Design, recruitment outcomes, and sample characteristics of the Strategies for Prescribing Analgesics Comparative Effectiveness (SPACE) trial

Erin E. Krebs, MD, MPH[a,b,] Agnes C. Jensen, MPH[a,] Sean Nugent, BA[a,] Beth DeRonne, PharmD[a,] Indulis Rutks, BS[a,] David Leverty, BA[a,] Amy Gravely, MA[a,] Siamak Noorbaloochi, PhD[a,b,] Matthew J. Bair[c,d,e,] Kurt Kroenke[c,d,e]

a) Center for Chronic Disease Outcomes Research, Minneapolis VA Health Care System; Minneapolis, MN; Postal address: One Veterans Drive (152) Minneapolis, MN 55417 USA
b) Department of Medicine, University of Minnesota Medical School, Minneapolis, MN; Postal address: 420 Delaware Street SE MMC 194 Minneapolis, MN 55455 USA
c) Center for Health Information and Communication, Roudebush VA Medical Center, Indianapolis, IN; Postal address: 1481 W. 10th St. Indianapolis, IN 46202 USA
d) Department of Medicine, Indiana University School of Medicine, Indianapolis, IN; Postal address: 1110 W. Michigan Street Long Hall 417 Indianapolis, IN 46202 USA
e) Regenstrief Institute, Inc., Indianapolis, IN; Postal address: 1101 West 10th Street Indianapolis, IN 46202 USA

Corresponding author:
Erin E. Krebs, MD, MPH
Minneapolis VA Health Care System (152)
One Veterans Drive
Minneapolis, MN 55417
Phone: 612-629-7559
Fax: 612-727-5699
Email: erin.krebs@va.gov

_____

ABSTRACT

This manuscript describes the study protocol, recruitment outcomes, and baseline participant characteristics for the Strategies for Prescribing Analgesics Comparative Effectiveness (SPACE) trial. SPACE is a pragmatic randomized comparative effectiveness trial conducted in multiple VA primary care clinics within one VA health care system. The objective was to compare benefits and harms of opioid therapy versus non-opioid medication therapy over 12 months among patients with moderate-to-severe chronic back pain or hip/knee osteoarthritis pain despite analgesic therapy; patients already receiving regular opioid therapy were excluded. Key design features include comparing two clinically-relevant medication interventions, pragmatic eligibility criteria, and flexible treat-to-target interventions. Screening, recruitment and study enrollment were conducted over 31 months. A total of 4491 patients were contacted for eligibility screening; 53.1% were ineligible, 41.0% refused, and 5.9% enrolled. The most common reasons for ineligibility were not meeting pain location and severity criteria. The most common study-specific reasons for refusal were preference for no opioid use and preference for no pain medications. Of 265 enrolled patients, 25 withdrew before randomization. Of 240 randomized patients, 87.9% were male, 84.1% were white, and age range was 21-80 years. Past-year mental health diagnoses were 28.3% depression, 17% anxiety, 9.4% PTSD, 7.9% alcohol use disorder, and 2.6% drug use disorder. In conclusion, although recruitment for this trial was challenging, characteristics of enrolled participants suggest we were successful in recruiting patients similar to those prescribed opioid therapy in usual care.

Keywords: opioid analgesics; analgesics; chronic pain; back pain; osteoarthritis

1. Introduction

Chronic musculoskeletal pain conditions are among the most prevalent and disabling of chronic diseases.[1] As a result of a decades-long expansion in opioid analgesic prescribing, treatment of chronic pain with long-term opioid therapy has become common practice. [2,3] Evidence has not kept up with practice. Although short-term trials suggest that opioids can produce small short-term reductions in pain intensity,[4,5] evidence for long-term effectiveness is largely absent. A systematic review conducted in 2014 and updated in 2016 found no controlled opioid trials that examined effects on pain, function or quality of life at one year or longer.[6] Evidence regarding harms of long-term opioid therapy is mostly derived from retrospective observational studies and few studies have published data on patient-reported adverse symptoms.

A 2014 National Institutes of Health Pathways to Prevention panel on the role of opioids in the treatment of chronic pain identified "a clear need for well-designed longitudinal studies of effectiveness and safety of long-term opioid use in the management of chronic pain."[7] The panel report noted barriers to conducting standard randomized controlled trials, such as difficulty recruiting and retaining participants, and suggested alternate and pragmatic study designs to address evidence gaps and inform clinical decision-making.

This manuscript describes the study protocol and recruitment outcomes for the Strategies for Prescribing Analgesics Comparative Effectiveness (SPACE) trial, a pragmatic randomized comparative effectiveness trial conducted in a VA health care system. The objective of the trial was to compare benefits and harms of opioid versus non-opioid medication therapy over 12 months. The study was designed to be maximally relevant to pain management in primary care practice and inform an important decision point in the care of patients with moderate-to-severe chronic pain—whether or not to initiate long-term opioid therapy. We chose non-opioid medication therapy as the comparison intervention for opioid therapy because we believed it to be the most clinically-relevant alternative pain management strategy.

2. Methods

2.1 Design overview

SPACE is a pragmatic randomized comparative effectiveness trial with masked outcome assessment comparing opioid therapy versus non-opioid medication therapy over 12 months for patients with chronic back pain or hip or knee osteoarthritis pain. Both interventions (opioid therapy and non-opioid therapy) were delivered using a telecare collaborative management model and a treat-to-target approach aiming for 30% improvement in pain and progress toward individualized goals.

Patients treated for back, hip, or knee pain at primary care clinics in the Minneapolis Veterans Affairs (VA) Health Care System were identified through automated searches of the electronic medical record (EMR) and screened by telephone for eligibility. Patients with chronic back pain or hip or knee osteoarthritis pain of at least moderate severity despite analgesic use were eligible. Eligible patients who provided informed consent and completed a baseline assessment visit were randomized to either the opioid therapy arm or the non-opioid therapy arm. Randomization was

at the individual level, restricted by primary pain location (back or hip/knee), and implemented using a computer program to conceal allocation.

Each arm included multiple medications that were available by prescription within the VA health care system during the study period. In both arms, pain medication management was provided for 12 months, with medication therapy tailored to individual patient preferences and adjusted within the assigned treatment arm to achieve therapeutic targets. Outcomes were evaluated by masked assessors at 3, 6, 9, and 12 months after enrollment. To offset costs of participation, participants were paid $100 in cash after completion of baseline, 6, and 12 month assessments and $20 by check after completion of the 3 and 9 month assessments. No incentives were provided for attendance at clinical intervention visits. The study was approved by the Minneapolis VA Institutional Review Board.

2.2 Pragmatic trial design

The intent for the SPACE trial was to be as pragmatic as possible, evaluating benefits and harms of opioid therapy in close to usual care conditions, and to maximize applicability to primary care practice settings. We considered pragmatic design principles outlined in the pragmatic-explanatory continuum indicator summary (PRECIS) in SPACE trial development.[8] Table 1 and Figure 1 show pragmatism of trial domains according to the updated PRECIS-2 tool.[9]

Table 1: PRECIS-2 scores for trial domains

	Domain	Score	Rationale
1	Eligibility criteria	5	We aimed to enroll a broad sample of primary care patients who would potentially receive either opioid or non-opioid medications in practice, including those with mental health and medical diagnoses that typically lead to exclusion from analgesic clinical trials.[10] Exclusion criteria were the minimum necessary to ensure patient safety and internal validity.
2	Recruitment path	1	Intensive recruitment efforts were necessary because of unappealing aspects of the trial, especially randomization to opioid or non-opioid therapy for 12 months.
3	Setting	5	We identified patients who attended a visit at multiple primary care clinics within a local health care system.
4	Intervention organization*	3	The most pragmatic approach would have been primary care prescribing of intervention drugs, but we didn't think this was feasible because of established analgesic prescribing habits and opinions.[11,12] To ensure participants received the treatment to which they were assigned, a study clinical pharmacist managed study medications, using an approach similar to pharmacist management of other chronic diseases in VA primary care.[13]
5	Flexibility: intervention delivery*	5	Both opioid and non-opioid therapy interventions were delivered with a high degree of flexibility in choice and dosing of medications.
6	Flexibility: intervention	5	We did not actively restrict non-study prescribing. Participants

	adherence*		were not withdrawn for receiving non-study medications. Participants were allowed nonpharmacological pain treatments. Patients who did not tolerate study medications were encouraged to continue in the study. We used standard clinical adherence monitoring practices.
7	Follow-up	2	Outcome assessment included extensive patient-reported data collection every 3 months to obtain long-term data and minimize missing data for patient-reported measures. We offered alternate response modes or short forms if needed.
8	Outcome	5	The primary outcome is long-term and highly relevant to patients with chronic pain.
9	Analysis	5	We used intent-to-treat analysis of data for all randomized participants with any follow-up outcomes.

*Both arms had the same approach to organization, flexibility, and adherence. Scores for each domain are rated from 1 (least pragmatic) to 5 (most pragmatic).

2.3 Eligibility

The target population was primary care patients with moderate-to-severe chronic back pain or hip or knee osteoarthritis pain despite analgesic therapy. Inclusion criteria were further defined as follows:

- *Moderate-severe pain intensity and interference with function*: The PEG scale, a 3-item multidimensional pain measure, was used to define pain severity.[14] Eligibility criteria were a) a score of ≥5 on the average pain intensity item and b) a mean score of ≥5 on the two functional items (interference with enjoyment of life; interference with general activity).
- C*hronic back pain or hip or knee osteoarthritis pain as the primary pain condition:* These conditions were selected because they are among the most common indications for long-term opioid therapy.[15-18] Chronic pain was defined as pain present nearly every day or every day for ≥6 months. The primary pain diagnosis was identified by patient self-report during the screening interview and confirmed by a targeted review of each patient's medical records. At least minimal documented evidence of the diagnosis (such as inclusion in a primary care problem list or visit note) was required. Patients with both chronic back pain and hip/knee osteoarthritis pain were stratified based on the most bothersome condition at the time of eligibility screening.
- *Despite analgesic therapy:* To ensure participants were appropriate for opioid therapy, eligibility criteria required ongoing pain despite use of at least one analgesic medication.

Patients were excluded if they were currently receiving long-term opioid therapy, had conditions that could interfere with outcome assessment, or had contraindications to all drugs in either arm. Specific exclusion criteria were as follows:

- *Current long-term opioid therapy:* EMR data were used to identify patients with any long-acting opioid (i.e., methadone, transdermal fentanyl, or any sustained-release opioid) dispensed in the prior three months or any prescription for ≥ 60 tablets/capsules of short-acting DEA schedule II or III opioid dispensed in the prior three months.
- *Conditions that could interfere with outcome assessment*: Patients were excluded if they met any of the following criteria: a) schizophrenia, bipolar disorder, or other psychotic disorder;

b) moderately severe cognitive impairment, defined as ≥2 errors on a brief cognitive screener;[19] c) anticipated back, knee, or hip surgery within 12 months; d) anticipated life expectancy of less than 12 months; or e) unable to complete outcome assessments for 12 months for any reason.

- *Contraindications to all drugs in either arm*: In general, contraindications for individual medications include known allergy, previous serious adverse effect, or failure of a previous adequate trial. Because of the diversity of drug classes in the non-opioid therapy arm, contraindications to all drugs in that arm were much less common than contraindications to all drugs in the opioid therapy arm. Class-level contraindications to opioid therapy were consistent with VA/DoD clinical practice guidelines[20] and included the following:
  - *Psychiatric instability*: This was defined as current uncontrolled severe depression, severe PTSD, or suicidal ideation and patients were excluded only if they had severe symptoms that were not actively treated. Patients with severe untreated symptoms on the initial screening interview were provided with information about accessing mental health care and offered the opportunity of rescreening later.
  - *Active substance use disorder (SUD) or history of opioid use disorder*: We did not exclude patients with SUD (other than opioid use disorder) in remission or those with recent illicit drug use but no apparent SUD. In addition to information obtained during the eligibility screening process, urine drug test (UDT) results obtained at the baseline assessment were considered in evaluation of this criterion. Patients who reported drug use or had a UDT positive for an illicit drug were asked if they were willing to discontinue drug use during the study. If they agreed, follow-up questions from the NIDA-modified ASSIST questionnaire[21] were asked and, for those with a positive UDT, follow-up testing was done prior to randomization. Participants who had a lower risk substance involvement score (0-3) on the ASSIST questionnaire and a negative follow-up UDT were randomized to a treatment group and continued in the study; those with a substance involvement score ≥ 4 and a negative follow-up UDT were considered by the PI on an individual basis.
  - *Documented or suspected prior misuse or diversion of prescribed controlled substances*
  - *Current regular benzodiazepine use*: Patients were excluded if they reported using any benzodiazepine more than once or twice in a typical week.
  - *Fibromyalgia and migraine headache*: Because opioid therapy may worsen these conditions, patients were excluded if they reported fibromyalgia or headaches interfering with activities ≥3 times a week.

2.4 Recruitment, eligibility screening, and enrollment

Primary care providers (PCPs) were informed of the study and asked to provide written authorization for the study to approach patients on their panels for participation. Over the recruitment period, 77 PCPs were contacted for permission; of these, 62 agreed, 8 refused, and 7 didn't respond. The PI's primary care panel was not contacted. Potentially eligible patients were identified from primary care panels through administrative database searches, which identified patients who had been given a relevant diagnosis at a primary care visit in the prior month and who were not currently dispensed long-term opioids.

Each month during the enrollment period, database searches were updated and a sample of patients from among those identified was selected for telephone screening. To maximize recruitment of potentially eligible women, we selected all women and a random sample of men for screening. After four months of recruitment, because many patients were being screened out and not eligible due to mild pain, we also added a database search selection preference for patients with a pain score ≥5 at the most recent visit. Specifically, men with pain score ≥5 were preferentially selected for screening, along with all women. If fewer than the target number of patients was identified for the month, additional men with pain scores of 0-4 were selected at random from among the remainder to reach the target number. At four months, we also determined that PTSD screening questions could be removed from the telephone eligibility screening interview due to low yield and we determined that recruitment efforts were exceptionally low-yield for patients > 80 years of age or living > 50 miles from the Minneapolis VA Medical Center, so we stopped selecting these patients for eligibility screening. Patients could also be referred directly from primary care to study eligibility screening.

Potentially eligible patients selected for screening were mailed an invitation letter describing the study and providing instructions for opting out of additional recruitment contacts. If they did not opt out, study personnel attempted to contact them by telephone to determine interest in participating and to conduct an eligibility interview. We made at least six attempts on two different days at three different times of the day before determining that a patient was unreachable. The screening interview assessed eligibility criteria in the following order: 1) pain condition and severity; 2) medication use; 3) depression and PTSD symptoms; 4) alcohol and drug use; 5) surgical plans and availability for 12 months; 6) cognitive screening test. Subsequently, the principal investigator completed a brief medical record review checklist to verify eligibility criteria. A face-to-face enrollment visit was scheduled for patients who were eligible and interested in participating.

At the enrollment study visit, study personnel completed the informed consent process with patients. After participants provided written informed consent, they were asked to complete baseline interviews, questionnaires, and testing, including urine and blood samples. If participants reported having an additional primary care or pain care provider outside the VA healthcare system, they were asked to provide authorization to contact non-VA health care providers for the duration of the study. We sent non-VA care providers a letter describing the study, a link to study registration at clinicaltrials.gov, and a request for past-year medical records. The first intervention visit with the pharmacist care manager was scheduled at the end of the enrollment visit.

2.5 Randomization, allocation concealment, and masking

Randomization was conducted at the first visit with the intervention pharmacist care manager; for most participants, this occurred approximately one week after the enrollment/baseline assessment visit. This delay allowed intervention clinicians to review urine drug testing (UDT) results prior to randomization. The pharmacist care manager completed a check of patients' understanding of the study and provided an opportunity to ask additional questions or express concerns prior to randomization.

A repeated measure completely randomized design with replication was used for the randomized treatment assignment and further control of subject variability. Randomization was restricted by primary pain location (back or hip/knee) as blocks. To ensure balanced numbers of participants with back pain or hip/knee pain, within-block randomizations were conducted in sequential 6-patient runs. To maintain allocation concealment, the pharmacist care manager randomized the patient using a study computer application program prepared by the study data team.

Masking participants to treatment assignment was not feasible due to the complexity of the medication therapy interventions. The study pharmacist care manager and supervising physician investigators were also aware of treatment assignment. Study medications were included in participants' EMR medication list, marked by an "INV" prefix (for "investigational"), and thus visible to all treating clinicians. Outcome assessors were masked to treatment arm assignment and did not enter the EMR. Study intervention clinicians were masked to outcome assessment data. Data team members were the only personnel with complete access to the study computer application and databases; other study personnel had restricted access appropriate to their role in the study.

2.6 Interventions

Participants were randomized to receive one of two active interventions, either opioid therapy or non-opioid medication therapy, for 12 months.

2.6.1 Overview of study interventions

In both intervention arms, participants received active medication management delivered using a telecare collaborative management model that has been previously evaluated in VA primary care.[22,23] The two arms differed only in the medications included. Each arm included multiple medications that were available by prescription within the VA health care system during the study period (see Table 2).

Table 2: Prescribing strategies and medications for intervention arms

Step	Opioid therapy arm	Non-opioid medication arm
First	Morphine IR*	Acetaminophen*
	Oxycodone IR	Naproxen
	Hydrocodone/acetaminophen	Ibuprofen
		Meloxicam
		Celecoxib
		Salsalate
		Etodolac
		Sulindac
		Diclofenac
		Diclofenac topical gel
Second	Morphine SR	Nortriptyline
	Oxycodone SA	Amitriptyline
		Gabapentin
		Lidocaine topical
		Capsaicin topical
Third	Fentanyl transdermal	Pregabalin
		Duloxetine
		Tramadol

*Preferred initial medication selection.

Pain medication therapy was tailored to individual patient preferences and adjusted within the assigned arm to achieve therapeutic targets. Participants entered the prescribing strategies at different levels, depending on their baseline medications and past medication histories. Their baseline medications may have been discontinued, adjusted to achieve an adequate therapeutic trial, or continued with the addition of adjunctive drugs, depending on their individual history of medication use (including dosing, scheduling, and adherence), therapeutic response, and adverse effects. In general, each step of the prescribing strategies was tried before moving to the next step. The study protocol outlined principles and preferred practices for medication prescribing decision points, which are outlined for each arm below.

Strategies employed to address medication tolerability problems depended on the adverse effect and patient preferences and included: 1) switching to a different medication; 2) lowering the medication dosage; 3) decreasing the frequency of dosing; 4) using medications on an as-needed only basis; or 5) prescribing medications for adverse symptom management.

When participants did not demonstrate a response to maximized therapy within their assigned arm, regimens were adjusted to reduce the risk for medication-related harms by switching from scheduled to as-needed dosing or reducing dosage (if applicable). Participants who desired discontinuation of all medications in their assigned arm were assisted with tapering to

discontinuation (if applicable) and transition back to their pre-enrollment medication regimen. Study clinicians did not manage or adjust off-protocol pain medications.

2.6.2 Opioid therapy

Participants assigned to the opioid therapy prescribing strategy were initially started on immediate-release (IR) opioids, either morphine IR, hydrocodone/acetaminophen, or oxycodone IR. Opioid dosage was titrated to achieve pain treatment goals, with adjustments approximately every 4 weeks, to a maximum daily dosage of 100 morphine-equivalent (ME) mg. For participants without a clear response at a daily dosage of 60 ME mg, rotation to a different opioid was considered. The opioid rotation process was guided by recommendations and equianalgesic dosing tables in the 2010 VA-DoD opioid guidelines.[20] When a simpler dosing regimen or more consistent analgesia was preferred, long-acting opioids were considered. Single opioid therapy was preferred, but dual therapy with both a scheduled sustained-action (SA) opioid and an as-needed IR opioid was considered according to participant needs and preferences. All participants assigned to the opioid therapy arm were prescribed a prophylactic bowel regimen to prevent constipation, most often docusate/sennosides or polyethylene glycol powder. Dose adjustments to these medications were made based on individual response.

2.6.3 Non-opioid medication therapy

Participants assigned to the non-opioid prescribing strategy were initially treated with a first-line analgesic drug—acetaminophen or a non-steroidal anti-inflammatory drug (NSAID) medication—unless first-line choices were clinically inappropriate, based on past medication trials and contraindications. Adjuvant oral therapies were usually added to an initial regimen, but were considered as initial options when first-line medications were deemed inappropriate. Topical products were considered when oral regimens were insufficient, poorly tolerated, or clinically less appropriate. Tramadol, an atypical opioid with mu opioid agonist effects and serotonin-norepinephrine effects that was not a controlled substance in the US when this study was launched, was included in the third-line and considered only when participants did not respond to other medications or combinations of medications. Failure to respond to an adequate trial of one drug in a given class did not eliminate other members of that drug class from consideration. Diclofenac 1% gel was added in the last year of the trial after it was added to the VA formulary; it was considered a first-line or adjuvant option when oral NSAIDs could not be used. A low-dose proton pump inhibitor was prescribed to participants who were started on an oral NSAID and met gastrointestinal prophylaxis criteria from the Arthritis Foundation quality indicator set for safe use of NSAIDs.[24]

2.6.4 Intervention delivery

Both interventions were delivered using a telecare collaborative management approach tested in the Stepped Care to Optimize Pain Care Effectiveness (SCOPE) study and described in detail previously.[22] This approach has two main components: optimized analgesic management by a care manager/physician team and symptom monitoring. For this trial, a single clinical pharmacist was the care manager and central intervention provider for both arms. Decisions about medication therapy were made during regular case management meetings of the pharmacist care

manager and study physician. Non-controlled medications were prescribed by the pharmacist care manager and controlled substances were prescribed by a study physician. Primary care clinicians were informed about patient progress and medication changes using care management notes entered in the EMR, which is consistent with usual consultant practice in the health care system.

Participants were scheduled for their first intervention visit with the pharmacist care manager approximately one week after study enrollment. During this visit, a detailed history was obtained of prescription medication, over-the-counter medication, and herbal, vitamin, and supplement use. Participants' preferences and concerns regarding pain medications were discussed and participants were asked to identify a goal pain score as well as at least one activity goal. Follow-up intervention visits were scheduled approximately every four weeks until a stable medication regimen was established and then every four-to-twelve weeks, based on participant preference. The protocol specified that visits should be conducted in-person at least every six months when possible; other follow-up visits were scheduled in-person or by telephone, depending on participant preference.

An interactive voice response (IVR) telephone system was used for symptom monitoring. IVR telephone calls were used for interval reassessment of pain and adverse effects and were usually scheduled seven days prior to intervention follow-up visits. The IVR script included the 3-item PEG pain scale, a patient global impression of pain change rating, the PHQ-4 item depression and anxiety screener, and questions about side effects, preference for medication change, and desire for a return call.[14,25] Due to technical limitations of the IVR system used in this study, response options for the PEG scale had to be modified. Instead of the validated 0-10 rating scale, a modified 1-7 rating scale was used. Automated symptom monitoring by IVR was used solely as a clinical tool to assist with monitoring and adjusting therapy. It was not used to assess research outcomes.

Decisions to change pain medications were made in a structured shared process. At each visit, symptom monitoring was reviewed and participants were asked about pain, sleep, side effects, and progress toward self-identified activity goals. Participants were encouraged to state their preferences regarding pain medications; these preferences, along with progress toward individualized goals, were the primary drivers of changes in therapy. When possible, participants were provided with options for medication adjustments and allowed to select their preferred option. Medication adjustments were made until pain management goals were achieved or all treatment options within the assigned arm were exhausted.

2.6.5 Clinical adherence monitoring

Multiple adherence monitoring strategies are recommended by clinical practice guidelines to reduce opioid-related harms.[20,26] These strategies include written medication agreements outlining responsibilities of patients and prescribers; checking state prescription drug monitoring programs (PDMP); using urine drug testing; and assessing behaviors potentially associated with non-adherence. In this trial, we applied recommended monitoring strategies as follows: 1) a written medication agreement based on the VA model opioid therapy agreement was reviewed and signed at the first intervention visit; 2) the state PDMP website was checked at each visit and

anytime there was a clinical reason for concern; 3) urine drug testing was done at baseline and anytime there was a clinical reason for concern; and 4) a checklist of potential non-adherence behaviors was completed by the pharmacist care manager at each visit.[27] We used the same approach to adherence monitoring in both arms to avoid introducing attention bias.

To assist with general medication adherence, all participants were provided with a seven-day pillbox. An updated medication list and letter was provided to participants after every medication change and an individualized dosing calendar was used for more complicated medication titrations and tapers.

2.7 Measures

Study measures included patient-reported outcomes at baseline and at 3, 6, 9, and 12 months and test-based outcomes (physical, cognitive, and laboratory testing) at baseline and at 6 and 12 months. A full battery of patient-reported outcomes were collected by in-person interview and written questionnaire at baseline and 6 and 12 months; a subset was collected by telephone interview at 3 and 9 months. To minimize missing primary outcome data, patients who were unable to attend in-person follow-up assessment visits were asked to complete at least a core set of measures by telephone or, if they were unable to complete telephone interviews, by mailed questionnaire. Table 3 displays measures and timing of administration.

Outcomes were collected by assessors who were masked to treatment arm assignment, including the study coordinator, one primary research assistant (RA), and one back-up RA. To establish rapport with participants, the coordinator and primary RA both participated in the baseline visit. The coordinator held training sessions with the RAs prior to and during data collection to ensure outcomes were assessed in a consistent manner.

Table 3: Measures and assessment schedule

Measure	Data source	Schedule (month)				
		0	3	6	9	12
Patient-reported measures						
Brief Pain Inventory (BPI) interference and severity scales[28,29]	Interview	X	X	X	X	X
Abbreviated Roland-Morris disability scale[30]	Interview	X		X		X
Global rating of change*[31]	Interview		X	X	X	X
Veterans RAND 12 Item Health Survey (VR-12)[32,33]	Interview	X	X	X	X	X
Modified adverse symptom checklist[34]	Interview	X	X	X	X	X
Multidimensional Fatigue Inventory[35,36]	Interview	X		X		X
PROMIS Sleep Disturbance short form[37]	Interview	X		X		X
Migraine Disability Assessment[38]	Interview	X		X		X
Patient Health Questionnaire (PHQ)-8[39]	Interview	X	X	X		X
Generalized Anxiety Disorder (GAD)-7[40]	Interview	X	X	X		X
Primary Care-PTSD screener[41]	Interview	X				X
Arizona Sexual Experience scale†[42]	Questionnaire	X		X		X
Alcohol Use Disorders Identification Test (AUDIT) †[43]	Questionnaire	X		X		X
Drug and tobacco use †	Questionnaire	X		X		X
Pain Medication Attitude Questionnaire†[44]	Questionnaire	X				X

Measure	Data source	Schedule (month)				
		0	3	6	9	12
Modified Prescribed Opioids Difficulties Scale (PODS)†‡[45]	Questionnaire	X				X
Treatment expectations and preferences‖	Interview	X				
Pain co-intervention checklist†	Questionnaire	X	X	X	X	X
Hospitalizations and emergency department visits	Interview/EMR		X	X	X	X
Falls	Interview		X	X		X
IU-TBANS cognitive test battery[46]	Interview	X		X		X
Physical tests						
Fullerton Advanced Balance scale[47]	Exam	X		X		X
Gait speed (6 meter)[48,49]	Exam	X		X		X
Chair stand test[50]	Exam	X		X		X
Grip strength[51]	Exam	X		X		X
Cold pressor testing[52]	Exam	X				X
Urine toxicology	Laboratory	X		X		X
Testosterone and estradiol	Laboratory	X				X

* "How would you describe your pain now, compared to when you started in our study?" Response options: much better, moderately better, a little better, no change, a little worse, moderately worse, much worse.

† Completed independently on paper questionnaire form.

‡ The PODS measure was adapted by changing "opiate" to "pain" medications in all questions.

‖ Developed specifically for this study as described in text.

2.7.1 Pain outcome measures

As recommended in expert consensus guidelines for pain clinical trials, we assessed multiple pain outcome domains.[53] The primary study outcome measure is the Brief Pain Inventory (BPI) interference scale, a seven-item measure of pain-related function. The main measure of pain intensity is the BPI severity scale. These BPI scales have been validated for use in numerous populations, including primary care patients with chronic pain.[28,29,53] The abbreviated 11-item Roland-Morris disability scale was used as a secondary measure of pain-related function.[30] A seven-grade patient-reported global rating of change question was used to assess patients' views of overall improvement or worsening in pain.[31]

2.7.2 Harm outcome measures

The primary measure of adverse effects is a modified adverse symptom checklist that assesses number and severity of common symptoms and participants' beliefs about whether or not symptoms are medication-related. The original checklist was developed to assess symptoms related to a variety of different medications among primary care patients.[34] For this study, we modified the checklist by adding five symptoms reflecting common patient-reported side effects of analgesic medications (i.e., memory problems, dry mouth, trouble concentrating, sweating, and weight gain)[54] not included in the original version. Secondary adverse outcomes included hospitalizations, emergency department visits, and falls. Hospitalizations and emergency department visits are identified by both medical record searches and patient report at outcome

assessment visits and counted from date of enrollment to 13 months after randomization. The World Health Organization system for standardized case causality assessment is used to determine whether the relationship between an event and analgesic use is probable or greater.[55] Falls were assessed by patient report at outcome assessment visits.

2.7.3 Additional secondary measures

Multiple secondary measures were assessed, as shown in Table 3. Drug and tobacco use were assessed with a single item screening test for drug use[56] and questions from the National Survey on Drug Use and Heath and NIDA-modified ASSIST.[21]

Participants were asked to complete ratings of treatment expectations and preferences. Items were based on previously published measures and modified for this study.[57,58] These included ratings of effectiveness and safety of the two study arms ("In general, how [effective/safe] do you consider [opioid medications/non-opioid medications] for long-term treatment of pain?" with response options 0 "not at all effective/safe" to 10 "most effective/safe possible"); expectations for personal benefit from study participation ("In terms of pain, how much improvement do you think is likely for you personally during this study?" with options 0 "no improvement" to 10 "a great deal of improvement"); preference for study arm assignment ("For this study, you will be assigned to a treatment group by chance. Now, imagine if you were given a choice between groups. Considering what you know so far, which treatment group would you choose?"); and strength of preference (You said you would prefer the _____ arm. How strong is your preference?" with options 0 "no preference" to 10 "strongest possible preference").

At in-person assessment visits, participants were asked to complete the Fullerton Advanced Balance scale, a 10-item test that evaluates multiple dimensions of balance in independent adults.[47,59] A chair stand test was conducted as the number of stands completed from a seated position in 30 seconds.[50] Gait speed was assessed with a timed six-meter walk protocol.[48,49] Grip strength was assessed with a Jamar hydraulic hand dynanometer.[51] Cold pressor testing was done by immersing the hand in a circulating cold water bath at 1° C for up to 5 minutes. [52]To assess cognition, the Indiana University Telephone-Based Assessment of Neuropsychological Status (IU-TBANS) was completed at in-person visits; this includes a battery of standard clinical instruments covering the following domains: attention, learning and recall, information processing speed, and verbal fluency.[46]

Because no single measure is diagnostic of opioid misuse, data from multiple sources were used to assess the occurrence of behaviors suggestive of opioid misuse or substance use disorder.[60] In addition to the measures in Table 3, clinical adherence data were collected at each intervention visit.

2.7.4 Process measures

The number and duration of pharmacist care manager visits and interactive voice response calls were extracted from electronic databases.

2.7.5 Qualitative interviews

Qualitative semi-structured follow-up interviews were conducted with a subset of 40 trial participants, including 10 responders and 10 non-responders from each arm. The purpose of these interviews was to gain a better understanding of participants' experiences in the trial and perceptions of the study interventions.

2.8 Statistical considerations

2.8.1 Sample size justification

The sample size (total n= 276) was based on estimated treatment effects on the primary outcome measure, the BPI Interference score. A between-group treatment difference of 1 point in the BPI Interference score represents a minimum meaningful intervention effect.[61] For our calculations, we assumed the standard deviation (SD) of the BPI Interference score in both arms would be 2.7, based on the intervention arm of a prior pain trial.[62] A sample size of 115 in each arm was estimated to provide 80% power to detect a 1-point difference in the mean BPI Interference score between groups, assuming 2-sided alpha of 0.05. To allow for up to 20% attrition, we aimed enroll 276 patients.

We estimated the planned sample size would provide 86% power to detect a 1-point difference in the BPI Severity score and 94% power to detect a 3-point difference in the Roland disability score. As a secondary measure of effectiveness, pain response was defined as a 30% improvement from baseline in the BPI Severity score and functional response as a 30% improvement in the BPI Interference score.[61] Power to detect a 20% difference in response rate (0.25 vs. 0.45) was estimated to be 0.86.

2.8.1 Statistical analysis

We will use an intent-to-treat analysis approach, including all participants in the arm to which they were originally assigned. We will check randomization-induced balance by comparing the observed baseline characteristics and potential confounders between the two arms. Medication use at the end of the trial in each arm will be presented in a descriptive table, including numbers using each drug and mean daily dose. For opioids, doses will be described as morphine-equivalent mg per day. For primary and secondary outcome measures, informative (non-parametric) summary statistics with corresponding standard deviations will be presented. Distributional assumptions for outcome measures and error terms will be examined using goodness-of-fit tests and diagnostic graphs.

We will initially compare group differences in scores on primary and secondary measures at each post-intervention time point, using a series of analyses of covariance (ANCOVA), which control for the baseline score as a covariate. For analyses of primary and secondary continuous measures (e.g., BPI scores, VR-12 scores), we will use mixed models of correlated within-block (i.e., back, hip/knee) repeated response measures. The models will have treatment group, categorical time points, and baseline values as fixed effects, with block and intercept as random effects. This Generalized Linear Mixed Modelling (GLMM) will be the conventional Gaussian mixed model for normally-distributed outcomes (e.g., BPI interference), a logistic mixed model with logit link

for the repeated-measures binary response outcomes (e.g., $\geq 30\%$ improvement in BPI interference from baseline), and a Poisson (or Negative Binomial) mixed model with log link for the count outcomes (e.g., number of hospitalizations). GLMM with the same set of fixed and random effects for other possibly skewed outcome measure distributions (e.g., gamma, beta, log-normal) will be used if the conventional transformation to normality is not satisfactory.

3. Results

3.1 Recruitment outcomes

Screening, recruitment and study enrollment were conducted over a 31-month period (June 2013-December 2016), with an average of eight new participants enrolled per month (range 2-14). Figure 2 shows results of the recruitment process. In summary, of the 4491 patients who were contacted by telephone, 2383 (53.1%) were ineligible, 1843 (41.0%) refused, and 265 (5.9%) enrolled. The most common reasons for ineligibility were not meeting pain location and severity criteria. Of the 1843 patients who refused, 696 (38%) cited no specific reason and 381 (21%) cited time constraints or distance. The most common study-specific reasons for refusal were preference for no opioid use (n=272, 15%) and preference for no pain medications (n= 229, 12%). Thirteen patients were counted as refusals because they did not show up for their enrollment appointment, after initially agreeing to participate.

Table 4 shows demographic and clinical characteristics of patients who were identified as potentially eligible by the pre-screening searches of the EMR, according to the screening process outcome. Those who were not selected for telephone screening are notably more often male, older, and more likely to be a veteran of World War II or the Korean War than the other groups; these differences are consistent with the preferential selection of women and the decision at four months to stop selecting patients > age 80 years or with home addresses > 50 miles from the medical center.

Table 4: Characteristics of patients selected for SPACE trial eligibility screening by screening outcome (n= 9,403)

	Not selected for telephone screening (n=3,836)	Unable to contact by telephone (n=1,076)	Screened and not eligible (n=2,383)	Screened and refused (n=1,843)	Enrolled (n=265)
Age in years, mean (SD)	67.4 ± 17.9	51.7 ± 15.9	58.1 ± 14.4	58.4 ± 15.0	57.5 ± 14.0
Sex, % male	96.7	85.9	89.0	90.5	87.9
Race, % white	85.6	77.8	81.0	82.0	84.1
Period of military service, %					
World War II, Korea	37.9	3.5	7.8	8.8	5.6
Vietnam War	35.8	36.8	50.7	50.2	50.6
Persian Gulf/Post-Vietnam	15.8	36.3	28.9	25.5	29.1
Iraq/Afghanistan	10.5	3.5	7.8	8.8	5.6
Qualifying pain diagnosis, %					

Back pain	52.0	57.3	54.9	58.9	63.4
Hip or knee osteoarthritis	55.3	51.2	51.9	50.0	44.9
Charlson comorbidity score, mean (SD) *	1.07 ± 1.6	0.48 ± 1.0	0.87 ± 1.4	0.90 ± 1.5	0.90 ± 1.3
Mental health diagnoses, * %					
Depression	17.2	24.0	28.9	23.2	28.3
Anxiety	10.0	18.1	18.2	14.5	17.0
PTSD	6.6	8.2	10.8	8.8	9.4
Alcohol use disorder	8.0	11.5	15.4	9.7	7.9
Drug use disorder	2.8	4.1	8.3	2.6	2.6

All values are from clinical and administrative databases.

* Based on ICD-9 and ICD-10 diagnoses in the prior 12 months.

3.2 Characteristics of participants

Of 265 veterans who enrolled in the study, 25 withdrew prior to randomization. Ten were withdrawn by the research team for likely substance use disorder after positive drug testing results (n=8) or inadequate understanding of the study (n=2). The remaining 15 chose to withdraw due to no longer interested (n=8), concerns about opioid use (n=5), transportation issues (n=1), and health problems (n=1).

Baseline characteristics of the 240 patients who were randomized are presented in Table 5. The majority of randomized participants were white and male; the age range was 21-80 years (mean age 58.3 years). When asked, prior to randomization, which medication group they would prefer, a majority (51.3%) of participants expressed no preference, 28.8% preferred the opioid group and 20.0% preferred the non-opioid group. Perceived medication effectiveness was rated 7.8 for opioids and 5.7 for non-opioids on a scale of 0 "not effective at all" to 10 "most effective possible." Perceived medication safety was rated 5.8 for opioids and 6.6 for non-opioids on a scale of 0 "not safe at all" to 10 "most safe possible". Participants reported high expectations for likely personal improvement during the study (mean 7.5 on a scale of 0 "no improvement" to 10 "a great deal of improvement").

Table 5: Baseline characteristics of 240 participants randomized in the SPACE trial

Age in years, mean ± SD	58.3 ± 13.7 (21-80)
Sex male, n (%)	208 (86.7)
Race/ethnicity, n (%)	
White	207 (87.0)
Black/African-American	18 (7.6)
Native American/Alaskan Native	4 (1.7)
Native Hawaiian/Pacific Islander	2 (0.8)
Hispanic/Latino	2 (0.8)
Asian	1 (0.4)
Multiple race/ethnicity categories	4 (1.7)

Education ≥ 4-year degree, n(%)	60 (25.0)
Employment, n(%)	
Employed for wages	81 (34.2)
Self-employed	14 (5.9)
Retired	99 (41.8)
Other	43 (18.1)
Primary qualifying pain diagnosis,[*] n (%)	
Back pain	156 (65.0)
Hip or knee pain	84 (35.0)
Substance use	
Current smoker, n (%)	38 (22.6)
AUDIT score ≥ 8, n (%)	5 (2.2)
Past year illicit drug use, n (%)	11 (4.9)
Pain measures	
BPI severity score, mean ± SD [0-**10**]	5.4 ± 1.3
BPI interference score, mean ± SD [0-**10**]	5.5 ± 1.9
Abbreviated Roland-Morris score, mean ± SD [0-**11**]	8.3 ± 2.2
Health related quality of life measures	
VR12 physical score, mean ± SD [**0**-100]	27.1 ± 8.1
VR12 mental score, mean ± SD [**0**-100]	47.6 ± 12.1
Mental health measures, n (%)	
Positive depression screen (PHQ-9 score ≥10)	53 (22.1)
Positive anxiety screen (GAD-7 score ≥10)	22 (9.2)
Positive PTSD screen (PC-PTSD ≥3)	50 (20.8%)
Arm preference, n (%)	
Unsure/no preference	123 (51.3)
Prefers opioid arm	69 (28.8)
Prefers non-opioid arm	48 (20.0)
Perceptions of treatment arms, mean ± SD	
Opioid effectiveness [**0**-10]	7.8 ± 2.0
Opioid safety [**0**-10]	5.8 ± 2.6
Non-opioid effectiveness [**0**-10]	5.7 ± 2.7
Non-opioid safety [**0**-10]	6.6 ± 2.8
Likely personal improvement, mean ± SD [**0**-10]	7.5 ± 1.9

All values are self-reported. For continuous scale scores, the minimum to maximum score range is given in brackets, with the worst possible score bolded.

*Participants with pain in both areas self-identified one location as primary.

4. Discussion

This trial is the first to evaluate long-term opioid therapy compared with long-term non-opioid medication therapy for chronic musculoskeletal pain. This study was designed as a pragmatic trial to overcome limitations common in prior opioid therapy trials. Key design features include comparing two clinically-relevant medication interventions that represent important competing

pain treatment approaches and incorporating flexibility into interventions to maximize relevance to clinical practice, benefits to participants, and, potentially, trial retention. The outcome assessment follow-up protocol was more structured (i.e., less pragmatic) in order to comprehensively assess multidimensional pain outcomes and conduct a detailed patient-reported and objective assessment of numerous potential opioid harms and adverse effects.

Experts have previously suggested that recruiting for a long-term trial of opioid therapy would likely be difficult; we found this prediction to be true. To enroll 265 patients, we attempted to contact 5567 and conducted telephone eligibility screening interviews with 4491. Our recruitment strategy allowed us to screen thousands of potentially eligible patients in a relatively efficient process. An additional advantage of this process is that we are able to describe characteristics of the population from which our study sample was drawn, thereby illuminating differences between the enrolled trial participants and non-enrolled patients.

Our pragmatic eligibility criteria were intended to enroll a study sample with characteristics similar to those of the patient population receiving opioids in primary care practice. Prior randomized controlled trials of opioid analgesics have typically excluded patients with current mental health diagnoses or a history of substance use disorder. In contrast, patients with these conditions often receive opioids in practice. For example, an observational study of VA patients with chronic pain receiving opioids in 2009-2011 reported diagnoses of PTSD in 20.9%, depression in 32%, and non-opioid substance use disorder in 7.8%.[16] Similarly, a study of VA patients who were new opioid recipients in 2011 reported diagnosed anxiety disorder in 9.4%, depressive disorder in 10.6%, and substance use disorder in 12.3%.[63] High psychiatric comorbidity (e.g., depression in 8.5%, anxiety in 6.9%) among opioid recipients was also demonstrated in a national database of health insurance claims of more than 10 million patients receiving opioid prescriptions.[64] Our rates of past-year PTSD (9.4%), anxiety disorder (17%), depression (28.3%), alcohol use disorder (7.9%) and drug use disorder (2.6%) diagnoses suggest that we were at least partially successful in recruiting patients similar to those receiving opioids in both VA and non-VA usual care.

In planning this trial, we considered an alternative design that would have included intervention delivery by patients' primary care clinicians. This approach would have been more pragmatic, but we decided against it for two main reasons. First, the alternative design would likely have necessitated clinician-level cluster randomization, which introduces additional risk of bias.[65,66] Second and most importantly, we were concerned about whether primary care clinicians would deliver the interventions with sufficient fidelity. Our decision was informed by a prior trial of two opioid prescribing strategies—stable dose versus liberal dose escalation—that used pain clinic providers to deliver the opioid prescribing interventions.[12] In that study, the treatment groups separated in terms of opioid dosing over time (17 mg/day mean difference between groups), but both groups had small-to-moderate dose escalations during the trial and, arguably, neither group received a true "stable dose" or "liberal dose escalation" prescribing strategy. Ultimately, we used research personnel for intervention delivery to maximize the likelihood that participants would receive the intervention to which they were assigned.

A major potential limitation of this study is that participants could not practicably be masked, due to the complexity and pragmatic nature of the medication interventions. Patient-reported

outcomes are thus subject to placebo and nocebo effects associated with expectations and perceptions. Pre-randomization arm preferences, treatment perceptions, and outcome expectations may provide potential insight into our eventual results. We find it interesting that participants had such high expectations of personal improvement at baseline. Less surprising to us, participants perceived opioids as having higher effectiveness for pain than non-opioid medications.

Acknowledgements

1.   Education IoMCoAPRC. Relieving pain in America: A blueprint for transforming prevention, care, education, and research. Washington, D.C.: National Academies Press (US); 2011.

2.   Boudreau D, Von Korff M, Rutter CM, et al. Trends in long-term opioid therapy for chronic non-cancer pain. *Pharmacoepidemiology and drug safety.* 2009;18(12):1166-1175.

3.   Caudill-Slosberg MA, Schwartz LM, Woloshin S. Office visits and analgesic prescriptions for musculoskeletal pain in US: 1980 vs. 2000. *Pain.* 2004;109(3):514-519.

4.   Abdel Shaheed C, Maher CG, Williams KA, Day R, McLachlan AJ. Efficacy, Tolerability, and Dose-Dependent Effects of Opioid Analgesics for Low Back Pain: A Systematic Review and Meta-analysis. *JAMA internal medicine.* 2016;176(7):958-968.

5.   da Costa BR, Nuesch E, Kasteler R, et al. Oral or transdermal opioids for osteoarthritis of the knee or hip. *The Cochrane database of systematic reviews.* 2014(9):CD003115.

6.   Chou R, Turner JA, Devine EB, et al. The effectiveness and risks of long-term opioid therapy for chronic pain: a systematic review for a National Institutes of Health Pathways to Prevention Workshop. *Annals of internal medicine.* 2015;162(4):276-286.

7.   Reuben DB, Alvanzo AA, Ashikaga T, et al. National Institutes of Health Pathways to Prevention Workshop: the role of opioids in the treatment of chronic pain. *Annals of internal medicine.* 2015;162(4):295-300.

8.   Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of clinical epidemiology.* 2009;62(5):464-475.

9.   Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *Bmj.* 2015;350:h2147.

10.   Chou R, Ballantyne JC, Fanciullo GJ, Fine PG, Miaskowski C. Research gaps on use of opioids for chronic noncancer pain: findings from a review of the evidence for an American Pain Society and American Academy of Pain Medicine clinical practice guideline. *The journal of pain : official journal of the American Pain Society.* 2009;10(2):147-159.

11.   Leverence RR, Williams RL, Potter M, et al. Chronic non-cancer pain: a siren for primary care--a report from the PRImary Care MultiEthnic Network (PRIME Net). *Journal of the American Board of Family Medicine : JABFM.* 2011;24(5):551-561.

12.   Naliboff BD, Wu SM, Schieffer B, et al. A randomized trial of 2 prescription strategies for opioid treatment of chronic nonmalignant pain. *The journal of pain : official journal of the American Pain Society.* 2011;12(2):288-296.

13.   Greer N, Bolduc J, Geurkink E, et al. *Pharmacist-Led Chronic Disease Management: A Systematic Review of Effectiveness and Harms Compared to Usual Care.* Washington (DC)2015.

14.   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.

15.   Butchart A, Kerr EA, Heisler M, Piette JD, Krein SL. Experience and management of chronic pain among patients with other complex chronic conditions. *The Clinical journal of pain.* 2009;25(4):293-298.

16.   Edlund MJ, Austen MA, Sullivan MD, et al. Patterns of opioid use for chronic noncancer pain in the Veterans Health Administration from 2009 to 2011. *Pain.* 2014;155(11):2337-2343.

17.   Goulet JL, Kerns RD, Bair M, et al. The musculoskeletal diagnosis cohort: examining pain and pain care among veterans. *Pain.* 2016;157(8):1696-1703.

18.   Von Korff M, Saunders K, Thomas Ray G, et al. De facto long-term opioid therapy for noncancer pain. *The Clinical journal of pain.* 2008;24(6):521-527.

19.   Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. *Medical care.* 2002;40(9):771-781.

20.   Group VDOTGW. Clinical Practice Guideline for Management of Opioid Therapy for Chronic Pain. In: Defense DoVADo, ed. Washington, D.C.2010.

21.   Abuse NIOD. Resource Guide: Screening for Drug Use in General Medical Settings. 2012; https://www.drugabuse.gov/publications/resource-guide-screening-drug-use-in-general-medical-settings/introduction.

22.   Kroenke K, Krebs E, Wu J, et al. Stepped Care to Optimize Pain care Effectiveness (SCOPE) trial study design and sample characteristics. *Contemporary clinical trials.* 2013;34(2):270-281.

23.   Kroenke K, Krebs EE, Wu J, Yu Z, Chumbler NR, Bair MJ. Telecare collaborative management of chronic pain in primary care: a randomized clinical trial. *Jama.* 2014;312(3):240-248.

24.   Saag KG, Olivieri JJ, Patino F, Mikuls TR, Allison JJ, MacLean CH. Measuring quality in arthritis care: the Arthritis Foundation's quality indicator set for analgesics. *Arthritis and rheumatism.* 2004;51(3):337-349.

25.   Kroenke K, Spitzer RL, Williams JB, Lowe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* 2009;50(6):613-621.

26.   Chou R, Fanciullo GJ, Fine PG, et al. Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. *The journal of pain : official journal of the American Pain Society.* 2009;10(2):113-130.

27.   Wu SM, Compton P, Bolus R, et al. The addiction behaviors checklist: validation of a new clinician-based measure of inappropriate opioid use in chronic pain. *Journal of pain and symptom management.* 2006;32(4):342-351.

28.   Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS. Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. *The Clinical journal of pain.* 2004;20(5):309-318.

29.   Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for chronic nonmalignant pain. *The journal of pain : official journal of the American Pain Society.* 2004;5(2):133-137.

30.   Stroud MW, McKnight PE, Jensen MP. Assessment of self-reported physical activity in patients with chronic pain: development of an abbreviated Roland-Morris disability scale. *The journal of pain : official journal of the American Pain Society.* 2004;5(5):257-263.

31.   Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H. Capturing the patient's view of change as a clinical outcome measure. *Jama.* 1999;282(12):1157-1162.

32.   Jones D, Kazis L, Lee A, et al. Health status assessments using the Veterans SF-12 and SF-36: methods for evaluating otucomes in the Veterans Health Administration. *The Journal of ambulatory care management.* 2001;24(3):68-86.

33. Selim AJ, Rogers W, Fleishman JA, et al. Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2009;18(1):43-52.

34. Weingart SN, Gandhi TK, Seger AC, et al. Patient-reported medication symptoms in primary care. *Archives of internal medicine.* 2005;165(2):234-240.

35. Lin JM, Brimmer DJ, Maloney EM, Nyarko E, Belue R, Reeves WC. Further validation of the Multidimensional Fatigue Inventory in a US adult population sample. *Population health metrics.* 2009;7:18.

36. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of psychosomatic research.* 1995;39(3):315-325.

37. Yu L, Buysse DJ, Germain A, et al. Development of short forms from the PROMIS sleep disturbance and Sleep-Related Impairment item banks. *Behavioral sleep medicine.* 2011;10(1):6-24.

38. Stewart WF, Lipton RB, Dowson AJ, Sawyer J. Development and testing of the Migraine Disability Assessment (MIDAS) Questionnaire to assess headache-related disability. *Neurology.* 2001;56(6 Suppl 1):S20-28.

39. Kroenke K, Spitzer RL, Williams JB, Lowe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *General hospital psychiatry.* 2010;32(4):345-359.

40. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine.* 2006;166(10):1092-1097.

41. Prins A, Ouimette P, Kimerling R, et al. The primary care PTSD screen (PC-PTSD): development and operating characteristics. *Prim Care Psychiatry.* 2003;9(1):9-14.

42. McGahuey CA, Gelenberg AJ, Laukes CA, et al. The Arizona Sexual Experience Scale (ASEX): reliability and validity. *Journal of sex & marital therapy.* 2000;26(1):25-40.

43. Fiellin DA, Reid MC, O'Connor PG. Screening for alcohol problems in primary care: a systematic review. *Archives of internal medicine.* 2000;160(13):1977-1989.

44. McCracken LM, Hoskins J, Eccleston C. Concerns about medication and medication use in chronic pain. *The journal of pain : official journal of the American Pain Society.* 2006;7(10):726-734.

45. Banta-Green CJ, Von Korff M, Sullivan MD, Merrill JO, Doyle SR, Saunders K. The prescribed opioids difficulties scale: a patient-centered assessment of problems and concerns. *The Clinical journal of pain.* 2010;26(6):489-497.

46. Unverzagt FW, Monahan PO, Moser LR, et al. The Indiana University telephone-based assessment of neuropsychological status: a new method for large scale neuropsychological assessment. *Journal of the International Neuropsychological Society : JINS.* 2007;13(5):799-806.

47. Rose DJ, Lucchese N, Wiersma LD. Development of a multidimensional balance scale for use with functionally independent older adults. *Archives of physical medicine and rehabilitation.* 2006;87(11):1478-1485.

48. Fransen M, Crosbie J, Edmonds J. Reliability of gait measurements in people with osteoarthritis of the knee. *Physical therapy.* 1997;77(9):944-953.

49. Terwee CB, Mokkink LB, Steultjens MP, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology.* 2006;45(7):890-902.

50. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Research quarterly for exercise and sport.* 1999;70(2):113-119.

51. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age and ageing.* 2011;40(4):423-429.

52. Mitchell LA, MacDonald RA, Brodie EE. Temperature and the cold pressor test. *The journal of pain : official journal of the American Pain Society.* 2004;5(4):233-237.

53. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2005;113(1-2):9-19.

54. Jonsson T, Christrup LL, Hojsted J, et al. Symptoms and side effects in chronic non-cancer pain: patient report vs. systematic assessment. *Acta anaesthesiologica Scandinavica.* 2011;55(1):69-74.

55. Centre UM. WHO-UMC system for standardised case causality assessment. 11. http://who-umc.org/Graphics/24734.pdf. Accessed 11/16/11.

56. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A single-question screening test for drug use in primary care. *Archives of internal medicine.* 2010;170(13):1155-1160.

57. George SZ, Robinson ME. Preference, expectation, and satisfaction in a clinical trial of behavioral interventions for acute and sub-acute low back pain. *The journal of pain : official journal of the American Pain Society.* 2010;11(11):1074-1082.

58. Linde K, Witt CM, Streng A, et al. The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. *Pain.* 2007;128(3):264-271.

59. Rutledge DN, Cherry BJ, Rose DJ, Rakovski C, Jones CJ. Do fall predictors in middle aged and older adults predict fall status in persons 50+ with fibromyalgia? An exploratory study. *Research in nursing & health.* 2010;33(3):192-206.

60. O'Connor AB, Turk DC, Dworkin RH, et al. Abuse liability measures for use in analgesic clinical trials in patients with pain: IMMPACT recommendations. *Pain.* 2013;154(11):2324-2334.

61. 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 : official journal of the American Pain Society.* 2008;9(2):105-121.

62. Kroenke K, Bair MJ, Damush TM, et al. Optimized antidepressant therapy and pain self-management in primary care patients with depression and musculoskeletal pain: a randomized controlled trial. *Jama.* 2009;301(20):2099-2110.

63. Mosher HJ, Richardson KK, Lund BC. The 1-Year Treatment Course of New Opioid Recipients in Veterans Health Administration. *Pain medicine.* 2016.

64. Quinn PD, Hur K, Chang Z, et al. Incident and long-term opioid therapy among patients with psychiatric conditions and medications: a national study of commercial health care claims. *Pain.* 2017;158(1):140-148.

65. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med.* 2009;6(5):e1000065.

66. Puffer S, Torgerson DJ, Watson J. Cluster randomized controlled trials. *J Eval Clin Pract.* 2005;11(5):479-483.

Figure 1: PRECIS-2 diagram*
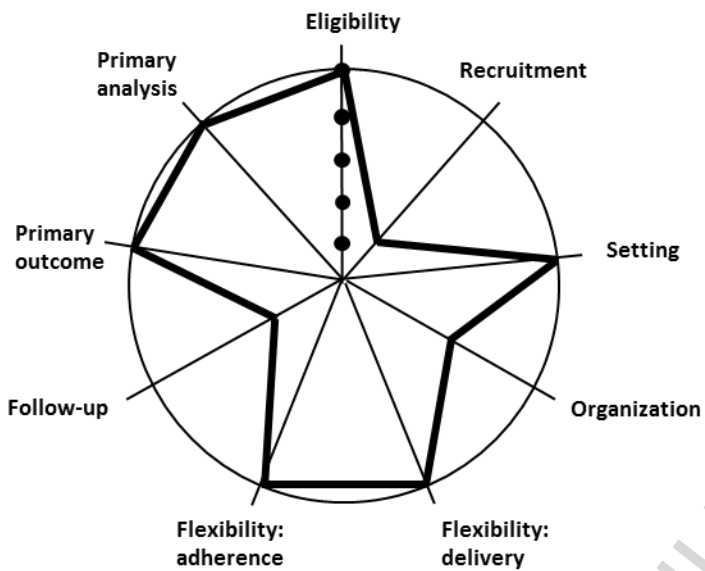
* The PRECIS-2 wheel is a visual representation of pragmatic design. Points near the rim of the circle represent more pragmatic design choices.

Figure 2: Screening and recruitment