# Randomized-controlled trial of virtual reality for chronic low back pain to improve patient-reported outcomes and physical activity

**Protocol Number: STUDY00000631**

**National Clinical Trial (NCT) Identified Number: NCT04409353**

**Principal Investigator: Brennan Spiegel, MD**

**Sponsor: NIAMS**

**Funded by: NIH**

**Version Number: v.8.1**

**Revision: 19 October 2023**

**Original: 28 April 2020**

## Summary of Changes from Previous Version:

| Affected Section(s) | Summary of Revisions Made | Rationale |
|---|---|---|
| 1.3, 8.1, 8.3.1 | Added Charlson Comorbidity Index to SoA and correction to SoA, harmonization of 8.1 with 1.1 objectives, error found in 8.3.1 | Charlson Comorbidy Index was requested by BACPAC other changes were corrected to text. |
| 2.3.1,6.3, 7.1, 8.2, 8.2.1, 9.4.7, 10.1.1.2, | • Stratified randomization by recruitment site.<br>• Added method for flagging patients with severe depression and SUD symptoms<br>• Revision showing AppliedVR server is HIPPA compliant<br>• Addition of cannabinoids to comedications comprehensive list of measures for confidentiality added from MOP<br>• Addition of mood disturbance and radicular pain added to sub-group analysis<br>• A description of additional language added to consent is provided. | These changes were made due to recommendations from the DSMB. |

| | | |
|---|---|---|
| **1.1, 1.3, 5.4, 5.5, 6.1.1, 6.2.4, 7.2, 8.2.1, 9.4.9,** | • **Replacing Charge 3 for Charge 4**<br>• **Additional surveys added to SoA**<br>• **Revision to Screen Failure definition**<br>• **Change in interval of review for the race/ethnicity data to monthly assessment**<br>• **Updated definition of Skills-based VR intervention**<br>• **Change in withdraw policy** | **These changes were made to provide the most up-to-date technology for the participants as well as update the feasibility of the protocol.** |
| **1.1, 1.3, 3.0, 8.2.1, 9.4.9, 8.1, 5.5** | • **Addition of tertiary objective of interest sleep quality (seconds of light/deep sleep) from the Fitbit**<br>• **An update was made to SoA showing increased frequency of Event assessment**<br>• **Needed to remove the AppliedVR tablet for data collection as that will not be available.**<br>• **Additional subgroups added to the subgroup analysis based on additional questions added to study**<br>• **The plan for retention has been updated.**<br>• **Information added for CURES application**<br>• **A revision in patient compensation was made** | **Changes were made to harmonize with the SAP and up-to-date information from the virtual reality vendor. We wanted to add in info related to our CURES application. We also improved the compensation to provide compensation earlier in the study.** |
| **3, 7.3, 8.2, 8.2.1, 8.3.3.1, 10.1.1.2** | • **Minor edits, updated language**<br>• **AE assessment frequency "weekly" instead of "biweekly".** | **Minor edits** |
| **5.5, 7.3, 8.2.1, 11** | • **Loss to follow-up definition revised**<br>• **Added info regarding AppliedVR VR usage data**<br>• **Addition of Appendix: Diversity, Inclusion, & Equity Supplement**<br>• **Updated 5.5 to reflect updated recruitment diversity benchmarks** | **Necessary changes to study operations, addition of supplemental funding initiatives as appendix** |
| **11.2** | • **Recruitment objectives modified to incorporate treating physician letter** | **Protocol reflects procedures and new hybrid recruiting strategy** |

| | | |
|---|---|---|
| | • **Diversity supplement interview methodology updated to reflect that we did not record interviews and we drew from PAB waitlist** | |
| **1, 8.2, 8.3.1** | • **Depression alert procedures updated to include actions to take when telephone assessment cannot be made; approved by DSMB**<br>• **Neck pain and rash added as a known risk, borrowing approved language from STUDY00001363.**<br>• **Table of contents page numbers updated** | **With DSMB guidance, created a contingency plan in the event that severely depressed individuals are lost to follow-up. List of known potential AEs updated to reflect oversight of known neck pain risk and to incorporate the incidence of allergy to face foam used in VR goggles.** |
| **5.5, 8.2, 10.1.1.2** | • **Onboarding procedures modified to take into account new instructional videos**<br>• **Compliance procedures clarified and modified to account for optional SMS text messaging via REDCap**<br>• **Remove 7 day wait after emailing recruitment letter**<br>• **Removal of procedure for mailing study flyer** | **Onboarding procedures now incorporate clear video instructions, and compliance monitoring updated to reflect current streamlined practices. General housekeeping on procedures to align with other VR trials.** |
| **7.2** | • **Modified definition of replaceable participant to exclude individuals who used VR intervention before withdrawing from study** | **Study adequately powered based on withdrawals to date, and analyses should include individuals who found the intervention unsatisfactory and subsequently withdrew** |
| **7.2, 12** | • **Clarified confusing language in replaceable participant definition from previous mod**<br>• **Addition of Appendix B: Digital Divide Supplement** | **Typos in language confused our new definition of a replaceable participant. Second sub-study added to appendix prior to initiation of new methods.** |
| **13** | • **Addition of Appendix C: Cultural Barriers to VR Supplement** | **Third sub-study added to appendix.** |

| | | |
|---|---|---|
| **13.2, 13.3** | • Addition of post-interview quantitative "Interviewer Sentiment Scale" to Appendix C<br>• Disclose payment and requirement to complete both interview and questionnaire | Directly capture effectiveness of AI interviewer |

*Abbreviations*

| | |
|---|---|
| AE | Adverse Event |
| CFR | Code of Federal Regulations |
| cLBP | Chronic Low Back Pain |
| CONSORT | Consolidated Standards of Reporting Trials |
| CBT | Cognitive Behavioral Therapy |
| CRF | Case Report Form |
| DSMB | Data and Safety Monitoring Board |
| eCRF | Electronic Case Report Forms |
| EHR | Electronic Health Record |
| GLP | Good Laboratory Practices |
| HIPAA | Health Insurance Portability and Accountability Act |
| IRB | Institutional Review Board |
| ITQ | Immersive Tendency Questionnaire |
| LBP | Lower Back Pain |
| MOOP | Manual of Operating Procedures |
| MRN | Medical Record Number |
| NCT | National Clinical Trial |
| NIH | National Institutes of Health |
| OHRP | Office for Human Research Protections |
| PI | Principal Investigator |
| QA | Quality Assurance |
| QC | Quality Control |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SOA | Schedule of Activities |
| UP | Unanticipated Problem |
| US | United States |
| VR | Virtual Reality |

# 1 TABLE OF CONTENTS

## STATEMENT OF COMPLIANCE

The trial will be carried out in accordance with International Conference on Harmonisation Good Clinical Practice (ICH GCP) and the following:

- United States (US) Code of Federal Regulations (CFR) applicable to clinical studies (45 CFR Part 46, 21 CFR Part 50, 21 CFR Part 56, 21 CFR Part 312, and/or 21 CFR Part 812)

National Institutes of Health (NIH)-funded investigators and clinical trial site staff who are responsible for the conduct, management, or oversight of NIH-funded clinical trials have completed Human Subjects Protection and ICH GCP Training.

The protocol, informed consent form(s), recruitment materials, and all participant materials will be submitted to the Cedars-Sinai Institutional Review Board (IRB) for review and approval. Approval of both the protocol and the consent form must be obtained before any participant is enrolled. Any amendment to the protocol will require review and approval by the IRB before the changes are implemented to the study. In addition, all changes to the consent form will be IRB-approved; a determination will be made regarding whether a new consent needs to be obtained from participants who provided consent, using a previously approved consent form.

## 1 PROTOCOL SUMMARY

### 1.1 SYNOPSIS

**Title:** Randomized-controlled trial of virtual reality therapy for chronic low back pain to improve patient-reported outcomes and physical activity.

**Study Description:**

This study will test the efficacy of an evidence-based virtual reality (VR) therapy program as a non-pharmacological supplement to management of chronic lower back pain. Study participants will be randomized to receive one of three VR programs: Skills-Based VR, Distraction VR, or Sham VR. In addition to a VR headset, all participants will receive a Fitbit Charge 4 watch. Study devices will be delivered to the participant's home with instructions for use via FedEx; participants will receive remote technical support. They will be followed for 90 days and complete Patient Reported Outcome (PRO) questionnaires to assess functional status, pain levels, and use of pain medications (including opioids). Participants will also be asked to provide consent/authorization to access medical records from their treating facility.

**Objectives:**

Primary Objective:

To assess the efficacy of immersive Skills-Based VR and Distraction VR in improving perceived pain from baseline to Day 30.

Secondary Objective:

To assess the efficacy of immersive Skills-Based VR and Distraction VR in improving perceived pain interference from baseline to Day 60 and Day 90.

To assess the efficacy of Skills-Based and Distraction VR in improving self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep from baseline to Day 90.

To assess the efficacy of Skills-Based and Distraction VR in improving self-reported perceptions of anxiety from baseline to Day 90.

To assess the efficacy of Skills-Based and Distraction VR in improving self-reported pain catastrophizing from baseline to Day 90.

To assess the efficacy of Skills-Based and Distraction VR in reducing use of opioids from baseline to Day 90.

Tertiary Objectives:

To assess the efficacy of Skills-Based and Distraction VR in improving self-reported physical function from baseline to Day 90.

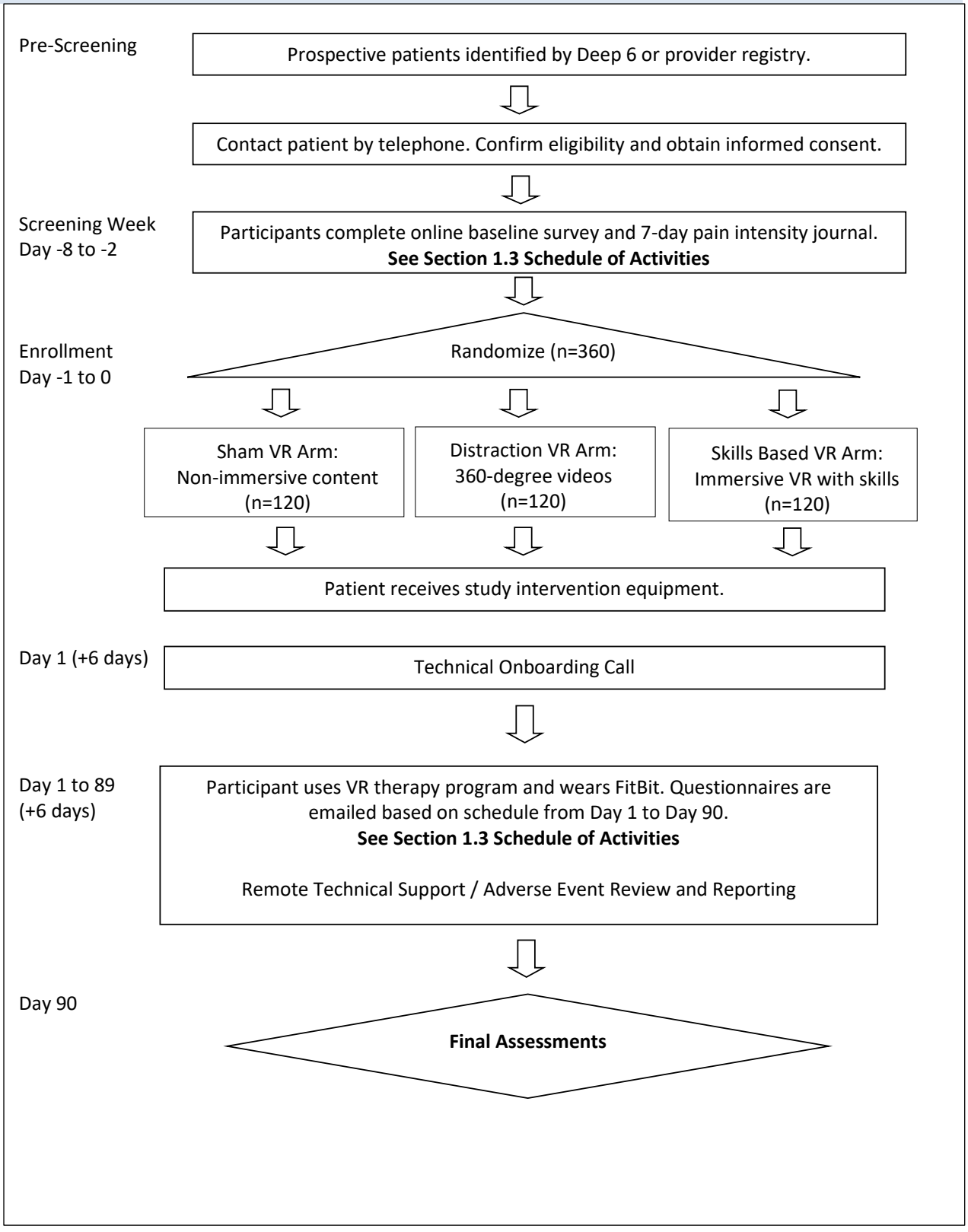
To assess the efficacy of Skills-Based and Distraction VR in improving self-reported depression from baseline to Day 90.

To assess the efficacy of Skills-Based and Distraction VR via patients' global impression of change (PGIC).

To assess the efficacy of Skills-Based and Distraction VR in improving wearable measures of physical activity from baseline to Day 90.

| | |
|---|---|
| | To assess the efficacy of Skills-Based and Distraction VR in improving wearable measures of sleep quality from baseline to Day 90.<br><br>To assess the effect of presence on the efficacy of Skills-Based and Distraction VR in improving measures of pain interference from baseline to Day 30.<br><br>To assess the effect of Immersive Tendencies on the efficacy of Skills-Based and Distraction VR in improving measures of pain interference from baseline to Day 30.<br><br>To assess the effect of dose on the efficacy of Skills-Based and Distraction VR in improving measures of pain interference from baseline to Day 30. |
| **Endpoints:** | Primary Endpoint: Change in PROMIS Pain Interference<br>Secondary Endpoints: Change in PROMIS Sleep Disturbance, PROMIS Anxiety, Pain Catastrophizing, MME usage. |
| **Study Population:** | Individuals ages 13 or older with an ongoing low back-pain problem that has persisted at least 3 months and has resulted in pain on at least half the days in the past 6 months. |
| **Phase:** | Phase 2 |
| **Description of Sites/Facilities Enrolling Participants:** | Three outpatient clinic sites, including the Cedars-Sinai Orthopaedic Clinic, Cedars-Sinai Medical Network Chronic Pain Program, and Attune Health, a musculoskeletal clinic affiliated with Cedars-Sinai. In addition, participants may enroll remotely. |
| **Description of Study Intervention:** | All participants will receive a virtual reality headset with one of three software programs:<br>- Immersive Skills-Based VR therapy<br>- Immersive Distraction VR therapy<br>Non-immersive Sham VR Therapy with 2-D videos |
| **Study Duration:** | 48 months |
| **Participant Duration:** | 90 days |

## 1.2 SCHEMA


Pre-Screening
Prospective patients identified by Deep 6 or provider registry.
Contact patient by telephone. Confirm eligibility and obtain informed consent.
Screening Week
Day -8 to -2
Participants complete online baseline survey and 7-day pain intensity journal.
See Section 1.3 Schedule of Activities
Enrollment
Day -1 to 0
Randomize (n=360)
Sham VR Arm:
Non-immersive content
(n=120)
Distraction VR Arm:
360-degree videos
(n=120)
Skills Based VR Arm:
Immersive VR with skills
(n=120)
Patient receives study intervention equipment.
Day 1 (+6 days)
Technical Onboarding Call
Day 1 to 89
(+6 days)
Participant uses VR therapy program and wears FitBit. Questionnaires are emailed based on schedule from Day 1 to Day 90.
See Section 1.3 Schedule of Activities
Remote Technical Support / Adverse Event Review and Reporting
Day 90
Final Assessments

## 1.3 SCHEDULE OF ACTIVITIES (SOA)

| **Procedures** | Pre-screening | Screening Week Day -8 to -2 | Enrollment Day -1 to 0 | Day 1 (+ 6 days) | Day 7 (+ 6 days) | Day 15 (+ 6 days) | Day 21 (+ 6 days) | Day 30 (+6 days) | Day 45 (+ 6 days) | Day 60 (+ 6 days) | Day 75 (+ 6 days) | Day 90 (+ 6 days) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Prospective patient identified (DEEP 6 or provider registry) | X | | | | | | | | | | | |
| Confirm eligibility and Informed Consent via telephone and REDCap | X | | | | | | | | | | | |
| Online Pain Intensity Journal (7 days) | | X | | | | | | | | | | X |
| Screening Week Online Baseline Survey (see below) | | X | | | | | | | | | | |
| Randomization | | | X | | | | | | | | | |
| Participant receives study intervention kit | | | X | | | | | | | | | |
| Technical onboarding call | | | | X | | | | | | | | |
| Intervention: Participant uses VR therapy program and wears Fitbit Charge 4 watch | | | | X | ---------- | | | | | | | X |
| Online survey: NIH HEAL Minimum Dataset | | X | | | | | | | | | | X |
| Immersive Tendency Questionnaire (ITQ)[1] and Motion Sickness Propensity Survey[2] | | X | | | | | | | | | | |
| Treatment expectation question | | X | | | | | | | | | | |
| Simulator Sickness Questionnaire (SSQ)[3] and Presence Survey | | | | X | | | | | | | | |
| Online survey: Primary outcome: PROMIS® Pain Interference | | X | | | X | X | X | X | X | X | X | X |
| Discontinuation of Treatment Questionnaire | | | | | X | X | X | X | X | X | X | X |
| PROMIS® Physical Function[4], Anxiety[4], Depression[4], Sleep disturbance[4]; Pain intensity/interference with Enjoyment of life/interference with General activity (PEG)[5], PCS-6[6] | | X | | | | X | | X | X | X | X | X |
| EMR data: Charlson Comorbidity Index[7], CURES and EMR: Prescription Data | | X | | | | | | | | | | X |
| Perceived study arm question, treatments in the last 90 days question | | | | | | | | | | | | X |
| Online survey: PHQ-2, GAD-2, PGIC, TAPS [1/2] | | X | | | | | | | | | | X |
| Remote technical support | | | | X | ---------- | | | | | | | X |
| End of study procedure question | | | | | | | | | | | | x |

| **Procedures** | Pre-screening | Screening Week Day -8 to -2 | Enrollment Day -1 to 0 | Day 1 (+ 6 days) | Day 7 (+ 6 days) | Day 15 (+ 6 days) | Day 21 (+ 6 days) | Day 30 (+6 days) | Day 45 (+ 6 days) | Day 60 (+ 6 days) | Day 75 (+ 6 days) | Day 90 (+ 6 days) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Event Assessment: AE, SAE and UP review and reporting | | | | X | X | X | X | X | X | X | X | X |

ITQ: Immersive Tendency Questionnaire
PEG: Pain, Enjoyment, General Activity
PCS-SF: Pain Catastrophizing Scale-Short Form 6
PHQ-2: Patient Health Questionnaire-2
GAD-2: Generalized Anxiety Disorder 2-item
PGIC: Patient Global Impression of Change
TAPS: Tobacco, Alcohol, Prescription Medication, and Other Substance Use
AE: Adverse event
SAE: Serious adverse event
UP: Unanticipated Problem

# 2 INTRODUCTION

## 2.1 STUDY RATIONALE

Chronic low back pain (cLBP) is a prevalent and costly condition that markedly impairs physical, emotional, and social function. The 2015 Global Burden of Disease Study estimated that the prevalence of cLBP increased by more than 17% between 2005 and 2015.[8] In 2010, low back pain ranked third in disability-adjusted life years in North America, and the prevalence is expected to increase further due to the aging population and rise in obesity rates.[8,9] The National Health Interview Survey found that more than a quarter of all workers reported low back pain in the prior 3 months.[10] More than half reported the pain caused disability related to self-care, work, or social activities.[11]

Diminished work productivity attributable to cLBP is conservatively estimated at more than $28 billion annually.[12] Escalating treatment expenses include an increase of 423% in the cost of opioid prescriptions for patients with spinal disorders from 1997-2004, increases of 307% in the volume of lumbar MRI and 231% in the number of spinal injections reimbursed by Medicare from 1994-2001, and a 220% increase in lumbar spinal fusion surgeries between 1990 and 2001.[13] The Back Pain Survey administered as part of the National Health and Nutrition Examination Survey (NHANES) found that opioids were the most commonly prescribed pain medication taken by patients with cLBP (18.8%), followed by antidepressants (17.8%).[14] More than three quarters of people using prescribed opioids took them long term. Yet, patients with cLBP often discover that opioids fall short in delivering meaningful pain reduction or improving health-related quality of life (HRQOL).[14] Additionally, opioids are associated with a host of adverse effects, including but not limited to fall risk, constipation, sedation, physical dependency, opioid use disorder, and drug related mortality.[15] Hence, there is a critical gap in pain management in cLBP; it is vital to address this evidence gap in a way that maximizes benefits for patients while minimizing harms from medical therapy.

## 2.2 BACKGROUND

The dynamic nature of clinical medicine, coupled with limited time to spend with individual patients, pose challenges to offering holistic care for patients with pain. Treatment of pain is often focused on pharmacological management, which can yield inconsistent and sub-optimal pain control. However, extensive data reveal that adjunctive non-pharmacological techniques, such as cognitive behavioral therapy and relaxation techniques, can modify cognitions and behaviors that influence the perception of pain. Therapeutic virtual reality (VR) has emerged as a promising and evidence-based treatment modality for both acute and chronic pain.[16-25] Users of VR wear a pair of goggles with a close-proximity stereoscopic screen that creates a sensation of being transported into lifelike, three-dimensional worlds that create a sense of “presence.” To date, VR has been used in numerous clinical settings to treat anxiety disorders, manage depression, support physical rehabilitation, and manage a wide range of acute and chronic pain syndromes.[21,26-30] For example, VR coupled with medication is effective in decreasing pain during bandage changes for severe burns as an alternative to opioids.[26,31] Similarly, VR reduces pain and provides positive distraction during routine procedures, such as intravenous line placements[29] and dental procedures.[27] Other studies reveal that VR helps manage chronic pain conditions such as complex regional pain syndrome,[32] lower back pain,[33,34] and chronic neck pain.[35] Multiple studies also demonstrate that VR offers clinical benefits in musculoskeletal pain, including cLBP. Recently, a randomized trial by Pozeg and colleagues revealed that VR is effective for managing

neuropathic pain from spinal injury.[36] Similarly, Jones and colleagues employed a VR-based distraction therapy to manage patients with a wide variety of chronic musculoskeletal pain syndromes, including cLBP, and documented a 33% reduction in pain before vs. after VR therapy.[33] In a separate study by Gromala and colleagues using a specialized VR game designed to empower patients with pain, VR reduced chronic musculoskeletal pain—again including cLBP patients—by 37% and outperformed a standard non-VR distraction intervention.[34] Thomas and colleagues evaluated a VR "dodgeball intervention" for cLBP and demonstrated that activating graded physical activity in a virtual game simulation improved lumbar spine flexion while reducing expectations of fear and harm vs. a non-VR control condition.[37] A more recent randomized comparative effectiveness study conducted by our group at Cedars-Sinai also demonstrated the analgesic benefits of VR for inpatient management of acute and chronic pain, including patients with cLBP, demonstrated a 24% reduction in pain that outperformed a 2D generic relaxation video.[38] Sixty-five percent of VR patients achieved a clinically significant pain response vs. 40% of controls (p=0.01; NNT=4).[38]

Beyond individual studies, there are now several meta-analyses evaluating the analgesia benefits of VR in both acute and chronic pain syndromes, and these studies reveal consistently significant benefits.[17-22] In one systematic review of randomized controlled inpatient VR trials published by the Cedars-Sinai group, we identified 11 studies comparing VR vs. a control condition in diverse populations.[21] VR was effective in most of these studies and well tolerated in all reports. However, the studies we identified were of mixed methodological quality. We concluded that the VR literature will benefit from larger, higher quality studies with a longer follow-up period, enhanced focus on long-term safety, and more robust efforts to identify patient-level predictors of efficacy, particularly in cLBP. In the meantime, the existing literature supports VR as a broadly effective therapy for both chronic and acute pain management across pain conditions, including cLBP.

The mechanisms of action (MOAs) of VR analgesia have been extensively evaluated over the past twenty years. Research indicates that VR reduces pain through four MOAs: First, by stimulating the visual cortex while engaging other senses, VR acts as a distraction to limit the user's processing of nociceptive stimuli.[39] The result is a form of "inattentional blindness" where the prefrontal cortex redirects attentional bandwidth to the virtual environment, leaving diminished ability to attend to pain signals outside the "spotlight of attention."[40] By overwhelming the visual, auditory, and proprioception senses, VR is thought to create an immersive distraction that restricts the brain from processing pain. Second, through gate control theory,[41] VR is thought to activate descending inhibitory pathways and inhibit spinal transmission of peripheral afferent pain signals.[42-44] Third, VR creates an illusion of time acceleration, effectively shortening the perception of pain episodes through its effects on prefrontal time perception.[45-47] For example, controlled trials reveal that VR reduces the perceived length of labor and delivery during childbirth, episiotomy repair, endoscopic procedures, and chemotherapy infusions by an average of 30-50%.[45-47] These effects have been demonstrated both clinically and experimentally. Hoffman and colleagues revealed that VR affects pain processing in the sensory and insular cortex, indicating it can reduce both the intensity of pain and the emotional response to pain.[48,49] Moreover, the investigators found that VR has the same fMRI effects as hydromorphone, and was equally effective at blocking acute pain as the powerful opioid.[48] Clinical trials also demonstrate reductions in sensory, cognitive, and affective components of pain, suggesting that the fMRI changes shown experimentally appear to translate into improved patient-centered outcomes across dimensions of pain. Investigators like Hoffman,[24,31,50,51] Rizzo,[52] Rothbaum,[53-55] and Bordnick,[56-58] among many others, are studying the neurobiological mechanisms of VR across a range of conditions.[21,30] Fourth, VR offers an immersive platform through which patients can develop and practice specific skills. This mechanism may underlie VR's established benefits in cognitive rehabilitation such as for phobia,[59-65] anxiety[66-68] and

depression[69,70] management, where skills may be learned in VR and durably extended outside of the virtual environments. In short, VR has multiple analgesic MOAs that translate into clinically important patient benefits. Importantly, the ubiquity of mobile high-performance computing has now reduced both the size and cost of VR devices such that mobile VR units are feasible for everyday use. Our team at CSMC has employed therapeutic VR in over 3,000 patients throughout inpatient and outpatient settings and found that it is feasible and practical for patients to use the equipment. Our latest data reveal that VR analgesia is effective across ages, with a greater benefit for those over the age of 65, possibly due to differences in expectations surrounding digital technologies(in press).We have also documented that VR is incrementally more effective for those with the highest pain scores (>7 out of 10 on numeric rating scale), demonstrating benefits in those with the most severe symptoms (in press). In short, the field of therapeutic VR has gained traction in the past several years on the strength of less expensive, more scalable, and higher quality VR equipment, meaningful advances in the clinical and translational science of therapeutic VR, dissemination of methodological guidance unique to VR trials, and evidence of patient acceptability of using the equipment. However, there has been no research to date evaluating longer-term outpatient use and safety of VR for cLBP, and there are important unanswered questions about the durability of VR efficacy that warrant further research.

In previous systematic reviews of VR RCTs, including our own analysis of the literature, many of the identified studies were limited by study design issues, such as improper control groups and incomplete PRO assessment.[21] PROs, such as functional status, HRQOL, and satisfaction with care, are necessary to determine whether VR is an effective, sustainable intervention for cLBP management.[71] Additionally, although evidence reveals that VR is effective as a short-term analgesic, it is less clear how or whether continued use of VR reduces the frequency, intensity, and experience of ongoing pain, predicts long-term change in other PROs, or reduces opioid use. It also remains unknown whether different forms of VR have varying efficacy in cLBP. Important unanswered clinical questions include: (1) do skills-based VR programs that leverage the immersive qualities of VR to teach exportable skills using cognitive behavioral therapy (CBT), guided meditation, and biofeedback-based breathing exercises outperform conventional VR that employs distraction? (2) Are there patient-level characteristics that predict a clinical response to VR? (3) Are there usage patterns or engagement characteristics that predict enhanced response to VR? (4) Can VR reduce pain while also reducing long-term opioid requirements? The current study will seek to answer these questions in cLBP. The long-term safety and tolerability of therapeutic VR is also largely untested. In short-term use, a minority of patients may develop transient dizziness, fatigue, or nausea—a syndrome called cybersickness—resulting from sensory mismatch between the visual and vestibular systems.[72-75] The prevalence of cybersickness has fallen with dramatic improvements in hardware and software. Other technical advances have reduced eye strain, minimized physical discomfort of wearing the headset, and reduced unnecessary visual motion.[74] Nonetheless, there are limited data monitoring the incidence of cybersickness over longer term usage, such as weeks or months. The current study will perform longer-term safety and tolerability assessments of therapeutic VR in cLBP.

## 2.3 RISK/BENEFIT ASSESSMENT

### 2.3.1 KNOWN POTENTIAL RISKS

This study poses minimal risk to participants. Immediate risks may include minor psychological distress from questionnaire items asking about health and employment status. There is also a small[76] short-term risk of transient risk of VR-related “cybersickness.” Cybersickness is transient vertigo, nausea, or

headache. It results from sensory mismatch between the visual and vestibular systems[72-75] and usually resolves within minutes of removing the VR headset. The prevalence of cybersickness has fallen with improvements in hardware and software. Technical advances have reduced eye strain, minimized physical discomfort of wearing a VR headset, and reduced unnecessary visual motion.[74] In the therapeutic setting, we also choose slow-moving scenes rather than highly kinetic visuals. As a result, cybersickness has become less prevalent and significant for people using VR.[74] In this study, each VR session lasts less than 20 minutes. Participants will be instructed to complete just one session during each use to reduce the risk of cybersickness.

There are no anticipated long-term physical risks from participating in this study. There is a small risk of breach of confidentiality associated with the electronic collection and transmission of protected health information. This risk will be minimized by following proper procedures for assuring data integrity and confidentiality.

The following is our list of study participant confidentiality safeguards:

- ***Electronic files –*** data identifying participants will be stored in status tracking logs within password-protected excel files on Cedars-Sinai encrypted shared drives e.
- ***Forms –*** survey forms and other pages containing personal identifying information will be saved within the REDCap system.
- ***Data listings -*** participant name, name code, hospital chart, record number, Social Security Number, or other unique identifiers will not be included in any published data listing.
- ***Data distribution*** - data listings containing PHI such as name, MRN, or other identifiers easily associated with a specific participant will not be distributed.
- ***Data disposal*** - computer listings that contain participant-identifying information will be disposed of in accordance with institutional policies and procedures, after study completion.
- ***Access*** - participant records will not be accessible to persons/institutions outside those listed on the HIPAA form signed by -the participant.
- ***Storage -*** study forms and related documents retained during and after study completion will be stored within a secure Box folder accessible only by approved study staff. Some documents, such as survey exports, may be locally stored on a Cedars-Sinai encrypted shared drive or computer desktop during data analysis.
- ***Passwords*** – Multi-factor password authentication will be required to access documents stored on the cloud and on local desktops at Cedars-Sinai.
- ***User Training -*** study staff with access to clinical computer systems are trained and certified to maintain confidentiality prior to authorization by the Cedars-Sinai IRB.
- ***System Testing –*** new computer systems used by clinical staff are processed by the Cedars-Sinai IT (EIS) to ensure the password-activated systems perform as intended.
- ***Certificate of Confidentiality*** – NIH funded research that involves human subjects and collects information which may identify a person is automatically

protected by a Certificate of Confidentiality to prevent forced disclosures (e.g., subpoenas).

- ***Privacy Breach Reporting*** – In the rare instance where an accidental release of identifiable PHI may occur, it will be reported to the Cedars-Sinai IRB and the Cedars-Sinai Privacy Office for adjudication and corrective actions.

### 2.3.2 KNOWN POTENTIAL BENEFITS

Potential immediate benefits include reduction of chronic low back pain and general improvement in psychological health. Potential long-term benefits include improved functionality, reduced opioid use, and improvements in overall physical and psychological health. This research will contribute to societal knowledge about the safety and efficacy of therapeutic VR.

### 2.3.3 ASSESSMENT OF POTENTIAL RISKS AND BENEFITS

Virtual reality hardware and software has advanced significantly in the last decade. Modern studies utilizing VR have found general satisfaction with devices among participants. The incidence and severity of VR-related side effects are low, and symptoms generally subside within minutes of taking off the headset.

Participation in the study may alleviate pain among individuals who have previously relied on opioids, potentially enabling them to reduce opioid use and related side effects. Sustained pain relief also may allow individuals to return to work faster, improve physical mobility, and enhance biopsychosocial health. As a result, the short- and long-term anticipated benefits of participation outweigh the minimal short-term risks.

# 3 OBJECTIVES AND ENDPOINTS

| OBJECTIVES | ENDPOINTS | JUSTIFICATION FOR ENDPOINTS |
|---|---|---|
| *Primary* | | |
| **To assess the efficacy of immersive Skills-Based VR and Distraction VR in improving perceived pain from baseline to Day 30.** The trial will be considered a success if there is statistical evidence of improvement in either VR group compared to sham VR. | The change from study baseline to Day 30 in pain interference as measured by the 8-item PROMIS Pain Interference scale is the primary endpoint. This scale measures the consequences of pain on relevant aspects of life, including the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities.<br>The study will be considered a success if there is a statistically significant difference of 5 points in the Pain Interference score between participants in either the Skills-Based or Distraction VR arm compared with the Sham VR arm. | The NIH Pain Consortium Research Task Force (RTF) draft standards for research on cLBP recommend a uniform minimal data set that includes self-reported measures of pain interference.[77]<br>The PROMIS PI scale is a validated instrument with excellent content validity, construct validity, and reliability in patients with chronic pain. Past work indicates that changes of 3.5 to 5.5 points in PROMIS-PI scores of people with LBP can be considered meaningful.[78,79] |
| Secondary | | |
| **To assess the efficacy of immersive Skills-Based VR and Distraction VR in improving perceived pain interference from baseline to Day 60 and Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to day 60 and day 90 in pain interference as measured by the 8-item PROMIS PI scale is a secondary endpoint. This scale measures the consequences of pain on relevant aspects of life, including the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities.<br>We will test for a statistically significant difference of 5 points in the PROMIS PI score from baseline, and compare differences between either VR group and control (sham) VR. | The NIH Pain Consortium RTF draft standards for research on cLBP recommend a uniform minimal data set that includes self-reported measures of pain interference.[77]<br>The PROMIS scale is a validated instrument with excellent content validity, construct validity, and reliability in patients with chronic pain.[79]<br>Past work indicates that changes of 3.5 to 5.5 points in PROMIS-PI scores of people with LBP can be considered meaningful.[78] |

| OBJECTIVES | ENDPOINTS | JUSTIFICATION FOR ENDPOINTS |
|---|---|---|
| **To assess the efficacy of Skills-Based and Distraction VR in improving self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to day 90 in sleep disturbance as measured by the 6-item PROMIS Sleep Disturbance scale is a secondary endpoint. This scale assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep. This includes perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of - and satisfaction with - sleep.<br>We will test for a statistically significant difference of 5 points in the PROMIS Sleep Disturbance score from baseline, and compare differences between either VR group and control (sham) VR. | The NIH Pain Consortium RTF draft standards for research on cLBP recommend a uniform minimal data set that includes self-reported measures of sleep disturbance.[77]<br>The PROMIS sleep disturbance item banks display strong measurement properties for assessing general aspects of sleep and sleep related impairments in cLBP patients.[80] |
| **To assess the efficacy of Skills-Based and Distraction VR in improving self-reported perceptions of anxiety from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in anxiety as measured by the 4-item PROMIS Anxiety scale is a secondary endpoint. This scale assesses self-reported perceptions of fear, anxious misery (worry, dread), hyperarousal, and somatic symptoms related to arousal.<br>We will test for a statistically significant difference of 5 points in the PROMIS anxiety score from baseline, and compare differences between either VR group and control (sham) VR. | While not included in the recommended minimal data set, the NIH Pain Consortium RTF considers anxiety a conceptually relevant construct associated with cLBP.[77]<br>The PROMIS Anxiety item banks have demonstrated strong measurement properties for assessing general aspects of anxiety in cLBP patients.[81] |
| **To assess the efficacy of Skills-Based and Distraction VR in improving self-reported pain catastrophizing from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in pain catastrophizing as measured by PCS SF-6 is a secondary endpoint.<br>We will test for a difference in rates of high catastrophizing as defined by a score of ≥7 on the PCS-SF6, and compare these differences between either VR group and control (sham) VR. | The NIH Pain Consortium RTF draft standards for research on cLBP recommend a uniform minimal data set that includes self-report measures of pain catastrophizing.[77] Further, changes in catastrophizing are associated with improvement in multidisciplinary pain treatment.[82] |

| OBJECTIVES | ENDPOINTS | JUSTIFICATION FOR ENDPOINTS |
|---|---|---|
| **To assess the efficacy of Skills-Based and Distraction VR in reducing use of opioids from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in weekly average opioid dosage calculated as a 7-day average of daily maximum milligrams morphine equivalent (MME) is a secondary endpoint. We will test for a greater than .5 SD in change from baseline, and compare differences between either VR group and control (sham) VR. | Despite some evidence from randomized controlled trials (RCTs) on the efficacy of opioids in the short-term treatment of low back pain, little evidence is available on long term efficacy and safety. Lowering MME as prescribed while lowering pain is an important goal of cLBP treatments.[83] |
| Tertiary/Exploratory | | |
| **To assess the efficacy of Skills-Based and Distraction VR in improving self-reported physical function from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in physical function as measured by the 6-item PROMIS Physical Function scale is an exploratory endpoint. This scale measures self-reported functioning of one's upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck, back), as well as instrumental activities of daily living.<br>We will test for a statistically significant difference of 5 points in the PROMIS Physical Function score from baseline, and compare differences between either VR group and control (sham) VR. | The NIH Pain Consortium RTF draft standards for research on cLBP recommend a uniform minimal data set that includes self-reported measures of physical function.[77]<br>The PROMIS physical function item banks display excellent measurement properties for assessing general aspects of physical function in cLBP patients.[80] |
| **To assess the efficacy of Skills-Based and Distraction VR in improving self-reported depression from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in physical function as measured by the 4-item PROMIS Depression scale is an exploratory endpoint. This scale measures self-reported negative mood (sadness, guilt), views of self (self-criticism, worthlessness), and social cognition (loneliness, interpersonal alienation), as well as decreased positive affect and engagement (loss of interest, meaning, and purpose).<br>We will test for a statistically significant difference of 5 points in | The NIH Pain Consortium RTF draft standards for research on cLBP recommend a uniform minimal data set that includes self-report measures of depression.[77]<br>The PROMIS depression item banks display excellent measurement properties for assessing general aspects of depression in cLBP patients.[79,80,84] |

| OBJECTIVES | ENDPOINTS | JUSTIFICATION FOR ENDPOINTS |
|---|---|---|
| | the PROMIS Depression score from baseline, and compare differences between either VR group and control (sham) VR. | |
| **To assess the efficacy of Skills-Based and Distraction VR via patients' global impression of change (PGIC).** The objective will be considered achieved if there is statistical evidence of higher rates of self-reported improvement in either VR group compared to control (sham) VR. | The overall effect of treatment from study baseline to Day 90 as measured by the PGIC is an exploratory endpoint. This scale measures self-reported belief regarding efficacy of treatment. A favorable response of 5-7 on the PGIC indicates significant improvement occurred over the course of the study. We will test for a statistically significant difference in PGIC responses between either VR group and control (sham) VR. | This exploratory endpoint will act as a check on internal validity for Primary and Secondary endpoints, and a general assessment of satisfaction with treatment.[79,80,84] The PGIC has established validity in measuring self-reported global change in cLBP trials.[85,86] |
| **To assess the efficacy of Skills-Based and Distraction VR in improving wearable measures of physical activity from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of measured improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in weekly total steps as measured by FitBit is an exploratory endpoint. Change in steps will also be examined for interaction effects on the primary outcome. | There is evidence that physical activity can improve mood, sleep, and general health, yet persons with chronic pain may limit their activity because of their pain.[87] Evidence suggests that activity monitors encourage increased function and correlate with improvements in mood among patients with chronic pain.[88,89] |
| **To assess the efficacy of Skills-Based and Distraction VR in improving wearable measures of sleep from baseline to Day 90.** The objective will be considered achieved if there is statistical evidence of measured improvement in either VR group compared to control (sham) VR. | The change from study baseline to Day 90 in sleep quantity (total minutes asleep as well as "sleep efficiency", see Section 4.6.3 for more info) as measured by FitBit is an exploratory endpoint. Changes in the seconds in sleep will also be examined for interaction effects on the primary outcome. | There is evidence that physical activity can improve mood, sleep, and general health, yet persons with chronic pain may limit their activity because of their pain.[87] Previous versions of the skills-based program used in this study showed improvements in the sleep subscale of a pain interference measure.[90] |
| Exploratory Subanalysis | | |
| **To assess subgroup analyses for the change from baseline in PROMIS-PI.** An MMRM | The impact of the following variables mentioned in section | See Section 9.4.7 for details. |

| OBJECTIVES | ENDPOINTS | JUSTIFICATION FOR ENDPOINTS |
|---|---|---|
| model will be used to test for treatment by subgroup interactions. | 9.4.7 will be assessed to changes in the primary outcome. | |

# 4 STUDY DESIGN

## 4.1 OVERALL DESIGN

This research will be conducted as randomized, double-blind, sham-controlled, 3-arm Phase 2 study. Two immersive VR arms use a skills-based VR therapy program, EaseVRx, and a distraction-based VR therapy program, EaseVRx-Distraction, while the sham VR arm uses a VR headset to deliver two-dimensional (non-immersive) content. The primary hypothesis is that participants randomized to either skills-based VR therapy or distraction VR therapy will report meaningful improvements in PROs, improved biometric outcomes, and reduced opioid use compared to participants receiving a non-immersive sham VR control intervention. It is currently planned as a single-site study. However, the self-administered nature of the intervention and remote data collection would accommodate additional sites from the BACPAC consortium if deemed necessary in time.

Review of the VR clinical research literature indicates that the principal source of bias in this research has been lack of a control arm that effectively blinds the participant to the treatment allocation. This is the first study, to our knowledge, in which participants randomized to the control arm will use a VR headset that enables participants to experience the novelty of the hardware but limits their viewing to content that is neither immersive nor interactive. The experience of using EaseVRx-Sham is similar to watching a large-screen TV. Rather than viewing 360-degree, 3D, interactive content, participants in the sham arm can choose among 2D nature experiences accompanied by music selected to be neither relaxing nor distracting. EaseVRx- Sham has the same number and duration of experiences as EaseVRx, and the functionality of the user interface to access the experiences is the same. Modifications were made to the appearance of the user interface in order to remove aspects that were added for therapeutic benefit. Thus, the use of this specially designed sham program enables us to isolate the effect of VR immersion.[35] We also expect that receiving a real VR headset will reduce attrition among participants randomized to the control arm.

## 4.2 SCIENTIFIC RATIONALE FOR STUDY DESIGN

The three-arm study design will allow us to determine whether a skills-based VR program that teaches exportable skills using CBT, guided meditation, and biofeedback-based breathing exercises outperforms conventional, distraction-based VR therapy in reducing cLBP. The 90-day study duration was chosen to assess whether VR therapy's efficacy extends beyond short-term analgesia and reduces the frequency and intensity of ongoing LBP through continuous use over a longer period. It also will enable us to evaluate longer-term change in other PROs and whether continued use of VR reduces opioid use. The

active control group (Sham VR) was designed to minimize bias related to the novelty of the hardware, addressing one of the deficits of prior studies.

### 4.3 JUSTIFICATION FOR DOSE

Prior research has shown that daily use of VR for 10 to 15 minutes is effective in managing pain, plus use as needed for breakthrough flares of pain. The skills-based VR therapy program, EaseVRx, is a standardized 56-day program consisting of scheduled daily virtual experiences. Each VR treatment experience lasts between 2-16 minutes, with an average duration of 6 minutes. To minimize the risk of cybersickness, participants are instructed to complete one VR treatment experience at a time. Given the low-risk nature of VR, they also may repeat experiences such as relaxation, breathing exercises, and games at other times during the day, in response to their pain. After the 56-day program is completed, participants in this arm are instructed to continue daily use, choosing from among the treatment experiences. Participants in the distraction-based VR (EaseVRx-Distraction) and the sham VR also will be instructed to use the VR headset at least once daily, choosing from among the experiences available in their devices.

### 4.4 END OF STUDY DEFINITION

A participant is considered to have completed the study if he or she has completed the online survey questionnaire that is emailed on Day 90, as shown in the Schedule of Activities (SoA), Section 1.3. The end of the study is defined as 96 days after the last participant is enrolled or when that participant completes the Day 90 survey questionnaire, whichever is sooner.

## 5 STUDY POPULATION

### 5.1 INCLUSION CRITERIA

In order to be eligible to participate in this study, an individual must meet all of the following criteria:

1. Able to provide consent to participate in research
2. Stated willingness to comply with all study procedures and availability for the duration of the study
3. Male or female, age 13 or older
4. An ongoing low back-pain problem that has persisted at least 3 months and has resulted in pain on at least half the days in the past 6 months.
5. Ability to comprehend spoken and written English
6. Owns a compatible android or iOS smartphone, or personal laptop or desktop computer (excluding tablets) to complete surveys and has access to email.

## 5.2 EXCLUSION CRITERIA

An individual who meets any of the following criteria will be excluded from participation in this study:

1. Presents with a condition that interferes with VR usage, including history of seizure, facial injury precluding safe placement of headset, significant visual or hearing impairment that impacts ability to see the VR images or follow audio instructions
2. Are being recommended for long-term hospitalization that would require more than three-week stay in the hospital
3. Received surgical procedure within the previous 8 weeks
4. Surgery is planned within the next 3 months
5. Is currently using a spinal cord stimulator
6. Has LBP attributable to a recognizable, specific pathology, including spinal infection, cancer, fracture, or inflammatory spondylopathies, consistent with the NIH Task force on Research Standards for cLBP
7. Previously participated in a VR clinical trial

Women who are currently pregnant or planning to become pregnant will not be excluded from the study

## 5.3 LIFESTYLE CONSIDERATIONS

NA

## 5.4 SCREEN FAILURES

All aspects of this study are conducted remotely, including the participants' use of a VR therapy headset and FitBit watch at home and collection of all PROs via survey questionnaires administered electronically. In addition, the VR headset has independent value. To ascertain their willingness and ability to respond to survey questionnaires delivered by email, potential participants are required to respond to a daily electronic, one-item, "Pain Diary" question and complete the baseline survey questionnaires over the course of a "Screening Week," following informed consent but prior to randomization. A screen failure is defined as a participant who completes fewer than 86 items on the large survey on the 4$^{th}$ day of the screener week as well less than 5 daily "pain diary questions. Participants are required to complete this by the 7$^{th}$ day of the screener week or within 48hrs of being sent a reminder email to complete the survey. Participants will be sent a manual reminder email if they complete less than 80% of the large survey on the 4$^{th}$ day of screener week as that patient population would not get reminder otherwise from the REDCap system. Age, race, and Hispanic ethnicity are some of the information that will be collected for all participants even if they fail to complete the Screening Week questionnaires. Individuals who inform investigators during the Screening Week that they are no longer able or willing to participate in the study for any reason or who are discovered to meet any of the exclusion criteria will be recorded. Individuals who satisfactorily complete the Pain Diary and baseline survey questionnaires will be enrolled and randomized.

## 5.5 STRATEGIES FOR RECRUITMENT AND RETENTION

The target sample size is 360. We expect a representative sample of the chronic low back pain patient population seen at Cedars-Sinai. Approximately 58% are female and 42% male. For the first half of the

study, we anticipated patients who identify as white to make up about 76% of the study population, with 12% black or African American, 5% Asian, and 6% who identify as multiracial. Approximately 11% were expected to be Hispanic or Latino. For the second half of the project, per supplementary funding to improve study diversity to better reflect the greater Los Angeles population (see Section 11), we aim to recruit a second-half cohort to be 35% non-Hispanic white, 17% non-Hispanic black, and 40% Hispanic. We expect to randomize 10 participants per month over 36 months.

Prospective participants will be referred by physicians at the designated clinical sites or identified using the Deep 6 AI® cohort building software, which uses natural language processing to search the Cedars-Sinai EHR for patients who meet the study's inclusion and exclusion criteria. We anticipate screening 800 potential participants in order to randomize 360. Recruitment of women is not expected to be a problem, because patients seeking care for cLBP are disproportionately female. Women who are pregnant or planning to become pregnant are not excluded from participating. Race and ethnicity of enrolled participants will be tracked monthly. If the proportion of minority participants appears to be lagging, the number contacted will be increased.

Eligible potential participants will be sent an IRB- approved recruitment letter explaining the study via email. For individuals identified via cohort builders who have seen a physician partnered with the study staff, a personalized letter bearing their treating physician's voice and signature will be used; all other cohort builder identified individuals will be sent a letter from the primary investigator. An informational study brochure will also be attached. The recipient may opt-out of further contact by replying to the email.

Research coordinators will monitor participant compliance with the survey questionnaires using an automated dashboard. Participants will be encouraged to remain in the study and complete the questionnaires using a combination of email reminders, telephone calls, and/or SMS text messages. from research coordinators. *Study coordinators will contact the participant following non-completion of a survey if >1 week has elapsed since the initial survey send date.* Technical support will be available throughout the study.

Participants will be eligible for up to $225 in Amazon gift card codes throughout the study after randomization. After completing 80% of the first month of surveys, participants will be sent a $25 Amazon gift card code. After completing 75% of the second month of surveys, participants will be sent another $25 Amazon gift card code. Once the participant completes 80% of the surveys in the 3rd month and returns the equipment, they will be sent a $175 Amazon gift card code along with recommendation for VR equipment that could be purchased after the end of the study. The $175 Amazon code compensation will not be released until the audiovisual headset and Fitbit device has been returned.

Participants will be emailed recommendations for hardware and software that can be used after the study once the equipment has been returned in the email with the Amazon codes.

A detailed description of the recruitment and retention process is in section 5.1 and 5.2 of the MOOP.

# 6 STUDY INTERVENTION

## 6.1 STUDY INTERVENTION(S) ADMINISTRATION

### 6.1.1 STUDY INTERVENTION DESCRIPTION

The two active interventions evaluated in this study are 360-degree, 3-D therapeutic visualization programs (EaseVRx and EaseVRx-Distration) delivered via a commercially available VR headset. The control intervention consists of 2-D visualizations delivered by the same VR headset. The three software programs were developed by AppliedVR. The EaseVRx programs are not currently regulated by the FDA. In September 2019, AppliedVR was awarded two grants by NIDA to conduct proof-of-concept studies to evaluate EaseVRx and EaseVRx-Distraction as opioid-sparing tools for acute and chronic pain. The results of these studies will inform the company's regulatory pathway with the FDA. Following are descriptions of the three programs:

**Skills-Based VR: EaseVRx**

EaseVRx was developed by AppliedVR in partnership with pain psychologist Beth Darnall, PhD, Associate Professor of Anesthesiology, Perioperative and Pain Medicine at Stanford University and an NIH- and PCORI-funded investigator in the use of VR for pain management. EaseVRx incorporates the evidence-based principles of CBT, mindful meditation, and physiologic biofeedback therapy using embedded biometric sensors. It combines psychoeducation, pain education, breathing training, relaxation exercises, and executive functioning games to provide a mind-body approach toward living better with chronic pain. The standardized, prescriptive, and reproducible 56-day program delivers a combination of skills training and CBT-related treatments through scheduled daily virtual experiences. The participants are allowed to complete as many sessions as they would like. In addition to the prescribed schedule of content the participant will have access to the full library of content. Each VR experience lasts between 2-16 minutes, with an average duration of 6 minutes. The VR treatment modules are designed to minimize triggers of emotional distress or cybersickness. There are 5 types of modules:

- Interoceptive: These modules are designed to help the user understand and perceive what is happening inside the body. They provide a biofeedback-like platform in which the changes in the observed environment reflect a progressively enhanced state of relaxation.
- Education: Help the user understand why the VR exercises are relevant to their pain, as well as teaching specific topics often used in pain psychology, including the neurobiology of pain, the role of mood and stress in pain, pain catastrophizing, activity pacing and setting goals. The goal is for the user to create self-management steps and a toolkit of strategies they can use to manage their response to pain.
- 360-degree videos: High-quality 360 videos with voiceovers, music, and sound effects that are designed to maximize relaxation and engagement of users.
- Game modules: Games are designed to maximize distraction and engagement, increasing the cognitive load on patients, and decreasing their perception of pain.
- Dynamic breathing: These modules are based on evidence-based biofeedback training designed to enhance awareness of one's physiological response to pain and to self-regulate that response. In a virtual world, the user experiences a gamified biofeedback session in which they are introduced to awareness of their breath via visualization in the form of air bubbles. In multiple sessions, the user receives increasingly challenging tasks to practice diaphragmatic breathing while interacting with the virtual environment. The user is also asked to pace their breath according to an expanding and contracting ring in the environment to slow the breath and create physiological changes associated with relaxation. The user's exhale is measured by the microphone embedded in the headset, offering biodata-enabled immersive therapeutics.

**Distraction-Based VR: EaseVRx-Distraction**

EaseVRx-Distraction has the same number of experiences, the same approximate duration of experiences, and the identical user interface as EaseVRx, with a linear prescribed sequence of experiences. The key difference is that instead of offering a variety of VR experiences including education, games, and breath biofeedback, EaseVRx Distraction only includes 360-degree videos - which are also present in EaseVRx. This is intended to remove the effect of education and skills-based training, while preserving the immersive experience of 360- degree VR.

**Sham VR: EaseVRx-Sham**

EaseVRx Sham software includes 2D nature footage accompanied by neutral music that is selected to be neither relaxing nor distracting, rather than 360-degree, 3D, interactive content specially selected for efficacy. The experience of using EaseVRx-Sham is similar to watching a large-screen TV, but it is not interactive or immersive. EaseVRx Sham has the same number and duration of experiences as EaseVRx, and the functionality of the user interface used to access the experiences is the same. Modifications were made to the appearance of the user interface in order to remove aspects that were added for therapeutic benefit.

The study intervention will be delivered using a commercially available VR headset, the PICO G2 4K (https://www.pico-interactive.com/us/G2_4K.html). The headset battery requires recharging after approximately 2.5 hours of use.

AppliedVR has created videos to help participants understand how to use the headset and cover trouble-shooting issues with the headset and software. The Cedars-Sinai study team also will provide technical support.

### 6.1.2 DOSING AND ADMINISTRATION

All participants are instructed to use the VR program at least once a day and as needed throughout the first eight weeks. Thereafter, the participant can use the program as needed. Section 4.3 provides the justification of dose.

## 6.2 PREPARATION/HANDLING/STORAGE/ACCOUNTABILITY

### 6.2.1 ACQUISITION AND ACCOUNTABILITY

The PICO VR headsets loaded with the EaseVRx software programs will be provided by AppliedVR. The devices will be managed by the research team at Cedars-Sinai Medical Center.

### 6.2.2 FORMULATION, APPEARANCE, PACKAGING, AND LABELING

N/A

### 6.2.3 PRODUCT STORAGE AND STABILITY

The PICO VR headsets will remain in the boxes supplied by AppliedVR until they are shipped to participants.

### 6.2.4 PREPARATION

Prior to shipping the VR headset to a study participant, the headset battery will be checked to make sure it is fully charged. Any headsets that are handled by staff will be sanitized by cleaning the fabric surfaces using Virex, the plastic housing using Sani-Wipes, and the glass lenses using alcohol-based lens cleaner.

## 6.3 MEASURES TO MINIMIZE BIAS: RANDOMIZATION AND BLINDING

We will allocate study participants using a random number generator to assign blocks of 3, 6, 9, or 12 to ensure there is an equal distribution in the EaseVRx skills-based group vs. the EaseVRx-Distraction group vs. the EaseVRx-Sham group. Participants, their clinical providers, and study statisticians will be blinded to the study arm. The groups will be labeled as A, B or C at random. A clinical research coordinator will ship the headset containing the assigned program to the participant and enter the group label into the log of enrolled participants. Datasets will be provided to the statistician using these group labels. The research coordinator who maintains the list of randomization assignments will not call participants on the telephone. Because the intervention is conducted in participants' homes and the data are collected remotely, the participants will not encounter each other and thus will be less likely to guess their study arm assignment. We anticipate there will be no circumstances during the study that require unblinding of an individual participant or a whole group because we do not expect any related SAEs to occur with this low risk intervention.

We have restructured the block randomization module within REDCap with stratification based on the 4 sites of recruitment (1 – pain clinic, 2 spine clinic, 3 – rheumatology clinic, 4 – Deep6). Randomization will be performed on a 1:1:1 basis across the 3 treatment groups and stratified across the 4 treatment sites with random block sizes (3, 6, 9, or 12). Allocation tables will be generated using STATA v16 software and uploaded into REDcap for patient allocation.

## 6.4 STUDY INTERVENTION COMPLIANCE

Data collection will be monitored daily by the study team. Participants initially will be sent a reminder email when they have completed no survey questionnaires 2 days after the survey is distributed. If the participant does not respond to the email two additional 48-hr reminders will be sent. S team member will call the participant on the next business day If no data are received. They will be asked why they have missed all three reminders and encouraged to continue to complete surveys as they are sent. If the participant does not complete any surveys in the nexty 72hrs they will be sent a final email with a link to last survey. The participant will not be contacted again until the end of study to request a return of equipment.

## 6.5 CONCOMITANT THERAPY

Participants in the study maybe receiving ongoing treatment for their chronic pain, as listed in the table below.

| Category | Specific therapy |
|---|---|
| Procedures | Injections (steroid, facet joints, radio frequency ablations, sacroiliac joint injections, Epidural injections, inter laminar or transforaminal) |
| | TENS unit |

| | |
|---|---|
| | Blocks (facet blocks, medial branch blocks, trigger point injections) |
| | Acupuncture |
| | |
| Physical Therapy | Physical Therapy and Aqua therapy |
| | Chiropractic care |
| | |
| Psychotherapy | CBT |
| | Biofeedback |
| | Screening and management of depression/anxiety |
| | |
| Medications | Benzodiazepine tapering program |
| | Opioid maintenance and tapering program |
| | Tylenol |
| | Muscle Relaxants |
| | Narcotics |
| | Gabapentin/pregabalin |
| | NSAIDS |
| | Cannabinoids |

### 6.5.1 RESCUE MEDICINE

NA

# 7 STUDY INTERVENTION DISCONTINUATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL

## 7.1 DISCONTINUATION OF STUDY INTERVENTION

We do not anticipate any events that would cause the study intervention to be discontinued.

## 7.2 PARTICIPANT DISCONTINUATION/WITHDRAWAL FROM THE STUDY

Participants are free to withdraw from the study at any time upon request. Participants request to withdraw will be asked to clarify whether they wish to discontinue the intervention but still complete the periodic survey questionnaires. In this case, the participant would not be considered withdrawn.

An investigator may withdraw a participant who develops any of the exclusions criteria in section 5. The reason for participant discontinuation or withdrawal from the study will be recorded.

Subjects who signed the informed consent form, were randomized and received the study intervention, and subsequently withdraw, or are withdrawn or discontinued from the study without engaging with the VR intervention will be replaced. Participants are considered to have engaged with the VR intervention if there is any self-reported VR usage in survey questionnaires or if any metadata are recorded in a participant's headset during their possession of the VR headset.

### 7.3 LOST TO FOLLOW-UP

A participant will be considered lost to follow-up if he or she fails to schedule an onboarding phone call within 30 days of device delivery or if he or she fails to complete the initial SSQ+ survey within 30 days of onboarding completion. For participants who complete the initial SSQ+ survey, fail to complete >80% of future surveys, and become unreachable by phone or email, the participant will be considered lost to follow-up 6 days following the email date of the participant's Day 90 survey.

If a participant fails to complete the Day 90 survey within six days, the following actions will be taken:

- The research coordinator will attempt to contact the participant and counsel them on the importance of completing surveys and returning the equipment to be eligible for the gift card.
- Before a participant is deemed lost to follow-up, the investigator or designee will make every effort to regain contact with the participant (where possible, 3 telephone calls and, if necessary, a certified letter to the participant's last known mailing address or local equivalent methods). Contact attempts will be documented in the participant's study record.
- Should the participant continue to be unreachable, he or she will be considered to have withdrawn from the study with a primary reason of lost to follow-up.

## 8 STUDY ASSESSMENTS AND PROCEDURES

### 8.1 EFFICACY ASSESSMENTS

We will capture all survey data via REDCap. REDCap is a secure web application for building and managing online surveys and databases. While REDCap can be used to collect virtually any type of data (including 21 CFR Part 11, FISMA, and HIPAA-compliant environments), it is specifically geared to support online or offline data capture for research studies and operations. If subjects do not complete the surveys, they will be sent a reminder prompt.

*Primary Outcome*

The change from study baseline to Day 30 in pain interference as measured by the 8-item PROMIS Pain Interference scale is the primary endpoint. The scale is rendered using a T-statistic, where a score of 50 represents the population mean and 10 points is a standard deviation (SD). Scoring of the instrument will occur in a SAS/STATA environment in which study statisticians are blinded to the study arm.

*Secondary Outcomes*

The change from study baseline to day 90 in pain interference as measured by the 8-item PROMIS PI scale is a secondary endpoint. The scale is rendered using a T-statistic, where a score of 50 represents the population mean and 10 points is a standard deviation (SD). Scoring of the instrument will occur in a SAS/STATA environment in which study statisticians are blinded to the study arm.

The change from study baseline to Day 90 in pain catastrophizing as measured by PCS SF-6 is a secondary endpoint. The Pain Catastrophizing scale (PCS) short form measures rates of high catastrophizing as defined by a score of ≥7. Scoring of the instrument will occur in a SAS/STATA environment in which study statisticians are blinded to the study arm.

The change from study baseline to Day 90 in anxiety as measured by the 6-item PROMIS Anxiety scale is a secondary endpoint. The scale is rendered using a T-statistic, where a score of 50 represents the population mean and 10 points is a standard deviation (SD). Scoring of the instrument will occur in a SAS/STATA environment in which study statisticians are blinded to the study arm.

The change from study baseline to day 90 in sleep disturbance as measured by the 6-item PROMIS Sleep Disturbance scale is a secondary endpoint. The scale is rendered using a T-statistic, where a score of 50 represents the population mean and 10 points is a standard deviation (SD). Scoring of the instrument will occur in a SAS/STATA environment in which study statisticians are blinded to the study arm.

The change from study baseline to Day 90 in 7-day average of daily maximum milligrams morphine equivalent (MME) is a secondary endpoint. Prescription data will be extracted from the EHR at baseline and again at Day 90 and may be supplemented by data from CURES. While daily MME of prescriptions is a calculated field in the EHR, calculation of the 7-day average will occur in a SAS/STATA environment in which study statisticians are blinded to the study arm.

CURES is the California Controlled Substance Utilization Review and Evaluation System, a database of Schedule II, III and IV controlled substance prescriptions dispensed in California. It is administered by the California Department of Justice (CA DOJ). We will request prescription data from the CURES database for each randomized patient from 90 days before enrollment until 90 days after the completion of the study. This will allow us to assess prescription changes that might be made right before or after the study ends. The data from CURES will only be used for research purposes and will comply with all requirements of the California Confidentiality of Medical Information Act and the Health Insurance Portability and Accountability Act of 1996 (HIPAA), including the HIPAA regulations in 45 Code of Federal Regulations parts 160 and 164. We will randomize the patients between 10/1/2020 to 9/5/2023. We will submit all consent forms with the request for data by 9/25/2023. The “CURES Access Period” will be designated from 9/1/2023 to 7/1/2024. We will provide CURES with the first name, last name, DOB, and copy of consent for every participant that we request information. The CA DOJ will be notified upon completion of the study, which is anticipated to be 7/30/2024. Per CURES regulations, the “CURES Data Destruction Date” will be 10/30/2029. On that date the dataset from CURES will be deleted and a notification will be sent to CURES. The CA DOJ will be provided with an advance copy of any manuscript or presentation that uses the CURES data, as well as a copy of final publications. The CA DOJ also will be notified when a team member is added to or removed from the study.

## 8.2 SAFETY AND OTHER PROCEDURES

Safety will be monitored by following up with participants for any adverse events (AE), assessment of adherence, monitoring of biometric information, and questionnaires by support staff.

Participant safety - Severe Depression & Substance Use Disorders

Study participants will complete the PROMIS Depression questionnaire at baseline and every two weeks throughout the study. Because the surveys are completed electronically, scoring takes place immediately. In a study comparing PROMIS Depression scores with the PHQ-9 commonly used to screen for depression, a t-score greater than 70 indicates severe depression. The REDCap system will be programmed to alert the study team via email if a participant’s score is higher than 70. A physician investigator will contact the participant and conduct a clinical assessment of suicide risk, provide education on the importance of a safe treatment plan, direct them to emergency services if indicated, and provide assistance with finding an appropriate treatment provider. If an individual meeting this

threshold can not be contacted by a physician investigator for a phone assessment, the individual will be emailed a list of available resources.

Substance use disorders will be screened using the TAPS 1&2 at baseline and then at week 12. The REDCap system has been calibrated to alert the team when the patient achieves a score >2 on the combined assessment. The research team will then email the patient a list of available resources.

### 8.2.1 PROCEDURES

Virtual Reality Headset data

Data from the PICO G2 4K device will be aggregated by AppliedVR's cloud-based software solution that implements robust industry standards to maintain secure databases and keep data private. The AppliedVR cloud server is a HIPAA compliant platform. The device account that corresponds with each device will be created using a clinical trial number that is only linked to the patient via our office list of patients enrolled in the study. AppliedVR does not collect personally identifiable information and does not collect IP addresses from synced participant devices in our database. Data will be stored and indexed in the AppliedVR server database whenever devices upload batches of analytic events, and the analytic events are timestamped. The data collected by the device will include time of use, date of use, and the module selected. Our database servers are IP firewalled and whitelisted such that they refuse any connection from IP addresses not preprogrammed by our team.

Fitbit data

Data from the Charge 4 device will be aggregated by Fitabase, a fully hosted, cloud-based software solution that implements robust industry standards to maintain secure databases and keep data private. Accordingly, the Charge 4 account that corresponds with each device will be created using an anonymous email address not linked to a real person. Fitabase does not collect personally identifiable information and does not collect IP addresses from synced participant devices. Data will be stored and indexed in the Fitabase SQL Server database in day total, hour, and minute-by-minute intervals. Our database servers are IP firewalled and whitelisted such that they refuse any connection from IP addresses not preprogrammed by our team.

Technical Support

Patients in all three arms will receive remote technical support from the research team. The idea is that issuing devices is usually insufficient to achieve behavior change; supporting those devices with high-touch yet scalable care is a vital component.

Patients will be provided with onboarding material as well as emails with a link to our study website. The study website was based on experience from previous remote VR therapy clinical trial and feedback from patients within it. On the website we will have instructional videos as well an extensive FAQ page. The patient will have an onboarding phone call once the device is delivered to their home. During this call, patients will be asked if they reviewed the onboarding video that was linked in the device tracking information email, and coordinators will answer any remaining questions and guide participants through the device or program if necessary. Contact information will be provided for technical support throughout the remainder of the participant's study enrollment.

Redcap survey data will also be monitored. We will assign two technical support staff members to monitor patients in all three arms. Patients will receive a telephone number and email to contact support staff as needed.

VR device metadata will be provided monthly by AppliedVR as CSV files containing timestamped records of content accessed at the individual device level. It will provide the amount of time the patient spent on the device and the modules that they completed.

## 8.3 ADVERSE EVENTS AND SERIOUS ADVERSE EVENTS

### 8.3.1 DEFINITION OF ADVERSE EVENTS (AE)

An adverse event (AE) is defined as any untoward medical occurrence associated with the use of an intervention in humans, whether or not considered intervention-related (21 CFR 312.32 (a)). Short-term AEs associated with the study include "cybersickness,"[76] a transient form of motion sickness that affects up to one in four people upon entering a VR environment. Cybersickness most commonly presents with a short-term feeling of dizziness that typically subsides quickly. There are no anticipated long-term AEs from participating in this study. A comprehensive list of all potential cybersickness symptoms can be found below:

- Nausea
- Fatigue
- Eyestrain
- Blurred vision
- Difficulty focusing
- Dizziness
- Vertigo (a sensation of spinning dizziness, as though the room or surrounding environment is spinning)
- Headache
- Fullness of the head
- Difficulty concentrating
- Postural instability

In addition, in rare situations, one may have an allergy to the facial interface that generates a rash, which typically resolves without any medication in under 24 hours. There is also the possibility of a participant experiencing neck pain as a result of using the headset for an extended period of time. Participants will be informed of ways to mitigate the possibility of neck pain while using the headset.

### 8.3.2 DEFINITION OF SERIOUS ADVERSE EVENTS (SAE)

A serious adverse event (SAE) or suspected adverse reaction is considered "serious" if, in the view of either the investigator or sponsor, it results in any of the following outcomes: death, a life-threatening adverse event, inpatient hospitalization or prolongation of existing hospitalization, a persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions. Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered serious when, based upon appropriate medical judgment, they may jeopardize the

participant and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. We do not expect any serious adverse events with a virtual reality intervention.

### 8.3.3 CLASSIFICATION OF AN ADVERSE EVENT

#### 8.3.3.1 SEVERITY OF EVENT

For adverse events (AEs) the following guidelines will be used to describe severity.

- **Mild** – Events require minimal or no treatment and do not interfere with the participant's daily activities.
- **Moderate** – Events result in a low level of inconvenience or concern with the therapeutic measures. Moderate events may cause some interference with functioning.
- **Severe** – Events interrupt a participant's usual daily activity and may require systemic drug therapy or other treatment. Severe events are usually potentially life-threatening or incapacitating. Of note, the term "severe" does not necessarily equate to "serious".

Research staff will call the patient if an event is noted in an event assessment that is sent on a weekly basis. Nothing more than mild side effects that do not require medical treatment are expected.

#### 8.3.3.2 RELATIONSHIP TO STUDY INTERVENTION

All adverse events (AEs) must have their relationship to study intervention assessed by the clinician who examines and evaluates the participant based on temporal relationship and his/her clinical judgment. The degree of certainty about causality will be graded using the categories below. In a clinical trial, the study product must always be suspect.

- **Related** – The AE is known to occur with the study intervention, there is a reasonable possibility that the study intervention caused the AE, or there is a temporal relationship between the study intervention and event. Reasonable possibility means that there is evidence to suggest a causal relationship between the study intervention and the AE.
- **Not Related** – There is not a reasonable possibility that the administration of the study intervention caused the event, there is no temporal relationship between the study intervention and event onset, or an alternate etiology has been established.

#### 8.3.3.3 EXPECTEDNESS

Research staff will be responsible for determining whether an adverse event (AE) is expected or unexpected. An AE will be considered unexpected if the nature, severity, or frequency of the event is not consistent with the risk information previously described for the study intervention. A detailed description of expectedness is provided in DSMP section 1.2.3.

### 8.3.4 TIME PERIOD AND FREQUENCY FOR EVENT ASSESSMENT AND FOLLOW-UP

The occurrence of an adverse event (AE) or serious adverse event (SAE) may come to the attention of study personnel during review of weekly event assessments by a study team member as well as on Day 1. We will employ a general question to allow the patient to provide an event description without bias by presupposing the nature of the event, as follows:

1) During the past 14 days, have you experienced anything uncomfortable, distressing or upsetting as a result of using the VR headset?
a. Yes
b. No
2) If Yes please describe what happened. A research staff member will follow-up up with you as soon as possible to learn more.
a. [open text]

A designated CRC will review the REDCap dashboard for completed AE assessments on a daily basis throughout the work week. Once they find a completed event assessment form, they will follow-up with a phone call. The following information be obtained from the patient during the call: onset of potential adverse event, event description, when they last used the VR headset, how long they used the headset, when they stopped using the headset, severity of the potential AE, outcome of the potential AE (see categories, below), time until resolution/stabilization of event, the perceived relationship of the event to the study intervention, whether or not the event was expected (Y/N), whether it was serious.

The outcome can be labeled using one of the following categories:
- Recovered, without treatment
- Recovered, with treatment
- Still Present, no treatment
- Still Present, being treated
- Residual effect(s) present-no treatment
- Residual effect(s) present-being treated

Research staff will record all reportable events with start dates occurring any time after informed consent is obtained until 7 (for non-serious AEs) or 30 days (for SAEs) after the last day of study participation. Please see section 1.3 in DSMP for instructions on how events will be followed for outcome information until resolution or stabilization.

### 8.3.5 ADVERSE EVENT REPORTING

Cedars-Sinai IRB
If a mild or moderate event occurs it does not need to be submitted to the Reportable New Information (RNI). If the event is an unexpected (not usually associated VR side effects) or severe event (requiring treatment), an AE will be submitted to the IRB as soon as possible but within at least 10 working days.

DSMB
All AEs, regardless of their severity, relatedness or expectedness, are reported in aggregate as part of the routine Data and Safety Monitoring Report to the NIAMS and DSMB, twice per year.

A detailed description of procedure for collecting event is in section 1.2.3 of the DSMP.

### 8.3.6 SERIOUS ADVERSE EVENT REPORTING

SAEs will be submitted to the Cedars-Sinai IRB via RNI as soon as possible by study coordinators, but no later than 10 business days from the Principal Investigator's or study team's awareness of the event,

incident, information or outcome. In addition, SAEs will be reported to the DSMB and NIAMS via the NIAMS executive secretary within 48 hours of becoming aware of the event.

#### 8.3.7 REPORTING EVENTS TO PARTICIPANTS

N/A

#### 8.3.8 EVENTS OF SPECIAL INTEREST

N/A

#### 8.3.9 REPORTING OF PREGNANCY

N/A

### 8.4 UNANTICIPATED PROBLEMS

#### 8.4.1 DEFINITION OF UNANTICIPATED PROBLEMS (UP)

The Office for Human Research Protections (OHRP) considers unanticipated problems involving risks to participants or others to include, in general, any incident, experience, or outcome that meets **all** of the following criteria:

- Unexpected in terms of nature, severity, or frequency given (a) the research procedures that are described in the protocol-related documents, such as the Institutional Review Board (IRB)-approved research protocol and informed consent document; and (b) the characteristics of the participant population being studied;
- Related or possibly related to participation in the research ("possibly related" means there is a reasonable possibility that the incident, experience, or outcome may have been caused by the procedures involved in the research); and
- Suggests that the research places participants or others at a greater risk of harm (including physical, psychological, economic, or social harm) than was previously known or recognized.

This definition could include an unanticipated adverse device effect, any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with, a device, if that effect, problem, or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application (including a supplementary plan or application), or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects (21 CFR 812.3(s)).

#### 8.4.2 UNANTICIPATED PROBLEM REPORTING

Study coordinators will report unanticipated problems (UPs) to the Cedars-Sinai IRB as a RNI submission as soon as possible by study coordinators, but no later than 10 business days from the Principal Investigator's or study team's awareness of the event, incident, information or outcome. UPs include SAEs and AEs, which are both unexpected and possibly related to the research as well as SAEs and AEs that meet the definition of a Research-Related Subject Injury (RRSI) – a medical condition that is caused by and/or directly related to the research study (i.e., the condition would not have existed "but for" the subject's participation in the study), and requires diagnosis or treatment.

The UP report will include the following information:

- Protocol identifying information: protocol title and number, PI's name, and the IRB project number;
- A detailed description of the event, incident, experience, or outcome;
- An explanation of the basis for determining that the event, incident, experience, or outcome represents an UP;
- A description of any changes to the protocol or other corrective actions that have been taken or are proposed in response to the UP.

UPs will also be reported to NIAMS and the DSMB via the NIAMS executive secretary within 48 hours of the PI becoming aware of the event.

### 8.4.3 REPORTING UNANTICIPATED PROBLEMS TO PARTICIPANTS

N/A

# 9 STATISTICAL CONSIDERATIONS

## 9.1 STATISTICAL HYPOTHESES

- **Primary Efficacy Endpoint(s):**

Primary Outcome: Pain interference over time as measured by PROMIS-Pain Interference 8a (T-Scored, Continuous).

Hypothesis: Immersive VR (Skills-based VR or Distraction VR) will lead to a greater improvement in pain interference than Sham VR over 30-day period.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy.

Additional comparison: If there is a statistically significant difference in Skills based VR vs Sham VR, then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy.

- **Secondary Efficacy Endpoint(s):**

Secondary Outcome: Pain interference over time as measured by PROMIS-Pain Interference 8a (T-Scored, Continuous).

Hypothesis: Immersive VR (Skills-based VR or Distraction VR) will lead to a greater improvement in pain interference than Sham VR at Day 60 and Day 90.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy.

Additional comparison: If there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy.

Secondary Outcome: Pain catastrophizing over time as measured by Pain Catastrophizing scale (PCS) short form.

Hypothesis: Skills-based VR will lead to a greater improvement in Pain catastrophizing than distraction-based VR or Sham VR.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Secondary Outcome: Anxiety over time as measured by PROMIS Anxiety.

Hypothesis: Skills-based VR will lead to a greater improvement in PROMIS Anxiety than distraction based VR or Sham VR.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Secondary Outcome: Quality of sleep over time as measured by PROMIS Sleep Disturbance.

Hypothesis: Skills-based VR will lead to a greater improvement in sleep quality than distraction based VR or Sham VR.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Secondary Outcome: Comparing the change from study baseline to Day 90 in weekly MME of prescribed medication.

Question: Which intervention is more effective in reducing opioid prescriptions of Morphine Milligram Equivalents (MME) or greater?

Hypothesis: We believe that Skills-based VR will have a statistically significant decrease in opioid use in comparison to distraction-based VR and Sham VR.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

## 9.2 SAMPLE SIZE DETERMINATION

Sample Size calculations

The primary goal of this aim is to test the efficacy of immersive VR in improving perceived pain. The trial will be a success if there is statistical evidence that either immersive VR group is better than control or sham VR. Preliminary studies showed that the PROMIS pain interference scale has a SD of 10. Assuming that the SD at baseline ($SD_0$) and at 30 days ($SD_1$) are similar and equal to 10 for this population, the variance for the difference in PROMIS from baseline to 30 days after intervention is $Var_{diff}= SD_0^2 +SD_1^2 - 2 \rho SD_0 SD_1$, where ρ is the correlation coefficient between measurements at baseline and 30 days. A conservative estimate of this variance is achieved when ρ = 0. Therefore, the estimate $SD_{diff} = (10^2 + 10^2)^{0.5} = 14.14$. Let $m_c$, $mv_1$, and $mv_2$ be the mean change in PROMIS score from baseline to 30 days for the control, VR1, and VR2 groups, respectively.

We estimate power by simulating 10000 trial replicates and testing the null hypothesis that Ho: $m_c = mv_1 = mv_2$ versus the alternative hypothesis that $H_1$: $m_c \neq mv_1$ or $m_c \neq mv_2$. To maintain the familywise error rate at 0.05, a two-sample t-test is used to compare the control arm to each VR arm and the test is declared statistically significant if the p-value of the two-sided test is less than 0.025. Under the alternative hypothesis that $| m_c - mv_1| = | m_c – mv_2| = 5$, data from 120 patients in each of the three arms achieve 83% power to detect a clinically meaningful effect 5 units in the PROMIS score. The actual type I error rate is 0.049. For the secondary outcome of comparison between the two VR arms, we test the null hypothesis that Ho: $mv_1 = mv_2$ versus the alternative hypothesis that $H_1$: $mv_1 \neq mv_2$ at the two-sided 0.05 level of significance if there is statistical evidence that both VR arms are better than control. Using the same assumptions as above and simulating 10000 trial replicates, then if both VR arms are better than the control arm, we can achieve 71.4% power to detect 5 units in PROMIS score between the two VR arms. This power was derived under the alternative hypothesis $| m_c - mv_1|= 5$ and $| mv_1 – mv_2| = 5$. The actual type I error rate for this conditional test is 0.02.

Power Estimation

We estimate power to assess the effect of VR type on treatment response using logistic regression, accounting for all possible confounding factors as described in Aim 1. The outcome variable in treatment response (binary) and the predictor of interest is VR type, traditional versus enhanced. Table 1 gives the minimum odds ratio that can be detected with 80% power with the two-sided 0.05 level of significance as a function of baseline probability of positive response when a patient is treated with traditional VR and $R^2$, the proportion of variability in the predictor of interest (VR type) that is explained by all relevant baseline covariates in the model using data from 120 patients in the traditional VR arm and 120 in the enhanced VR arm. For example, data from 240 patients achieve 80% power to detect an odds ratio of 2.19 if the probability of positive response when a patient is treated with traditional VR is 0.3 and 10% of the variability in VR type is explained by all other baseline covariates in the model. These odds ratios vary between 2.01 and 2.8 and are clinically meaningful. Therefore, we have enough power to test statistical significance of predictors of interest in the multivariable logistic regression model.

**Table 1.** Minimum detectable odds ratio as a function of the proportion of variability in VR type variable that is explained by all other relevant covariates in the model and the baseline probability of positive treatment response.

| **Baseline probability** | $R^2$ | | | |
|---|---|---|---|---|
| | 0.0 | 0.1 | 0.2 | 0.4 |
| 0.2 | 2.26 | 2.36 | 2.48 | 2.82 |
| 0.3 | 2.11 | 2.19 | 2.30 | 2.61 |
| 0.5 | 2.01 | 2.18 | 2.30 | 2.64 |

## 9.3 POPULATIONS FOR ANALYSES

All efficacy and safety data summaries and analyses will be performed by study arm using an Intent-to-Treat (ITT) population defined as all randomized patients.

The number of patients identified as candidates will be reported, as will the number screened (i.e. consented). The number and percentage of patients randomized, patient population (ITT), and treatment status (completed, discontinued/withdrew) will be summarized both by treatment group and overall. Reasons for discontinuation/withdrawal will be presented.

An exploratory, per protocol (PP) analysis will focus on patients who use the assigned intervention on at least 50% of days during the first 30-day period. Usage meta-data on the headsets will facilitate sample definition in this population. If headset meta-data is not available, we will rely on weekly self-report usage surveys.

## 9.4 STATISTICAL ANALYSES

### 9.4.1 GENERAL APPROACH

All statistical analyses will be performed jointly by the Cedars-Sinai Biostatistics Core and the Cedars-Sinai Center for Outcomes Research and Education (CS_CORE) using SAS® software version 9.3 or higher (SAS Institute, Cary, NC, USA), R pack3.5.0 (R Foundation for Statistical Computing, Vienna, Austria)or Stata software version 14 or higher (StataCorp LLC, College Station, TX, USA).
Continuous variables will be summarized using descriptive statistics (N, mean, standard deviation, minimum, median, and maximum). Categorical variables will be summarized using the number and percentage of patients in each category. Data will be summarized with respect to patient demographic and baseline characteristics both across the study and by study-arm. The efficacy endpoints, safety assessments, and other outcome results for each treatment group will be summarized descriptively unless otherwise indicated. In addition, statistical model estimates of least squares means, treatment differences, p-values and 95% confidence intervals will also be provided where relevant. The fit of linear models will be assessed using residual plots and/or other diagnostic plots as appropriate. The fit of logistic models will be assessed using Hosmer-Lemeshow goodness-of-fit and/or receiver operating characteristic (ROC) curves as appropriate. All statistical tests will be 2-sided and performed at the 0.05 level of significance unless stated otherwise. Baseline is defined as the value at the screening week (Day -8 to Day -2) for all parameters, unless specified otherwise.

### 9.4.2 ANALYSIS OF THE PRIMARY EFFICACY ENDPOINT(S)

The goal of this analysis is to test the effectiveness of immersive VR in improving measures of pain interference. The trial will be a success if there is statistical evidence that either immersive VR group is better than sham VR.
The primary efficacy endpoint is the change from study baseline to Day 30 in pain interference as measured by PROMIS-PI t-score. The primary analysis will compare each treatment group separately to the control group using a linear mixed model repeated measures (MMRM) analysis.[1] The repeated measures are the change from baseline PROMIS-PI t-score obtained at Days 7, 15, and 21, respectively. The model will include fixed categorical effects for treatment, week, treatment by week interaction, and the baseline PROMIS-PI t-score as a continuous covariate. We will employ the Stata *xtmixed* command

with restricted maximum likelihood estimation (REML) and an unstructured within-patient covariance structure for this model. We will evaluate the assumptions of the model, including normality, using residual and other diagnostic plots of model fit.

From this model, we will estimate least squares means, standard errors, treatment differences in least squares means, and 95% confidence intervals for each time period. Primary inference will be based on the treatment comparison of least squares means for Day 30, and a p-value will be presented for this time period only. The null hypothesis is that the mean difference in the primary endpoint between the treatment groups and the sham control group is zero, versus the alternative hypothesis that these differences are not zero. The hypotheses can be expressed as follows:

$$H0_a: \mu SB - \mu control = 0 \text{ versus } H1_a: \mu SB - \mu control \neq 0$$

$$H0_b: \mu D - \mu control = 0 \text{ versus } H1_b: \mu D - \mu control \neq 0$$

Where μSB refers to the mean change from baseline to Day 30 in Promis-PI t-score in the Skills-based VR treatment group, μcontrol refers to the mean change from baseline to Visit Day 30 in Promis-PI t-score in the sham treatment group, and μD refers to the mean change from baseline to Day 30 in Promis-PI t-score in the Distraction VR treatment group. The test will be performed using the final MMRM model with a two-sided significance level of 5%. Estimated least squares means for change from baseline and the observed absolute values (± SE) by treatment group will be plotted over time.

### 9.4.3 ANALYSIS OF THE SECONDARY ENDPOINT(S)

*PROMIS-Pain Interference*

The change from study baseline to Day 60 and 90 in pain interference as measured by PROMIS-Pain Interference is a secondary endpoint. This analysis will compare the treatment groups, separately, to the control group using a MMRM analysis. The repeated measures are the change from baseline PROMIS-Pain Interference score obtained for Day 60 and 90, respectively. The model will include fixed categorical effects for treatment, week, treatment by week interaction, and the baseline PROMIS-Pain Interference score as a continuous covariate. We will employ the Stata *xtmixed* command with REML and an unstructured within-patient covariance structure for this model. We will evaluate the assumptions of the model, including normality, using residual and other diagnostic plots of model fit.

From this model, we will estimate least squares means, standard errors, treatment differences in least squares means, and 95% confidence intervals for each time point. Primary inference will be based on the treatment comparison of least squares means for Day 90, and a p-value will be presented for this time period only. The null hypothesis is that the mean difference in the primary endpoint between the experimental arms and the sham arm is zero, versus the alternative hypothesis that this difference is not zero.

*PROMIS-Sleep Disturbance*

The change from study baseline to Day 90 in sleep disturbance as measured by PROMIS-Sleep Disturbance is a secondary endpoint. This analysis will compare the treatment groups, separately, to the control group using a MMRM analysis. The repeated measures are the change from baseline PROMIS-Sleep Disturbance score obtained at Days 15, 30, 60, and 90, respectively. The model will include fixed categorical effects for treatment, week, treatment by week interaction, and the baseline PROMIS-Sleep Disturbance score as a continuous covariate. We will employ the Stata *xtmixed* command with REML and an unstructured within-patient covariance structure for this model. We will evaluate the assumptions of this model, including normality, using residual and other diagnostic plots of model fit.

From this model, we will estimate least squares means, standard errors, treatment differences in least squares means, and 95% confidence intervals for each time point. Primary inference will be based on the treatment comparison of least squares means for Day 90 from this model, and a p-value will be presented for this time period only. The null hypothesis is that the mean difference in the primary endpoint between the experimental arms and the sham arm is zero, versus the alternative hypothesis that this difference is not zero.

*PROMIS-Anxiety*
The analysis approach for this endpoint mirrors the above description for PROMIS-Anxiety

*PCS SF-6*
The change from study baseline to Day 90 in pain catastrophizing as measured by PCS SF-6 is a secondary endpoint. This analysis will compare the treatment groups, separately, to the control group using a MMRM analysis. The repeated measures are the change from baseline PCS SF-6 score obtained for Day 15, 30, 60, and 90, respectively. The model will include fixed categorical effects for treatment, week, treatment by week interaction, and the baseline PCS SF-6 score as a continuous covariate. We will employ the Stata *xtmixed* command with REML and an unstructured within-patient covariance structure for this model. We will evaluate the assumptions of this model, including normality, using residual and other diagnostic plots of model fit.

From this model, we will estimate least squares means, standard errors, treatment differences in least squares means, and 95% confidence intervals for each time point. Primary inference will be based on the treatment comparison of least squares means for Day 90 from this model, and a p-value will be presented for this time period only. The null hypothesis is that the mean difference in the primary endpoint between the experimental arms and the sham arm is zero, versus the alternative hypothesis that this difference is not zero.

The change from study baseline to Day 90 in weekly MME of prescribed medication is a secondary endpoint. Descriptive summary statistics will be presented for weekly MME of prescribed medication at two timepoints: baseline and Day 90. Analysis of this endpoint will compare the difference in weekly MME of prescribed medication between the treatment groups and control group at baseline and Day 90 using analysis of covariance (ANCOVA), adjusting for baseline MME. We will assess the assumptions underpinning the ANCOVA model graphically and will undertake appropriate transformation of the weekly average MME outcome as deemed appropriate. If no suitable transformations can be found, bootstrapped confidence intervals for the between-group differences in weekly MME will be produced. Both adjusted and unadjusted between-group comparisons will be presented with 95% confidence intervals.

### 9.4.4 SAFETY ANALYSES

All safety analyses will be descriptive. No statistical testing will be performed.

### 9.4.5 BASELINE DESCRIPTIVE STATISTICS

We will summarize demographic and baseline characteristics both by treatment group and overall. Any time-to-event counts will be measured from the day of randomization. Body Mass Index (BMI) will be calculated according to: BMI=weight (kg)/(height (m))2. Age will be calculated according to: Age=(date of event-birth date+1)/365.25. Weekly average opioid dosage will be calculated as a 7-day average of daily maximum milligrams morphine equivalent (MME) as prescribed. Mean activity measures will be calculated as weekly averages (e.g. weekly steps, weekly sleep minutes). Activity classification (i.e.

sedentary versus light versus moderate, etc.) will be provided by FitBit and reported as such. Zip code will be matched to median income using census data as an aggregate measure of socio-economic status.

### 9.4.6 PLANNED INTERIM ANALYSES

No formal interim analysis or interim statistical testing for treatment comparisons is planned.

### 9.4.7 SUB-GROUP ANALYSES

Subgroup analyses are planned for the change from baseline in PROMIS-PI. An MMRM model will be used to test for treatment by subgroup interactions. Interactions with a significance level of less than 10% will be considered potentially important and flagged for further assessment. In general, the models will include fixed categorical effects for treatment, week, treatment by week interaction, subgroup, and treatment by subgroup interaction. The p-values of interaction terms will be presented, as well as the least squares means and 95% confidence intervals by treatment and subgroup classification factor. Descriptive statistics of the observed and change from baseline PROMIS-PI t-score will also be presented by treatment and week within each subgroup.

Subgroup analyses will be performed for the following subgroups:

- Dosage of VR (minutes per week)
- Previous experience with VR
- ITQ score
- Presence score
- Patient comorbidities
- History of spinal surgery
- Pain severity and duration
- Sociodemographics (i.e. age, sex, race, ethnicity, workers comp, marital status, education)
- Other treatments (see section 6.5)
- TAPS-1, TAPS-2
- Mood disturbances (e.g., current depressive disorder or not, severely depressed mood vs. non-depressed, etc.)
- Radicular vs. non-radicular back pain
- SSQ Cutoff (Score of >15)
- Treatment Expectation Question
- Perception of arm of study question
- Motion Sickness Propensity Survey

The primary MMRM models will be re-fit to include terms describing subgroups. The change from baseline PROMIS-PI score obtained at Day 30 will be compared between the treatment groups using MMRM analysis. The repeated measures are the change from baseline PROMIS-PI score obtained for Day 7, 15, 21, and 30, respectively. The model will include fixed categorical effects for treatment, week, treatment by week interaction, relevant subgroup, treatment by subgroup interaction, and the baseline PROMIS-PI score as a continuous covariate. The estimated least squares means and 95% confidence intervals on the treatment comparison for Day 7, 15, 21, and 30, will be presented in forest plots for each subgroup of relevance.

See Section 9.4.9 under exploratory analysis for more detail.

### 9.4.8 TABULATION OF INDIVIDUAL PARTICIPANT DATA

We will tabulate individual participant data according to the Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM) as specified by the Data Integration, Algorithm Development and Operations Management Center (DAC).

### 9.4.9 EXPLORATORY OUTCOMES

A few exploratory hypotheses have been considered:

Exploratory Outcome: Global impression of change with pain treatment using PGIC
Question: Which intervention is associated with patient's perceived change in condition due to treatment?
Dependent variable: patient's perceived change in condition as assessed by PGIC
Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Exploratory Outcome: Quality of sleep over time by biometric data on Fitbit Charge 4 (total minutes asleep and sleep efficiency per night))
Hypothesis: Skills-based VR will lead to a greater improvement in sleep quality than distraction based VR or Sham VR.
Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Exploratory Outcome: Depression over time as measured by PROMIS Depression (version 4)
Hypothesis: Skills-based VR will lead to a greater improvement in PROMIS Depression than distraction-based VR or Sham VR.
Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Exploratory Outcome: Increased physical function over time as measured by total steps using biometric data
Hypothesis: Skills-based VR will lead to a greater improvement in physical activity than distraction-based VR or Sham VR.
Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy

Exploratory Outcome: Improvement in physical function as measured by PROMIS physical function (version 6b)
Hypothesis: Skills-based VR will lead to a greater improvement in physical activity than distraction-based VR or Sham VR.

Comparisons: Superiority for both of the following (1) Distraction VR therapy vs. sham VR therapy; (2) skills-based VR therapy vs. sham VR therapy, if there is a statistically significant difference in Skills based VR vs Sham VR then a third test will be completed (3) Skills-based VR therapy vs Distraction VR therapy See Section 4.6.3 of the SAP for details on the assessment of each the above outcomes.

# 10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

## 10.1 REGULATORY, ETHICAL, AND STUDY OVERSIGHT CONSIDERATIONS

### 10.1.1 INFORMED CONSENT PROCESS

#### 10.1.1.1 CONSENT/ASSENT AND OTHER INFORMATIONAL DOCUMENTS PROVIDED TO PARTICIPANTS

Consent forms describing in detail the study intervention, study procedures, and risks are given to the participant and electronic documentation of informed consent is required prior to starting screener week procedures.

#### 10.1.1.2 CONSENT PROCEDURES AND DOCUMENTATION

All eligible participants identified either by their treating physician or by Deep6 will be contacted by a research coordinator using an IRB approved recruitment email that will explain the study. The patient will be able to reply to the communication with a request to opt-out of further communication. A research coordinator will telephone individuals who do not opt out to discuss their interest in study participation. An IRB-approved script will explain the study in lay language, including the purpose, procedures, and potential risks of the study and research rights as a participant. Study Coordinators will ensure the identity of the person on the phone call is indeed the prospective participant that was intended. REDCap will be used to email the IRB approved electronic consent form to the patient. Participants will have the opportunity to carefully review the electronic consent form and ask questions prior to signing. Participants will complete their consent forms electronically using the 21 CFR 11-compliant online platform, REDCap. A child assent form will be included as part of the informed consent to allow for adolescent participants to be enrolled. Every subject will be informed of the approximate time to complete the consent process. A copy of the informed consent documents will be provided to participants for their records via email. The informed consent will be signed by the patient and the designated investigator before the participant undergoes any study-specific procedures. The rights and welfare of the participants will be protected, and it will be emphasized that their medical care will not be adversely affected in any way if they decline to participate in this study.

All patients will be able to opt-out of the study via email or by phone. We request removal of an official wait time between sending the email/letter and contacting subjects by phone. Our justification is that, to date, in the rural chronic pain study alone, 153 recruitment letters have been sent to Cedars-Sinai patients. However, none of them have contacted our study staff prior to the 7-day period, and of the 33 individuals who were sent a recruitment letter and later declined, only two declined by email. 98.7% of recruitment letters have not obviated the need for a follow-up recruitment call, and this response rate compares to our other trials. Most individuals do not recall seeing or reading the letter when finally reached by phone.

A verbal explanation will be provided in terms suited to the participant's comprehension of the purposes, procedures, and potential risks of the study and of their rights as research participants. Participants will have the opportunity to carefully review the electronic consent form and ask questions prior to signing. The participants should have the opportunity to discuss the study with their family or surrogates or think about it prior to agreeing to participate. The participant will sign the informed consent document prior to starting the screening week surveys. Participants must be informed that participation is voluntary and that they may withdraw from the study at any time, without prejudice.

The informed consent states that any changes in pain management should be made with the participants treating physician.

### 10.1.2 STUDY DISCONTINUATION AND CLOSURE

We do not anticipate any events that would warrant study discontinuation and closure.

### 10.1.3 CONFIDENTIALITY AND PRIVACY

Participant privacy is strictly held in trust by the participating investigators, their staff, and their interventions. Therefore, the study protocol, documentation, data, and all other information generated will be held in strict confidence. No information concerning the study or the data will be released to any unauthorized third party without prior written approval of the sponsor.

Representatives of the IRB or regulatory agencies may inspect all documents and records required to be maintained by the investigator for the participants in this study. The clinical study site will permit access to such records as needed.

Data in the study is collected in three ways: in real-time, at infrequent intervals throughout the study, and through medical record queries. Real-time data, including biometrics and survey data delivered via mobile device, will be stored on secure servers hosted by Amazon Web Services (AWS) and will contain only a unique identifier for each participant. Virtual reality adherence data will also be tracked in real-time and hosted on secure servers by AppliedVR; a separate unique ID will be assigned to each participant. Data collected at infrequent intervals throughout the study, such as entry, 14-day interval assessments, and exit questionnaires will be stored on secure CSMC servers with unique ID's for each participant. Data collected from medical record numbers, such as opioid prescriptions and physician history, will reside on secure CSMC servers and an ID will be assigned to each individual in order to abstract PHI/PII and the medical record number. Each dataset will utilize different unique ID's and a list linking each unique ID to each participant will be stored internally on the secured CSMC network. The linking list allows a researcher with access to the secured files to merge all data using statistical software, while maintaining data confidentiality.

To minimize risk of breaches in confidentiality associated with the access and recording of protected health information, study staff will be assigned unique passwords and usernames to access secure servers. Additionally, identifiable information for participants will be obfuscated using unique ID numbers and a linking list will be held in a secure location.

A detailed list of procedures that protect confidentiality is in section 20.1 of the MOOP.

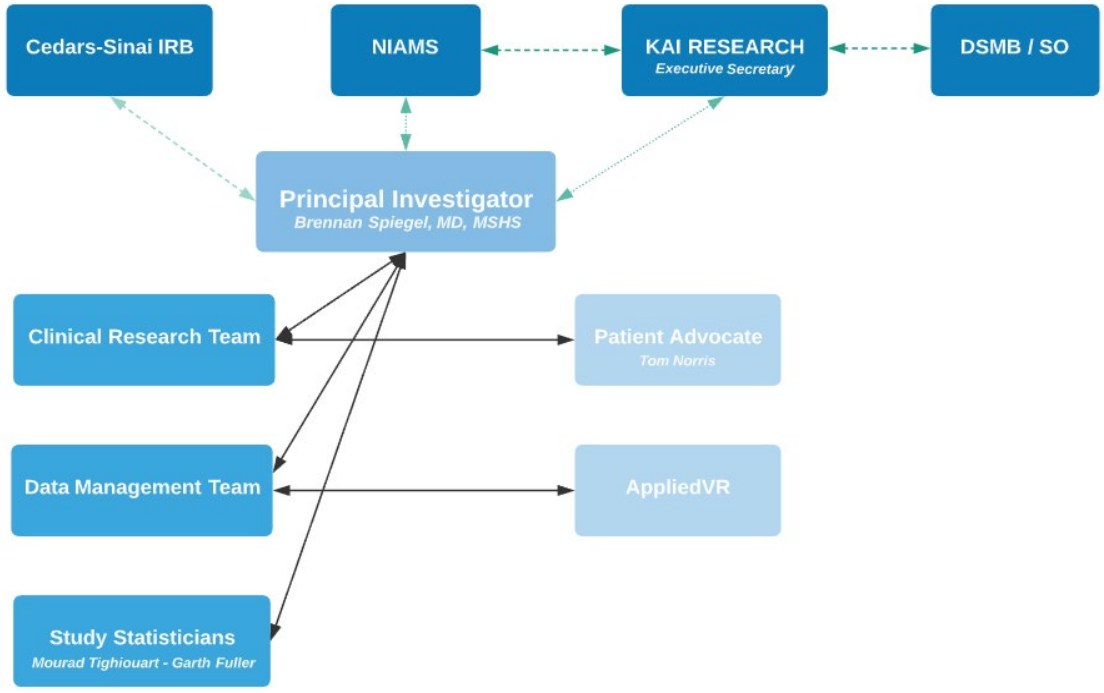
### 10.1.4 FUTURE USE OF STORED SPECIMENS AND DATA

N/A

### 10.1.5 KEY ROLES AND STUDY GOVERNANCE

| **Principal Investigator** |
|---|
| Brennan Spiegel, MD, MSHS<br>Director of Health Services Research |
| Cedars-Sinai Medical Center |
| 116 N. Robertson Blvd, Suite 800<br>Los Angeles, CA 90048 |
| 310.423.6784 |
| Brennan.Spiegel@cshs.org |
| |

As shown in Figure 1, below, the PI oversees the grant and interacts with all the key stakeholders, including NIAMS, KAI Research and associated DSMB, the Cedars-Sinai IRB, and members of the Clinical Research Team, Data Management Team, and the Study Statisticians.




### 10.1.6 SAFETY OVERSIGHT

Safety oversight will be under the direction of a DSMB composed of individuals with the appropriate expertise. The DSMB will be appointed by and serve as advisory to the NIAMS. Members of the DSMB should be independent from the study conduct and free of conflict of interest, or measures should be in place to minimize perceived conflict of interest. The DSMB will meet at least semiannually to assess safety and efficacy data on each arm of the study. The DMSB will operate under the rules of an approved charter that will be written and reviewed at the organizational meeting of the DSMB. At this time, each data element that the DSMB needs to assess will be clearly defined. The NIH will provide its input to specify the study sponsor National Institutes of Health staff.

### 10.1.7 CLINICAL MONITORING

As the intervention in this trial is not greater than minimal risk, a designated member of the study team will conduct periodic monitoring every 6 months and review safety data such as AEs reported, using DSM reports attached in DSMP. The compiled reports will be discussed amongst the study team at the earliest opportunity. These reports will also be submitted to executive secretary and NIAMS via the semiannual DSMB report. The study team will compile with any requested action by the DSMB.

### 10.1.8 QUALITY ASSURANCE AND QUALITY CONTROL

Our quality assurance (QA) and control (QC) procedures are designed to support adherence to the protocol, obtain complete follow-up data for all participants enrolled, and establish and maintain high standards for validity in collected data ahead of analysis. We approach QA as processes and safeguards for the prevention of data errors, and QC as mitigation efforts reducing impact from errors that have occurred during data capture and/or processing. While much of the QA effort occurs ahead of study launch, both efforts will require real-time and periodic tasks conducted through the duration of the study to ensure consistent checks of data integrity, completeness, and correctness.
QA via error prevention will focus largely on the REDCap system. Before data is entered into REDCap, study personnel will:

- Indicate units in question stems and use field validation whenever possible;
- Define and enforce range minimum/maximum where applicable;
- Ensure uniformity in date formatting across all REDCap date fields;
- Program pre-defined multiple-choice fields whenever possible (as opposed to free text);
- Indicate specific, standardized choices to identify data as intentionally missing (e.g “Not Applicable,” “Don’t Know”) as opposed to blank fields;
- Enforce skip/no-skip logic where appropriate;
- Standardize assignment of raw values. (i.e. if “Yes” is coded as a ‘1’ in one field, it should be coded as ‘1’ in all other project fields); and
- Use case and punctuation consistently across all field labels.

Each aspect of the REDCap data collection system will be tested before actual study data is collected. Study personnel will enter mock data into REDCap forms, serving as the “User” for acceptance testing. We will document the success or failure of a) the user interface for data entry, b) the on-line univariate and range data validation checks, and c) custom functions and coding. This mock data will be exported as SAS/STATA datasets by the research coordinator and subjected to QC procedures.
Our fundamental QC approach prioritizes error detection. Automated QC will occur in a SAS/STATA programming environment and will be tested/iterated on mock data before being employed monthly to exported datasets from both REDCap and the EHR. These QC programs will target potential data anomalies including:

- Missing data or forms;
- Out-of-range or erroneous data;
- Inconsistent, improperly formatted, or out-of-range dates; and
- Fields on a "completed form" not completed.

Once the study begins, routine QC reports will be prepared monthly (with 2 weeks lead-time before proposed monthly uploads to DAC) by the study statisticians. These reports will describe target and actual enrollment, eligibles screened with reasons for screen failure, and participant disposition (enrolled; active, completed, discontinued treatment, and lost to followup). These reports will also provide proportions of forms completed/missing, as well as summaries of problems identified by QC processes. Changes to QC programming will be documented and re-tested for accuracy.

Finally, daily monitoring for data completeness will be undertaken by the study team as a QA effort. Patients will initially be sent a reminder email when they have not submitted data after 2 days. If no response is given two additional 48-hr reminders will be sent. Research staff will attempt to contact the patient up to 3 times in one week.

### 10.1.9 DATA HANDLING AND RECORD KEEPING

#### 10.1.9.1 DATA COLLECTION AND MANAGEMENT RESPONSIBILITIES

Study staff will develop, test, and maintain the data capture system using a web-based data collection system, REDCap, as the primary source of data entry and storage. Developed by Vanderbilt University, REDCap is a software toolset and workflow methodology for electronic collection and management of research and clinical trial data. The REDCap system provides a secure, web-based application that is flexible and provides: 1) an intuitive interface for users to enter data and have real time validation rules (with automated data type and range checks) at the time of entry; 2) HIPAA compliant audit trails for tracking page views, data manipulation and export procedures; 3) record locking and electronic signature functions; 4) control over user rights to view and manipulate data; 5) a report builder capable of querying patient records; and 6) automated export procedures for SAS/STATA datasets.

The REDCap system complies with all applicable guidelines to ensure patient confidentiality, data integrity, and reliability. Quality assurance and control procedures will be applied to ensure the completeness, validity, and accuracy of the study database. The MOOP describes data collection processes, database development procedures, quality control processes, and reporting in greater detail.

EHR data will be retrieved by the Research Informatics and Scientific Computing Core (RISCC) at Cedars-Sinai, which facilitates interaction between the research community and the Epic data warehouse and production environment. We will collaborate with RISCC to design queries for portions of participants' records as defined in the data dictionary. The data extracts issuing from these queries will be delivered monthly as CSV files via Box, a HIPAA compliant cloud content management system. These will then be converted to SAS/STATA data sets and subjected to QC procedures in a SAS/STATA environment.

VR device metadata will be provided monthly by AppliedVR as CSV files containing timestamped records of content accessed at the individual device level. These records will feature a device-ID unique to each patient. A linking list containing study-IDs and device-IDs will be stored on a secure server behind the Cedars-Sinai firewall, and only associated with other study data following database lock. In order to monitor weekly adherence, study-staff will access a subset of the metadata limited to study-ID and timestamps of usage. These data will be subjected to QC procedures in a SAS/STATA environment.

Deidentified Fitbit data will be stored by a research-grade, IRB-approved service (Fitabase, San Diego, CA) as minute level data from Fitbit servers for all participants. This data will be retrieved as CSV files and linked to study data in a secure SAS/STATA environment before being subjected to QC procedures.

In order to maintain standard data definitions and structures, the data dictionary and proposed database structures reflect the Standard Clinical Data Interchange Standards Consortium's (CDISC) Study Data Tabulation Model (SDTM). The data elements and relationships described by this standard will facilitate information exchange with DAC. Furthermore, we will employ DAC's forthcoming approach to data cleanliness and completeness standards and notation.

#### 10.1.9.2 STUDY RECORDS RETENTION

In compliance with Protection of Human Subjects regulations, records related to the conduct of this trial, including but not limited to source documentation, informed consent forms, essential study documentation, and documentation of IRB activities, will be retained by the Investigator for a period of 3 years following the official close of the study.

### 10.1.10 PROTOCOL DEVIATIONS

A protocol deviation is any noncompliance with the clinical trial protocol. The noncompliance may be either on the part of the participant, the investigator, or the study site staff. As a result of deviations, corrective actions are to be developed by the site and implemented promptly.

These practices are consistent with ICH GCP:

- 4.5 Compliance with Protocol, sections 4.5.1, 4.5.2, and 4.5.3
- 5.1 Quality Assurance and Quality Control, section 5.1.1
- 5.20 Noncompliance, sections 5.20.1, and 5.20.2.

Protocol deviations/violations impacting participant safety will be reported to NIAMS and the DSMB Safety Officer through the NIAMS Executive Secretary within 48 hours of the investigator becoming aware of the event; all other deviations/violations that do not impact participant safety can be reported as part of the routine DSMB meeting report. The investigator will also report deviations within 5 working days of identification of the protocol deviation to the IRB. All deviations must be addressed in study source documents and reported to NIAMS. Protocol deviations must be sent to the reviewing Institutional Review Board (IRB) per their policies. The principal investigator is responsible for knowing and adhering to the IRB requirements.

### 10.1.11 PUBLICATION AND DATA SHARING POLICY

This study will be conducted in accordance with the following publication and data sharing policies and regulations:

This trial will be registered at ClinicalTrials.gov, and results information from this trial will be submitted to ClinicalTrials.gov. In addition, every attempt will be made to publish results in peer-reviewed journals. Data from this study may be requested from other researchers 3 years after the completion of the primary endpoint by contacting the Cedars-Sinai Center for Outcomes Research and Education.

We will also comply with the data sharing policy at NIH under NOT-OD-08-033. We will submit the final manuscripts to the NIH National Library of Medicine PubMed central for archiving upon acceptance for publication.

### 10.1.12 CONFLICT OF INTEREST POLICY

The independence of this study from any actual or perceived influence is critical. The Cedars-Sinai Identifying and Disclosing Researchers' Financial Interests in Research Policy: Human Research Protection Program requires the Principal Investigator and all members of the study team to disclose their, their spouses'/domestic partners', and dependent children's financial interests in the research regardless of the source of funding. Additional disclosure requirements may apply to investigators involved in Federally-funded research as described in the Cedars-Sinai Management of Industry Relations and Conflicts of Interest Program Policy: Corporate Integrity Program. The Cedars-Sinai Institutional Review Board (IRB) retains authority to determine if the research is allowed to proceed under the terms of the management plan developed by IR/COI considering the potential impact on the process and documentation of informed consent and the equitable selection of subjects. Review of disclosed COIs associated with a human research protocol by IR/COI is requested through submission of the CS-IRB COI Disclosure Form, which details IRB-supported guidelines on methods to manage, mitigate, or eliminate COIs.

## 10.2 ADDITIONAL CONSIDERATIONS

NA

# 11 APPENDIX A: RESEARCH SUPPLEMENT – ADVOCATING FOR UNDERREPRESENTED PATIENT POPULATIONS IN RESEARCH

## 11.1 BACKGROUND AND APPROACH

Black and Hispanic patients are up to five times less likely to have access to digital health information than non-Hispanic whites. In addition to disparities in access and adoption of digital health interventions, there are known disparities in the incidence and reporting of pain by racial and ethnic minorities. This Diversity, Inclusion and Engagement Supplement aims to develop a framework to advance diversity and inclusion efforts for future digital health trials at our medical center and beyond. Further, it will enhance the parent study by seeking to increase the proportion of participants with historically less access to and familiarity with digital technologies while enhancing overall participant racial and ethnic diversity.

### Approach

The scientific approach of this supplemental grant will be inspired by the **NIH Stage Model for Behavioral Intervention Development** (**Figure 1**).

**First**, we will conduct focus groups with non-Hispanic Black and Hispanic patients with cLBP to identify the most important content areas to be included in our two sets of culturally tailored study materials, described below. **Then**, we will employ insights from our patient partners to craft culturally tailored and engaging study materials and will then iteratively optimize the materials through cognitive de-briefing interviews, described below. **Finally**, we will assess whether implementation of culturally tailored study materials leads to a higher proportion of non-Hispanic Black and Hispanic participants when coupled with a novel study-participant cohort building tool, discussed below.

## 11.2 AIMS AND METHODOLOGY

**Aim 1: Tailor recruitment materials for non-Hispanic Black and Hispanic patients with cLBP in partnership with representative patient advisory boards (PABs).**

To achieve this aim, we will update current study materials, including informational flyers, consent forms, email templates, VR cognitive debriefing manuals, and instructions for use of assigned hardware and software to better reflect the knowledge, attitudes, beliefs, and preferences of non-Hispanic Black and Hispanic study participants. Our goal will be to create two separate sets of culturally tailored materials, one per target group, to supplement existing non-tailored materials. This will be accomplished through a set of representative PAB focus groups to generate themes and concepts for study materials, integration of PAB input to modify materials, and iterative optimization of the materials through one-on-one cognitive debriefing interviews.

***PAB Focus Groups:*** We will conduct foundational research with representative groups of non-Hispanic Black and Hispanic individuals with cLBP to inform development of culturally tailored study materials. We will convene 4 focus groups, including 2 groups each of non-Hispanic Black and Hispanic participants, each comprising 4-6 individuals for a total of 16-24 focus group participants. PAB members will be recruited through a combination of former study participants from the parent study, including both completers and participants who withdrew, along with additional participants naïve to the study identified through the Cedars-Sinai Patient Engagement Office and the Cedars-Sinai Medical Network. We will also seek participants from outside of Cedars-Sinai through patient advocacy groups as well as online and print ads. Given the high prevalence of cLBP in the population at large, the high prevalence of non-Hispanic Blacks and Hispanic populations in Southern California, and our successful track record of exceeding study recruitment goals to date, we do not anticipate challenges identifying up to 24 participants to join the PAB focus groups. In constructing the groups, we will seek to diversify the PABs across age, sex, education, and SES strata, where possible, to reflect different life experiences and perspectives within the target populations.

PAB focus groups will be held online using Zoom and will capture audio (not video) recordings of the proceedings after obtaining participant consent**.** Before the focus groups, we will develop a guide with participant instructions, open-ended think-aloud exercises, and scripted probes. Participants will also receive PDF or print copies of all current study materials prior to the focus groups (according to participant preference to receive materials by email or paper mail), and they will be asked to review the materials in advance of the meeting. A licensed neurosurgeon and/or PhD social scientist will moderate each focus group, with assistance from other members of study staff who will facilitate and take notes. Each focus group is expected to last approximately 90 minutes. Audio recordings will be transcribed by a third party (Keystrokes, Santa Monica, CA) for subsequent analysis by the study team.

Participants will initially be asked to react to the existing study materials in their own words and without prompting. The moderator will then follow scripted probes to expand on themes, concepts, and specific language or visuals suggested by the participants. Through group interaction, we will seek to identify unique relevant language with the goal of later incorporating the PAB input into modifying and culturally tailoring the materials, including both visual and textual elements. We will conduct multiple groups to ensure that interactions of a single group do not bias any one conclusion and to provide greater generalizability of our findings.

Upon completion of the focus groups, trained social scientists will qualitatively analyze the transcripts using ATLAS.ti software, including coding of patient language and classification of

vocabulary into major and minor concepts. The evaluation process will generate key words, phrases, and quotes to be used in developing tailored study materials for each group. To be considered credible, concepts for adoption will ideally be raised by more than one participant and by participants in more than one group, although individual decisions will be made using the discretion of the research team and with recognition that materials will undergo subsequent modifications based on the cognitive debriefings. We will use ATLAS.ti to generate code count histograms within major and minor concepts and will develop a network to depict a framework describing the breadth and depth of concepts to consider integrating into the study materials. The result will be a detailed "blueprint" for updating the study materials.

***Updating Study Materials:*** We will next develop and refine tailored study materials. After conducting the PAB focus groups, abstracting and organizing themes and concepts, and collecting suggestions for language and visuals, we will develop draft study materials for each targeted group: one set for non-Hispanic Blacks, and one for Hispanic participants, with cLBP. In developing the materials, we will aim to not exceed a sixth-grade reading level based on the validated "simple measure of gobbledygook" (SMOG) calculator, an approach we also followed in creating item banks for the NIH Patient Reported Outcome Measurement Information System (PROMIS). As with any study materials, we will also seek to minimize ambiguity or cognitive difficulty, avoid multi-barreled questions, and use language that is as concise and simply worded as possible while attempting to use common English words and avoiding slang unless considered appropriate or warranted by PAB members.

***Cognitive Debriefing Interviews:*** Once the study materials have been modified using input from the PAB focus groups, we will conduct patient cognitive debriefing interviews to assess for content validity, clarity, comprehensiveness and appropriateness of the resulting materials. We will prepare a scripted interview guide to elicit patient feedback on the draft materials. One-on-one interviews will be conducted via telephone with the same moderator who conducted the PAB meetings. For these individual interviews, we will sequentially recruit as many as 10 participants from each targeted group, drawing from PAB participants and the PAB waitlist, and will evaluate perceptions about the verbiage, images, and appearance of the draft materials. These interviews will not berecorded.. We will then iteratively update the materials until they are optimized, as evidenced by receiving fewer actionable suggestions over the course of the interviews (with modifications made after each interview). If we encounter a substantial number of actionable suggestions for improvement after 10 interviews, indicating lack of thematic saturation, then we will conduct up to 5 additional interviews, drawing again from PAB focus group participants. Through this process, we will fine-tune the materials and be prepared to deliver them as part of Aim 2 of the supplemental research, described below.

**Aim 2. Oversample non-Hispanic Black and Hispanic participants in parent study using a cohort building tool housed within the electronic health record (EHR).**

After developing two sets of culturally tailored study materials, we will next seek to deliver those materials to eligible study participants. This, in turn, will require us to identify enough non-Hispanic Black and Hispanic participants with cLBP to meet our target goals, described later in this section. To achieve this aim, we will employ Deep6 AI® (Pasadena, CA**)**, a natural language processing (NLP) cohort building tool, and enlist Cedars-Sinai's Enterprise Information Services (EIS), to oversample non-Hispanic Black and Hispanic participants from throughout Cedars-Sinai and our participating recruitment sites. Deep6 specifically integrates with our EHR to search both structured and unstructured data to build cohorts for clinical trials, and EIS will conduct SQL queries based on exclusionary criteria, race and ethnicity, and visits with recruiting physicians to capture additional participants not included in the Deep6 sample. Although Deep6 is already part of our protocol, it is not currently being used to specifically target study

participants by race or ethnicity. Further, it can only provide a maximum of 7,500 qualifying patients per pull. In this supplement, we will oversample target populations using Deep6 and EIS, then deliver the appropriate set of tailored recruitment materials developed in Aim 1 to these identified prospective participants. We will monitor outcomes by comparing actual recruitment against target oversampling goals, as discussed, below.

## 11.3 REVISED RECRUITMENT TARGETS

***Current and Revised Recruitment Targets:*** Currently, the target sample size for the parent study is 360 participants. Regarding Hispanic participants, although we are meeting our original recruitment goal which was set at 11% to reflect U.S. Census data, there is a much higher prevalence of Hispanic/Latinx individuals in Southern California where the 2019 Los Angeles County Census revealed a 48% prevalence of this group. **We intend to implement these new recruitment materials and strategies once we have randomized 50% of the overall study cohort.** Because certain participants are eligible for replacement if they are withdrawn after randomization, this will occur at approximately **participant #200** based on current rates of withdrawal**.** Specifically, beginning at this point, we will seek to recruit a second-half study sample that is **17% non-Hispanic black** (previous target 12%), **40% Hispanic** (previous target 11%), and **35% non-Hispanic white** (previous target 76%).

## 11.4 STRATEGIES TO ENSURE TARGETS ARE MET

***Monitoring Recruitment Progress:*** The research team will meet weekly during the study to monitor progress towards achieving target goals and will track progress using tables and visual dashboards. If the performance line falls below the target line for three consecutive weeks, then we will initiate a three-pronged strategy to course-correct as needed:

Strategy 1—Adjust Deep6 and EIS search algorithms and sample rates: If recruitment tracking reveals projections that we are falling short of target goals, then we will consult with Deep6 NLP and EIS specialists to fine-tune the parameters of the search algorithm. For example, if race/ethnicity fields are inaccurate or incomplete for some records, there may be other viable approaches such as NLP searches of clinician-inputted text within notes indicating race or ethnicity. Further, if we find specific groups are disproportionately unable to complete the screening process, we will further oversample the ratio of these participants contacted within the cohort builder lists to ensure enrollment rates meet our new target proportions.

Strategy 2—Interviews of non-enrolled patients: If necessary, we will seek consent to interview non-Hispanic Black and Hispanic patients who were invited to participate in the study, received the tailored recruitment materials, but nonetheless opted not to enroll in the trial. We will conduct individual cognitive debriefing interviews using the same techniques described under Aim 2, but here focus on individuals who experienced the recruitment materials first-hand. We will update recruitment materials further if these interviews elicit specific and actionable feedback that is incremental to the Aim 1 results.

Strategy 3—Interviews of patients who withdrew from the study: In addition to interviewing patients who declined to enroll, we will also seek consent to interview non-Hispanic Black and Hispanic participants who enrolled in the study but later dropped out before completion despite receiving tailored post-enrolment materials. These study participants may have a different

perspective from those who never enrolled and will also have experienced a broader range of study materials beyond the recruitment flyer and consent form (e.g., surveys, VR debriefing manual, instructions for hardware/software usage). We will again use the same cognitive interviews techniques previously described, and will further modify study materials, as needed, based on these additional de-briefing interviews.

# 12 APPENDIX B: RESEARCH SUPPLEMENT – UNDERSTANDING PATIENT PREDICTORS OF RESPONSE (3UH3AR076573-03S2)

## 12.1 BACKGROUND AND APPROACH

Although digital health technologies are now widely available for both therapeutic and monitoring applications, there are wide variations in patient knowledge, attitudes, beliefs and preferences regarding the uptake and effectiveness of digital health interventions. In addition, there are sociodemographic variations in willingness to engage in digital health studies, both for chronic pain and other common disorders. A recent study published in the Journal of Racial and Ethnic Health Disparities in February of 2021 found that older, less educated, and economically disadvantaged Black and Hispanic patients were up to five times less likely to have access to digital health information. Additionally, a 2016 study exploring the trends of uptake by Medicare beneficiaries found that digital health information has not yet fully reached the elderly. There have been few efforts to systematically examine patient-level predictors of digital health uptake and benefit among diverse patients with chronic pain.

This substudy will identify specific patient characteristics that influence enrollment in and completion of the parent study by employing mixed methods to examine variations in demographics, engagement, and benefit among diverse participants. The project will also offer mentorship, research training, and authorship experience to support early career development for Lindsey Ross, MD, MHDS.

## 12.2 AIMS AND METHODOLOGY

**Aim 1**: **To identify explicit and latent variables associated with membership in subgroups of participants in the parent RCT of VR therapy for cLBP.**

The Digital Divide (DD) can be explained as separating patients into overarching subgroups or "classes" – those who enroll in and complete studies like the parent VR study and subsequently use the technology in their daily lives to manage pain, and those who do not. Our hypothesis is twofold: first, that there are patient-level, explicitly measured variables associated with enrollment in and successful completion of these types of studies (Hypothesis$_1$). Second, the DD in study completeness will be further explained by latent (unobservable) "classes" of patients (Hypothesis$_2$); membership in these latent classes is predicated upon unobservable or latent patient-level variables. Analyzing the explicit variables identified in $H_1$ in conjunction with the identified latent classes, we hope to gain a better understanding of what explicit measures or combinations of measures may be missing or need further consideration for inclusion in studies testing VR for cLBP, or other related technologies or outcomes.

To test $H_1$, we will identify important patient-level (measured) variables associated with enrollment in and the successful completion of the parent RCT, as well as those associated with the decision not to enroll. Specifically, we will implement a two-part regression model. The first part with a logistic

regression analysis modeling a binary outcome of enrollment, then the second part with a generalized linear regression for modeling study completeness.

To test $H_2$, explicit variables including demographic variables and patient reported outcomes will be used in a latent class analysis (LCA) aimed at determining whether there are latent classes of patients and, if so, what explicit measures may be associated with class membership. LCA uses underlying patterns in the data (e.g., unobserved characteristics often associated with knowledge, attitudes, and beliefs) to identify latent classes (unidentified subgroups) of patients. Latent classes help guide qualitative investigation of surrounding phenomenology and determination of explicit variables that predict membership in a latent class.

**Aim 2: Explicate patient experiences on both sides of the Digital Divide (DD), including understanding how they arrived on a given side, by describing phenomenology surrounding enrollment in and successful completion of an RCT studying VR as treatment for cLBP.**

Semi-structured focus groups will be conducted separately among classes on either side of the DD as well as among the intent-to-treat sample (those randomized) and participants who chose not to enroll or were otherwise withdrawn or discontinued prior to randomization to investigate and attempt to identify the underlying constructs perpetuating membership in one vs. the other groups of patients. We will investigate the influence of knowledge, attitudes, and beliefs on the use of novel technology such as VR and wearable biosensors (as the parent study uses Fitbit) in the management of cLBP, ascertain difficulties and limitations with the use of digital technology, and explore areas for improvement in designing trials of this kind. Separately, from among the listed groups of participants, we aim to recruit N=24 (4 focus groups of up to 6 members each). Focus group data will include and qualitative data elicited through discussions characterizing experiences of using VR and describing how these factors have or have not influenced patients to preclude, decrease, continue, or increase use.

Participants who were contacted for participation in the parent study will first be approached by email recruitment letter, then a phone call, for participation in this study. Participants expressing interest will be contacted again with a list of possible dates and times to participate. Prior to the scheduled focus group, participants will also be emailed an information sheet describing the risks and benefits of participating in the sub-study, and at the start of the focus group session, present participants will be asked to verbally confirm that they have read and understood the information sheet. Participants will be reimbursed with a $100 Amazon e-gift card for their participation once they have completed the focus group, which should take no longer than 60 minutes.

## 12.3 CAREER DEVELOPMENT PLAN

The supplemental funding mechanism included a career development and mentoring plan, centered around holistic growth as a future physician scientist. The primary goals for Dr. Ross' career development plan include (1) learning to independently lead and coordinate a scientific research project; (2) communicate and manage a large research team; and (3) produce high-quality research manuscripts and presentations of findings on national platforms.

The research timeline, presented below, includes the primary goals of improving scientific research knowledge and production of manuscripts and long-term goals of ultimately completing a K08 grant application by 2024.

**Research Timeline**

| Year | Year 1 | Year 2 | Year 3 | Year 4 |
|---|---|---|---|---|
| Skills | Data management and methods skill building | Quantitative/Qualitative analysis methods | Manuscript production | Presentations and future funding |
| Aims | DEI supplemental qualitative and quantitative methods and manuscript production<br><br>Preparation for data analysis to start at completion of parent study | Aim 1: Initiate semi-structured interviews<br><br>Aim 2: Complete data collection and analysis | Complete manuscript of supplemental data<br><br>Apply for additional funding with preliminary data | Complete manuscript assistance for parent data grant<br><br>Submit K08 with supplemental data |
| Mentorship | Dr. Brennan Spiegel | | | |

# 13 APPENDIX C: RESEARCH SUPPLEMENT – INVESTIGATING THE BARRIERS TO IMPLEMENTATION AND THE APPROPRIATENESS OF VR CONTENT ACROSS DIVERSE CULTURES: A RETROSPECTIVE MIXED METHODS STUDY

## 13.1 BACKGROUND AND APPROACH

Digital therapeutics such as virtual reality (VR) technologies have the potential to expand our approach to chronic pain management. While this novel modality continues to be investigated, there are discrepancies across individuals of differing sociodemographic backgrounds in their engagement and use of VR. A recent study published in the *Journal of Racial and Ethnic Health Disparities* found that Black and Hispanic patients were up to five times less likely to have access to digital health information. Another recently published article in *The Lancet* stressed the need for designing digital health applications to meet the needs of women – especially with racial or ethnic minority backgrounds. Other studies reported similar differential uptake, interaction, and use of virtual reality based on geography and age. Beyond general digital medicine, virtual reality as an end consumer product has been found to be subject to similar disparities in uptake and use across multiple sociodemographic factors.

Understanding barriers to uptake of novel therapeutic technologies is essential for formulating solutions that can establish equity and inclusiveness when deploying VR in healthcare to a diverse population. Failure to address these sociodemographic disparities can result in disadvantaged groups being unable to optimally utilize alternative non-opioid digital treatments and thus may require more traditional treatments including opioids. To achieve equity, we must verify that the digital therapeutic content is appropriate and accessible across diverse cultures—an investigation that is instrumentally important for the development of future VR applications. To that end, in this supplemental, mixed methods study we seek to correlate PRO data collected in the parent study with newly collected and subsequently quantized qualitative data acquired from individual participant interviews.

### Approach

An NIH commissioned report on a “best practices” approach of mixed methods research (MMR) described MMR as “focusing on research questions that call for real-life contextual understandings, multi-level perspectives, and cultural influences…It is a systematic and rigorous form of inquiry that uses

methods of data collection such as in-depth interviews, ethnographic observation, and review of documents."

In this supplemental study, we will conduct MMR employing NIH best practice methodology to better understand the barriers to VR implementation and the appropriateness of VR content across diverse cultures. This process will follow an embedded mixed methods design (QUAL(quan)) in which we will sample up to 35 individuals who have completed their participation in the skills-based VR therapy treatment arm of either the parent study (STUDY00000631) or another concurrently running study that employs the same intervention in its treatment arm (STUDY00001262). We plan to sample from this additional source (STUDY00001262) to ensure the maximal sampling of the targeted subgroups as identified by this supplement.

Potential subjects will be identified and considered eligible for participation in this study based on their current timepoint of participation within the respective study they are currently enrolled in (STUDY00001262 or STUDY00000631). For the parent study (STUDY00000631), this is defined as any participant in the active treatment arm that has completed their day 60 PROMIS Pain Interference survey. For the concurrently running study (STUDY00001262), this is defined as any participant in the active treatment arm that has completed their week 8 PROMIS Pain Interference survey. The rationale for these criteria is the time-equivalency of data points, as this measure shall be used in further analyses within this supplemental study. The other criterion for eligibility is that the individual is 18 years or older. All participants that meet these criteria will be considered for participation in this sub-study.

All identified potential participants will be contacted via phone calls (see phone script in local files). Within this call, subjects will be informed on the goals and methods of this study and their anonymity and compensation should they agree to participate. Finally, they will be informed on the possibility of being assigned to an AI-powered interviewer and the data protection measures that have been emplaced for such a modality. At the end of the call, verbal consent will be obtained from the participant and said consent will be documented within research notes. An interview time and date will be scheduled with the participant and a follow-up confirmation email will be sent to them directly after the call that has the information sheet for this study attached. Two days before their scheduled interview, participants will be sent a reminder email that restates the date and time of their interview and includes the link to the Zoom call on which the interview will be held.

These individuals will subsequently be randomized into two groups to participate in a one-on-one (1:1) semi-structured interview exploring their beliefs and experiences of the VR intervention during their time in the parent study. Considering the subject and focus of this interview, no PHI is intended to be asked of the participants nor recorded for the purposes of this study. The first group will have standard 1:1 semi-structured interviews led by a trained social scientist. Participants in the second group will also have 1:1 semi-structured interviews, but these interviews will be led by an artificial intelligence (AI)-powered virtual social scientist. This AI interviewer is a version of GPT4 that has been iteratively refined by the study team to follow best practices in conducting 1:1 semi-structured interviews through prompt parameters. The refined prompt parameters and set interview questions will be automatically given to GPT-4 upon each initialization of the GPT4-powered AI interviewer. These prompt parameters and interview guide are set or "locked", as it is not intended for the purposes of this investigation to implement revisions or improvements of the AI interviewer based on the data received during any of the interviews. Regarding collected data, all information collected during the interview (e.g., interview transcripts) and all data uploads/downloads between the interview software and GPT4 will be encrypted, transmitted, and stored in Microsoft Azure services with whom we have an executed business agreement that all data storage will meet or exceed HIPAA requirements.

Transcripts created by this process will be coded and analyzed independently by both AI and human social scientists. This analysis will follow a general, open-coding approach wherein first level codes will be created inductively from the data with refined second-level codes created from the first level. Results yielded from these analyses will be compared qualitatively to assess the thematic coverage, breadth and overlap between the results generated by the AI and human social scientists.

A later analysis will use only the second-level codes generated by the human social scientists. The human-generated second-level codes will be quantized by calculating their frequencies of occurrence in each transcript. These code frequencies will be used to identify qualitative factors that may be correlated with the participants' PRO data from the parent study. Additionally, we will compare how these significant factors and the overarching qualitative themes differ across the demographic groups of interest. Successful completion of this supplemental study will guide future VR development and research to be more inclusive of and sensitive to diverse racial/ethnic, socioeconomic, geographic, gender, and age populations. Finally, we will evaluate the feasibility of using an AI-powered virtual social scientist to both collect and analyze qualitative data.

## 13.2 AIMS AND METHODOLOGY

**Aim 1: Understand any barriers to implementing therapeutic VR technologies across diverse populations and learn how cultural differences influence perception of the VR content.**

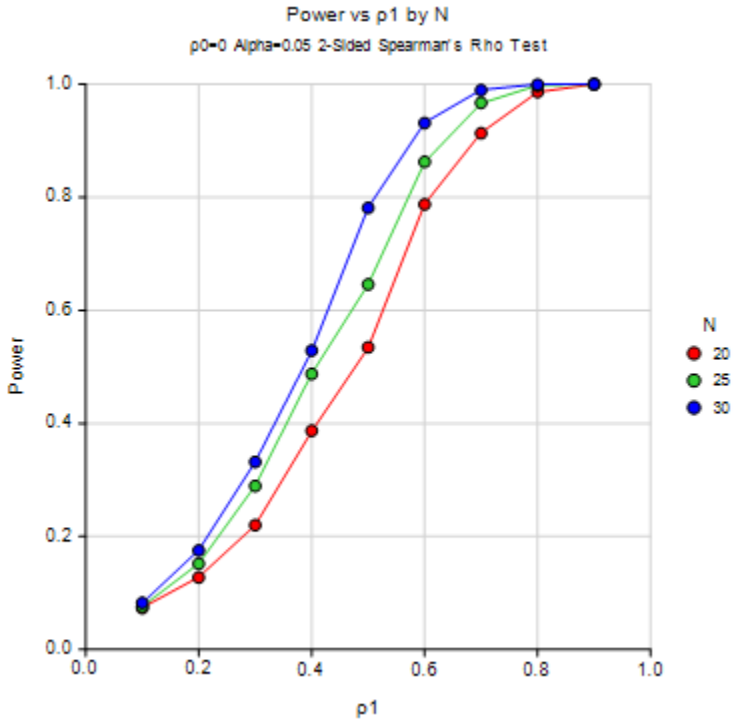
To achieve this aim, we will interview a subset of participants by employing purposive sampling; participants who were enrolled in the skills-based VR arm of the parent study will be chosen based on criteria to maximize diversity across race/ethnicity, socioeconomic brackets, geography (based on zip-code: rural and non-rural), gender, and age. Interviews will be performed via Zoom video conferencing platform, and only the audio (not video) of the interviews will be captured. Interviews will be semi-structured to accommodate open, constructive dialogue. Prior to the first interview, we will develop a guide for interviewers that includes participant instructions, open-ended questions, and scripted probes to expand on themes. Interview questions will be tailored to target what barriers interviewees experienced while using VR and how they perceived the VR content in terms of their own cultural framework. Each interview is expected to last no more than 60 minutes. The audio recordings will be transcribed for subsequent analyses. Each subgroup will attempt to include at least five participants; to accommodate this diversity, at least 30 subjects will be recruited. Recruitment will continue until 1) at least 30 subjects have been interviewed and 2) thematic saturation has been established as defined by 3 consecutive interviews with no new emergent themes.

Upon completion of the interviews, trained social scientists will perform a general qualitative analysis via the open coding of the transcripts using ATLAS.ti software, including coding of participant language and classification of vocabulary into major and minor concepts. The evaluation process will generate recurring key words, phrases, and quotes, which can further be divided and compared among the subgroups. All participants will be assigned to five subgroups: race/ethnicity, socioeconomic bracket, geography, gender, and age. For each of these subgroups, we will use ATLAS.ti to generate code count histograms within major and minor concepts and will develop a network to depict a framework describing the breadth and depth of concepts. The result will identify differences among the subgroups and provide feedback for developing future, more inclusive VR content.

**Aim 2. Explore the relationships of how diverse groups differentially use VR technologies and how PROs correlate with quantized qualitative data.**

To achieve this aim, we will quantize the qualitative data derived from interview transcripts by calculating the frequency of qualitative code reoccurrences within each participants' interview transcripts. Then, we will correlate the entire sample's quantized count data with their corresponding PRO scores as recorded in the parent study. To account for the non-normal, discrete nature of the count data, we will calculate Spearman's rho to assess the strength and directionality of the relationship between the constructed codes and the PRO scores (see Figure 1 below for power justification). Codes that are found to have a significant correlation with the PRO scores will be compared between the demographic subgroups of interest. This shall be performed through a combination of visual analysis and Kruskal-Wallis tests (when sample size allows).

**Figure 1: Power Estimate for Spearman Correlation at n = 30**




Finally, we will complete the qualitative analysis process by constructing overarching cross-group themes that illustrate the views, opinions, and experiences of utilizing the skills-based VR intervention. This process will also include an exploration of how both the significantly correlated qualitative codes and the cross-group themes differ in context and content across the demographic subgroups of interest.

By following this process, we hope to identify demographic-specific factors that both promote and inhibit the therapeutic efficacy of skills-based VR pain management. Awareness of these relationships will be necessary for improving the equity and inclusiveness of future VR therapeutics.

**Aim 3. Investigate the feasibility of AI to conduct 1:1 semi-structured interviews.**

To achieve this aim, participants will be randomly assigned to one of two groups that are differentiated by the type of interview they will receive. In one group, participants will have a 1:1 semi-structured interview via Zoom with a human social scientist that is trained in qualitative interview approaches (A.C.). The other group shall undergo the same process but with a virtual AI-powered interviewer whose avatar and voice will come across through Zoom. Both groups will be interviewed using the same set of probe questions for a maximum time per interview of 60 minutes. Only the audio from the interviews will be recorded and subsequently transcribed. Upon completion of the interviews and subsequent transcription process, all the transcripts will undergo qualitative data analysis by both AI and trained human social scientists independently. As such, this qualitative data analysis is expected to yield two sets of results that are respective to the analyst type (AI and human). Following completion of the interview, participants will be asked to complete a 10-item “Interviewer Sentiment Scale” questionnaire, delivered through REDCap. These results will be compared qualitatively for differences in thematic breadth and depth as well as the comparison of inductively created codes between the two results sets.

### 13.3 STRATEGIES TO ENSURE TARGETS ARE MET

***Monitoring Recruitment Progress:*** The research team will meet weekly for the duration of the study to monitor progress towards achieving target goals and will track progress using tables and visual dashboards. During these meetings, potential interviewees will be identified based on their subgroup profile. The meeting will also include discussions on addressing any unexpected barriers to participant recruitment. Should we have difficulties in recruiting participants, we will contact and attempt to recruit patients who withdrew from their respective study. These study participants may have a different perspective from those who completed their participation in the study and will also have views, opinions, and rationales as to why they chose to discontinue their role in the study. We will again use the same 1:1 interview techniques previously described and will include the data collected from these participants' transcripts within the main qualitative and subsequent mixed-methods analysis as described above. To motivate participation, we will offer reimbursement with a $100 Amazon e-gift card for completion of the interview and follow-up questionnaire.

## 14 PROTOCOL AMENDMENT HISTORY

| Version | Date | Description of Change | Brief Rationale |
|---|---|---|---|
| 2 | Revisions 6/8/20 | General Updates | Clarifications and consistency with other study documents. |
| 3.1 | 8/31/20 | Necessary updates to sample size estimation & randomization. Additional edits incorporated from DSMB recommendations. | Revisions recommended from DSMB. |
| 3.2 | 9/16/20 | Charge 2 replaced with Charge 4, increased interval of review of race/ethnicity, updated description of the VR program. | Changes made to reflect up-to-date technology and update feasibility of protocol. |
| 4 | 10/23/20, 11/10/20 | Necessary updates to SoA table, included additional sub-groups to sub-group analysis, updated objectives section, removed AppliedVR tablet from data collection for will not be available. Revisions to patient compensation, information added regarding CURES application. | Changes were made to harmonize with the SAP and up-to-date information from the virtual reality vendor. |
| 5 | 5/20/21 | Edits made to update AE assessment frequency, minor edits. | Minor edits. |
| 6 | 11/5/21 | Revised loss to follow-up definition, added info regarding AppliedVR VR usage data.<br>Edits made to diversity benchmarks and addition of DIE supplement. | Overhaul in study second half recruitment strategies require updating parent study protocol; too much harmonization between the two to necessitate brand new IRB project. |
| 6.1 | 4/25/22 | Recruitment objectives modified to incorporate treating physician letter. Diversity supplement interview methodology updated to reflect that we | Protocol reflects procedures and new hybrid recruiting strategy. |

| | | | |
|---|---|---|---|
| | | did not record interviews and we drew from PAB waitlist. | |
| 6.2 | 7/11/2022 | Depression alert procedures updated to include actions to take when telephone assessment cannot be made; approved by DSMB. Neck pain and rash added as a known risk, borrowing approved language from STUDY00001363. Table of contents page numbers updated. | With DSMB guidance, created a contingency plan in the event that severely depressed individuals are lost to follow-up. List of known potential AEs updated to reflect oversight of known neck pain risk and to incorporate the incidence of allergy to face foam used in VR goggles. |
| 6.3 | 10/24/2022 | Onboarding procedures modified to take into account new instructional videos. Compliance procedures clarified and modified to account for optional SMS text messaging via REDCap. Remove 7 day wait after emailing recruitment letter.<br>Removal of procedure for mailing study flyer. | Onboarding procedures now incorporate clear video instructions, and compliance monitoring updated to reflect current streamlined practices. General housekeeping on procedures to align with other VR trials. |
| 6.4 | 7/14/2023 | Modified definition of replaceable participant to exclude individuals who used VR intervention before withdrawing from study. | Study adequately powered based on withdrawals to date, and analyses should include individuals who found the intervention unsatisfactory and subsequently withdrew. |
| 7.0 | 8/29/2023 | Clarified confusing language in replaceable participant definition from previous mod.<br>Addition of Appendix B: Digital Divide Supplement. | Typos in language confused our new definition of a replaceable participant.<br>Second sub-study added to appendix prior to initiation of new methods. |
| 8.0 | 9/12/2023 | Addition of Appendix C: Cultural Barriers to VR Supplement. | Third sub-study added to appendix. |
| 8.1 | 10/17/23 | Addition of post-interview quantitative “Interviewer Sentiment Scale” to Appendix C | Directly capture effectiveness of AI interviewer |
| | | | |
| | | | |

## 15 REFERENCES

1. Witmer BG, Singer MJ. Measuring presence in virtual environments: A presence questionnaire. *Presence.* 1998;7(3):225-240.

2. Vinson NG, Lapointe J-F, Parush A, Roberts S. Cybersickness induced by desktop virtual reality. In: *Proceedings of Graphics Interface 2012.*2012:69-75.

3. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology.* 1993;3(3):203-220.

4. Cella D, Choi SW, Condon DM, et al. PROMIS® adult health profiles: efficient short-form measures of seven health domains. *Value in health.* 2019;22(5):537-544.

5. Krebs EE, Lorenz KA, Bair MJ, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *Journal of general internal medicine.* 2009;24(6):733-738.

6. George SZ, Lentz TA, Zeppieri Jr G, Lee D, Chmielewski TL. Analysis of shortened versions of the Tampa Scale for Kinesiophobia and Pain Catastrophizing Scale for patients following anterior cruciate ligament reconstruction. *The Clinical journal of pain.* 2012;28(1):73.

7. Sundararajan V, Henderson T, Perry C, Muggivan A, Quan H, Ghali WA. New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality. *Journal of clinical epidemiology.* 2004;57(12):1288-1294.

8. Hurwitz EL, Randhawa K, Yu H, Côté P, Haldeman S. The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. *European Spine Journal.* 2018;27(6):796-801.

9. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet (London, England).* 2012;380(9859):2197-2223.

10. Kennedy J, Roll JM, Schraudner T, Murphy S, McPherson S. Prevalence of persistent pain in the U.S. adult population: new data from the 2010 national health interview survey. *The journal of pain : official journal of the American Pain Society.* 2014;15(10):979-984.

11. Medicine Io. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research.* Washington, DC: The National Academies Press; 2011.

12. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time and cost due to common pain conditions in the US workforce. *Jama.* 2003;290(18):2443-2454.

13. Deyo RA, Mirza SK, Turner JA, Martin BI. Overtreating chronic back pain: time to back off? *Journal of the American Board of Family Medicine : JABFM.* 2009;22(1):62-68.

14. Shmagel A, Ngo L, Ensrud K, Foley R. Prescription Medication Use Among Community-Based U.S. Adults With Chronic Low Back Pain: A Cross-Sectional Population Based Study. *The journal of pain : official journal of the American Pain Society.* 2018;19(10):1104-1112.

15. Crofford LJ. Adverse effects of chronic opioid therapy for chronic musculoskeletal pain. *Nature Reviews Rheumatology.* 2010;6(4):191.

16. Wiederhold BK, Gao K, Sulea C, Wiederhold MDJC, Behavior,, Networking S. Virtual reality as a distraction technique in chronic pain patients. 2014;17(6):346-352.

17. Scheffler M, Koranyi S, Meissner W, Strauss B, Rosendahl JJB. Efficacy of non-pharmacological interventions for procedural pain relief in adults undergoing burn wound care: a systematic review and meta-analysis of randomized controlled trials. 2018;44(7):1709-1720.

18. Chan E, Foster S, Sambell R, Leong PJPo. Clinical efficacy of virtual reality for acute procedural pain management: a systematic review and meta-analysis. 2018;13(7).

19. Riva G, Wiederhold BK, Mantovani FJC, Behavior,, Networking S. Neuroscience of virtual reality: From virtual exposure to embodied medicine. 2019;22(1):82-96.

20. Luo H, Cao C, Zhong J, Chen J, Cen YJWR, Regeneration. Adjunctive virtual reality for procedural pain management of burn patients during dressing change or physical therapy: A systematic review and meta-analysis of randomized controlled trials. 2019;27(1):90-101.

21. Dascal J, Reid M, IsHak WW, et al. Virtual reality and medical inpatients: a systematic review of randomized, controlled trials. 2017;14(1-2):14.

22. Indovina P, Barone D, Gallo L, Chirico A, De Pietro G, Giordano AJTCjop. Virtual reality as a distraction intervention to relieve pain and distress during medical procedures. 2018;34(9):858-877.

23. Garrett B, Taverner T, McDade PJJmi. Virtual reality as an adjunct home therapy in chronic pain management: an exploratory study. 2017;5(2):e11.

24. Hoffman HG, Doctor JN, Patterson DR, Carrougher GJ, Furness III TAJP. Virtual reality as an adjunctive pain control during burn wound care in adolescent patients. 2000;85(1-2):305-309.

25. Benham S, Kang M, Grampurohit NJOo, participation, health. Immersive virtual reality for the management of pain in community-dwelling older adults. 2019;39(2):90-96.

26. Carrougher GJ, Hoffman HG, Nakamura D, et al. The effect of virtual reality on pain and range of motion in adults with burn injuries. 2009;30(5):785-791.

27. Furman E, Jasinevicius TR, Bissada NF, Victoroff KZ, Skillicorn R, Buchner MJTJotADA. Virtual reality distraction for pain control during periodontal scaling and root planing procedures. 2009;140(12):1508-1516.

28. Garrett B, Taverner T, Masinde W, Gromala D, Shaw C, Negraeff MJTCjop. A rapid evidence assessment of immersive virtual reality as an adjunct therapy in acute pain management in clinical practice. 2014;30(12):1089-1098.

29. Gold JI, Kim SH, Kant AJ, Joseph MH, Rizzo ASJC, Behavior. Effectiveness of virtual reality for pediatric pain distraction during IV placement. 2006;9(2):207-212.

30. Malloy KM, Milling LSJCpr. The effectiveness of virtual reality distraction for pain reduction: a systematic review. 2010;30(8):1011-1018.

31. Hoffman HG, Patterson DR, Carrougher GJJTCjop. Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: a controlled study. 2000;16(3):244-250.

32. Sato K, Fukumori S, Matsusaki T, et al. Nonimmersive virtual reality mirror visual feedback therapy and its application for the treatment of complex regional pain syndrome: an open-label pilot study. 2010;11(4):622-629.

33. Jones T, Moore T, Choo JJPo. The impact of virtual reality on chronic pain. 2016;11(12).

34. Jin W, Choo A, Gromala D, Shaw C, Squire P. A Virtual Reality Game for Chronic Pain Management: A Randomized, Controlled Clinical Study. Paper presented at: MMVR2016.

35. Sarig-Bahat H, Weiss PLT, Laufer YJS. Neck pain assessment in a virtual environment. 2010;35(4):E105-E112.

36. Pozeg P, Palluel E, Ronchi R, et al. Virtual reality improves embodiment and neuropathic pain caused by spinal cord injury. 2017;89(18):1894-1903.

37. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski SJTJoP. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. 2016;17(12):1302-1317.

38. Tashjian VC, Mosadeghi S, Howard AR, et al. Virtual reality for management of pain in hospitalized patients: results of a controlled trial. 2017;4(1):e9.

39. Li A, Montaño Z, Chen VJ, Gold JIJPm. Virtual reality and pain management: current trends and future directions. 2011;1(2):147-157.

40. McCaul KD, Malott JMJPb. Distraction and coping with pain. 1984;95(3):516.

41. Katz J, Rosenbloom BNJPR, Management. The golden anniversary of Melzack and Wall's gate control theory of pain: Celebrating 50 years of pain research and management. 2015;20(6):285-286.

42. Gold JI, Belmont KA, Thomas DAJC, Behavior. The neurobiology of virtual reality pain attenuation. 2007;10(4):536-544.

43. Keefe FJ, Huling DA, Coggins MJ, et al. Virtual reality for persistent pain: a new direction for behavioral pain management. 2012;153(11):2163.

44. Hoffman HG, Chambers GT, Meyer III WJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. 2011;41(2):183-191.

45. JahaniShoorab N, Zagami SE, Nahvi A, et al. The effect of virtual reality on pain in primiparity women during episiotomy repair: a randomize clinical trial. 2015;40(3):219.

46. Frey DP, Bauer ME, Bell CL, et al. Virtual reality analgesia in labor: the VRAIL pilot study—a preliminary randomized controlled trial suggesting benefit of immersive virtual reality analgesia in unmedicated laboring women. 2019;128(6):e93-e96.

47. Chirico A, Lucidi F, De Laurentiis M, Milanese C, Napoli A, Giordano AJJocp. Virtual reality in health system: beyond entertainment. a mini-review on the efficacy of VR during cancer treatment. 2016;231(2):275-287.

48. Hoffman HG, Richards TL, Van Oostrom T, et al. The analgesic effects of opioids and immersive virtual reality distraction: evidence from subjective and functional brain imaging assessments. 2007;105(6):1776-1783.

49. Hoffman HG, Richards TL, Bills AR, et al. Using fMRI to study the neural correlates of virtual reality analgesia. 2006;11(1):45-51.

50. Hoffman HG, Garcia-Palacios A, Patterson DR, et al. The effectiveness of virtual reality for dental pain control: a case study. 2001;4(4):527-535.

51. Hoffman HG, Patterson DR, Carrougher GJ, Sharar SRJTCjop. Effectiveness of virtual reality–based pain control with multiple treatments. 2001;17(3):229-235.

52. Rizzo AA, Difede J, Rothbaum BO, et al. VR PTSD exposure therapy results with active duty OIF/OEF combatants. Paper presented at: MMVR2009.

53. Maples-Keller JL, Bunnell BE, Kim S-J, Rothbaum BOJHrop. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. 2017;25(3):103.

54. Maples-Keller JL, Price M, Rauch S, Gerardi M, Rothbaum BOJBt. Investigating relationships between PTSD symptom clusters within virtual reality exposure therapy for OEF/OIF veterans. 2017;48(2):147-155.

55. Maples-Keller JL, Yasinski C, Manjin N, Rothbaum BOJN. Virtual reality-enhanced extinction of phobias and post-traumatic stress. 2017;14(3):554-563.

56. Bordnick PS, Graap KM, Copp HL, Brooks J, Ferrer MJC, Behavior. Virtual reality cue reactivity assessment in cigarette smokers. 2005;8(5):487-492.

57. Bordnick PS, Traylor A, Copp HL, et al. Assessing reactivity to virtual reality alcohol based cues. 2008;33(6):743-756.

58. Bordnick PS, Traylor AC, Carter BL, Graap KMJRoswp. A feasibility study of virtual reality-based coping skills training for nicotine dependence. 2012;22(3):293-300.

59. Maskey M, Rodgers J, Grahame V, et al. A Randomised Controlled Feasibility Trial of Immersive Virtual Reality Treatment with Cognitive Behaviour Therapy for Specific Phobias in Young People with Autism Spectrum Disorder. *Journal of autism and developmental disorders.* 2019;49(5):1912-1927.

60. Gujjar KR, van Wijk A, Kumar R, de Jongh A. Efficacy of virtual reality exposure therapy for the treatment of dental phobia in adults: A randomized controlled trial. *Journal of anxiety disorders.* 2019;62:100-108.

61. Rahani VK, Vard A, Najafi M. Claustrophobia Game: Design and Development of a New Virtual Reality Game for Treatment of Claustrophobia. *J Med Signals Sens.* 2018;8(4):231-237.

62. Minns S, Levihn-Coon A, Carl E, et al. Immersive 3D exposure-based treatment for spider fear: A randomized controlled trial. *Journal of anxiety disorders.* 2019;61:37-44.

63. Tardif N, Therrien CE, Bouchard S. Re-Examining Psychological Mechanisms Underlying Virtual Reality-Based Exposure for Spider Phobia. *Cyberpsychol Behav Soc Netw.* 2019;22(1):39-45.

64. Hong YJ, Kim HE, Jung YH, Kyeong S, Kim JJ. Usefulness of the Mobile Virtual Reality Self-Training for Overcoming a Fear of Heights. *Cyberpsychol Behav Soc Netw.* 2017;20(12):753-761.

65. Botella C, Fernandez-Alvarez J, Guillen V, Garcia-Palacios A, Banos R. Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review. *Curr Psychiatry Rep.* 2017;19(7):42.

66. Oing T, Prescott J. Implementations of Virtual Reality for Anxiety-Related Disorders: Systematic Review. *JMIR Serious Games.* 2018;6(4):e10965.

67. Chad R, Emaan S, Jillian O. Effect of virtual reality headset for pediatric fear and pain distraction during immunization. *Pain Manag.* 2018;8(3):175-179.

68. Navarro-Haro MV, Modrego-Alarcon M, Hoffman HG, et al. Evaluation of a Mindfulness-Based Intervention With and Without Virtual Reality Dialectical Behavior Therapy((R)) Mindfulness Skills Training for the Treatment of Generalized Anxiety Disorder in Primary Care: A Pilot Study. *Front Psychol.* 2019;10:55.

69. Niki K, Okamoto Y, Maeda I, et al. A Novel Palliative Care Approach Using Virtual Reality for Improving Various Symptoms of Terminal Cancer Patients: A Preliminary Prospective, Multicenter Study. *J Palliat Med.* 2019;22(6):702-707.

70. Atkinson AB, Kennedy AL, Carson DJ, Hadden DR, Weaver JA, Sheridan B. Five cases of cyclical Cushing's syndrome. *Br Med J (Clin Res Ed).* 1985;291(6507):1453-1457.

71. Hufford MR, Shiffman S. Methodological issues affecting the value of patient-reported outcomes data. *Expert Rev Pharmacoecon Outcomes Res.* 2002;2(2):119-128.

72. Gavgani AM, Nesbitt KV, Blackmore KL, Nalivaiko E. Profiling subjective symptoms and autonomic changes associated with cybersickness. *Auton Neurosci.* 2017;203:41-50.

73. Kim YY, Kim HJ, Kim EN, Ko HD, Kim HT. Characteristic changes in the physiological components of cybersickness. *Psychophysiology.* 2005;42(5):616-625.

74. Dużmańska N, Strojny P, Strojny A. Can Simulator Sickness Be Avoided? A Review on Temporal Aspects of Simulator Sickness. *Front Psychol.* 2018;9:2132.

75. LaViola Jr JJJASB. A discussion of cybersickness in virtual environments. 2000;32(1):47-56.

76. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-Administered Skills-Based Virtual Reality Intervention for Chronic Pain: A Randomized Controlled Pilot Study. *JMIR Form Res.* 2020.

77. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on research standards for chronic low back pain. *Phys Ther.* 2015;95(2):e1-e18.

78. Amtmann D, Kim J, Chung H, Askew RL, Park R, Cook KFJJopr. Minimally important differences for Patient Reported Outcomes Measurement Information System pain interference for individuals with back pain. 2016;9:251.

79. Licciardone J, Worzer WE, Hartzell MM, Kishino N, Gatchel RJJJoABR. An overview of the Patient-Reported outcomes measurement information system (PROMIS) for assessing chronic low back pain patients. 2017;22(2):e12057.

80. Lapin B, Davin S, Stilphen M, Benzel E, Katzan IL. Validation of PROMIS CATs and PROMIS Global Health in an Interdisciplinary Pain Program for Patients With Chronic Low Back Pain. *Spine (Phila Pa 1976).* 2020;45(4):E227-E235.

81. Cook KF, Jensen SE, Schalet BD, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *J Clin Epidemiol.* 2016;73:89-102.

82. Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. *J Consult Clin Psychol.* 2001;69(4):655-662.

83. Deyo RA, Von Korff M, Duhrkoop D. Opioids for low back pain. *BMJ.* 2015;350:g6380.

84. Schalet BD, Hays RD, Jensen SE, Beaumont JL, Fries JF, Cella D. Validity of PROMIS physical function measured in diverse clinical samples. *J Clin Epidemiol.* 2016;73:112-118.

85. Field J, Newell DJC, therapies m. Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. 2012;20(1):17.

86. Farrar JT, Young JP, Jr., LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain.* 2001;94(2):149-158.

87. Sundararaman LV, Edwards RR, Ross EL, Jamison RN. Integration of Mobile Health Technology in the Treatment of Chronic Pain: A Critical Review. *Reg Anesth Pain Med.* 2017;42(4):488-498.

88. Ryan CG, Gray HG, Newton M, Granat MHJMt. The relationship between psychological distress and free-living physical activity in individuals with chronic low back pain. 2010;15(2):185-189.

89. Huijnen IP, Verbunt JA, Peters ML, et al. Do depression and pain intensity interfere with physical activity in daily life in patients with Chronic Low Back Pain? *Pain.* 2010;150(1):161-166.

90. 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.

91. Fox J, Bailenson J, Binney JJPT, Environments V. Virtual experiences, physical behaviors: The effect of presence on imitation of an eating avatar. 2009;18(4):294-303.

92. Hoffman HG, Sharar SR, Coda B, et al. Manipulating presence influences the magnitude of virtual reality analgesia. *Pain.* 2004;111(1-2):162-168.

93. Bulu STJC, Education. Place presence, social presence, co-presence, and satisfaction in virtual worlds. 2012;58(1):154-161.