intervention—5.6 (0.9) Comparator—5.8 (0.7) Diff Δ*: 3 wk: -0.7 7 wk: -1.1	NR **Physical performance (3 wk, 7 wk)** 6-minute walk Berg Balance Scale
AR Intervention Trials						
Elshazly, 2016[66] Some concerns Saudi Arabia	≥3 months of OA, able to walk ≥30 ft without assistance, and not in any sports or physical therapy; mean ages 58-60, sex/gender NR	Game involving standing and taking steps on virtual platform (device NR) N = 20 (20) NR; 8 weeks	Sensorimotor training (SMT) N = 20 (20) NR; 8 weeks Conventional walking program (control) N = 20 (20) NR; 8 weeks	**WOMAC (total)** Baseline means (SD): Intervention—71.7 (3.4) SMT—71.7 (2.8) Control—71.9 (3.1) Diff Δ (Intervention-SMT)*: 4 wk: -13.5 8 wk: -19.5 Diff Δ (Intervention-control)*: 4 wk: -14.1 8 wk: -29.9	**VAS** Baseline means (SD): Intervention—6.8 (0.9) SMT—6.6 (1.2) Control—6.68 (0.84) Diff Δ (Intervention-SMT)*: 4 wk: -1.9 8 wk: -1.8 Diff Δ (Intervention-control)*: 4 wk: -1.8 8 wk: -2.0	NR **Quality of Life (4 wk, 8 wk)** CDC Health Related Quality of Life **Physical Performance (4 wk, 8 wk)** Position sense

Author, Year RoB Country	Key Participant Characteristics	XR Intervention N Randomized (N Analyzed) Setting; Duration	Comparator(s) N Randomized (N Analyzed) Setting; Duration	Outcomes		
				Pain-Related Functioning	**Pain Intensity/Severity**	**Adverse Events & Other Eligible Outcomes**
Lin, 2020[67] Some concerns Taiwan	Knee OA (ACR criteria), Kellgren and Lawrence score ≥ 2, able to walk > 15 m, and not needing NSAIDS; mean ages 56-58 yr, 43-60% female	Games involving interaction with virtual targets through lower limb and trunk movements (via sensor pad for feet), in addition to temperature therapy and TENS N = 40 (40) Clinic; 4 weeks	Temperature therapy, TENS, and conventional exercise program (stretching, stabilization exercises, etc.) N = 40 (40) Clinic; 4 weeks	**WOMAC-physical function‡** Baseline means (SD): Intervention—505.1 (328.4) Comparator—581.0 (383.8) Diff Δ*: 2 wk: 39.1 4 wk: 60.3 8 wk: 90.0 16 wk: 82.8	**WOMAC-pain** Baseline means (SD): Intervention—161.2 (114.7) Comparator—170.2 (121.3) Diff Δ*: 2 wk: 0.7 4 wk: 3.9 8 wk: -19.9 16 wk: 4.9	No adverse effects observed in either group **Quality of Life (2-16 wk)** WHOQOL-BREF **Physical Performance (2-16 wk)** Biodex stability system 10 m walk time Stair ascent, descent time
Mete, 2022[68] High Turkey	Knee OA, Kellgren and Lawrence stages 2-3; median ages 57-60 yr, 77-88% women	Games involving control of on-screen avatars through knee flexion and extension, via special device (MarVAJED) with sensors for joint positions and provided auditory and visual feedback, in addition to comparator treatment N = 32 (30) Clinic; 6 wk	Conventional treatment with ultrasound, TENS, temperature therapy, and muscle strengthening exercises N = 32 (30) Clinic; 6 wk	**WOMAC (total)** Baseline medians (IQR): Intervention—19.7 (18.2, 21) Comparator—15.1 (9.3, 18) Diff Δ NC†	**WOMAC–pain** Baseline medians (IQR): Intervention—6 (5.37, 7.12) Comparator—4.5 (4.3, 6) Diff Δ NC† **VAS (at rest)** Baseline medians (IQR): Intervention—32.2 (20.8, 40.0) Comparator—36.3 (30.0, 40.0) Diff Δ NC†	NR **Kinesiophobia (6 wk)** TSK **Physical performance (6 wk)** Pedalo Balance Score Knee flexion & extension ROM Knee proprioception at 30, 60° Peak torque of knee flexion & extension at 120, 240°

Author, Year RoB Country	Key Participant Characteristics	XR Intervention N Randomized (N Analyzed) Setting; Duration	Comparator(s) N Randomized (N Analyzed) Setting; Duration	Outcomes		
				Pain-Related Functioning	Pain Intensity/Severity	Adverse Events & Other Eligible Outcomes
Nambi, 2020c[69] Some concerns Saudi Arabia	Male soccer players with post-traumatic OA ≥ 3 mo following ACL injury (verified by orthopedic surgeon) and pain rating 4-8; mean ages 22-23 yr	Games using ProKin system that required knee movements to interact with visual targets N = 20 (18-20) Clinic; 4 wk	Sensorimotor training (SMT) N = 20 (18-20) Clinic; 4 wk Standard exercise program (control) N=20 (19-20) Clinic; 4 wk	**WOMAC (total)** Baseline means (SD): Intervention—72.3 (4.2) SMT—72.5 (4.5) Control—71.2 (3.8) Diff Δ (Intervention-SMT)*: 4 wk: -22.1 8 wk: -9.8 3 mo: -14.0 Diff Δ (Intervention-control)*: 4 wk: -29.3 8 wk: -31.0 3 mo: -25.1	**VAS** Baseline means (SD): Intervention—7.2 (0.5) SMT—7.4 (0.4) Control—7.3 (0.4) Diff Δ (Intervention-SMT)*: 4 wk: -2.3 8 wk: -0.8 3 mo: -0.8 Diff Δ (Intervention-control)*: 4 wk: -3.1 8 wk: -1.6 3 mo: -3.2	NR NR

Notes. * Diff Δ calculated by review team, unable to standardize as no SD for change reported.

† Diff Δ not reported and cannot be calculated using provided results.

‡ WOMAC total scores NR.

Abbreviations. ACL=anterior cruciate ligament; ACR=American College of Rheumatology; CDC=Center for Disease Control and Prevention; Diff Δ=difference in change scores; IQR=interquartile range; mo=months; NC=not calculable; NR=not reported; NSAIDS=non-steroidal anti-inflammatory drugs; OA=osteoarthritis; RoB=risk of bias; ROM=range of motion; SD=standard deviation; SMT=sensorimotor training; TENS=transcutaneous electrical nerve stimulation; TSK=Tampa Scale of Kinesiophobia; VAS=Visual Analog Scale; wk=weeks; WOMAC=Western Ontario and McMaster Universities Arthritis Index; WHOQOL-BREF=World Health Organization Quality of Life Brief Version; yr=years.

VR Intervention Trial

VR physical activity may result in better pain-related functioning, and greater decreases in pain intensity, compared with conventional therapy (low COE, Table 16). Ozlu, 2023[65] was conducted in Turkey and compared 3 weeks of games (using Oculus glasses) that required lateral trunk movements (*ie,* catching fish or bananas) plus conventional therapy (n = 41) with conventional therapy only (n = 41). Conventional therapy involved ultrasound imaging as well as transcutaneous electrical nerve stimulation (TENS). Participants were middle-aged men and women who had no cognitive impairments. This trial was rated high for RoB due to concerns in every domain, including randomization, deviations from the intended intervention, missing data, and selective reporting bias.

For pain-related functioning, WOMAC was assessed at baseline, 3 weeks, and 7 weeks, showing greater improvement in the VR group at both timepoints (Diff Δ –8.7 at 3 weeks and -5.1 at 7 weeks). Regarding pain intensity, VAS scores were measured at the same time points and demonstrated similarly greater reductions in the VR group (Diff Δ –0.7 at 3 weeks and -1.1 at 7 weeks). This study also assessed physical performance, finding essentially no differences between groups on the 6-minute walk test (Diff Δ –0.9 m at both time points) but greater improvement in the VR group on the Berg Balance Scale (Diff Δ 1.9 at 3 weeks and 2.9 at 7 weeks).

Table 16. Certainty of Evidence: VR Physical Activity for Chronic Knee Pain

Outcome Outcome Measure	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects			Certainty	What Happens
		VR Physical Activity	Ultrasound & TENS	Diff Δ		
Pain-Related Functioning or Interference WOMAC	7 weeks N = 82 1 RCT[65]	-5.2	-0.2	**-5.0**	⊕⊕◯◯ Low[a]	VR physical activity may result in better pain-related functioning, compared to ultrasound and TENS.
Pain Intensity or Severity VAS	7 weeks N = 82 1 RCT[65]	-1.5	-0.4	**-1.1**	⊕⊕◯◯ Low[a]	VR physical activity may result in decreased pain intensity compared to ultrasound and TENS.

Notes.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. Downgraded 2 levels for study limitations (study rated high for risk of bias).

Abbreviations. CI=confidence interval; Diff Δ=difference in change scores; No.=number; RCT=randomized controlled trial; TENS=transcutaneous electrical nerve stimulation; VAS=Visual Analogue Scale; VR=virtual reality; WOMAC=Western Ontario and McMaster Universities Arthritis Index.

AR Intervention Trials

The evidence is very uncertain on the effects of AR physical activity on pain-related functioning, pain intensity, and adverse events, compared with conventional therapy (very low COE; Table 17). Four

studies evaluated AR physical activity for chronic knee pain due to osteoarthritis. AR exercise programs lasted 4-8 weeks and were all compared with conventional therapy involving exercise and other techniques.[66-69] All studies were very small (total n = 40-80). Studies were conducted in the Middle East (k = 3)[66,68,69] and Taiwan.[67] Three trials included middle-aged and older men and women,[66-68] while the fourth included only young male soccer players.[68] This latter study also specified that these athletes had post-traumatic osteoarthritis after an injury to the anterior cruciate ligament.[68] Three studies were rated some concerns for RoB[66,67,69] and 1 was rated high RoB.[68] There were methodological concerns in most domains, including randomization, deviations from the intended intervention, missing data, and selective reporting bias. All studies assessed pain-related functioning and pain intensity, but only 1 study reported on adverse events.[67]

Pain-Related Functioning or Interference, and Pain Intensity or Severity

All 4 studies used WOMAC to evaluate pain-related functioning. Three trials[66,68,69] reported total WOMAC scores (though 1 study only reported medians [IQR]), and the fourth study[67] provided domain scores only. In general, all the groups improved over time, with the AR group having greater improvement in WOMAC total scores (*eg*, Diff Δ -13.5 at 4 weeks and -19.5 at 8 weeks, compared with sensorimotor training).[66] All studies also assessed pain intensity, using either VAS or WOMAC-pain subscale. Two trials showed greater reductions in VAS for the AR group (Diff Δ -0.8 to -3.1),[66,69] and 1 trial found inconsistent results over time using WOMAC-pain scale (Diff Δ -19.9 to 4.9).[67] The fourth trial reported only medians (IQR) for pain intensity.[68]

Adverse Events

Only 1 study evaluated adverse events; none were detected in either group during the 4-week duration of the intervention.[67]

Other Outcomes

Two studies examined quality of life: 1 used the Centers for Disease Control and Prevention (CDC) Health Related Quality of Life scale,[66] and the other used the World Health Organization Quality of Life Brief Version (WHOQOL-BREF).[67] There were generally small improvements in all the groups and no clear differences between groups (*eg*, Diff Δ 0.5-1.9 for WHOQOL-BREF physical domain scores).[67] One study assessed kinesiophobia using TSK but only provided medians (IQR).[68] Three trials evaluated physical performance using a variety of measures, including 10-meter walk, balance and position sense, and ROM.[66-68] All groups generally improved in these measures but there were inconsistent results in terms of which group did better. For example, Lin, 2020[67] showed that the control group had greater reduction in time for 10-meter walk at 2 weeks (Diff Δ 1.1 s), no substantial differences between groups at 4 and 8 weeks (Diff Δ -0.1 s at both time points), and the AR group had better times at 16 weeks (Diff Δ -0.9 s).

Table 17. Certainty of Evidence: AR Physical Activity for Chronic Knee Pain

Outcome Outcome Measure	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects			Certainty	What Happens
		AR Physical Activity	Conventional Therapy	Difference		
Pain-Related Functioning or Interference WOMAC	2-4 months N = 180 3 RCTs[66,67,69]	-57.0*	-27.2 to -37.6*	**Diff Δ: -19.5 to -29.9***	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of AR physical activity compared to conventional therapy.
Pain Intensity or Severity VAS; WOMAC-pain	2-4 months N = 180 3 RCTs[66,67,69]	-3.9*	-1.9 to -2.1*	**Diff Δ: -1.8 to -2.0***	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of AR physical activity on pain intensity compared to conventional therapy.
Adverse Events	Mean 4 weeks N = 80 1 RCT[67]	0	0	**0**	⊕◯◯◯ Very low[c,d,e]	The evidence is very uncertain on the effect of AR physical activity on adverse events compared to conventional therapy.

Notes. * Values for mean change in AR physical activity, conventional therapy (2 groups: sensorimotor training and walking program), and Diff. Δ taken from Elshazly et al[66] at 8 weeks.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. Downgraded 2 level for study limitations (studies rated some concerns and high for risk of bias).

b. Downgraded 2 levels for indirectness (high prevalence of women in some studies; one trial with only young male athletes).

c. Downgraded 1 level for study limitations (study rated some concerns for risk of bias).

d. Downgraded 1 level for indirectness (authors do not describe how they measured adverse events).

e. Downgraded 1 level for imprecision (no events were detected in either arm).

Abbreviations. AR=augmented reality; CI=confidence interval; Diff Δ=difference in change scores; no.=number; RCT=randomized controlled trial; VAS=Visual Analogue Scale; WOMAC=Western Ontario and McMaster Universities Arthritis Index.

KQ1 OTHER CONDITIONS

We identified 11 trials that were single studies evaluating the use of VR ($k = 5$) or AR ($k = 6$) interventions for a variety of chronic pain conditions. Trial characteristics and main findings are summarized in Table 18. Detailed characteristics and results are provided in Appendix I. Below, we first describe the trials evaluating VR interventions. Then we provide the results from trials of AR interventions.

Table 18. Summary of Findings for KQ1 Other Conditions

Author, Year RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Pain Intensity or Severity	Adverse Events & Other Eligible Outcomes
VR Psychological Skills Trials						
Cueno, 2023[70] High USA	Chronic migraine Severe psychiatric comorbidities were excluded; mean ages 42-43 yr, 83% women	Biofeedback training using Oculus Go mobile headset and heartrate monitor N = 25 (14) Home; 12 wk	Wait-list control N = 25 (22) N/A; 12 wk	**MIDAS** Baseline means (SD): Intervention—100.4 (72.6) Comparator—77.6 (64.7) Diff Δ (3 mo): -9.7*	**Headache days/month** Baseline means (SD): Intervention—23.7 (5.6) Comparator—25.4 (5.8) Diff Δ (3 mo): -0.4*	VR group: Nausea 29%, dizziness 22% (comparator group NR) **Pain catastrophizing (12 wk)** • CAP
Chuan, 2023[71] Some Concerns Australia	Cancer-related neuropathic pain Independent in most ADLs; psychiatric comorbidities (not stabilized with treatment) were excluded; mean ages 56-63 yr, 64% women	Progressive muscle relaxation and guided pain visualization using Oculus Rift S N = 19 (19) Clinic; 3 mo	Short nature videos viewed on Oculus Rift S N = 20 (20) Clinic; 3 mo	**BPI-interference** Baseline means (SD): Intervention—4.7 (2.4) Comparator—4.1 (2.7) Diff Δ: 1 mo: -1.0* 3 mo: -1.6*	**BPI-intensity** Baseline means (SD): Intervention—4.9 (1.1) Comparator—4.4 (1.9) Diff Δ: 1 mo: -0.5* 3 mo: -0.2*	Intervention: 21% nausea, 21% dizziness, 21% eyestrain Control: 25% nausea, 20% dizziness, 40% eyestrain **Quality of life (1 mo, 3 mo)** • QLQ-C30 (all domains) **Opioid use** • Average MME in previous week
Darnall, 2020[72] High USA	Back pain & fibromyalgia Adults with self-reported chronic, nonmalignant low back pain or fibromyalgia for ≥ 6 months with average pain intensity >4 (scale NR) over past month; mean ages NR; 26-33% female	Pain self-management strategies using Oculus Go VR headset N = 35 (25) Home; 22 days	Pain self-management strategies through audio recordings N = 39(29) Home; 22 days	**DVPRS** (item scores only – Pain Stress, Pain Mood, Sleep, and Activity Interference)	**DVPRS – Pain** Baseline means (SD): VR—8.4 (3.5) Audio only—4.5 (1.8) Diff Δ (3 wk): -4.5	24% of VR group had any nausea or motion sickness (comparator group not assessed) **Pain global change (3 wk)** • PGIC

Author, Year RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Pain Intensity or Severity	Adverse Events & Other Eligible Outcomes
VR Other Trials						
Reynolds, 2022[73] Some Concerns New Zealand	Metastatic breast cancer related pain, anxiety, and/or fatigue Experiencing pain, fatigue, and/or anxiety in the past week; mean ages 51-53 yr, 100% women	Sequence of 2 pleasant distracting experiences using Pico Goblin VR headset: 1st—Ripple 2nd—Happy Place N = 20 (38) Home; 9 days, 1-wk washout, 9 days	Inverse sequence of 2 experiences using Pico Goblin VR headset: 1st—Happy Place 2nd—Ripple N = 18 (38) Home, 9 days, 1-wk washout, 9 days	**BPI (total score)‡** Baseline means (95% CI): Happy Place—37.7 (31.2, 44.2) Ripple—42.1 (35.5, 48.6) 9-day mean difference: -2.2*	**NR**	Some participants reported feeling “claustrophobic” or “a bit dizzy/nauseous." (text comments from acceptability survey, rates NR) **Quality of life (9 days)** • EuroQoL-5D-5L (index and VAS general health)
Wankhade, 2022[74] High India	Frozen shoulder Stage 2-3 primary or idiopathic frozen shoulder, excluded any post-operative history of shoulder injuries, diabetes, and/or rheumatoid arthritis; mean age & % women NR	Physical therapy using Oculus Rift N = 25 (25) Clinic; 2 wk	Conventional physical therapy N = 25 (25) Clinic; 2 wk	**Shoulder Pain and Disability Index** Baseline mean (SD) NR Diff Δ: 16.3*	**NRS** Baseline mean (SD) NR Diff Δ: 1.0*	NR **Physical performance (2 wk)** • ROM shoulder flexion, extension, abduction, adduction, internal & external rotation
AR Physical Activity Trials						
Ambrosino, 2020[75] High Italy	Rheumatoid arthritis Enrolled for 4 wk intensive orthopedic and rheumatologic rehabilitation, excluded malignancy, intolerance to exercise, and pregnancy; mean ages 27-28 yrs, 35-40% male	Exercises with Nintendo Wii-Fit N = 20 (20) Clinic & Home; 12 wk	Conventional rehabilitation N = 20 (20) Clinic & Home; 12 wk	**Health Assessment Questionnaire (HAQ)** Baseline mean (SD): Intervention:1.8 (0.2) Comparator: 1.8 (0.3) Diff Δ (12 wk): -0.6*	NR	NR NR

Author, Year RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Pain Intensity or Severity	Adverse Events & Other Eligible Outcomes
Ditchburn, 2020[76] High United Kingdom	Musculoskeletal pain in ≥ 2 locations Able to walk unassisted for at least 0.5 miles, musculoskeletal pain in ≥2 joints for >12 weeks, excluding systemic conditions that may cause pain or self-reported injuries that contra-indicate exercise; mean ages 70-72 yr, 65-82% female	Games using the IREX system (volleyball, sharkbait, formula racing, snowboard, and bird & balls) N = 27 (27) Clinic; 6 wk	Exercises matched to games N = 27 (27) Clinic; 6 wk	NR	NRS Baseline mean (SD): Intervention: 3.0 (1.9) Control: 3.3 (2.8) Diff Δ (6 wk): -1.0*	"There were no adverse events, reactions, or reports of motion sickness amongst participants…" **Quality of life (6 wk)** • MAPS questionnaire **Physical performance (6 wk)** • Postural sway with eyes open & closed
Gouveia e Silva, 2020[77] Some concerns Brazil	Post-polio syndrome Diagnosis by consensus of Halstead and Rossi, no other rehabilitation or physical exercise during intervention; mean ages 55-56 yr; 50% female	Nintendo Wii games (bowling, boxing, golf, tennis) N = 19 (19) Clinic; 7 wk	Movement that mimics that of the Wii games N = 20 (20) Clinic; 7 wk	NR	**VAS** Baseline means (SD): Intervention: 6.2 (3.0) Control: 6.9 (1.6) Diff Δ: 7 wk: 0.0 11 wk: -0.1	Later upper limb muscle pain (after 1st session): AR group 15%, control group 10% **Physical performance (7 wk, 11 wk)** • Box and Block (dexterity) • Functional Reach Assessment (balance)

Author, Year RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Outcomes Pain Intensity or Severity	Outcomes Adverse Events & Other Eligible Outcomes
Karahan, 2016[78] Some Concerns Turkey	Ankylosing spondylitis Lack of regular exercise during the previous 6 months, excluded cardiopulmonary or other serious comorbidities; mean ages 36-37 yrs, 14-21% female	Games using Microsoft Kinect N = 28 (28) Clinic; 8 wk	Usual care N = 29 (29) Clinic; 8 wk	**Bath Ankylosing Spondylitis Functional Index (BASFI)** Baseline means (SD): Intervention—3.7 (1.5) Control—3.9 (1.6) 8-week Diff Δ: Intervention: -0.8 Control: 0.0	**VAS** Baseline means (SD): Intervention—4.9 Control—5.1 8-week Diff Δ: Intervention: -1.3 Control: -0.1	NR **Quality of life (8 wk)** • Ankylosing Spondylitis Quality of Life (ASQOL) questionnaire
AR Embodiment Trials						
Lewis, 2021[79] High UK	Complex Regional Pain Syndrome (CRPS) Identified from the CRPS UK network registry, excluded co-morbidity that might influence CRPS symptoms (*ie*, stroke, diabetes, fibromyalgia); mean age 52 yr, 65% female	Mediated virtual reality (MVR). (MIRAGE system) N = 23 (21 analyzed) Clinic; 6 wk	Sham N = 22 (18 analyzed) Clinic; 6 wk	NR	**NRS** Baseline mean (SD) Intervention: 5.6 (3.3) Comparator: 5.7 (3.4) Diff Δ: -0.05*	NR NR
Rothgangel, 2018[80] Some concerns Germany	Phantom limb pain Unilateral amputation with average PLP intensity 3/11 and ≥1 episode of PLP in past week; sufficient	Mirror therapy with iPad N = 26 (22) Home, clinic; 6 mo	No additional therapy N = 24 (19) Home, clinic; 6 mo	**PDI** Baseline means (SD): iPad: 30.5 (16.5) Mirror therapy: 23.6 (18.2)	**NRS** Baseline means (SD): iPad: 5.9 (1.9) Mirror therapy: 5.4 (2.4)	NR **Quality of life (10 wk, 6 mo)** • EuroQoL-5D-5L

Author, Year RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Pain Intensity or Severity	Adverse Events & Other Eligible Outcomes
	cognition, motor and cognitive skills; excluded stroke, severe mental disorders; mean ages 60-63 yr, 56-80% male		Traditional mirror therapy N = 26 (21) Home, clinic; 6 mo	Control: 32 (20.1) 10-week Diff Δ: iPad vs mirror therapy: 5.1* iPad vs control: 3.9* 6 mo: iPad vs mirror therapy: 3.6* iPad vs control: 0.9*	Control: 5.8 (2.1) Diff Δ: 10 wk: iPad vs mirror therapy: 0.5* iPad vs control: 0.4* 6 mo: iPad vs mirror therapy: 0.9* iPad vs control: -0.5*	**Pain global change (10 wk, 6 mo)** • GPE

Notes. * Difference in mean change calculated by review team, unable to calculate SMD

‡ Data abstracted from graphs using Plotdigitizer

Abbreviations. AR=augmented reality; BPI=Brief Pain Inventory; CRPS=Complex Regional Pain Syndrome; DASS-SF=Depression Anxiety Stress Scales – Short Form; Diff Δ =difference in mean change; EuroQoL-5D-5L=European Quality of Life-5 questionnaire; FACIT=Functional Assessment of Chronic Illness Therapy; GPE=global perceived effect; HAQ=Health Assessment Questionnaire; IREX=Interactive Rehabilitation and Exercise; MIDAS=Migraine Disability Assessment Test; MME=morphine milligram equivalent; N/A=not applicable; NR=not reported; NRS=Numeric Rating Scale; oMEDD=oral morphine equivalent daily dose; PDI=Pain Disability Index; PGIC=Pain Global Impression of Change scale; QLQ-C30= The European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30; RoB=risk of bias; SD=standard deviation; USA=United States of America; VAS=Visual Analog Scale; VR=virtual reality; yr=year.

VR Intervention Trials

Three trials evaluated VR psychological skills interventions for migraine,[70] cancer-related neuropathic pain,[71] and self-report chronic back pain or fibromyalgia.[72] A fourth trial examined VR distraction for pain, anxiety, and/or fatigue due to metastatic breast cancer,[73] and the last trial used VR physical activity intervention for frozen shoulder.[74]

VR Psychological Skills

Cueno, 2023[70] was conducted in the US and compared biofeedback training using an Oculus Go mobile VR headset (paired with a Polar H10 heart rate monitor, $n = 25$) to wait-list control ($n = 25$). VR group participants experienced a beach or hilltop setting with accompanying music to help cue breathing and attain individualized optimal respiratory rate. Participants were primarily young and middle-aged women (mean ages 42-43 years, 83% women). This study was rated high RoB, primarily due to the high attrition disproportionately affecting the VR group. For pain-related functioning, Migraine Disability Assessment (MIDAS) was assessed at baseline and 12 weeks. MIDAS scores improved in both arms with somewhat greater reduction in the VR group (Diff Δ -9.7). The number of headache days per month also decreased for both groups, with little difference between groups (Diff Δ -0.4). Adverse events were only assessed for the VR group: 29% experienced nausea and 22% experienced dizziness. Pain catastrophizing was also assessed with the Concerns about Pain Scale (CAP).

The second VR psychological skills study, Chuan, 2023,[71] examined VR psychological skills for cancer-related neuropathic pain. The study was conducted in Australia and compared a program teaching progressive muscle relaxation and pain visualization techniques (using Oculus Rift S headsets; $n = 19$) with a VR control consisting of short nature videos viewed through the same VR headsets ($n = 20$). Participants were functionally independent in most daily activities. Pain-related functioning was assessed with BPI-Interference at baseline, 1 month, and 3 months, showing slightly greater reduction in the VR psychological skills group (Diff Δ −1.0 at 1 month and −1.6 at 3 months). Similar results were seen for pain intensity, measured using BPI-Intensity. Adverse events were assessed through reported symptoms of cybersickness as well as the tolerability of the interventions. There were no significant differences in the rates of nausea (21% vs 25%, $p = 1.00$), dizziness (21% vs 20%, $p = 1.00$), or eyestrain (21% vs 40%, $p = 0.30$) between the VR psychological skills and control groups. One participant from the control group withdrew from the study due to severe headaches. Quality of life was assessed with European Organization for the Research and Treatment of Cancer Core Quality of Life questionnaire (EORTC QLQ-C30) and median dose of opioids (average in the past week) was 0 morphine milligram equivalents per day at baseline and follow-up in both groups.

In the third VR psychological skills study, Darnall, 2020[72] compared a pain education and self-management training (delivered using Oculus Go headset; $n = 35$) to similar skills training via audio recordings ($n = 39$) for participants with self-reported chronic low back pain or fibromyalgia. This trial was conducted in the US and was rated as high RoB due to concerns with the randomization process, deviations from the intended intervention, missing outcome data, and selective reporting bias. At baseline and 21 days, DVPRS was used to assess pain intensity and pain-related interference in a variety of domains. Greater reductions were seen in the VR psychological skills group (*eg*, DVPRS pain intensity Diff Δ -4.5 and activity interference Diff Δ -1.33). Adverse events were assessed through a survey on day 22. A quarter of participants (6 of 25) in the VR group experienced any nausea or

motion sickness. Using the PGIC, 84% of VR group and 62% in the control group reported improvement; 0% in the VR group and 3% in the control group reported worsening pain.

VR Other

Reynolds, 2022[73] examined VR distraction intervention for middle-aged women with fatigue and/or anxiety related to metastatic breast cancer. This crossover trial was conducted in New Zealand and compared 2 randomization schedules for the Happy Place application (a relaxing, animated camping scene with soothing music) and Ripple application (3 VR 360° nature scenes) using the Pico Goblin VR headset. Group 1 (n = 20) was randomized to Ripple, then Happy Place and Group 2 (n = 18) was randomized to Happy Place, then Ripple. There was a 1-week washout period in between interventions in both groups. This trial was assessed as some concerns RoB due to concerns of deviations from the intervention and missing outcome data. Pain-related functioning was assessed with BPI at baseline, day 7, and day 9. There was a slightly greater reduction in BPI for the Ripple application (Diff Δ –2.2). Adverse events were measured throughout with an open-ended survey question, with some participants reporting feelings of claustrophobia and/or nausea while using the headset. Authors measured quality of life through the EuroQoL-5D-5L Index, and there were similar improvements in both groups.

In the second study, Wankhade, 2022[81] compared VR physical activity (using Oculus Quest; n = 25) with conventional physical therapy (n = 25) for participants with frozen shoulder patients (mean age and sex NR). This trial was rated as high RoB due to concerns with the randomization process, deviations from the intended intervention, missing outcome data, and selective reporting bias. Pain-related functioning was assessed with the Shoulder Pain and Disability Index (SPADI) at baseline and 2 weeks, and only within-group differences were reported. SPADI increased for both groups, with greater change in the VR group (Diff Δ 16.3). Pain intensity was measured with the NRS, and the VR group had greater increases in pain (Diff Δ 1.1). Physical performance was assessed through various ROM measures.

AR Intervention Trials

We identified 6 trials examining AR interventions, 4 of which were physical activity programs for musculoskeletal pain,[76] rheumatoid arthritis,[75] ankylosing spondylitis,[78] and post-polio syndrome.[77] The remaining 2 both evaluated AR embodiment interventions, 1 for complex regional pain syndrome (CRPS),[79] and the other for phantom limb pain (PLP).[80]

AR Physical Activity

Ambrosino, 2020[75] compared in-home exergaming using Nintendo Wii-Fit (n = 20) with conventional rehabilitation (n = 20) for rheumatoid arthritis. AR group participants were asked to play each of 5 games for 10 minutes each day at home over 8 weeks, after an intensive 4-week rehabilitation in clinic. The control group were advised to continue the rehabilitation exercises at home after the same 4-week in-clinic program. The study was conducted in Italy and included young men and women. This trial was rated as some concerns for RoB due to concerns about the randomization process, outcome measurement, and selective reporting bias. Pain-related functioning was assessed with the Health Assessment Questionnaire (HAQ), with greater improvement in the AR group post-intervention (Diff Δ -0.6).

The second trial, Ditchburn, 2020,[76] evaluated AR physical therapy using the IREX system (n = 27) versus traditional gym-based exercises (n = 27) for musculoskeletal pain in multiple locations. In the

AR group, participants played each of 5 games 3 times per session, twice weekly for 6 weeks. The control group performed gym-based exercises that were similar, also twice weekly for 6 weeks. This study was conducted in the UK and included predominantly older women. This trial was rated high RoB due to concerns with the randomization process, deviations from the intended intervention, missing outcome data, and selective reporting bias. Authors measured current pain intensity and pain experienced within the past 30 days using the NRS. For current pain intensity, the AR group had greater reductions post-intervention (Diff Δ -1.0). Authors stated, "There were no adverse events, adverse reactions, or reports of motion sickness…" Additionally, physical performance was assessed as postural control, measured as Center of Pressure displacement and velocity.

In the third trial, Gouveia e Silva, 2020[77] examined the effect of games using Nintendo Wii for post-polio syndrome. The AR group played Nintendo Wii games (boxing, bowling, tennis, and golf; $n = 19$) and the control group were asked to perform exercises that mimicked the movements of Nintendo games, but without the interactive video game interface ($n = 20$). Both groups exercised for a total of 7 weeks, with 4 additional weeks follow-up afterwards. This study was conducted in Brazil and included middle-aged men and women with post-polio syndrome who were not undergoing other rehabilitation or exercising regularly. This trial was rated some concerns RoB due to concerns about adherence to the intervention and missing outcome data. Authors reported that both groups showed similar improvements in pain, measured by the VAS (Diff Δ 0.0 at 7 weeks and -0.1 at 11 weeks). Authors evaluated late upper limb muscle pain after the first session only, reporting that 15% of the AR group and 10% of the control group experienced this. Physical performance was assessed with the Functional Reach Assessment (FRA) and Box and Block (BB) tests. Both groups improved on these assessments, with the AR group showing greater improvement on both tests (*eg*, Diff Δ 1.0 for FRA and 5.3 for BB at 7 weeks).

The fourth study, Karahan, 2016,[78] compared exercises using Microsoft Kinect 360 ($n = 28$) to usual care ($n = 29$) for ankylosing spondylitis. The AR group performed exercises 5 times per week for 8 weeks. This study was conducted in Turkey and included mostly young men. Eight (29%) participants in the AR group and 7 (24%) participants in the control group were taking anti-TNF α medication. This study was rated as some concerns RoB due to concerns about missing outcome data and selective reporting bias. Pain-related functioning was assessed with the Bath Ankylosing Spondylitis Functional Index (BASFI) and pain intensity with the VAS, both at baseline and 8 weeks. The AR group showed improvements in both BASFI and VAS at 8 weeks, whereas the control group did not change from baseline (Diff Δ -0.8 for BASFI and -1.2 for VAS). Quality of life was measured with the Ankylosing Spondylitis Quality of Life (ASQOL), and similarly, only the AR group showed any improvement (Diff Δ −2.8).

AR Embodiment

Lewis, 2020[79] compared the impact of manipulated ($n = 23$) versus non-manipulated ($n = 21$) images of the affected hand in participants with CRPS. This trial was conducted in the UK and recruited participants from CPRS network registry and clinics. This trial was rated high RoB, in large part due to concerns regarding missingness of outcome data and deviations from the intervention. Authors used the MIRAGE system and required that participants place both hands resting palm down into the 2 apertures of the system. With their hands resting on the flat surface, the participants viewed a real-time digital image of their hands through a 'window-like' surface above and perpendicular to the apertures. Participants in both groups viewed their hand images during 1-minute sessions for a maximum of 5 sessions, with 4 occurring over 4 weeks and a final session during the sixth week. In the intervention

group sessions, study personnel manipulated the appearance of the hand based on how participants desired their hand to appear. Pain intensity was assessed using NRS at baseline and before and after each session. There were no clear differences between groups (Diff Δ –0.05).

In the second study, Rothgangel, 2018[80] examined the effect of mirror therapy using an iPad (n = 26), compared with traditional mirror therapy (n = 26) and no therapy (n = 24) on phantom limb pain. This study was conducted in Germany. Participants were predominantly older men. This trial was assessed as some concerns RoB due to concerns about the randomization process, adherence to the intervention, and missing outcome data. During the first 4 weeks, both iPad and mirror therapy groups performed exercises with the intact limb in front of a mirror and were instructed to also perform the exercises with the phantom limb as soon as they perceived voluntary, pain-free movement. At the last session, participants in the iPad group (n = 26) were given the tablet, instructions, and training materials, and encouraged to use it as often as they wished. In the traditional mirror therapy group (n = 25), participants were encouraged to perform mirror therapy as often as they wished. In the comparator group (n = 24), participants performed the same exercises during the first 4 weeks without a mirror and instructed not to do exercises with their phantom limb. Afterwards, they were encouraged to perform self-delivered exercises with the intact limb as often as they wished. Pain-related functioning was measured with the Pain Disability Index at 10 weeks and 6 months, with the traditional mirror therapy group and control group having greater reductions than the iPad group at both time points (*eg*, Diff Δ 5.1 and 3.6 at 10 weeks and 6 months, comparing iPad with mirror therapy). Similar results were seen for pain intensity, as measured with NRS at the same time points. Quality of life was assessed with the EuroQoL-5D, and pain global change was measured by the Global Perceived Effect (GPE); these were similar between groups and did not change substantially on follow-up.

POST-SURGICAL PAIN & REHABILITATION (KQ2)

We identified 7 trials evaluating XR interventions for rehabilitation after knee or hip replacement surgery. Two trials examined VR interventions[82,83] and 5 studies used AR interventions[84-88] for rehabilitation after knee replacement (1 of these also included participants with hip replacements). All AR interventions involved physical activity. Trial characteristics and main findings are summarized in Table 19. Detailed trial characteristics and findings are found in Appendix K. Below, we first describe findings for VR intervention trials and then present findings for AR studies.

Table 19. Summary of Findings for Post-Surgical XR Interventions

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Adverse Events & Other Eligible Outcomes
VR Intervention Trials						
Fuchs, 2022[82] High Israel	Patients with osteoarthritis undergoing unilateral TKA; mean age 70 yrs, 52-63% women	Nature or music film watched on Oculus, while undergoing continuous passive motion therapy N = 30 (30) Hospital; 2 days	Continuous passive motion therapy N = 25 (25) Hospital; 2 days	**WOMAC** Baseline means (SD): Intervention—36.4 (15.1) Comparator—34.5 (17.0) Diff Δ (6 mo)*: 1.1	*VAS* Baseline medians (IQR): Intervention—6 (5-8) Comparator—6 (6-8) Diff Δ NR†	NR NR
Jin, 2018[83] High China	Patients with osteoarthritis undergoing TKA; mean age 66 yrs, 55-61% women	Game using Oculus headset, rowing a boat using knee flexion N = 33 (33) Hospital; NR	Passive flexion of knee using arms N = 33 (33) Hospital; NR	**WOMAC** Baseline (SD): Intervention—45.0 (5.1) Comparator—44.2 (5.7) Diff Δ*: 1 mo: -3.9 3 mo: -4.7 6 mo: -5.6	**VAS** Baseline (SD): Intervention—7.4 (1.1) Comparator—7.4 (1.3) Diff Δ*: 3 days: -0.3 5 days: -0.5 7 days: -0.5	NR **Physical performance (3, 7, 14 days)** • ROM
AR Intervention Trials						
Eichler, 2019[84] High Germany	Patients with osteoarthritis after TKA or THA; mean ages 53-57 yrs, 49-54% women	Exercises using Microsoft Kinect sensor, as demonstrated by an avatar N = 56 (48) Home; 3 months (after 3 wk inpatient rehab)	Usual care N = 55 (39) Home; NR (after 3 wk inpatient rehab)	**WOMAC** Baseline (SD) Intervention—26.4 (18.5) Comparator—24.8 (16.4) Standardized Diff Δ (3 mo)‡: -0.29	NR	NR **Quality of life (3 mo)** • SF-36 **Physical performance (3 mo)** • 6-minute walk • Stair Ascend test • Five Times Chair Rise test

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning	Pain Intensity/Severity	Adverse Events & Other Eligible Outcomes
Janhunen, 2023[85] Some concerns Finland	After first primary TKA; mean ages 66-67 yrs; 63-64% women	Games using Microsoft Kinect, involved similar movements as home PT N = 25 (21) Home; 16 wk	Standard PT N = 27 (25) Home; 16 wk	**OKS** Baseline (SD) Intervention—26.7 (6.7) Comparator—26.9 (6.5) Standardized Diff Δ (4 mo)‡: 0.32, p=0.27	**VAS** Baseline (SD) Intervention—57.1 (18,3) Comparator—54.2 (21.6) Standardized Diff Δ (4 mo)‡: -0.39, p=0.18	NR **Physical performance (2, 4 mo)** • TUG • Short Physical Performance Battery • Muscle force flexion, extension • ROM flexion, extension
Piqueras, 2013[86] High Spain	After primary TKA, with active ROM flexion 80º and extension –10º, without signs of stiffness, and able to walk (walking aid ok); mean age 73 yrs, 72% women	Screen and leg movement sensors to instruct and monitor knee exercises N = 90 (68) Home; 2 wk	Conventional PT N = 91 (65) Clinic; 2 wk	NR	VAS Baseline (SD): Intervention—3.8 (2.01) Comparator—4.3 (1.93) Standardized Diff Δ‡: 2-wk: -0.05, p=0.804 3-mo: 0.22, p=0.28	NR **Physical performance (2 wk, 3 mo)** • TUG • Quadriceps, hamstring strength • ROM flexion, extension
Prvu Bettger, 2020[87] Some concerns USA	TKA for non-traumatic conditions and expected to discharge home; mean age 65 yrs, 60-65% women	Virtual telehealth system (VERA) to demonstrate exercises (with avatar) and monitor performance N = 153 (140) Home; NR	Conventional PT care N = 153 (140) Home; NR	**KOOS** Baseline (SD): Intervention—37.0 (12.0) Comparator—36.0 (13.0) Diff Δ*: 6 wk: -1.8 12 wk: 1.4	**NRS** Baseline (SD): Intervention—5.2 (2.1) Comparator—5.7 (2.0) Diff Δ (12 wk)*: 0.2	Number of falls (12 wk): Intervention—19.4% Comparator—14.6% Difference 4.8% (-2.6, 12.3) **Physical performance (6 wk)** • ROM extension, flexion

Author, Year RoB Country	Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes		
				Pain-Related Functioning	**Pain Intensity/Severity**	**Adverse Events & Other Eligible Outcomes**
Shim, 2023[88] Some concerns Korea	Post-TKA and discharged home; mean ages 68-72 yrs; 75-82% women	Exercises using Microsoft Kinect N = 28 (27) Home; 12 wk	Conventional rehabilitation N = 28 (27) Home; 12 wk	**WOMAC** Baseline (SD): Intervention—83.1 (13.0) Comparator—81.1 (14.4) Diff Δ NC†	**NRS** Baseline (SD): Intervention—5.7 (2.1) Comparator—5.5 (2.2) Diff Δ NC†	NR **Quality of life (3 mo)** • EuroQoL-5D **Physical performance (3, 12, 24 wk)** • 4-meter gait speed • Berg balance scale • Quadriceps strength • Hamstring strength • ROM

Notes. * Diff Δ calculated by review team, unable to standardize as no SD for change reported.

† Diff Δ not reported and cannot be calculated using provided results.

‡ Standardized Diff Δ calculated by review team.

Abbreviations. Diff Δ=difference in change scores; EuroQoL-5D=European Quality of Life-5 dimensions KOOS=Knee Injury and Osteoarthritis Outcome Score; mo=month; NC=not calculable; NRS=Numeric Rating Scale; OKS=Oxford Knee Score; PT=physical therapy; RoB=risk of bias; ROM=range of motion; SD=standard deviation; SF-36=36-item Short Form health survey; TKA=total knee arthroplasty; THA=total hip arthroplasty; TUG=timed up and go; VAS=Visual Analog Scale; VR=virtual reality; wk=week; WOMAC=Western Ontario and McMaster Universities Arthritis Index.

VR Intervention Trials

Both VR trials evaluated short-term interventions in the immediate post-surgical period while participants were still hospitalized, comparing these with standard physical therapy for rehabilitation after primary total knee replacement.[82,83] Both studies included older men and women (mean ages 66-70 years, 53-58% women) previously diagnosed with osteoarthritis. Fuchs, 2022[82] was conducted in Israel and compared watching a nature or music film on Samsung Gear VR headsets while undergoing continuous passive motion (CPM), physiotherapy (n = 30), or the same CPM physiotherapy without VR (n = 25) on post-operative days 1-2. Jin, 2018[83] was conducted in China and compared daily boat-rowing exercises requiring knee flexion via VR headset (Mide Technology Inc.) (n = 33) with conventional physical therapy, also involving knee flexion (n =33). These exercises began on post-operative day 1 but total duration of therapy was not reported. Both trials were rated high RoB due to concerns in randomization and allocation procedures, appropriateness of statistical analyses, and missing data.

The evidence is very uncertain on the effects of VR for pain-related functioning or pain intensity when compared with non-VR rehabilitation (low COE, Table 20). Neither trial reported on adverse events. Both VR studies evaluated change in WOMAC as the measure of pain-related functioning. Fuchs, 2022[82] found that WOMAC scores similarly worsened for both groups at 6 months (Diff Δ=1.1), while Jin, 2018[83] reported improvement in WOMAC scores in both groups at 1, 3, and 6 months, with greater reductions in WOMAC in the VR group (*eg*, Diff Δ = −5.6 at 6 months). Both VR studies also assessed pain intensity using the VAS. Fuchs, 2022[82] measured VAS before and after therapy sessions on post-operative days 1 and 2, but only reported medians (IQR). Jin, 2018[83] assessed VAS on post-operative days 1, 3, 5, and 7, with slightly greater reductions in the VR group (*eg*, Diff Δ = −0.5 at day 7, compared with day 1). Jin, 2018[83] also evaluated knee ROM before surgery and on post-operative days 3, 7, and 14.

Table 20. Certainty of Evidence: VR for Post-Surgical Pain & Rehabilitation

Outcome Outcome Measure	Follow-up No. of Participants (Studies)	Anticipated Absolute Effects			Certainty	What Happens
		VR	Comparator	Diff Δ		
Pain-Related Functioning or Interference WOMAC	6 months N = 121 2 RCTs[83,89]	-23.5*	-17.9*	-5.6*	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of VR on pain-related functioning when compared to non-VR rehabilitation.
Pain Intensity or Severity VAS	2-7 days N = 121 2 RCTs[83,89]	-3.5*	-3.0*	-0.5*	⊕◯◯◯ Very low[a,c]	The evidence is very uncertain on the effect of VR on pain intensity when compared to non-VR rehabilitation.

Notes. * Values for mean change in VR and comparator groups, and Diff. Δ calculated from data reported in Jin, 2018.[83]
GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. Downgraded 2 levels for study limitations (study rated high risk of bias).

b. Downgraded 1 level for inconsistency (direction of effects inconsistent across studies).

c. Downgraded 1 level for indirectness (pain only measured at 1 week post-surgery in 1 study; other study only reported medians).

Abbreviations. CI=confidence interval; Diff Δ=difference between groups in mean change scores; RCT=randomized controlled trial; VAS=Visual Analog Scale; VR=virtual reality; WOMAC=Western Ontario and McMaster Universities Arthritis Index.

AR Interventions

Three trials evaluated AR physical activity interventions using Microsoft Kinect,[84,85,88] while the other 2 trials used other AR sensor technologies that also monitored participant movements and displayed these in digital environments.[86,87] AR intervention duration was 2 weeks to 4 months, and all 5 trials compared AR physical activity against usual care or standard rehabilitation. One study was conducted in the US,[85] 1 in Asia,[88] and the others in Europe.[84,86,87] Studies included 52-306 middle-aged and older men and women (mean ages 53-73 years, 49-82% women). Three studies were rated some concerns for RoB due to a range of issues, including potential bias in outcomes assessment, concerns about adherence, and missing data.[85,87,88] The other 2 were rated high RoB due to high drop-out and missing data.[84,86]

The evidence is very uncertain on the effects of AR physical activity interventions on pain-related functioning, pain intensity, and adverse events, compared with standard rehabilitation (very low COE, Table 21). Pain-related functioning was evaluated in 4 studies, using WOMAC, Oxford Knee Score (OKS), and Knee Injury and Osteoarthritis Outcome Score (KOOS).[84,85,87,88] Outcomes at 3-4 months were inconsistent across studies. For example, Eichler, 2019[84] found improvement in both groups at 3 months, with slightly greater reduction in WOMAC in the AR group (standardized Diff Δ -0.29), but Prvu-Bettger, 2020[87] showed slightly greater reduction in KOOS in the control group at 3 months (Diff Δ -1.4).

Similarly, pain intensity was assessed in 4 trials using VAS or NRS, and there were inconsistent results across studies.[85-88] While Janhunen, 2023[85] and Shim, 2023[88] reported greater reductions in pain intensity in the AR group (*eg*, standardized Diff Δ -0.39 at 4 months),[85] Piqueras, 2013[86] showed greater reductions in the control arm at 3 months (standardized Diff Δ 0.22). Prvu-Bettger, 2020[87] found very similar improvements in both groups at 3 months using NRS (Diff Δ 0.2) and using the VAS, assessed at baseline, 2 months, and 4 months. Both groups improved in VAS a similar amount (standardized Diff Δ -0.39, $p = 0.18$). The Kinect group had a baseline mean score of 57.1 (SD 18.3) with mean reduction of 36.3 at 4 months, while the control group baseline mean was 54.2 (SD 21.6) and decreased by 26.7 over this time.

Only 1 trial reported on adverse events. Prvu-Bettger, 2020[87] assessed the proportion of participants who experienced any falls during the 12 weeks post-discharge, finding 19% (27/139) of AR group and 15% (20/137) of the control group had this event. However, they did not attempt to determine if these falls were related to the treatment or another health condition. No other types of events or symptoms were assessed.

Two studies assessed quality of life. Shim, 2023[88] reported EuroQoL-5D index scores at baseline, 3, 12 and 24 weeks. Both groups showed some improvement at follow-up, with no clear difference between groups. Eichler, 2019[84] reported 36-item Short Form Survey (SF-36) physical and mental component scores (PCS and MCS, respectively) at baseline and 3-month follow-up. In both groups, PCS improved

at 3 months but similarly there were no clear differences between groups (standardized Diff Δ -0.04). The AR group mean baseline PCS was 33.8 (SD 7.6) with mean improvement of 10.7 at 3 months, while the control group baseline PCS was 33.3 (SD 7.9) with a mean improvement of 11.1 (SD 7.2). In both groups, MCS did not change at 3 months and there were no clear differences between groups (standardized Diff D Δ -0.24). The AR group mean baseline MCS was 54.8 (SD 10.6) with a change of -2.5 (SD 12.4) at 3 months, while the control group baseline MCS was 53.9 (SD 11.8) with a change of 0.1 (SD 8.5).

All 5 trials evaluated physical performance using a variety of measures, including TUG, 6-minute walk, Short Physical Performance Battery (SPPB), and ROM at knee, among others. In general, there were either no differences in improvements between groups or slightly greater improvement in the AR groups, but there were also some inconsistent results. For example, Eichler, 2019[84] found slightly greater distances for 6-minute walk in the AR group at 3 months (standardized Diff Δ 0.16). However, while Janhunen, 2023[85] showed greater reduction in times for TUG in the AR group at 4 months (standardized Diff Δ -0.71, $p = 0.04$), Piqueras, 2013[86] reported a greater improvement in the control group at 3 months (standardized Diff Δ 0.51, $p = 0.020$).

Table 21. Certainty of Evidence: AR Physical Activity for Post-Surgical Pain and Rehabilitation

Outcome Outcome Measure(s)	Follow-Up No. of Participants (Studies)	Anticipated Absolute Effects			Certainty	What Happens
		AR Physical Activity	Comparator	Difference		
Pain-Related Functioning or Interference WOMAC; OKS; KOOS	3-4 months N = 525 4 RCTs[84,85,87,88]	12.1*	9.8*	Stand. Diff Δ: -0.32 SD*	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of AR on pain-related functioning when compared to standard rehabilitation.
Pain Intensity or Severity VAS; NRS	4 months N = 595 4 RCT[84,85,87,88]	-36.3*	-26.7*	Stand. Diff Δ: -0.39 SD*	⊕◯◯◯ Very low[a,b]	The evidence is very uncertain on the effect of AR on pain intensity when compared to standard rehabilitation.
Adverse Events	3 months N = 276 1 RCT[87]	19.4%†	14.6%†	4.8% more (2.6 fewer to 12.3 more)†	⊕◯◯◯ Very low[c,d,e]	The evidence is very uncertain on the effect of AR on falls risk compared with standard rehabilitation.

Notes. * Values for mean change in VR and comparator groups, and Stand Diff. Δ from data reported in Janhunen, 2023.[85] Pain-related functioning was assessed by OKS and pain intensity with VAS.

† Values are proportion with any falls for VR and comparator groups, and difference between groups from Prvu Bettger, 2020.[87]

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations:

a. Downgraded 2 levels for study limitations (studies rated some concerns and high RoB).

b. Downgraded for inconsistent results across studies.

c. Downgraded 1 level for study limitations (studies rated some concerns RoB).

d. Downgraded 1 level for indirectness (authors only reported falls without assessment of cause, and no other adverse events).

e, Downgraded 1 level for imprecision (confidence interval includes fewer events and substantially more events).

Abbreviations. CI=confidence interval; Diff Δ=Between-group difference in mean change scores (Intervention-Comparator); KOOS=Knee Injury and Osteoarthritis Outcome Score; NRS=numeric rating scale; OKS=Oxford Knee Score; VAS=visual analogue score; WOMAC=Western Ontario and McMaster Universities Arthritis Index.

KQ2 OTHER CONDITIONS

Four studies (6 articles) evaluated XR interventions for a variety of pain conditions; these included 2 VR studies for neck pain[90] and work-related injuries,[91] and 2 AR studies, both for post-stroke rehabilitation.[92-95] Trial characteristics and main findings are summarized in Table 22. Detailed characteristics and results are provided in Appendix K.

VR Interventions

Sarig Bahat, 2020[90] compared VR physical activity (Oculus Rift with a built-in 3D tracker, $n = 22$) to conventional physical therapy ($n = 23$) for flight-associated neck pain in Israeli pilots (mean ages 28-30, 95% men). Three different VR modules (ROM, velocity, and accuracy) were used to improve head control. Participants were asked to use the VR at least 5 minutes per day, 4 times per week for 4 weeks. This study was rated as some concerns RoB due to concerns with adherence to the intervention and missing outcome data. Pain-related functioning was assessed at baseline, post-intervention (4 weeks), and 7 months, but only medians (IQR) were reported. The NDI at 7 months was 9 (6, 18) for VR physical activity group and 18 (6, 26) for control group; baseline medians (IQR) were similar. VAS was used to assess average pain intensity during the past week at baseline, 4 weeks, and 7 months. The control group had greater reductions in pain (Diff Δ 11.9 at 4 weeks, 11.6 at 7 months). Additionally, authors reported mean scores of various physical performance measures at baseline and 4 weeks (*eg*, ROM flexion and extension, isometric strength flexion and extension, global peak velocity).

Abd-Elsayed, 2021[91] was a retrospective cohort study conducted in the US and included participants who used a VR psychological skills program for work-related injuries (and who were receiving workers' compensation, $n = 36$). This study was rated as critical risk of bias due to presenting only uncontrolled pre-post results. This study also assessed self-reported use of opioids by participants.

AR Interventions

Both AR studies evaluated AR physical activity interventions for post-stroke rehabilitation of upper limb functioning. The first study, Rodriguez-Hernandez, 2021a (and other articles),[92,93,96] was conducted in Spain and compared exercises enabled by a variety of AR devices (Microsoft Kinect, Hand-Tutor glove, and 3D Tutor; $n = 23$) with conventional rehabilitation ($n = 23$). Both groups completed 150 minutes of therapy per day, for 5 consecutive days per week for 3 weeks. This study was rated high RoB due to concerns about deviations from intended interventions and missing outcome data. Quality of life was assessed with EuroQoL-5D (only domain scores reported) at baseline and 3 months. This study also assessed several physical performance measures, including the Fugl-Meyer Assessment-Upper Extremity and Action Research Arm Test.[93,96] The AR physical activity group

showed greater improvement on both physical performance tests, for example Diff Δ 10.1 at 3 weeks and 8.9 at 3 months on the Fugl-Meyer Assessment.

The second study, Taveggia, 2016,[94] was conducted in Italy and compared an AR physical activity program using Armeo Spring exoskeleton device (Hocoma Inc., Zurich, Switzerland) for sensing movements around the shoulder, elbow, and wrist joints and displaying these movements in a digital virtual environment (n = 27) with conventional rehabilitation (n = 27). Both groups received therapy for 1 hour per day, 5 days per week for 6 weeks. This study was rated some concerns for ROB due to concerns about deviations from intended interventions and missing outcome data. Pain-related functioning was not assessed. Pain intensity was measured using VAS at baseline, 6 weeks, and 12 weeks, showing that the AR physical activity group had greater reductions in pain intensity (Diff Δ −1.1 at 6 weeks, -1.9 at 12 weeks). This study reported that no adverse events were detected in either group. Physical performance was measured with the Motricity Index, and the AR group showed greater improvements at follow-up (Diff Δ 6.3 at 4 weeks, 37.9 at 7 months).

Table 22. Summary of Findings for KQ2 Other Conditions

Author, Year Study Design RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Outcomes Pain Intensity or Severity	Outcomes Adverse Events & Other Eligible Outcomes
VR Intervention Studies						
Sarig Bahat, 2020[90] RCT Some Concerns Israel	Flight-associated neck pain Air force pilots with average neck pain ≥20/100 on VAS during past week; mean ages 28-30 yr, 91% men	Exercises aimed at improving ROM, velocity, and accuracy using Oculus Rift N = 22 (18) Home; 4 wk	Conventional physical therapy N = 23 (17) NA; NA	**NDI** Median (IQR): Intervention: Baseline—15 (12,22) 4 wk—10 (6,26) 7 mo—9 (6,18) Comparator: Baseline—16 (10,20) 4 wk—16 (8,20) 7 mo—18 (6,26)	**VAS** Baseline means (SD): Intervention—36.4 (22.9) Comparator—49.5 (21.1) Diff Δ: 4 wk: 11.9 7 mo: 11.6*	NR **Physical performance (4 wk)** • ROM flexion, extension, rotation bilaterally • Isometric strength flexors, extensors • Global peak velocity, mean velocity, time to peak velocity, & accuracy
Abd-Elsayed, 2021[91] Cohort Critical USA	Acute or chronic workplace and injuries receiving workers' compensation; mean age 45 yrs, 56% female	Using PICO headset for modules on coping skills and pain education, and telephone consultations with range of clinicians N = N/A (36) Home; 90 days	N/A	NR	**VAS** Baseline mean: 6.0† 12 wk mean: 5.4† Δ: -0.6	NR **Opioid use (90 day)** • Proportion with increase or decrease in opioids, or cessation
AR Intervention Trials						
Rodriguez-Hernandez, 2021a[92]; Rodriguez-Hernandez, 2021b[96];Rodriguez-	Post-stroke rehabilitation ≤ 6 mo since stroke, upper limb involvement, dependence in ADLs, absence of other serious and disabling	Microsoft Kinect, Hand-Tutor glove, and 3D Tutor devices for sensing and displaying movements to participants and therapists	Conventional rehabilitation N = 23 (20) Clinic; 3 weeks	NR	NR	NR **Quality of life (3 mo)** • European Quality of Life-5 dimensions (EuroQoL-5D) domain scores (mobility, selfcare, daily activities,

Author, Year Study Design RoB Country	Pain Condition Key Participant Characteristics	Intervention N Randomized (N Analyzed) Setting; Duration	Comparator N Randomized (N Analyzed) Setting; Duration	Outcomes Pain-Related Functioning or Interference	Pain Intensity or Severity	Adverse Events & Other Eligible Outcomes
Hernandez, 2023[93] RCT High Spain	pathologies; mean age 63 yr, 19% women	N = 23 (23) Clinic; 3 weeks				pain/discomfort, anxiety/depression) **Physical performance (3 mo)** • Fugl-Meyer Assessment Upper Extremity • Action Research Arm Test
Taveggia, 2016[94] RCT Some concerns Italy	Post-stroke rehabilitation 0.5-12 mo post-stroke (1st episode), with self-reported functional impairments of upper extremity, no peripheral nerve injury or MSK condition of affected limb, no contracture or invasive treatment for spasticity in past 6 mo	Armeo Spring exoskeleton device for sensing arm and hand movements, displays movements in digital environment N = 7 (27) Clinic; 6 weeks	Conventional rehabilitation N = 27 (27) Clinic; 6 weeks	NR	**VAS** Baseline means (SD): Intervention: 4.5 (1.5) Comparator: 4.2 (2.0) Diff Δ: 6 wk: -1.1* 12 wk: -1.9*	No adverse events reported in either group **Physical performance (3 mo)** • Motricity Index

Notes. * Diff Δ calculated by review team, unable to standardize as no SD for change reported.

† No SD reported; these are means before using VR headset (article also reported post-use means and those were lower but had the similar change from baseline to 12 weeks, Δ =-0.7).

Abbreviations. ADL=activities of daily living; AR=augmented reality; Diff Δ=difference in mean change scores; IQR=interquartile range; mo=month; MSK=musculoskeletal; N/A=not applicable; NDI=Neck Disability Index; NR=not reported; RCT=randomized controlled trial; RoB=risk of bias; ROM=range of motion; SD=standard deviation; VAS=Visual Analog Scale; VR=virtual reality; wk=week.

DISCUSSION

In summary, trials most commonly evaluated XR interventions involving physical activity, and most often compared XR to standard, non-XR exercise interventions or conventional rehabilitation. XR physical activity interventions appear to have benefit for some conditions (*eg*, chronic neck pain), but the evidence is very uncertain for others (*eg*, chronic low back pain). XR psychological skills interventions were also evaluated for several different conditions, most often compared to a variety of non-active interventions, including usual care or VR sham. VR psychological skills may have some benefit for chronic low back pain, but these interventions were not compared with active treatments. While pain-related functioning and pain intensity were well reported by included studies, adverse events were not frequently addressed, and when reported, focused mostly on the XR group.

Summary of Key Findings

Chronic Low Back Pain (KQ1)

- 6 studies evaluated different types of VR interventions for chronic low back pain, with variable comparators (VR sham, usual care); only 1 VR study reported on adverse events (15% experienced VR-associated dizziness).
- The effect of VR embodiment (compared with either non-VR physical therapy or VR sham) on pain-related functioning and pain intensity is very uncertain (very low COE).
- VR psychological skills, compared with VR control or usual care, may result in greater improvement in pain-related functioning and pain intensity (low COE), but the evidence on adverse events is very uncertain (very low COE).
- All 16 AR studies evaluated AR physical activity interventions, compared most often with non-AR physical activity ($k = 10$); only 2 studies addressed adverse events, with both reporting no events detected.
- The effects of AR physical activity (compared with non-AR physical activity, medications, or usual care) on pain-related functioning, pain intensity, and adverse events are very uncertain (very low COE).

Neck Pain (KQ1)

- 6 trials evaluated XR physical activity interventions (5 VR, 1 AR) for chronic neck pain, all compared with non-XR physical activity programs.
- VR physical activity interventions, compared with non-VR physical activity, may result in little to no difference in pain-related functioning at 3-6 weeks (low COE) and the evidence is very uncertain for effects at 3-4 months (very low COE); for pain intensity, VR physical activity interventions may result in greater improvement at 3-6 weeks but little to no difference at 3-4 months (low COE).
- AR physical activity intervention, compared with non-AR physical activity, may improve pain-related functioning and reduce pain intensity (low COE).

- The evidence is very uncertain on adverse effects of XR physical activity interventions for neck pain (very low COE); 5 studies evaluated adverse events, but 3 only reported events for the XR group and 2 studies did not detect events in either group.

Fibromyalgia (KQ1)

- 5 trials evaluated AR interventions for fibromyalgia, with 4 of these being physical activity programs using Wii or Kinect; no trial reported on adverse events.
- The effects of AR physical activity interventions (compared with either non-AR physical activity or usual care) on pain-related functioning and pain intensity are very uncertain (very low COE).
- The effect of AR-enhanced CBT, compared with usual care, on pain-related functioning is very uncertain (very low COE).

Chronic Knee Pain (KQ1)

- 5 trials evaluated XR physical activity interventions (1 VR, 4 AR) for chronic knee pain, all compared with non-XR rehabilitation programs; only 1 trial reported on adverse events.
- A VR physical activity intervention may result in better pain-related functioning and less pain at 7 weeks, compared with standard rehabilitation (low COE).
- The effects of AR interventions (compared with standard rehabilitation) on pain-related functioning, pain intensity, and adverse events are very uncertain (very low COE).

Post-Surgical Pain & Rehabilitation (KQ2)

- 7 trials evaluated XR physical activity interventions (2 VR, 5 AR) for rehabilitation after knee or hip replacement surgery, all compared with non-XR standard rehabilitation; only 1 trial reported on adverse events.
- The effects of XR physical activity interventions (compared with standard rehabilitation) on pain-related functioning, pain intensity, and adverse events are very uncertain (very low COE).

Other Conditions (KQ1 & KQ2)

- 15 studies evaluated a range of XR interventions (7 VR, 8 AR) for a variety of other pain conditions; these were all small (total n = 36-75), half were physical activity interventions (k = 8), and a quarter were psychological skills programs (k = 4).

Limitations

Defining and separating VR and AR interventions can be challenging, and existing frameworks for XR technologies sometimes differ in where boundaries are drawn. Although most previous systematic reviews of XR interventions for pain did not stratify results by level of immersion, we sought to operationalize the distinction between full immersion (*ie*, VR) and partially immersive experiences (*ie*, AR) in order to provide greater clarity on benefits and harms. To ensure we addressed a broad range of clinically relevant XR interventions, we included some AR interventions that many would consider minimally immersive. We also categorized XR intervention types into broad categories that led to grouping together interventions with varying schedules, durations, and content. We limited eligibility to English-language studies, and thus did not include or review non-English studies.

EVIDENCE GAPS & FUTURE RESEARCH

The evidence on XR interventions for chronic pain is hampered by serious methodological concerns (half of eligible studies were rated high for RoB) and by the small size of most study samples. Three-quarters of ongoing (or recently completed) trials on XR interventions for chronic pain were also quite small, with expected total $n < 100$. Due to concerns regarding the limits of randomization to achieve balance in very small trials (with respect to baseline measures and unmeasured confounding), we elected to calculate the between-group differences in change scores (Diff Δ), instead of directly comparing follow-up scores. Although we undertook this strategy to provide the most informative interpretation of study findings, this approach cannot eliminate risks of bias that would be successfully addressed by randomization in sufficiently large trials. The preponderance of small pilot studies using convenience samples is consistent with the early state of the science in the emerging field of XR interventions for pain care.

XR can be a means of delivery of a range of clinical interventions and types of therapies. To identify the highest-value contributions of XR to pain care, future research is needed to clarify how VR and AR may improve pain outcomes, including the *mechanisms* by which XR technologies may enhance the benefits of specific types of pain therapies. For the most part, included studies were not designed to address this question, often because XR interventions were not compared with analogous *non*-XR intervention types. For example, while several XR psychological skills interventions used fully immersive VR technology, only 1 of these studies used a non-XR psychological skills comparator, so it remains unclear what additional benefits or harms may be attributable to XR technology. In contrast, while many XR physical activity interventions were compared to non-XR physical activity, many of these evaluated less immersive AR devices, and none compared VR and AR interventions to better understand the relative effects of higher levels of immersion. In some cases, it may be more straightforward to envision how immersion may contribute to the key mechanism of a pain therapy. For example, AR embodiment could help participants imagine a movable limb (in place of an amputation) in a real-world setting, leading to less phantom limb pain. Even in this case, however, there is a lack of studies demonstrating added value of XR technology when compared to established interventions using analogous mechanisms (*eg*, conventional mirror therapy).[97-99] We found 1 small eligible study comparing AR mirror therapy with conventional mirror therapy for phantom limb pain, which showed that conventional mirror therapy had greater effects on primary outcomes than AR mirror therapy.[80]

One commonly proposed general mechanism for XR benefits is increased patient engagement with varying interventions. To understand whether patient engagement is affected positively or negatively by different forms of XR technology, it would be important to evaluate patient engagement (including adherence and patient experience) in comparing XR interventions with the analogous *non*-XR interventions. Current evidence also does not address XR technology acceptance across diverse patient populations or evaluate how this may impact intervention effects. Although indicators of patient XR acceptance and experience were beyond the scope of this review, we notably found few studies evaluating XR interventions for older adults with conditions such as chronic low back pain and neck pain that are highly prevalent in older age groups. This is an important gap given potentially greater barriers in technology literacy and acceptance in older adults, although early qualitative studies suggest that barriers to technology use and acceptance may be overestimated among some older adults with chronic pain.[100] It may also be important to apply conceptual frameworks for assessing technology acceptance for XR interventions to better understand how this contributes to patient experience and uptake of interventions.[101,102] Future work will need to evaluate acceptance of XR technology among

diverse populations, investigate associated intervention engagement and adherence, and examine outcomes and implementation resources for populations that have high prevalence of chronic pain conditions but may face greater barriers—internal or external—to technology adoption.

Included studies rarely evaluated adverse events using rigorous methods that allowed for direct comparisons between XR and non-XR interventions. This is a critical gap to address in future research, as adverse events are an important component of the patient experience and often impact whether someone will start or continue an intervention. At a minimum, adverse events should be assessed systematically and reported for each arm, and involve participant interviews with open-ended questions and/or checklists.[103] In addition to information on whether the adverse event led to discontinuation of the treatment, studies should report the rates of serious adverse events (usually defined as events that are life-threatening, requiring hospitalization, or resulting in persistent disability). Accurate observation of serious adverse event rates will also require substantially larger studies. Furthermore, it will be important to examine whether adverse events for XR interventions vary for different subgroups of patients, such as by sex or gender. For example, there is some evidence that VR-related nausea and motion sickness are more common among women than men, and that this may be driven by greater incompatibility of many VR devices with the observed range of interpupillary distances among women.[104]

Intervention dose and duration varied widely across included studies, as has also been noted by previous reviews of XR interventions for pain.[105] Minimal effective dose and duration are likely to vary by intervention type and mechanism. Distraction, for example, may have temporary effects that do not require multiple frequent sessions,[105] while the benefits of physical activity accumulate over repeated sessions. Overall, XR interventions lasted from several weeks to several months, with studies often not providing the rationale for length, number, frequency, or duration of XR sessions. Studies of some analogous non-XR interventions also have similar limitations; for example, past reviews have found similar dose and duration variability among non-XR physical activity therapies for chronic pain.[106] Non-XR psychological skills interventions for chronic pain also vary in dose and duration, although there are many examples of effective low-intensity low-cost interventions.[107-109] Future studies on XR psychological skills interventions should consider applying methods developed for analogous non-XR psychological interventions to provide critical information on minimum effective dose and duration, and the impact of participant adherence on these values.[110-114]

Limited understanding of XR's impacts on pain therapy mechanisms and intervention adherence also makes it difficult to differentiate when XR is a key active component of the therapeutic intervention versus enhancing the effects of another therapy (*eg*, by increasing engagement). This distinction between core intervention components and modifiable peripheral components is fundamental to implementation science, making implementation science frameworks potentially helpful even in this early phase of XR pain research.[115,116] As is often true when incorporating new technologies into various existing interventions, there will likely be situations in which XR technology acts more as a core component of the intervention itself, and cases in which XR technology augments the effects of an intervention as an implementation strategy. Intentionally designing hybrid implementation-effectiveness studies as part of the research continuum may help clarify XR contributions earlier and in more pragmatically applicable ways.

Use of XR interventions in clinical settings is also particularly dependent on implementation contexts, making implementation research key to effective rollout for XR interventions with demonstrated benefits. Existing recommendations for XR research have focused on content development through

user centered-design, early testing for feasibility and acceptability, and design of clinical trials to demonstrate efficacy.[117] Although some recommendations for implementation exist in patient populations such as brain injury,[118] further work using implementation science frameworks is needed to identify how best to integrate XR technologies into the clinical space and optimize pain outcomes. Clinicians often have a wide range of interest and experience in using XR technologies, as patients do, and substantial facility investment may be required for clinician training. Remote or home use of XR interventions for self-management, particularly with clinician support, may also provide meaningful benefits. Employing an implementation science framework (*eg*, RE-AIM) during development phases and beyond could speed the translation of XR technologies to the end user. Although cost considerations were beyond the scope of this review, an evaluation of resources needed for XR treatments for pain should consider both the cost of treatments themselves (particularly compared with analogous non-XR therapies) and what is needed for successful implementation (*eg*, staff training materials and time, logistical support for distributing XR devices to patients). As VHA has begun to pilot XR interventions across different clinical settings, including the use of RelieVRx in outpatient treatment of chronic low back pain,[14] these implementation and hybrid evaluations should be considered as important next steps for understanding the real-world effectiveness and value of these interventions.

In summary, XR technology has considerable potential as part of a comprehensive plan for pain treatment. Given possibilities for home use and remote monitoring, and the increasing affordability of some XR technologies, XR interventions may address some common patient barriers to access and use of non-drug therapies for pain.[119] But it remains unclear how and under which circumstances XR adds the most benefit and the least risks for pain treatments. Evaluating benefits and risks in generalizable ways will require larger studies that include more diverse populations (particularly those who may experience more barriers to technology use, such as older adults and rural populations*)*. Additionally, future work should compare XR interventions to analogous non-XR intervention types; investigate mechanisms and added value of XR technology; systematically evaluate adverse events; and examine participant experiences, including attitudes toward XR, as well as barriers and facilitators of access and use.

CONCLUSIONS

Evidence on benefits and harms of XR interventions to treat or prevent chronic pain is limited due to methodological concerns, small study size, and lack of reporting on adverse events. XR physical activity interventions may have benefits for some conditions (*eg*, chronic neck pain) but the evidence is very uncertain for others (*eg*, chronic low back pain and post-surgical pain and rehabilitation). XR psychological skills interventions may also have some benefit for chronic low back pain, but studies did not compare to analogous non-XR treatments. Future work is needed to better understand how and by what mechanisms XR interventions may impact pain outcomes, particularly in more diverse populations and settings. Larger studies and application of implementation frameworks are important next steps for advancing this field.

REFERENCES

1. Rikard SM, Strahan AE, Schmit KM, Guy GP, Jr. Chronic Pain Among Adults - United States, 2019-2021. *MMWR Morb Mortal Wkly Rep*. Apr 14 2023;72(15):379-385. doi:10.15585/mmwr.mm7215a1
2. Kligler B, Bair MJ, Banerjea R, et al. Clinical Policy Recommendations from the VHA State-of-the-Art Conference on Non-Pharmacological Approaches to Chronic Musculoskeletal Pain. *J Gen Intern Med*. May 2018;33(Suppl 1):16-23. doi:10.1007/s11606-018-4323-z
3. Becker WC, DeBar LL, Heapy AA, et al. A Research Agenda for Advancing Non-pharmacological Management of Chronic Musculoskeletal Pain: Findings from a VHA State-of-the-art Conference. *J Gen Intern Med*. May 2018;33(Suppl 1):11-15. doi:10.1007/s11606-018-4345-6
4. Sterling M, de Zoete RMJ, Coppieters I, Farrell SF. Best Evidence Rehabilitation for Chronic Pain Part 4: Neck Pain. *J Clin Med*. Aug 15 2019;8(8)doi:10.3390/jcm8081219
5. Rice D, McNair P, Huysmans E, Letzen J, Finan P. Best Evidence Rehabilitation for Chronic Pain Part 5: Osteoarthritis. *J Clin Med*. Oct 24 2019;8(11)doi:10.3390/jcm8111769
6. PAIN MANAGEMENT. In: Administration DoVA-VH, editor. Washington, DC2009.
7. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. Oct 29 2011;378(9802):1560-71. doi:10.1016/S0140-6736(11)60937-9
8. Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*. Apr 2000;85(3):317-332. doi:10.1016/S0304-3959(99)00242-0
9. Milgram PK, Fumio. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*. 1994;E77-D(No. 12 December 1994)
10. Rauschnabel PA, Reto Felix, Chris Hinsch, Hamza Shahab, Florian Alt. What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior*. 2022;133(August)doi:https://doi.org/10.1016/j.chb.2022.1072890747-5632
11. Goudman L, Jansen J, Billot M, et al. Virtual Reality Applications in Chronic Pain Management: Systematic Review and Meta-analysis. *JMIR Serious Games*. May 10 2022;10(2):e34402. doi:10.2196/34402
12. Dreesmann NJ, Su H, Thompson HJ. A Systematic Review of Virtual Reality Therapeutics for Acute Pain Management. *Pain Manag Nurs*. Oct 2022;23(5):672-681. doi:10.1016/j.pmn.2022.05.004
13. Osumi M, Inomata K, Inoue Y, Otake Y, Morioka S, Sumitani M. Characteristics of Phantom Limb Pain Alleviated with Virtual Reality Rehabilitation. *Pain Med*. May 1 2019;20(5):1038-1046. doi:10.1093/pm/pny269
14. VA Diffusion Marketplace. https://marketplace.va.gov/
15. DistillerSR. DistillerSR Inc. https://www.distillersr.com/
16. Cochrane Handbook for Systematic Reviews of Interventions. Updated August 22, 2023. https://training.cochrane.org/handbook
17. Sterne JAC, Savovic 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
18. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. Oct 12 2016;355:i4919. doi:10.1136/bmj.i4919
19. metafor: Meta-Analysis Package for R. https://cran.r-project.org/web/packages/metafor/index.html

20. Sil S, Dahlquist LM, Thompson C, et al. The effects of coping style on virtual reality enhanced videogame distraction in children undergoing cold pressor pain. Psychosocial & Personality Development 2840. *Journal of Behavioral Medicine*. 2014;37(1):156-165. doi:https://dx.doi.org/10.1007/s10865-012-9479-0
21. Werbart A. Emile, or on devastation: When virtual boundlessness meets inner emptiness. Psychoanalytic Therapy 3315. *The Psychoanalytic Quarterly*. 2014;83(1):71-96. doi:https://dx.doi.org/10.1002/j.2167-4086.2014.00077.x
22. GRADEpro GDT. https://www.gradepro.org/
23. Maddox T, Liesl Oldstone, Charisse Y. Sparks, Josh Sackman, Alexis Oyao, Laura Garcia, Roselani U. Maddox, Kelsey Ffrench, Heidy Garcia, Takisha Adair, Ann Irvin, David Maislin, Brendan Keenan, Robert Bonakdar, Beth D. Darnall,. In-Home Virtual Reality Program for Chronic Lower Back Pain: A Randomized Sham-Controlled Effectiveness Trial in a Clinically Severe and Diverse Sample. *Mayo Clinic Proceedings: Digital Health*. 2023;1(4):563-573. doi:https://doi.org/10.1016/j.mcpdig.2023.09.003
24. Harvie DS, Kelly J, Kluver J, Deen M, Spitzer E, Coppieters MW. A randomized controlled pilot study examining immediate effects of embodying a virtual reality superhero in people with chronic low back pain. *Disability and rehabilitation Assistive technology*. 2022:1-8. doi:https://dx.doi.org/10.1080/17483107.2022.2129846
25. Yilmaz Yelvar GD, Cirak Y, Dalkilinc M, Parlak Demir Y, Guner Z, Boydak A. Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*. 2017;26(2):538-545. doi:https://dx.doi.org/10.1007/s00586-016-4892-7
26. Kammler-Sücker KI, Löffler A, Flor H. Effects of personalized movement models in virtual reality on pain expectancy and motor behavior in patients with chronic back pain: a feasibility study. Article. *Virtual Reality*. 2023;doi:10.1007/s10055-023-00800-4
27. Eccleston C, Fisher E, Liikkanen S, et al. A prospective, double-blind, pilot, randomized, controlled trial of an "embodied" virtual reality intervention for adults with low back pain. *Pain*. 2022;163(9):1700-1715. doi:https://dx.doi.org/10.1097/j.pain.0000000000002617
28. Groenveld TD, Smits MLM, Knoop J, et al. Effect of a Behavioral Therapy-Based Virtual Reality Application on Quality of Life in Chronic Low Back Pain. Article. *The Clinical journal of pain*. 2023;39(6):278-285. doi:10.1097/AJP.0000000000001110
29. Garcia LM, Birckhead BJ, Krishnamurthy P, et al. An 8-Week Self-Administered At-Home Behavioral Skills-Based Virtual Reality Program for Chronic Low Back Pain: Double-Blind, Randomized, Placebo-Controlled Trial Conducted During COVID-19. *Journal of medical Internet research*. 2021;23(2):e26292. doi:https://dx.doi.org/10.2196/26292
30. Garcia LM, Birckhead BJ, Krishnamurthy P, et al. Three-Month Follow-Up Results of a Double-Blind, Randomized Placebo-Controlled Trial of 8-Week Self-Administered At-Home Behavioral Skills-Based Virtual Reality (VR) for Chronic Low Back Pain. *The journal of pain*. 2022;23(5):822-840. doi:https://dx.doi.org/10.1016/j.jpain.2021.12.002
31. Garcia L, Birckhead B, Krishnamurthy P, et al. Durability of the Treatment Effects of an 8-Week Self-administered Home-Based Virtual Reality Program for Chronic Low Back Pain: 6-Month Follow-up Study of a Randomized Clinical Trial. *Journal of medical Internet research*. 2022;24(5):e37480. Erratum in: J Med Internet Res. 2022 Jun 8;24(6):e40038 PMID: 35675658 [https://www.ncbi.nlm.nih.gov/pubmed/35675658]. doi:https://dx.doi.org/10.2196/37480

32. Maddox T, Garcia H, Ffrench K, et al. In-home virtual reality program for chronic low back pain: Durability of a randomized, placebo-controlled clinical trial to 18 months post-treatment. *Regional Anesthesia and Pain Medicine*. 2022:e22104093. doi:https://dx.doi.org/10.1136/rapm-2022-104093
33. Maddox T, Sparks C, Oldstone L, et al. Durable Chronic Low Back Pain Reductions to 24-Months Post-Treatment for An Accessible, 8-Week, In-Home Behavioral Skills-Based Virtual Reality Program: A Randomized Controlled Trial. *Pain Med*. May 23 2023;doi:10.1093/pm/pnad070
34. Anam S, Uchino E, Misawa H, Suetake N. Automatic bone boundary detection in hand radiographs by using modified level set method and diffusion filter. 2013:51-55.
35. Nambi G, Alghadier M, Kashoo FZ, et al. Effects of Virtual Reality Exercises versus Isokinetic Exercises in comparison with Conventional Exercises on the Imaging Findings and Inflammatory Biomarker Changes in Soccer Players with Non-Specific Low Back Pain: A Randomized Controlled Trial. *International journal of environmental research and public health*. 2022;20(1)doi:https://dx.doi.org/10.3390/ijerph20010524
36. Fatoye F, Gebrye T, Mbada CE, et al. Cost effectiveness of virtual reality game compared to clinic based McKenzie extension therapy for chronic non-specific low back pain. *British journal of pain*. 2022;16(6):601-609. doi:https://dx.doi.org/10.1177/20494637221109108
37. Nambi G, Abdelbasset WK, Elsayed SH, et al. Comparative Effects of Isokinetic Training and Virtual Reality Training on Sports Performances in University Football Players with Chronic Low Back Pain-Randomized Controlled Study. *Evidence-based complementary and alternative medicine : eCAM*. 2020;2020:2981273. doi:https://dx.doi.org/10.1155/2020/2981273
38. Afzal MW, Ahmad A, Mohseni Bandpei MA, Gilani SA, Hanif A, Waqas MS. Effects of virtual reality exercises and routine physical therapy on pain intensity and functional disability in patients with chronic low back pain. *JPMA The Journal of the Pakistan Medical Association*. 2022;72(3):413-417. doi:https://dx.doi.org/10.47391/JPMA.3424
39. Li Z, Yu Q, Luo H, et al. The Effect of Virtual Reality Training on Anticipatory Postural Adjustments in Patients with Chronic Nonspecific Low Back Pain: A Preliminary Study. *Neural plasticity*. 2021;2021:9975862. doi:https://dx.doi.org/10.1155/2021/9975862
40. Sato T, Shimizu K, Shiko Y, et al. Effects of Nintendo Ring Fit Adventure Exergame on Pain and Psychological Factors in Patients with Chronic Low Back Pain. *Games for health journal*. 2021;10(3):158-164. doi:https://dx.doi.org/10.1089/g4h.2020.0180
41. Nambi G, Abdelbasset WK, Alsubaie SF, et al. Short-Term Psychological and Hormonal Effects of Virtual Reality Training on Chronic Low Back Pain in Soccer Players. *Journal of sport rehabilitation*. 2021;30(6):884-893. doi:https://dx.doi.org/10.1123/jsr.2020-0075
42. Nambi G, Abdelbasset WK, Alrawaili SM, Alsubaie SF, Abodonya AM, Saleh AK. Virtual reality or isokinetic training; its effect on pain, kinesiophobia and serum stress hormones in chronic low back pain: A randomized controlled trial. *Technology and health care : official journal of the European Society for Engineering and Medicine*. 2021;29(1):155-166. doi:https://dx.doi.org/10.3233/THC-202301
43. Kim T, Lee J, Oh S, Kim S, Yoon B. Effectiveness of Simulated Horseback Riding for Patients With Chronic Low Back Pain: A Randomized Controlled Trial. *Journal of sport rehabilitation*. 2020;29(2):179-185. doi:https://dx.doi.org/10.1123/jsr.2018-0252
44. Zadro JR, Shirley D, Simic M, et al. Video-Game-Based Exercises for Older People With Chronic Low Back Pain: A Randomized Controlledtable Trial (GAMEBACK). *Physical therapy*. 2019;99(1):14-27. doi:https://dx.doi.org/10.1093/ptj/pzy112
45. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and Safety of a Virtual Reality Dodgeball Intervention for Chronic Low Back Pain: A Randomized Clinical

Trial. *The journal of pain*. 2016;17(12):1302-1317. doi:https://dx.doi.org/10.1016/j.jpain.2016.08.011
46. Monteiro-Junior RS, de Souza CP, Lattari E, et al. Wii-Workouts on Chronic Pain, Physical Capabilities and Mood of Older Women: A Randomized Controlled Double Blind Trial. *CNS & neurological disorders drug targets*. 2015;14(9):1157-64.
47. Yoo JH, Kim SE, Lee MG, et al. The effect of horse simulator riding on visual analogue scale, body composition and trunk strength in the patients with chronic low back pain. *International journal of clinical practice*. 2014;68(8):941-9. Comment in: Int J Clin Pract. 2014 Aug;68(8):931 PMID: 25074333 [https://www.ncbi.nlm.nih.gov/pubmed/25074333]. doi:https://dx.doi.org/10.1111/ijcp.12414
48. Kim S-S, Min W-K, Kim J-H, Lee B-H. The Effects of VR-based Wii Fit Yoga on Physical Function in Middle-aged Female LBP Patients. *Journal of physical therapy science*. 2014;26(4):549-52. doi:https://dx.doi.org/10.1589/jpts.26.549
49. Oh HW, Lee MG, Jang JY, et al. Time-effects of horse simulator exercise on psychophysiological responses in men with chronic low back pain. *Isokinetics and Exercise Science*. 2014;22(2):153-163. doi:https://dx.doi.org/10.3233/IES-140533
50. Mbada CE, Makinde MO, Odole AC, et al. Comparative effects of clinic- And virtual reality-based McKenzie extension therapy in chronic non-specific low-back pain. Article. *Human Movement*. 2019;20(3):66-79. doi:10.5114/hm.2019.83998
51. Cetin H, Kose N, Oge HK. Virtual reality and motor control exercises to treat chronic neck pain: A randomized controlled trial. *Musculoskeletal science & practice*. 2022;62:102636. doi:https://dx.doi.org/10.1016/j.msksp.2022.102636
52. Nusser M, Knapp S, Kramer M, Krischak G. Effects of virtual reality-based neck-specific sensorimotor training in patients with chronic neck pain: A randomized controlled pilot trial. *Journal of rehabilitation medicine*. 2021;53(2):jrm00151. doi:https://dx.doi.org/10.2340/16501977-2786
53. Tejera DM, Beltran-Alacreu H, Cano-de-la-Cuerda R, et al. Effects of Virtual Reality versus Exercise on Pain, Functional, Somatosensory and Psychosocial Outcomes in Patients with Non-specific Chronic Neck Pain: A Randomized Clinical Trial. *International journal of environmental research and public health*. 2020;17(16)doi:https://dx.doi.org/10.3390/ijerph17165950
54. Sarig Bahat H, Croft K, Carter C, Hoddinott A, Sprecher E, Treleaven J. Remote kinematic training for patients with chronic neck pain: a randomised controlled trial. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*. 2018;27(6):1309-1323. doi:https://dx.doi.org/10.1007/s00586-017-5323-0
55. Sarig Bahat H, Takasaki H, Chen X, Bet-Or Y, Treleaven J. Cervical kinematic training with and without interactive VR training for chronic neck pain - a randomized clinical trial. *Manual therapy*. 2015;20(1):68-78. doi:https://dx.doi.org/10.1016/j.math.2014.06.008
56. I R, M R, S E, S K, A RZ. A Novel Virtual Reality Technique (Cervigame R) Compared to Conventional Proprioceptive Training to Treat Neck Pain: A Randomized Controlled Trial. *Journal of biomedical physics & engineering*. 2019;9(3):355-366. doi:https://dx.doi.org/10.31661/jbpe.v0i0.556
57. Polat M, Kahveci A, Muci B, Gunendi Z, Kaymak Karatas G. The Effect of Virtual Reality Exercises on Pain, Functionality, Cardiopulmonary Capacity, and Quality of Life in Fibromyalgia Syndrome: A Randomized Controlled Study. *Games for health journal*. 2021;10(3):165-173. doi:https://dx.doi.org/10.1089/g4h.2020.0162

58. Carvalho MSd, Carvalho LC, Menezes FdS, Frazin A, Gomes EdC, Iunes DH. Effects of Exergames in Women with Fibromyalgia: A Randomized Controlled Study. *Games for health journal*. 2020;9(5):358-367. doi:https://dx.doi.org/10.1089/g4h.2019.0108
59. Villafaina S, Collado-Mateo D, Dominguez-Munoz FJ, Fuentes-Garcia JP, Gusi N. Benefits of 24-Week Exergame Intervention on Health-Related Quality of Life and Pain in Women with Fibromyalgia: A Single-Blind, Randomized Controlled Trial. *Games for health journal*. 2019;8(6):380-386. doi:https://dx.doi.org/10.1089/g4h.2019.0023
60. Collado-Mateo D, Dominguez-Munoz FJ, Adsuar JC, Garcia-Gordillo MA, Gusi N. Effects of Exergames on Quality of Life, Pain, and Disease Effect in Women With Fibromyalgia: A Randomized Controlled Trial. *Archives of physical medicine and rehabilitation*. 2017;98(9):1725-1731. doi:https://dx.doi.org/10.1016/j.apmr.2017.02.011
61. Collado-Mateo D, Dominguez-Munoz FJ, Adsuar JC, Merellano-Navarro E, Gusi N. Exergames for women with fibromyalgia: a randomised controlled trial to evaluate the effects on mobility skills, balance and fear of falling. *PeerJ*. 2017;5:e3211. doi:https://dx.doi.org/10.7717/peerj.3211
62. Martin-Martinez JP, Villafaina S, Collado-Mateo D, Perez-Gomez J, Gusi N. Effects of 24-week exergame intervention on physical function under single- and dual-task conditions in fibromyalgia: A randomized controlled trial. *Scandinavian journal of medicine & science in sports*. 2019;29(10):1610-1617. doi:https://dx.doi.org/10.1111/sms.13502
63. Villafaina S, Borrega-Mouquinho Y, Fuentes-Garcia JP, Collado-Mateo D, Gusi N. Effect of Exergame Training and Detraining on Lower-Body Strength, Agility, and Cardiorespiratory Fitness in Women with Fibromyalgia: Single-Blinded Randomized Controlled Trial. International *journal of environmental research and public health*. 2019;17(1)doi:https://dx.doi.org/10.3390/ijerph17010161
64. Garcia-Palacios A, Herrero R, Vizcaino Y, et al. Integrating virtual reality with activity management for the treatment of fibromyalgia: acceptability and preliminary efficacy. *The Clinical journal of pain*. 2015;31(6):564-72. doi:https://dx.doi.org/10.1097/AJP.0000000000000196
65. Ozlu A, Unver G, Tuna HI, Menekseoglu AK. The Effect of a Virtual Reality-Mediated Gamified Rehabilitation Program on Pain, Disability, Function, and Balance in Knee Osteoarthritis: A Prospective Randomized Controlled Study. *Games for health journal*. 2023;12(2):118-124. doi:https://dx.doi.org/10.1089/g4h.2022.0130
66. Elshazly FAA, Nambi GS, Elnegamy TE. Comparative study on virtual reality training (VRT) over sensory motor training (SMT) in unilateral chronic osteoarthritis—A randomized control trial. *Int J Med Res Health Sci*. 2016;5(8):7-16.
67. Lin Y-T, Lee W-C, Hsieh R-L. Active video games for knee osteoarthritis improve mobility but not WOMAC score: A randomized controlled trial. *Annals of physical and rehabilitation medicine*. 2020;63(6):458-465. doi:https://dx.doi.org/10.1016/j.rehab.2019.11.008
68. Mete E, Sari Z. The efficacy of exergaming in patients with knee osteoarthritis: A randomized controlled clinical trial. *Physiotherapy research international : the journal for researchers and clinicians in physical therapy*. 2022;27(3):e1952. doi:https://dx.doi.org/10.1002/pri.1952
69. Nambi G, Abdelbasset WK, Elsayed SH, Khalil MA, Alrawaili SM, Alsubaie SF. Comparative effects of virtual reality training and sensory motor training on bone morphogenic proteins and inflammatory biomarkers in post-traumatic osteoarthritis. *Scientific reports*. 2020;10(1):15864. doi:https://dx.doi.org/10.1038/s41598-020-72587-2
70. Cuneo A, Yang R, Zhou H, et al. The Utility of a Novel, Combined Biofeedback-Virtual Reality Device as Add-on Treatment for Chronic Migraine: A Randomized Pilot Study. *The Clinical journal of pain*. 2023;doi:https://dx.doi.org/10.1097/AJP.0000000000001114

71. Chuan A, Hatty M, Shelley M, et al. Feasibility of virtual reality-delivered pain psychology therapy for cancer-related neuropathic pain: a pilot randomised controlled trial. *Anaesthesia*. 2023;78(4):449-457. doi:https://dx.doi.org/10.1111/anae.15971
72. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-Administered Skills-Based Virtual Reality Intervention for Chronic Pain: Randomized Controlled Pilot Study. *JMIR formative research*. 2020;4(7):e17293. doi:https://dx.doi.org/10.2196/17293
73. Reynolds LM, Cavadino A, Chin S, et al. The benefits and acceptability of virtual reality interventions for women with metastatic breast cancer in their homes; a pilot randomised trial. *BMC cancer*. 2022;22(1):360. Comment in: Int J Surg. 2022 Oct;106:106940 PMID: 36152921 [https://www.ncbi.nlm.nih.gov/pubmed/36152921]. doi:https://dx.doi.org/10.1186/s12885-021-09081-z
74. Wankhade S, Phansopkar P, Chitale N. Effect of virtual reality aided physical therapy in adjunct to traditional therapy in frozen shoulder patients. Article. *Journal of Medical Pharmaceutical and Allied Sciences*. 2022;11(4):5167-5171. doi:10.55522/jmpas.V11I4.1314
75. Ambrosino P, Iannuzzi GL, Formisano R, et al. Exergaming as an Additional Tool in Rehabilitation of Young Patients with Rheumatoid Arthritis: A Pilot Randomized Controlled Trial. *Games for health journal*. 2020;9(5):368-375. doi:https://dx.doi.org/10.1089/g4h.2019.0167
76. Ditchburn J-L, van Schaik P, Dixon J, MacSween A, Martin D. The effects of exergaming on pain, postural control, technology acceptance and flow experience in older people with chronic musculoskeletal pain: a randomised controlled trial. *BMC sports science, medicine & rehabilitation*. 2020;12:63. doi:https://dx.doi.org/10.1186/s13102-020-00211-x
77. Gouveia E Silva EC, Lange B, Bacha JMR, Pompeu JE. Effects of the Interactive Videogame Nintendo Wii Sports on Upper Limb Motor Function of Individuals with Post-Polio Syndrome: A Randomized Clinical Trial. *Games for health journal*. 2020;9(6):461-471. doi:https://dx.doi.org/10.1089/g4h.2019.0192
78. Karahan AY, Tok F, Yildirim P, Ordahan B, Turkoglu G, Sahin N. The Effectiveness of Exergames in Patients with Ankylosing Spondylitis: A Randomized Controlled Trial. *Advances in clinical and experimental medicine : official organ Wroclaw Medical University*. 2016;25(5):931-936. doi:https://dx.doi.org/10.17219/acem/32590
79. Lewis JS, Newport R, Taylor G, Smith M, McCabe CS. Visual illusions modulate body perception disturbance and pain in Complex Regional Pain Syndrome: A randomized trial. *European journal of pain (London, England)*. 2021;25(7):1551-1563. doi:https://dx.doi.org/10.1002/ejp.1766
80. Rothgangel A, Braun S, Winkens B, Beurskens A, Smeets R. Traditional and augmented reality mirror therapy for patients with chronic phantom limb pain (PACT study): results of a three-group, multicentre single-blind randomized controlled trial. *Clinical rehabilitation*. 2018;32(12):1591-1608. doi:https://dx.doi.org/10.1177/0269215518785948
81. Fascio E, Vitale JA, Sirtori P, Peretti G, Banfi G, Mangiavini L. Early Virtual-Reality-Based Home Rehabilitation after Total Hip Arthroplasty: A Randomized Controlled Trial. Article. *Journal of Clinical Medicine*. 2022;11(7)1766. doi:10.3390/jcm11071766
82. Fuchs L, Kluska A, Novak D, Kosashvili Y. The influence of early virtual reality intervention on pain, anxiety, and function following primary total knee arthroplasty. *Complementary therapies in clinical practice*. 2022;49:101687. doi:https://dx.doi.org/10.1016/j.ctcp.2022.101687
83. Jin C, Feng Y, Ni Y, Shan Z. Virtual reality intervention in postoperative rehabilitation after total knee arthroplasty: A prospective and randomized controlled clinical trial. *International Journal of Clinical and Experimental Medicine*. 2018;11(6):6119-6124.

84. Eichler S, Salzwedel A, Rabe S, et al. The effectiveness of telerehabilitation as a supplement to rehabilitation in patients after total knee or hip replacement: randomized controlled trial. *JMIR rehabilitation and assistive technologies*. 2019;6(2):e14236.
85. Janhunen M, Katajapuu N, Paloneva J, et al. Effects of a home-based, exergaming intervention on physical function and pain after total knee replacement in older adults: a randomised controlled trial. *BMJ open sport & exercise medicine*. 2023;9(1):e001416. doi:https://dx.doi.org/10.1136/bmjsem-2022-001416
86. Piqueras M, Marco E, Coll M, et al. Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthoplasty: a randomized controlled trial. *Journal of rehabilitation medicine*. 2013;45(4):392-396.
87. Prvu Bettger J, Green CL, Holmes DN, et al. Effects of Virtual Exercise Rehabilitation In-Home Therapy Compared with Traditional Care After Total Knee Arthroplasty: VERITAS, a Randomized Controlled Trial. *The Journal of bone and joint surgery American volume*. 2020;102(2):101-109. doi:https://dx.doi.org/10.2106/JBJS.19.00695
88. Shim GY, Kim EH, Lee SJ, et al. Postoperative rehabilitation using a digital healthcare system in patients with total knee arthroplasty: a randomized controlled trial. Article. *Archives of Orthopaedic and Trauma Surgery*. 2023;doi:10.1007/s00402-023-04894-y
89. Golubnitschaja O, Liskova A, Koklesova L, et al. Caution, "normal" BMI: health risks associated with potentially masked individual underweight-EPMA Position Paper 2021. *The EPMA journal*. 2021;12(3):243-264. doi:https://dx.doi.org/10.1007/s13167-021-00251-4
90. Bahat HS, German D, Palomo G, Gold H, Nir YF. Self-Kinematic Training for Flight-Associated Neck Pain: a Randomized Controlled Trial. *Aerospace medicine and human performance*. 2020;91(10):790-797. doi:https://dx.doi.org/10.3357/AMHP.5546.2020
91. Abd-Elsayed A, Hussain N, Stanley G. Combining Virtual Reality and Behavioral Health to Promote Pain Resiliency: Analysis of a Novel BioPsychoSocial Modality for Solving Pain in the Workplace. *Pain and therapy*. 2021;10(2):1731-1740. doi:https://dx.doi.org/10.1007/s40122-021-00333-1
92. Rodriguez-Hernandez M, Criado-Alvarez J-J, Corregidor-Sanchez A-I, Martin-Conty JL, Mohedano-Moriano A, Polonio-Lopez B. Effects of Virtual Reality-Based Therapy on Quality of Life of Patients with Subacute Stroke: A Three-Month Follow-Up Randomized Controlled Trial. *International journal of environmental research and public health*. 2021;18(6)doi:https://dx.doi.org/10.3390/ijerph18062810
93. Rodríguez-Hernández M, Polonio-López B, Corregidor-Sánchez AI, Martín-Conty JL, Mohedano-Moriano A, Criado-Álvarez JJ. Can specific virtual reality combined with conventional rehabilitation improve poststroke hand motor function? A randomized clinical trial. Article. *Journal of NeuroEngineering and Rehabilitation*. 2023;20(1)38. doi:10.1186/s12984-023-01170-3
94. Taveggia G, Borboni A, Salvi L, et al. Efficacy of robot-assisted rehabilitation for the functional recovery of the upper limb in post-stroke patients: a randomized controlled study. *European journal of physical and rehabilitation medicine*. 2016;52(6):767-773.
95. Ossmy O, Mukamel R. Using Virtual Reality to Transfer Motor Skill Knowledge from One Hand to Another. *Journal of visualized experiments : JoVE*. 2017;(127)doi:https://dx.doi.org/10.3791/55965
96. Rodriguez-Hernandez M, Polonio-Lopez B, Corregidor-Sanchez AI, Martin-Conty JL, Mohedano-Moriano A, Criado-Alvarez JJ. Effects of Specific Virtual Reality-Based Therapy for the Rehabilitation of the Upper Limb Motor Function Post-Ictus: Randomized Controlled Trial. *Brain Sci*. Apr 28 2021;11(5)doi:10.3390/brainsci11050555

97. Soler MD, Kumru H, Pelayo R, et al. Effectiveness of transcranial direct current stimulation and visual illusion on neuropathic pain in spinal cord injury. *Brain : a journal of neurology*. 2010;133(9):2565-77. doi:https://dx.doi.org/10.1093/brain/awq184
98. Morkisch NJ, Katrin; dos Santos, Luara Ferreira; Dohle, Christian. Modelling of therapeutic action during mirror therapy: A qualitative analysis as basis for technical solutions. *Current Directions in Biomedical Engineering*. 2017;3(No.1):45-48. doi:https://doi.org/10.1515/cdbme-2017-0010
99. Moseley LG, Gallace A, Spence C. Is mirror therapy all it is cracked up to be? Current evidence and future directions. *Pain*. Aug 15 2008;138(1):7-10. doi:10.1016/j.pain.2008.06.026
100. Dy M, Olazo K, Lyles CR, et al. Usability and acceptability of virtual reality for chronic pain management among diverse patients in a safety-net setting: a qualitative analysis. *JAMIA Open*. Oct 2023;6(3):ooad050. doi:10.1093/jamiaopen/ooad050
101. Jones C, Miguel Cruz A, Smith-MacDonald L, Brown MRG, Vermetten E, Bremault-Phillips S. Technology Acceptance and Usability of a Virtual Reality Intervention for Military Members and Veterans With Posttraumatic Stress Disorder: Mixed Methods Unified Theory of Acceptance and Use of Technology Study. *JMIR Form Res*. Apr 21 2022;6(4):e33681. doi:10.2196/33681
102. Sagnier C L-EE, Lourdeaux D, Thouvenin I, Valléry G. User Acceptance of Virtual Reality: An Extended Technology Acceptance Model. *International Journal of Human-Computer Interaction*. 06 Jan 2020 2020;36(11):993-1007. doi:https://doi.org/10.1080/10447318.2019.1708612
103. Allen EN, Chandler CI, Mandimika N, Leisegang C, Barnes K. Eliciting adverse effects data from participants in clinical trials. *Cochrane Database Syst Rev*. Jan 16 2018;1(1):MR000039. doi:10.1002/14651858.MR000039.pub2
104. Stanney K, Fidopiastis C, Foster L. Virtual Reality Is Sexist: But It Does Not Have to Be. *Front Robot AI*. 2020;7:4. doi:10.3389/frobt.2020.00004
105. Baker NA, Polhemus AH, Haan Ospina E, et al. The State of Science in the Use of Virtual Reality in the Treatment of Acute and Chronic Pain: A Systematic Scoping Review. *The Clinical journal of pain*. 2022;38(6):424-441. doi:https://dx.doi.org/10.1097/AJP.0000000000001029
106. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev*. Sep 28 2021;9(9):CD009790. doi:10.1002/14651858.CD009790.pub2
107. Richmond H, Hall AM, Copsey B, et al. The Effectiveness of Cognitive Behavioural Treatment for Non-Specific Low Back Pain: A Systematic Review and Meta-Analysis. *PLoS One*. 2015;10(8):e0134192. doi:10.1371/journal.pone.0134192
108. Khoo EL, Small R, Cheng W, et al. Comparative evaluation of group-based mindfulness-based stress reduction and cognitive behavioural therapy for the treatment and management of chronic pain: A systematic review and network meta-analysis. *Evid Based Ment Health*. Feb 2019;22(1):26-35. doi:10.1136/ebmental-2018-300062
109. Heapy AA, Higgins DM, Goulet JL, et al. Interactive Voice Response-Based Self-management for Chronic Back Pain: The COPES Noninferiority Randomized Trial. *JAMA Intern Med*. Jun 1 2017;177(6):765-773. doi:10.1001/jamainternmed.2017.0223
110. Wells SY, Morland LA, Hurst S, et al. Veterans' reasons for dropping out of prolonged exposure therapy across three delivery modalities: A qualitative examination. *Psychol Serv*. Aug 2023;20(3):483-495. doi:10.1037/ser0000714

111. Ankawi B, Piette, J.D., Buta, E. et al. Adherence to Daily Interactive Voice Response Calls for a Chronic Pain Intervention. *Journal of Technology in Behavioral Science*. 2022;(7):343–350. doi:https://doi.org/10.1007/s41347-022-00254-6
112. Meis LA, Noorbaloochi S, Hagel Campbell EM, et al. Sticking it out in trauma-focused treatment for PTSD: It takes a village. *J Consult Clin Psychol*. Mar 2019;87(3):246-256. doi:10.1037/ccp0000386
113. Murphy JL, Cordova MJ, Dedert EA. Cognitive behavioral therapy for chronic pain in veterans: Evidence for clinical effectiveness in a model program. *Psychol Serv*. Feb 2022;19(1):95-102. doi:10.1037/ser0000506
114. Beehler GP, Loughran TA, King PR, et al. Patients' perspectives of brief cognitive behavioral therapy for chronic pain: Treatment satisfaction, perceived utility, and global assessment of change. *Fam Syst Health*. Jun 2021;39(2):351-357. doi:10.1037/fsh0000606
115. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. Aug 7 2009;4:50. doi:10.1186/1748-5908-4-50
116. Damschroder LJ, Hagedorn HJ. A guiding framework and approach for implementation research in substance use disorders treatment. *Psychol Addict Behav*. Jun 2011;25(2):194-205. doi:10.1037/a0022284
117. Birckhead B, Khalil C, Liu X, et al. Recommendations for Methodology of Virtual Reality Clinical Trials in Health Care by an International Working Group: Iterative Study. *JMIR Ment Health*. Jan 31 2019;6(1):e11973. doi:10.2196/11973
118. Brassel S, Power E, Campbell A, Brunner M, Togher L. Recommendations for the Design and Implementation of Virtual Reality for Acquired Brain Injury Rehabilitation: Systematic Review. *J Med Internet Res*. Jul 30 2021;23(7):e26344. doi:10.2196/26344
119. Becker WC, Dorflinger L, Edmond SN, Islam L, Heapy AA, Fraenkel L. Barriers and facilitators to use of non-pharmacological treatments in chronic pain. *BMC Fam Pract*. Mar 20 2017;18(1):41. doi:10.1186/s12875-017-0608-2
120. Sarig Bahat H, Hadar D, Treleaven J. Predictors for Positive Response to Home Kinematic Training in Chronic Neck Pain. *Journal of manipulative and physiological therapeutics*. 2020;43(8):779-790. doi:https://dx.doi.org/10.1016/j.jmpt.2019.12.008
121. Chin S, Cavadino A, Akroyd A, et al. An Investigation of Virtual Reality Nature Experiences in Patients With Metastatic Breast Cancer: Secondary Analysis of a Randomized Controlled Trial. *JMIR cancer*. 2022;8(3):e38300. doi:https://dx.doi.org/10.2196/38300
122. Tsai L, Milburn PA, Cecil CLt, Lowry PS, Hermans MR. Comparison of Recurrence and Postoperative Complications Between 3 Different Techniques for Surgical Repair of Idiopathic Hydrocele. *Urology*. 2019;125:239-242. doi:https://dx.doi.org/10.1016/j.urology.2018.12.004