PROJECT DESCRIPTION

1. Principal Investigator (Co-PI): Alicia Heapy, Ph.D.
Co-PI (off-site): John Piette, PhD – VA Ann Arbor Healthcare System

Title: Patient-Centered Pain Care Using Artificial Intelligence and Mobile Health Tools: Phase 2

2. Purpose:
The current study will evaluate a new approach for delivering cognitive behavioral therapy (CBT) pain management that may substantially improve our ability to systematically adapt CBT to patients' unique needs while ensuring that scarce clinician resources focus on patients who need more intensive treatment. Specifically, we will use **Reinforcement Learning (RL)**, a type of **Artificial Intelligence (AI)** to ensure that Veterans with chronic pain receive the intensity and type of therapy they need. Based on patients' daily feedback, AI-CBT will make recommendations to potentially step-up the intensity of each patient's CBT follow-up from automated calls to more engaging live telephone therapy sessions (15 minutes or 45 minutes) as needed. The specific aims of the study are as follows:

1. Using a randomized non-inferiority design, we will determine whether an AI adaptive program of telephone-based, personalized and stepped pain CBT that includes brief therapy sessions and automated calls as well as standard one hour treatment sessions (AI-CBT) can achieve equal outcomes compared to a standard CBT program of 10 hour-long sessions;
2. We will use a budget impact analysis to quantify the difference in therapist time associated with AI- CBT relative to standard CBT as well as differences in overall ambulatory and inpatient service use;
3. We will determine whether AI-CBT results in greater patient satisfaction and engagement, due to its more patient-centered approach that automatically focuses on "what works" for each Veteran.

Our *central hypothesis* is that a pain management program that uses AI algorithms and regular feedback collected via IVR about Veterans' physical activity, CBT skill practice, and pain-related physical functioning and that automatically adapts treatment delivery will achieve outcomes that are as good as standard approaches (or even better), but use substantially less clinician time.

3. Background:
Musculoskeletal disorders are highly prevalent among VA patients, with chronic back pain the most frequently reported type.[1,2] Among OEF/OIF Veterans, back pain and other musculoskeletal conditions are the most prevalent of all diagnosed medical and psychiatric conditions.[3] VA data suggest an annualized increase in the prevalence of low back pain of 4.8% per year due to factors such as an aging population and increasing prevalence of obesity.[1,4] Chronic low back pain is associated with work interruption, emotional distress, and risky health behaviors such as substance use.[5] Emerging evidence suggests that chronic pain compromises successful treatment and management of other chronic conditions.[6] For all of these reasons, increasing access to effective, convenient treatments for chronic low back pain is a national VA priority.[7] Cost of treating back pain in VA is $2.2 billion annually.[8]

Cognitive and Behavioral Interventions to Improve Pain Management
CBT is the most widely-accepted evidence-based psychological treatment for chronic pain.[8] CBT is informed by theory recognizing that patients' beliefs, attitudes and coping styles play central roles in determining their experiences of pain.[9] CBT is an attractive alternative to pharmacotherapy because impacts on functioning can last long after treatment is discontinued, and CBT does not entail the negative side effects of opioids. The goal of CBT for pain is to assist patients in developing an adaptive problem solving approach to pain management, and CBT targets both reductions in pain symptoms as well as their associated disability and emotional distress. The VA Evidence-Based Psychotherapy (EBP) program uses a CBT for pain that contains 10 hour-long sessions delivered weekly. Sessions provide pain education, teach and encourage the practice of pain self-management skills, and promote productive and pleasurable activity and exercise. Skills address both cognitive processes (e.g., catastrophizing) and behaviors (e.g., relaxation). A meta-analysis found moderate to large effects of CBT in improving pain related outcomes,[10] and a Cochrane systematic review reached nearly identical conclusions.[11]

Prior Research on Adapting Treatment to Patients' Individual Needs

Another foundational area of research for the proposed study is the theory of tailored health communication, which suggests that patients are more likely to internalize health messages when those messages are relevant to them personally.[12] The state-of-the-science in tailoring uses surveys to identify patients' needs, health beliefs, learning styles, cultural context, and other factors prior to crafting health related messages targeting behavioral changes. The data needed to tailor health messages is substantial, and many patients may not be willing or able to accurately report that information at program outset.[13] For example, Drs. Kerns and Heapy found that CBT skills training was no more effective when skill presentation was tailored according to what patients thought they wanted before initiating treatment.[14] Also, previous tailoring systems typically have tailored based on static patient traits, rather than on updated information about patients' status or treatment response. In the proposed study, we will tailor Veterans' pain CBT services using IVR reported feedback about their: pain-related physical functioning measured objectively via pedometer step counts, perceived functioning scores, and progress with CBT skill practice. Based on this real-time feedback,AI-CBT will personalize each patient's course of treatment automatically in order to achieve the greatest benefits for the population while using clinical resources as efficiently as possible.

Mobile Health (mHealth) Approaches to Self-Management Support

Because mHealth services have low marginal costs, they can cost-effectively reach large numbers of patients between face-to-face encounters to provide self-management support.[15] More than 50 studies have demonstrated that patients can provide reliable and valid information about psychiatric symptoms and substance abuse disorders via IVR and other mobile health technology.[16-19] Trials suggest that mHealth interventions can improve self-management behaviors[20,21] and may improve outcomes of chronic illness care.[22] The benefits of standard CBT diminish after patients discontinue therapy, and maintenance interventions delivered via IVR sustain those improvements in symptoms and self-management skills.[22,23] Because mHealth services are so promising, VA is making significant investments in development through the national Office of Telehealth.[24] Despite their potential, mHealth interventions typically deliver simplistic series of messages based on pre-determined "if-then" rules and deterministic protocols. As a result, mHealth interactions can feel "robotic" to users and many of them may drop out of treatment.[25] By way of analogy, no patient would continue to see a physician if that provider failed to adapt their management approach or goals based on the patient's response to prior treatments. We propose a model for taking advantage of the cost and accessibility benefits of mHealth services, while ensuring that these powerful tools are integrated systematically with personal, professional care from trained VA CBT therapists.

4. Significance:

The proposed team includes VA researchers with a long-history of Veteran-centered health services research on mobile health, and non-VA experts with unique expertise in AI and tailored messaging that can catalyze the development of novel tools for improving chronic pain management. The use of AI to tailor patients' treatment plan is unprecedented in chronic disease management. These approaches have become widely used in other areas (e.g., online advertising) because the powerful AI algorithms can significantly improve the user-centered experience and services' impact on target outcomes. Because this approach can optimize each patient's treatment response with the most efficient use of scarce therapist time, it could dramatically increase the number of Veterans with access to pain CBT given constraints on VHA budgets. Success with this system will provide a foundation for a broader program of research on AI supported interventions to improve chronic illness care for Veterans with high priority chronic conditions such as depression, diabetes, and obesity.

5. Research Plan:

Overview

This will be a randomized non-inferiority study comparing standard CBT for chronic pain to a strategy that uses mobile health technology and artificial intelligence in conjunction with CBT therapists to deliver evidence-based, stepped pain therapy. Patients in both groups will receive CBT delivered via telephone by pain CBT therapists. For patients in the standard CBT group, the therapist will deliver 10, 45 minute-long CBT sessions based on content used throughout VA. Patients randomized to the AI-CBT treatment group, named the *Responsive Efficient Accessible Chronic pain Technology (REACT) program,* will be asked to report their pedometer-measured step-counts, pain-related functioning, and CBT skill practice via five-minute daily IVR calls, and will receive a pre-recorded therapist feedback message at the end of the week. Some of those IVR calls also will include reminders regarding the dates and modalities for upcoming CBT sessions. Based on patients' IVR feedback, AI-CBT will make recommendations to potentially step-up the intensity of each

patient's CBT follow-up using more engaging live telephone therapy sessions (15 minutes or 45 minutes). Based on experience gained from each patient's own history and the overall population of patients, the AI-CBT engine will seek to optimize the population's total improvement in functioning while maintaining each patient at the least resource-intensive mode of CBT delivery. Outcomes will be measured via telephone survey at three and six months post recruitment, and additional data will be collected via clinical records. We will use data from therapists' activity logs and administrative files to conduct a budget impact analysis. De-identified survey data will be entered at VA Connecticut or paper copies will be mailed using VA approved carrier and tracking number to the Ann Arbor VA site for entry by approved staff. Additional data to aid translation of study findings into practice will be collected via qualitative interviews with CBT therapists/team members, and patients with various levels of program response.

Conceptual Framework

The intervention we will evaluate is based clinically on widely-adopted and evidence-based models of CBT for VA pain management (described above),[26] and it links those concepts with a strategy for personalized stepped care using **reinforcement learning (RL)**. RL is a field of artificial intelligence that allows an "intelligent agent" to learn what treatment choices work best in order to optimize a measurable outcome (termed the system's "reward"; see Figure 1). The process used to optimize treatment choices in RL mimics the way that humans learn skills such as riding a bicycle, i.e., through systematic adaptation and generalization accompanied by targeted trials of new behaviors with measurable outcomes. RL algorithms similar to those we will apply in the proposed study are the basis of online consumer targeting programs such as Netflix, Google, and Amazon.com,[27] where a service learns automatically how to deliver information that is most relevant to each user. In the current trial, the RL agent will be a computer system that makes weekly recommendations or "action choices" for each patient with respect to the mode and intensity of CBT that the patient should receive. Those recommendations will be based on the patient's progress, the progress of similar patients, and other contextual information for that action choice.

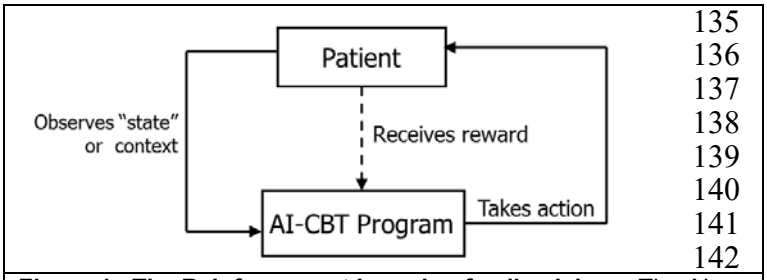

Figure 1. The Reinforcement Learning feedback loop. The AI-CBT actions are the 3 CBT session types; the "reward" is IVR-reported pedometer step-counts, and "state" data is IVR-collected information on patients' CBT skill practice & pain-related functioning.

Potential actions the AI-CBT program will recommend include: a standard 45 minutes telephone CBT therapy session, a 15 minute telephone CBT therapy session, and an IVR automated therapy session designed to teach and reinforce skill-based learning. Fifteen minutes was chosen to be consistent with the time increments of the health and behavior CPT codes (15, 30, 45 and 60) used to bill for behavioral interventions for chronic pain. Content for each session type will be based on standard VA CBT for pain management, modified by a Panel of Experts to be most effective given the length and mode of each contact. The AI-CBT agent's recommendation regarding which action to take will be based on each patient's IVR-reported pedometer step counts (i.e., the "reward") as well as other "state" information (Figure 1) also collected via IVR. Importantly, the RL algorithm will learn not only based on each patient's own treatment response, but will incorporate experience from the response of other patients who have similar characteristics and response trajectories as indicated by the "state space." Based on this feedback loop, the RL engine will modify the probability distribution across treatment choices and make recommendations for each patient each week. Because actions will be probabilistic rather than "hard-wired," the AI-CBT program will avoid potentially over-reactive treatment changes that can result when therapists attempt to tailor care non-systematically or using deterministic flow diagrams. AI CBT patients will begin with an IVR automated therapy session. Based on their progress as measured by feedback on the "reward" and "state" features, patients who progress toward functional goals will be kept in less resource intensive options, and patients who need more intensive follow-up will be moved automatically to more time-intensive, therapist-delivered CBT.

Patient Identification and Recruitment

Criteria for Opt Out Letters:

Patients with diagnoses of chronic back pain receiving care in facilities affiliated with the VA Ann Arbor Healthcare System and the VA Connecticut Healthcare System will be identified via VA Informatics and Computing Infrastructure (VINCI). Eligible patients will have back-pain-related diagnoses including low back and spine conditions, and nerve compression (ICD-9 codes 724.01, 724.02, 724.03, 724.09, 724.1, 724.2, 724.3, 724.4 and 724.5 or ICD-10 codes associated with back pain) and a score of ≥ 4 (indicating moderate pain) on the 0-10 Numerical Rating Scale on at least two separate outpatient encounters in the past year.

Study Eligibility Criteria:

Screening of interested and potentially eligible patients will be conducted by medical record review and telephone interview using validated measures. Inclusion criteria include: (1) at least moderate pain-related disability as determined by a score of 5+ on the Roland Morris Disability Questionnaire at baseline; (2) at least moderate musculoskeletal pain as indicated by a pain score of ≥ 4 on the Numeric Rating Scale; (3) pain on at least half of the days of the prior six months as reported on the Chronic Graded Pain Scale; and (4) a touch-tone cell or land line phone (5) a confirmation of back pain documented in the medical record. Exclusion criteria include: (1) active psychotic symptoms, suicidality, severe depressive symptoms (PHQ-9 score of >20), substance use disorder or dependence, active manic episode or poorly controlled bipolar disorder as identified by Mini International Neuropsychiatric Interview,[29] or severe depression identified by chart review of diagnoses and mental health treatment notes; (2) life threatening conditions that could impede participation, such as COPD requiring oxygen or cancer requiring chemotherapy; (3) cognitive impairment defined by a score of ≤5 on the Six-Item screener;[30] (4) sensory deficits that would impair participation in telephone calls; and (5) current CBT or surgical treatment related to back pain (6) Patient does not plan on getting most of medical care for pain in the next year at VA. After obtaining agreement from patients' primary care providers (an email will be sent asking providers if we may send their patients the recruitment letter. Similar to an already approved study, this will be an opt-out email to the providers meaning if they wish for more information we will provide it or if they request their patients do not receive this letter we will not send it. However, we will inform them in the email and at a presentation at a monthly primary care staff meeting that if they do not respond we will assume that we have their permission and send the letter, to the Veterans informing them about the study and inviting participation. Veterans who do not opt-out by postage-paid response card will be called by research staff to explain the study, conduct screening, and solicit their involvement. VA and other investigators have found that this opt-out procedure is less burdensome to patients[31] and results in a larger sample that is more representative of vulnerable patients[32] than opt-in procedures (see Human Subjects). A Waiver or Written Informed Consent (WWIC) will be requested to complete this telephone eligibility screening before informed consent has been signed to join the trial. Willing, eligible Veterans will be given or mailed the consent and HIPAA form along with a postage paid return envelope. We will not begin study procedures until we have received the signed consent and HIPAA forms back. We have used this same process in numerous prior IRB-approved VA studies and have found that it is an efficient and effective way to recruit large samples of Veterans without requiring an in-person recruitment visit. The study coordinator will track the percentage of eligible Veterans who enroll and will actively solicit reasons for declining. This information will be used to assess the intervention's reach as described in the implementation portion of the application (Aim 3).

Recruitment methods will include flyers posted in patient care areas and from clinician referrals at the VACHS main campuses (West Haven and Newington) and community-based outpatient clinics (CBOCs), clinician referrals, Craigslist under the "Volunteers" section of the site, flat screen informational televisions around the hospital, Good Morning VA Connecticut, the PRIME Center newsletter, VA Connecticut's Facebook page, a VA public affairs press release, vet centers and numerous locations within the greater New Haven area such as libraries and grocery stores (upon their approval). We will also provide the approved flyer to veteran coordinators at local universities and colleges for posting or distribution. Finally, we will recruit patients at a Pain Education table positioned outside the Patient Education room in Building 2, 1st floor. This table will be staffed by study research staff including research assistants and/or co-investigators. We plan to include approved recruitment flyers and study description cards at the table and information about chronic pain and chronic pain management, pertinent VACHS resources, and other currently approved research studies conducted by our group. Screening of interested and potentially eligible patients will be conducted by medical record review and telephone interview using validated measures as described above, using a WWIC to

complete the eligibility screening measures before Veterans are asked to join the trial. Willing, eligible Veterans will be given or mailed the consent form along with a postage page return envelope in order to enroll in the trial.

Pilot- We will recruit up to 5 participants to pilot the IVR and AI system. These participants will not be included in the final analysis. We will call the pilot participants at approximately week 4 to request feedback about their experiences with the program/system. If we identify any system issues during their participation they will be corrected prior to enrolling patients into the main trial**.**

Randomization

After completing baseline assessments, patients will be randomized to AI-CBT or standard telephone CBT by the research staff. The randomization sequence and opaque randomization envelopes for recruiters will be generated prior to recruitment by the study statistician using a random number generator. Analysts and the investigators will be blinded to patients' group assignment until initial outcome analyses are completed. To ensure balance across treatment arms in potential modifiers of intervention effect, randomization will be conducted within blocks defined by site (Michigan versus Connecticut).

Common Elements of Standard and AI-CBT

Both CBT conditions will consist of 10 treatment modules delivered over 10 weeks. The same therapist at each site will provide treatment to patients in both groups. In each arm, the 10-week course of therapy will include an introductory module, followed by eight pain coping skills training modules and concluding with a session emphasizing skill consolidation and relapse prevention. The introductory module will present the biopsychosocial model which explains how chronic pain leads to dysfunction and provides a rationale for the efficacy of pain coping skills. The eight skills that will be presented were selected based on their efficacy in improving pain outcomes and their appeal to patients in prior trials. These include sessions on: physical activity, behavioral activation, pacing, sleep hygiene, and relaxation. Other modules will address common maladaptive cognitions including pain catastrophizing and fear of movement or kinesiophobia. Using procedures developed in two previous VA studies, during sessions 2-9 participants will be assigned a goal related to a newly presented skill (e.g., "practice relaxation exercise for 20 minutes daily") and a daily walking goal (e.g., average daily steps over the prior week plus 10%). As participants progress through treatment, they will continue to practice prior goals. In order to maintain consistency in the behavioral and cognitive restructuring targets of CBT between the two treatments, participants in both groups will be assigned the same skill practice goals and the same formula will be used for assigning steps goals.

Patient and Therapist Materials. Patients in both treatment conditions will use a handbook that was developed from materials used in prior trials, refined for use with IVR CBT through an HSR&D-funding Short Term Project Award, and currently in use in our HSR&D-funded IIR of IVR-based CBT for low back pain (see appendices). The handbook is written at the 6th grade reading level so that it can be used with or without the guidance of a therapist. The handbook will be identical for both conditions except that the AI-CBT handbook will contain additional information that describes the three AI modes (45 minute, 15 minute, and IVR sessions) and how to prepare for each type of session. The therapist manuals will be adapted from materials developed for our IVR-based CBT for Chronic Low Back Pain trial. The AI-CBT section will detail specific guidelines for each treatment mode: 45 minute, 15 minute, and IVR.

Pedometers for Monitoring Physical Activity. All patients will be given a pedometer and a log for monitoring their step counts. We will use an Omron HJ-320. Patients will be given or mailed a pedometer after completing their baseline assessment and returning their consent form.

Standard Telephone CBT (control)

Control patients will receive telephone CBT consisting of 10 weekly modules delivered via 45 minute telephone contacts with a therapist. The format of each session will include: (1) review of patients' pedometer logs and coping skill practice, (2) review of previous material and correction of misunderstandings of the information, (3) assignment of new step count goals and discussion of the new skills-based material, and (4) discussion of specific step and skill practice goals. Positive feedback and praise will be offered for any skill practice and step goal efforts and accomplishment. Barriers to practice or goal completion will be identified and problem-solving techniques will be used to address them.

Cognitive Behavioral Therapy Supported by Artificial Intelligence (AI-CBT), *titled for patients the "Responsive Efficient Accessible Chronic pain Technology (REACT) program"*

Daily IVR Reports. We will collect reports about patients' pedometer-measured step counts, CBT skill practice, and pain-related functioning via daily five-minute IVR calls. If the initial call is missed, the system will automatically try again 15 minutes later and again 1 hour later. We have successfully used these methods in studies achieving high patient response rates. Step counts will measure activity over the prior 24 hours, and patients will report their skill practice using a 0 -10 scale. Pain-related functioning will be assessed using the PEG, a three item scale assessing pain intensity and interference.[33] Practice of target behaviors will be measured using a 0 (not at all accomplished) to 10 (completely accomplished) scale and items that we have successfully used in our ongoing trial if IVR CBT. Questions measured by the telephone IVR system can be seen in the Table 1 below.

Table 1: IVR Questions

IVR QUESTION FIELDS
Pain management skill adherence
Pain intensity
Self efficacy
Pain interference
Sleep
Affect/Mood
Meaningful Activity adherence
Prior skill practice
True/False quiz

Step counts will be used in the "reward function" that the RL algorithm will seek to optimize; and functioning and skill practice reports will be used as "state" information (see Appendix A) that the system will take into account when making decisions that optimize patients' treatment course. The AI system will be able to accommodate missing IVR reports, and patients who fail to complete more than 50% of the daily IVR calls in a two week period will be called by a research associate to trouble shoot problems and encourage compliance with feedback. In addition to being the source of data with which AI-CBT will personalize each patient's course of treatment, web-based reports of data from IVR calls will be used to inform therapists of participants' progress. These data will be particularly important for informing the abbreviated 15 minute therapist sessions where a priority will be placed on the efficient use of treatment time, and during the IVR sessions where the entire session will be pre-recorded. Once a week the IVR call will include a brief message alerting patients of the date, time, and modality for their subsequent week's session.

Special Circumstances for Telephone Minutes

Veterans with concerns of telephone minute allotments may have minutes purchased by the study via gift certificates. This will allow these participants to engage in this telephone intervention to the same degree as participants who do not have minute limit concerns. These gift certificates will be distributed in an adhoc basis if a participant presents concerns to study staff. Staff will purchase minutes from the VA canteen with gift certificates and then distribute to participant.

AI-CBT Action Recommendations. After week 1, session options will include: (a) 45 minute live telephone therapist sessions; (b) 15 minute live telephone therapist sessions; and (c) IVR treatment sessions. To avoid scheduling conflicts, AI-CBT patients will be assigned a one-hour block of time each week in which both they and the CBT therapist are available for treatment. This same time slot will be used for either the 45 minute -long therapist sessions, the 15 minute therapist sessions or the IVR CBT sessions. Each Monday morning, the CBT therapists will receive a list of AI-CBT personalized treatment recommendations for that week for each

patient. By noon on Monday, the therapist will have a finalized schedule of which patients require what types of contact that week, and which patients need to have a summary of the therapist's comments and recommendations recorded for the week's IVR CBT therapy call.

During *Week 1*, AI-CBT patients will look through their handbook and they will review the program's goals and processes, and the standard introductory material contained in session 1 of the standard CBT. At the end of the week the AI-CBT patients will receive a therapist pre-recorded feedback message along with an alert of the date, time and modality of their subsequent week's session. The *45 minute AI-CBT Sessions* will be identical to the control condition and will follow the same progression of content used for control patients. *15 Minute Telephone CBT Sessions* will mirror the content of the 45 minute sessions, though in a compressed form. The therapist manual and patient handbook will emphasize the importance of using the session time efficiently and using a consistent format that includes: reviewing the patient's daily IVR reports, clarifying the current week's adaptive pain coping skill in the patient handbook, and setting goals for skill practice and step counts for the coming week. Prior to the session, therapists will review patients' daily IVR reports. If participants have not been successful in meeting step or skill practice goals, the therapist will help the participant address barriers to goal attainment. The therapist then will ask the patient to describe the current week's adaptive pain coping skill as a brief check of their understanding, and will review goals for the coming week with a discussion of anticipated barriers. Remaining time will be used to review the skill and to encourage the patient to read their patient handbook. Much of the content for the *IVR CBT sessions* has been developed and implemented as part of our ongoing IVR-based CBT for Chronic Low Back Pain trial (see Appendix D). Our experience in that trial suggests that patients complete the IVR sessions more than 90% of the time and that satisfaction rates are high. During these sessions, patient will receive a 2-5 minute pre-recorded feedback message from their therapist, during which the therapist will review the patient's IVR-reported changes in step counts, functioning, and skill practice. Reinforcement will be provided for effort, and improvements will be noted. IVR messages will include a review of the pain coping skill practice and step goals for the coming week, and participants will have the option of leaving a message for their therapist via the IVR system should they have a question. Therapists can leave a response message, also on the IVR system.

The AI Engine. Patients' IVR-reported step counts, skill practice, and pain-related functioning will be accessed by the AI engine daily to update the probabilities the system uses to determine which treatment to recommend the next week. We will use a state-of-the-art "contextual bandit" AI algorithm (LinUCB) designed to make careful choices while learning quickly from a patient's treatment response as well as the experience of other patients with similar characteristics (see Appendix A for technical details).[34-37] With increased interactions, the system will learn to tailor decisions more effectively to maximize population-level improvements in functioning while minimizing clinician time. In this way, AI-CBT will function similar to the best clinicians, who learn from experience within and across patients. In the context of the trial, this means that patients enrolled early will likely receive less personalized CBT courses that are relatively similar to the standard CBT approach (i.e., a greater number of 45 minute-long sessions), while patients enrolled later will receive services that are more personalized and include a greater frequency of 15 minute therapist sessions and IVR sessions. To maximize the efficiency of this "learning curve": (1) patients will be recruited over a longer period than would potentially be necessary, so that AI-CBT can gain as much experience as possible from patients recruited first and apply that knowledge to patients entering the program later; and (2) patients will be randomized with a greater "N" in the AI-CBT group in order to maximize the system's experience (see power calculation). As part of our evaluation, we will compare outcomes across randomization groups separately for early versus later enrollees, in order to test the hypothesis that AI-CBT will result in greater efficiency over time. These analyses also will allow us to project program benefits if AI-CBT were implemented in VA for thousands of Veterans and multiple years of experience.

CBT Treatment Fidelity will be assessed using the Yale Adherence and Competence Scale (YACS),[38] a validated scale that assesses therapist adherence and competence in delivering manualized behavioral therapy. A trained rater will rate audiotapes of 30% of all CBT therapist sessions to assure that treatment is consistent with the manual and will provide corrective feedback to the therapists whenever drift occurs.

Role of the Expert Panel. The AI-CBT program will be supervised with ongoing input from an Expert Panel comprised of experts in: pain management, CBT for chronic pain, clinical trials using behavioral interventions, adaptation of psychotherapy for telephone delivery, and IVR (see Appendix E and letters of support). The Panel will be directed by Dr. Robert Kerns (Co-I) who has led many large and geographically-dispersed panels

of experts in pain research and treatment, and will ensure that the panel meetings are efficient and effective. The Panel will meet regularly by teleconference during the start-up period to review and revise the treatment materials and refine the AI algorithm to reflect any constraints that should be put into place to limit the choices the algorithm can make, e.g., "if the patient's activity level decreases more than 20% two weeks in a row, recommend two 45 minute-long CBT session regardless of what the prior week's contact was." After start-up, the Panel will teleconference quarterly and in ad hoc sessions as needed.

Measurement & Analytic Plan

Overview. We have selected outcome measures based on recommendations from the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommendations. Endpoint measures are consistent with CONSORT guidelines recommending that non-inferiority trials use outcomes similar to those used in efficacy studies. In addition to primary and secondary outcomes (defined below), we will examine treatment satisfaction, treatment credibility, patient engagement and dropout, and goal accomplishment. Process and outcome data will be collected via the following sources: *Patient Surveys:* Baseline, three-month, and six-month surveys will be conducted at the VA or via telephone by trained research assistants or via mail, sent with postage paid return envelopes. *Qualitative Interviews:* These will be conducted with purposive samples of patients in the AI-CBT group at follow-up. We will target patients who demonstrate significant improvement, patients who were very satisfied with AI-CBT, patients without significant improvement, patients who were dissatisfied, and patients who dropped out of the intervention. *CBT Therapist Logs:* Logs will be used to track therapist time spent in patient treatment, attempting to reach patients, and key information about those interactions. *IVR:* The AI-CBT IVR system will capture information about intervention patients' pedometer-measured step-counts, pain related functioning, CBT skill practice, and missed data reporting. *Administrative and Clinical Data Systems:* These will be used to track patients' use of other VA inpatient and outpatient services for pain management, mental health, and medical care.

Primary Outcome. The 24-item Roland Morris Disability Questionnaire (RMDQ) is an IMMPACT endorsed measure[40] of pain-related disability for persons with chronic back pain. Strong evidence supports the RMDQ's reliability, validity, and responsiveness to change during trials.[41]

Secondary Outcomes. ***Global Pain Intensity*** **will be assessed using the Numeric Rating Scale (NRS-I) an IMMPACT-recommended 11-point numeric rating scale of pain severity.[39]** ***Pain-Related Interference*** **will be measured using the Brief Pain Inventory (BPI).[57] Pain-related functioning will be assessed using the PEG, a three item scale assessing pain intensity and interference.[33] Depression Symptom Severity will be assessed using the 9-item Patient Health Questionnaire (PHQ-9) a widely used measure with excellent internal consistency and stability.[28]** ***The Patient Global Perception of Change*** **scale is a single item measure that quantifies a participant's overall perception of improvement since beginning treatment and the clinical importance of that improvement. Participants indicate improvement on a 7 point "much worse" to "much better" scale. This is a well-validated measure recommended by IMMPACT.[40] Finally, we will use the** ***Veterans SF-12 to assess health-related quality of life*****. This measure has demonstrated good internal consistency and is strongly correlated with socioeconomic status and morbidities.[42]**

Resource Use (Aim 2). ***Intervention Costs.*** **Therapists will use a log to** record time spent in intervention-related activities for a random 20% of all treatment days. Time records will be combined with wage data from the VA Financial Management System to estimate intervention-specific personnel costs. Technology costs of the AI-CBT program include fixed costs (e.g., software development and computer maintenance) plus variable costs (e.g., minute costs for IVR calls). One-time fixed start-up costs will be reported separately. *VA Inpatient and Outpatient Service Use* data will be obtained from the Musculoskeletal Diagnoses Cohort (MSD), a project currently underway as part of the VA Connecticut Healthcare System's CREATE. The MSD is developing validated algorithms for using VA electronic health record data to identify utilization events, comorbid conditions, receipt of opioid medications, and pain screening results, for patients with pain-related diagnoses. Information on non-VA admissions will be collected by the patient survey. To mitigate recall bias, we will use a two-time frame method that asks about utilization over the past 6 months and over the past 2 months with more weight given to the shorter timeframe.[43]

Treatment Satisfaction and Engagement (Aim 3). For patients in the AI-CBT group, we will calculate *IVR adherence* as we have in the past,[44] i.e., as the proportion of days during which an assessment was attempted in which one was successfully completed and the number of weeks during which the patient completed at least four out of seven requested IVR reports. Participants' judgments of *Treatment Credibility* will be assessed using a reliable questionnaire adapted from Borkovec and Nau.[45] Treatment credibility has been shown to be significantly associated with treatment satisfaction, engagement in treatment, and number of sessions attended. The *Pain Treatment Satisfaction Scale* of the Patient Outcomes Questionnaire will be used to assess patient satisfaction with various domains of pain care.[46] This 5-item measure shows good internal consistency and significant associations with staff and patient ratings of patient improvement. *Attendance in "Live" Telephone CBT Sessions and Program Dropout* In order to understand reasons for treatment dropout, we will attempt to reach samples of patients with low levels of engagement for qualitative interviews. Participants will rate their *Continued Skill Use* at follow-up for each of the target behaviors emphasized in the CBT program on a 0 (not at all accomplished) to 10 (completely accomplished) scale. As described above, *Daily IVR calls* will be used to collect data in the AI-CBT condition regarding pedometer measured step-counts, CBT skill practice, and pain-related functioning using pre-recorded questions we have used successfully in our prior studies.

Demographics and Other Covariates Measured at Baseline. We will measure patients' baseline *Sociodemographic and Pain Characteristics* that have been shown to be associated with treatment outcomes, e.g., age, gender, education level, racial/ethnic background, marital status, occupational status, pain duration, and number and location of pain sites. We also will gather data on participants' level of health literacy.[47] *Pain Classification* will be derived through a systematic evaluation of each enrolled participant's EMR using an assessment tool based on clinical guidelines for diagnosing and treating back pain.[48] Pain classes will include non-specific back pain, back pain with a radicular component, or back pain associated with other specific spinal causes. A nurse practitioner, who is supported by the West Haven COIN and who has been trained to use this tool and is using it in our trial IIR 09-058. *Psychiatric and Substance Abuse Comorbidities* will be measured using medical record diagnoses and mental health encounters. Additional self-report information will be collected using subscales of the Mini International Neuropsychiatric Interview (MINI)[29] related to mood and substance abuse disorders. *Pain Medication Use* will be assessed through patient survey and a review of computerized pharmacy records. Pain medication will be coded as non-steroidal anti-inflammatory, non-narcotic analgesics, narcotic analgesics, and benzodiazepines and other sedative/hypnotics. *Distance from VA* will be calculated and used as a measure of geographic access. The *Pain Catastrophizing Scale* is a 13-item self-report scale that examines thoughts and feelings people may experience when they are in pain including rumination, magnification, and helplessness.[49] Finally, *Pain-related Fear* will be measured using the Tampa Scale of Kinesiophobia-revised (TSK-R), which has two subscales (Fear of Harm/Activity Avoidance and Pathophysiological Beliefs) and has been shown to be sensitive to treatment-related change.

Sample Size and Power Calculation

Our primary outcome analyses are consistent with the CONSORT Statement on Reporting of Non-inferiority and Equivalence Randomized Trials.[50] The sample size was calculated using the Non-Inferiority Test module available in the statistical software PASS 2008. To ensure that the AI-CBT program retains a clinically acceptable effect, the non-inferiority margin was set at 2 points on the Roland Morris Disability Questionnaire (RMDQ).[41] A two point difference (or difference in reduction) in the RMDQ is considered to be a minimally clinically significant effect.[51] The power calculation was based on a significance level of 0.025, with a power of 90%, when postulating a true difference in group means of 0 and a standard deviation of the outcome of 4.5 in both groups. Specifically, if we denote by D the true difference in mean RMDQ scores (at 12 weeks) between the AI-CBT and standard CBT groups, with the non-inferiority margin set at 2, we plan on testing the null hypothesis H0: D>=2 versus the alternative hypothesis H1: D<2, which amounts to a one-sided, two-sample t-test. Thus, if the null hypothesis is rejected, it can be concluded that AI-CBT is non-inferior to CBT. We will test this hypothesis at a one-sided .025 significance level based on a confidence level where we will reject the null hypothesis when the upper bound of the two-sided 95% confidence interval for D is less than 2. To ensure that the AI-CBT algorithm has as much information as possible to learn quickly how best to tailor patients' course of therapy, we will disproportionately randomize patients to the AI-CBT group in an allocation ratio of 1.37:1 for the AI-CBT: standard CBT groups, respectively(of the total sample approximately 60% in the AI-CBT group and 40% in the Standard CBT group). Assuming this ratio, we will have 90% power to detect non-inferiority with a total sample size of 221 patients, or 128:93. To account for a 20% drop-out rate in both groups, we will enroll 278 patients (160 in the AI-CBT group, 118 in CBT group).

Baseline Comparability, Reach, and Representativeness. We will examine baseline differences across groups in measures of study endpoints as well as other potential prognostic indicators, such as patients' age, comorbid diagnoses, and history of pain treatment. Any differences across groups in baseline characteristics will be controlled statistically in analyses comparing outcomes. The RE-AIM framework is a methodology for systematically considering all strengths and weaknesses of an intervention in order to better guide program planning.[52] To evaluate reach, we will ask patients who decline study participation whether they would be willing to provide informed consent to participate in a brief survey that identifies their reasons for declining participation and the characteristics that differentiate them from enrollees.

Analysis of Endpoints (addressing Specific Aim 1). We will test for non-inferiority of AI-CBT compared to standard telephone CBT at post treatment by comparing the upper limit of the two-sided 95% confidence interval for the difference in the mean RMDQ scores, calculated as AI-CBT minus standard CBT, to the pre-set non-inferiority margin of 2 points in RMDQ and will conclude that AI-CBT is non-inferior if the upper limit is less than 2. Because intent-to-treat analysis can raise the risk of type I error in a non-inferiority trial,[50] we will conduct both a per protocol and intent-to-treat analysis. The "per protocol" group assignment will be defined as completing four or more CBT sessions (either IVR or "live" sessions); however, we will revisit this definition with the Expert Panel prior to the start of the trial. We will declare AI-CBT non-inferior to standard CBT only if AI-CBT is shown to be non-inferior using both the intention-to-treat and per-protocol analysis sets. We expect that RMDQ scores will be normally distributed. If not, we will use transformations to achieve normality. We also will develop a two-level linear mixed-effects model that uses RMDQ follow-up scores at both assessment times (three and six months) as the dependent variables; treatment group, time and the treatment by time interaction as categorical explanatory variables; and baseline RMDQ score as a continuous covariate. If this model shows no significant time by group effect, we can drop the interaction term and test for the time-averaged non-inferiority of AI-CBT compared to standard CBT between 3 and 6 months. This model will also allow for adjustment for design-related factors (e.g., site and age). An unstructured variance-covariance matrix will be used to model the error variance. Secondary outcomes including pain intensity, emotional functioning, global perception of change, and quality of life will be analyzed in a manner similar to that used for the primary outcome. Analyses of all other outcomes will be conducted on an intent-to-treat basis.

Intensity of Service Use (Specific Aim 2). We will compare service utilization by category (e.g., CBT therapist time, PCP visits, and pharmacy use) between groups. We will conduct a budget impact analysis[53] and will include the cost of the intervention (personnel, supplies, CBT therapist training, and IVR fixed/variable costs) as well as costs for specific medical care services likely to be affected. Data from CBT therapists time records will be combined with wage data from the VA Financial Management System to produce estimates of intervention-specific personnel costs. Costs associated with the use of specific medical care services, such as medications, will be obtained from the Decision Support System (DSS) files. Cost analysis will be conducted in accordance with the guidance provided by Mauskopf et al.[53] including the use of sensitivity analysis and scenarios that allow for varying assumptions about intervention uptake, compliance or component costs. All resource use and cost comparisons will be adjusted for any observed differences in baseline characteristics. Because costs of resource utilization are usually skewed, alternative modeling techniques (e.g., log-transformed costs, negative binomial regression) will be used.

Intervention Engagement and Satisfaction with Care (Specific Aim 3). We will conduct extensive analyses of the process of intervention delivery in both arms. We will monitor the proportion of telephone CBT sessions that are completed, and we will determine the patient and session characteristics associated with patients' reports of skill practice. Patients in the AI-CBT group will report their satisfaction with aspects of the intervention (e.g., whether it provided information useful for achieving behavioral targets), and we will assess the correlation between satisfaction ratings and measures of: intervention engagement, patients' baseline characteristics, and changes in pain-related functioning. Differential dropout across groups will be examined using Kaplan-Meier curves and survival models.

Preplanned Subgroup Analysis. Because AI-CBT will continue to learn patterns in patients' experience throughout the intervention period, we hypothesize that the second 50% of patients randomized will show an even larger difference in clinician time than the first 50%, while still demonstrating non-inferiority in pain-related outcomes. Differences in pain related functioning and in clinician treatment time across treatment groups will be tested in this subgroup analysis after stratifying the sample into early versus later recruits.

Approach to Missing Data. If more than 15% of a covariate is missing, we will use multiple imputation methods based on the SAS MI Procedure. We will check if the pattern of missingness is monotonic (i.e., if patients missing data at 3 months also have missing data at 6 months) and use a Markov chain Monte Carlo method that assumes multivariate normality to impute missing values. We will impute five sets of data and combine the imputed results using the SAS MIANALYZE procedure to obtain a valid estimate of the confidence limit and treatment effect.[54] When data are missing for items within scale scores, we will use recommended imputation procedures rather than deleting patients list-wise from the analysis.

Mediators and Moderators of Intervention Effects. We will use multivariate modeling to identify the mechanisms through which the intervention achieves effects on outcomes and differential effects across subgroups.[55] Initial models will include only treatment group as the predictor. Subsequent nested models will introduce potential mediators (such as the number of completed therapist sessions), and we will evaluate changes in the relationship between experimental condition and outcomes before and after covariates are introduced. Analyses of effect moderation will focus on baseline pain severity and comorbid diagnoses using standard approaches to evaluate interactions between these covariates and patients' experimental condition.[55] Significant interactions will be interpreted by plotting regression lines for predicted outcomes of patients with high and low values of the moderator.

Evaluating the Reliability of Patients' IVR Reports. We will evaluate the integrity of IVR-reported step counts and functioning by examining associations between IVR reports and baseline characteristics that the literature suggests would be associated with patients' functioning (e.g., baseline SF-12 scores, comorbid medical diagnoses, and age). We also will examine serial correlations across IVR reports under the assumption that all correlations between scores and proximal scores should be positive and roughly of equal magnitude controlling for the time difference between reports.

RE-AIM[53] Dimensions of Intervention Adoption, Implementation, and Maintenance. Adoption will be evaluated by examining variation in study participation and intervention engagement across sociodemographic and clinical subgroups of eligible patients. For example, we will determine whether older patients or those with less education have more difficulty responding to queries about their step counts or pain-related functioning via IVR. Adoption at the provider level will be monitored by recording the proportion of providers who are willing to have their patients participate and providers' reasons for not participating. Implementation and maintenance will be evaluated through semi-structured questions at follow-up designed to identify program characteristics that might be a barrier to patients' use of the intervention in other settings and the intervention characteristics that patients feel would make it more valuable to others with chronic pain. We also will meet with clinicians in each site to gauge their willingness to adopt and maintain a similar intervention, and the ways such a system can be designed to best complement existing services.

Qualitative Interviews and Mixed-Methods. We will use audio-taped interviews with patients, CBT therapists and/or team members to provide a context for interpreting intervention effects and suggest additional subgroup analyses. Interviews will be transcribed verbatim, and we will enter the transcripts into NVivo for file storage and selective retrieval. Using accepted techniques,[56] Drs. Piette and Heapy will independently read transcripts, approaching the data with analytic categories in mind, but identifying other categories in the data. An iterative process will be used until agreement is reached on categories and their definitions, after which we will develop a coding template and enter it into NVivo as a tree diagram.

6. References

1. Johannes CB, Le TK, Zhou X, Johnston JA, Dworkin RH. The Prevalence of Chronic Pain in United States Adults: Results of an Internet-Based Survey. *J Pain.* 2010;11:1230-1239.
2. Crowe M, Whitehead L, Jo Gagan M, Baxter D, Panckhurst A. Self-management and chronic low back pain: a qualitative study. *J AdvNurs.* 2010;66:1478-1486.
3. Walker, Robyn L. U.S. War Veterans and Pain. painedu.org . 2012.
4. Kerns RD, Philip EJ, Lee AW, Rosenberger PH. Implementation of the Veterans Health Administration National Pain Management Strategy. *Translational Behav Med.* 2011.
5. Spelman J, Hunt S, Seal K, Burgo-Black A. Post Deployment Care for Returning Combat Veterans. *J Gen Intern Med.* 2012;27:1200-1209.

6. Buckenmaier, CC and Gallagher, RM. War on Pain: New Strategies in Pain Management for Military Personnel and Veterans. 2011. Federal Pratitioner Supplement.
7. Lew HL, Otis JD, Tun C, Kerns RD, Clark ME, Cifu DX. Prevalence of chronic pain, posttraumatic stress disorder, and persistent postconcussive symptoms in OIF/OEF veterans: Polytrauma clinical triad. *J Rehabil Res Dev.* 2009;46:697-702.
8. Gironda RJ, Clark ME, Ruff RL et al. Traumatic brain injury, polytrauma, and pain: Challenges and treatment strategies for the polytrauma rehabilitation. *Rehabil Psychol.* 2009;54:247-258
9. Kerns, R. D. Pain among returning service members from OEF/OIF. 2007.
10. Chou R, Huffman LH. Nonpharmacologic Therapies for Acute and Chronic Low Back Pain: A Review of the Evidence for an American Pain Society/American College of Physicians Clinical Practice Guideline. *Ann Intern Med.* 2007;147:492-504.
11. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA. Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain physician.* 2009;12:E35.
12. Nicholas MK, Asghari A, Corbett M et al. Is adherence to pain self-management strategies associated with improved pain, depression and disability in those with disabling chronic pain? Eur *J Pain.* 2011.
13. Dobscha SK, Corson K, Perrin NA et al. Collaborative Care for Chronic Pain in Primary Care. *JAMA.* 2009;301:1242-1252.
14. Sauro, Measuring Usability with the System Usability Scale (SUS) (2011) http://www.measuringusability.com/sus.php
15. Kroenke K, Bair MJ, Damush TM et al. Optimized Antidepressant Therapy and Pain Self-management in Primary Care Patients With Depression and Musculoskeletal Pain. *JAMA.* 2009;301:2099-2110.
16. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA. Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician.* 2009;12:E35.
17. Institute of Medicine. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. 2011. Institute of Medicine.
18. Lamb SE, Hansen Z, Lall R et al. Group cognitive behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. *The Lancet.* 2010;375:916-923.
19. Smeets RJEM, Vlaeyen JWS, Hidding A, Kester ADM, van der Heijden GJMG, Knottnerus JA. Chronic low back pain: Physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain.* 2008;134:263-276.
20. Williams DA, Kuper D, Segar M, Mohan N, Sheth M, Clauw DJ. Internet-enhanced management of fibromyalgia: A randomized controlled trial. *Pain.* 2010;151:694-702.
21. Kerns RD, Rosenberg R, Jamison RN, Caudill MA, Haythornthwaite J. Readiness to adopt a self-management approach to chronic pain: the Pain Stages of Change Questionnaire (PSOCQ). *Pain.* 1997;72:227-234.
22. Gersh E, Arnold C, Gibson SJ. The Relationship Between the Readiness for Change and Clinical Outcomes in Response to Multidisciplinary Pain Management. *Pain Medicine.* 2011;12:165-172.
23. Kerns RD, Rosenberg R. Predicting responses to self-management treatments for chronic pain: application of the pain stages of change model. *Pain.* 2000;84:49.
24. Jensen MP, Nielson WR, Kerns RD. Toward the development of a motivational model of pain self-management. *J Pain.* 2003;4:477-492.
25. Biller N, Arnstein P, Caudill MA, Federman CW, Guberman C. Predicting completion of a cognitive-behavioral pain management program by initial measures of a chronic pain patient's readiness for change. *Clin J Pain.* 2000;16:352-359.
26. Keefe FJ. Cognitive behavioral therapy for managing pain. The Clinical Psychologist 1996;49:4-5.
27. Lihong L, Chu W, Langford J, Schapire RE. A contextual bandit approach to personized news article recommendation. WWW 2010:661-70.
28. Kroenke, K. & Spitzer, R. (2002). The PHQ-9: A New Depression Diagnostic and Severity Measure. Psychiatric Annals, 39(9),1-7.
29. Sheehan DV, Lecrubier Y, Harnett-Sheehan K, et al. The Mini International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview. Journal of Clinical Psychiatry 1998;59:22-33.
30. Tariq SH, Tumosa, Chibnall JT, Perry MH, Morley JE. Comparison of the Saint Louis University mental status examination and the Mini Mental state examination for detecting dementia and mild neurocognitive disorder - a pilot study. American Journal of Geriatric Psychiatry 2006;14:900-10.

31. Singleton P, Wadsworth M. Consent for the use of personal medical data in research. BMJ 2006;333:255-8.
32. Junghans C, Feder G, Hemingway H, Timmis A, Jones M. Recruiting patients to medical research: double blind randomised trial of "opt-in" versus "opt-out" strategies. BMJ Online 2005;331:940+.
33. Krebs EE, Lorenz KA, Bair MJ, Damush TM, Wu J, Sutherland JM, Asch SM, Kroenke K. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *J Gen Intern Med.* 2009; 6:733-8.
34. Li L, Chu W, Langford J, Schapire RE. A contextual-bandit approach to personalized news article recommendation. WWW 2010;Apr 26-30:661-70.
35. Abbasi-Yadkori Y, Pal D, Szepesvari C. Improved algorithms for linear stochastic bandits. In Advances in Neural Information Processing Systems (NIPS). 2011.
36. Chu W, Li L, Reyzin L, Schapire RE. Contextual bandits with linear payoff functions. In International Conference on Artificial Intelligence and Statistics Available at: http://machinelearningwustledu/mlpapers/paper_files/AISTATS2011_ChuLRS11pdf 2011:208-14.
37. Auer P. Using confidence bounds for exploitation-exploration trade-offs. Journal of Machine Learning Research 2002;3:397-422.
38. Carroll KM, Nich C, Silfry RL, et al. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. Drug and Alcohol Dependence 2000;57:225-38.
39. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. The Journal of Pain 2008;9:105-21.
40. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;1:9-19.
41. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine 2000;25:115-24.
42. Kazis LE, Lee A, Spiro A, et al. Measurement comparisons of the medical outcomes study and veterans SF-36® health survey. Health Care Financing Review 2004;25:43.
43. Bhandari A, Wagner T. Self-reported utilization of health care services: improving measurement and accuracy. Medical Care Research and Review 2006;63:217-35.
44. Piette JD, Rosland AM, Marinec NS, Striplin D, Bernstein SJ, Silveira MJ. Engagement with automated patient monitoring and self-management support calls: experience with a thousand chronically-ill patients. Medical care 2013;51:216-23. PMID: 23222527.
45. Borovec TD, Nau CD. Credibility of analog therapy rationales. Experimental Psychology 1972;3:257-69.
46. Clark ME, Gironda RJ, Young RW. Development and validation of the Pain Outcomes Questionnaire-VA. Journal of Rehabilitation Research and Development 2003;40:381-96.
47. Chew LD, Bradley KA, Boyko E, J. Brief questions to identify patients with inadequate health literacy. Family Medicine 2004;36:588-94.
48. Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 2007;147:478-91.
49. Sullivan MJL, Bishop S, Pivik J. The pain catastrophizing scale: development and validation. Psychological Assessment 1995;7:524-32.
50. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJW. Reporting of noninferiority and equivalence trials: an extension of the CONSORT statement. JAMA 2006;295:1152-60.
51. Bombardier M, Hayden J, Beaton DE. Minimal clinical importance difference: low back pain outcome measures. Rheumatology 2001;28:431-8.
52. Glasgow RE, McKay HG, Piette JD, Reynolds KD. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? Patient Educ Couns 2001;44:119-27.
53. Mauskopf JA, Sullivan SD, Annemans L. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good resarch practices--budget impact analysis. Value Health 2007;10:336-47.
54. Rubin DB. Multiple imputation after 18+ years (with discussion). Rubin, DB (1996) Multiple imputation after 18+ years (with discussion) Journal of the American Statistical Association, 91, 473-489 1996;91:473-89.

55. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51:1173-82.
56. Miller W, Crabtree B. Clinical research: a multimethod typology and qualitative roadmap. In: Crabtree B, Miller, WL, ed. Doing qualitative research. Thousand Oaks: Sage Publications; 1999.
57. Keller et. al, (2004). Validity of the Brief Pain Inventory for Use in Documenting the Outcomes of Patients With Noncancer Pain. *Clinical Journal of Pain, 20(5), 309-318.*