Clinical Trial Protocol

Title:	Low dose naltrexone for chronic pain in osteoarthritis and inflammatory arthritis
Protocol Number:	IRB #3062 (VA Boston Healthcare System)
ClinicalTrials.gov:	NCT03008590
Investigational Product:	Naltrexone 4.5 mg
Protocol Version:	5 (September 9, 2019)

Principal Investigator	Paul Monach, MD, PhD
PI Address:	Rheumatology Section, VA Boston Healthcare System 150 South Huntington Ave. Jamaica Plain, MA 02130
PI Telephone:	857-364-5552
PI email:	paul.monach@va.gov

Trial Sponsor:	VA ORD
AE Reporting:	Paul Monach Study coordinator, Britte Beaudette-Zlatanova
Protocol and AE Review:	IRB of VA Boston Healthcare System, DMC (VA Central Office)
Funding:	CSR&D Merit Review Award

Table of Contents

1. Synopsis

Title	Low dose naltrexone for chronic pain in osteoarthritis and inflammatory arthritis
Clinical Phase	II
Investigators	Principal Investigator: Paul A. Monach, MD, PhD
Study Principal Investigator	Paul A. Monach, MD, PhD
Co-Investigators	Eugene Bacorro, MD Maureen Dubreuil, MD Maneet Kaur, MD Antonio Lazzari, MD Caryn Libbey, MD John Otis, PhD
IND Holder / Sponsor	Not applicable (exempt)
Funding	VA CSR&D Merit Review Award
Accrual Objective	60 enrolled subjects, to achieve 54 subjects completing protocol. Patients with osteoarthritis (minimum 30) or inflammatory arthritis (rheumatoid arthritis or non-axial spondyloarthritis, minimum 10) will be enrolled.
Accrual Period	18 months
Study Duration	24 months
Study Design	Randomized, double-blind, cross-over, placebo-controlled trial to evaluate the efficacy of naltrexone 4.5 mg daily in reducing chronic pain in patients with persistent pain from arthritis. Additional medical problems and medications given to control pain must be stable for 8 weeks. At study enrollment, each patient will be randomized to receive either oral naltrexone for 8 weeks followed by oral placebo for 8 weeks, or placebo for 8 weeks followed by naltrexone for 8 weeks. The patient will be blinded to the cross-over time. The primary outcome measures and most secondary outcome measures are patient-reported and will be recorded weekly by patients. Other outcome measures will be collected at 3 in-person visits.
Primary Objective	Reduction in interference of pain with activities, with naltrexone compared to placebo
Primary Outcome Measures	Brief Pain Inventory (BPI) (pain interference)
Secondary Objectives	Reduction in severity of pain with naltrexone Reduction in fatigue with naltrexone. Improvement in quality of life with naltrexone Monitoring of mood during naltrexone treatment Assessment of changes in disease activity during naltrexone treatment Assessment of factors associated with improvement: neuropathic pain, type of arthritis, biomarkers of inflammation

	Safety of naltrexone in patients with rheumatic diseases
Secondary Outcome Measures	Brief Pain Inventory (all other sub-scales than pain interference) painDETECT (neuropathic pain) Brief Fatigue Inventory PROMIS-29 (HRQoL survey) Beck Depression Inventory-II Clinical Global Impression of Severity (CGI-S) and Improvement (CGI-I) Scale Use of "as-needed" analgesic medications Disease activity: DAS28, BASDAI Biomarkers associated with inflammation Adverse events, including grade 1 neuropsychiatric events
Inclusion Criteria	Patients must meet all of the following criteria in order to be eligible for enrollment: • One or more of the following chronic conditions: osteoarthritis, rheumatoid arthritis, peripheral spondyloarthritis • Average daily pain interference with function (average of the 7 parts of question 9 on the BPI) rated at least 4 on a scale of 0-10, and no higher than 9 • No increase in medication in the past 8 weeks made with the expectation of improving pain • No plan to start another medication or a non-pharmacologic treatment regimen likely to affect pain during the next 16 weeks • Age at least 18 • Capable of informed consent, and willingness to comply with study procedures, including receipt of weekly phone calls from the study coordinator
Exclusion Criteria	• Use of opioids including tramadol, in the past 7 days • Pregnant, breast feeding, or unwilling to engage in contraceptive practices if sexually active and capable of conceiving • Schizophrenia, bipolar disorder, or poorly controlled depression or anxiety • Previous use of low-dose naltrexone for more than 8 weeks or in the past 2 weeks • Back pain described by the patient as greater in severity than arthritic pain in all peripheral locations • Significant kidney disease, defined as glomerular filtration rate < 30 ml/min • Liver cirrhosis. There is no specific screening procedure to exclude cirrhosis. • Peripheral neuropathy described by the patient as greater in severity than arthritic pain. There is no specific screening procedure. • Plan to have surgery during the next 16 weeks • Other qualitative circumstances that the investigator feels would make

<table>
<tr><td></td><td>the patient a poor candidate for this clinical trial, such as an unstable social situation or unreliable transportation</td></tr>
<tr><td>Study Treatment</td><td>Patients will undergo a blinded randomization to receive naltrexone or placebo during the first treatment period, and will receive the opposite during the second treatment period:
<table>
<tr><td>Group</td><td>Weeks 1-8</td><td>Weeks 9-16</td></tr>
<tr><td>1</td><td>Low-Dose Naltrexone</td><td>Placebo</td></tr>
<tr><td>2</td><td>Placebo</td><td>Low-Dose Naltrexone</td></tr>
</table>
Naltrexone will be provided at 4.5 mg in a capsule, to be taken daily. Placebo capsules will be made to be indistinguishable from naltrexone.
Use of any opioid agonist, including tramadol, is an exclusion criterion for enrollment, and these drugs must also be avoided during the trial.</td></tr>
<tr><td>Data Collection Schedule</td><td>The primary outcome measure and most secondary outcome measures will be reported by patients weekly using paper forms.

In-person study visits: screening/baseline and weeks 8 and 16. Adverse events and disease activity will be assessed at the in-person visits.</td></tr>
<tr><td>Statistical Analysis</td><td>Pain interference on the Brief Pain Inventory, an average of scores of 7 questions, each measured on a scale of 0-10, will be analyzed as the primary outcome measure.

The summary of each patient's response to LDN compared to placebo will be made using linear contrasts. Thus, for each patient:
$d = a_{n1}x_{n1} + a_{n2}x_{n2} \ldots + a_{n8}x_{n8} - a_{p1}x_{p1} - a_{p2}x_{p2} \ldots - a_{p8}x_{p8}$
where each "a" indicates a weight assigned to that observation on the basis of fitting the pre-post crossover data from all patients with a cubic spline or two linear splines; n indicates treatment with naltrexone; p indicates treatment with placebo; x indicates magnitude of pain interference; and numbers indicate weeks on the treatment. The first two observations following cross-over will be dropped in each patient. Thus, in group 1, weights of a_{p1} and a_{p2} will be 0, and in group 2, weights a_{n1} and a_{n2} will be 0. Weights will be adjusted in individual patients in the event of missing data; the weight given to the summary value of each patient's data (d) in the full analysis, however, will be equal. In each patient, the sum of all weights will be set to 0. The distribution of d among all patients will then be compared to a null distribution by t-test.

The minimum effect size (ES) that characterizes treatments that are regarded as being effective for reducing pain in OA is about 0.3 compared to oral placebo. Estimating that the response during placebo will have a standard deviation $\sigma = 2.0$ on a scale of 0-10, the minimum absolute mean difference between LDN and placebo that we wish to detect is $\delta = ES * \sigma = 0.3 * 2 = 0.6$. The design has 80% power to detect a mean difference between treatments of 0.6 or larger with a two-sided type I error of 5%. This allows for up to 4 of 30 patients in each group to drop out before 12 weeks.

The primary analysis will include osteoarthritis and inflammatory arthritides together to maximize power. Osteoarthritis and inflammatory arthritis will also be analyzed separately because they have different pathophysiologies.

Effects of covariates on response to treatment will be analyzed in two ways:
1. Linear regression, where response (d) is the dependent variable, and</td></tr>
</table>

	independent variables include age, sex, diagnosis of OA or IA, group assignment (1 or 2), baseline pain severity, baseline CRP (with or without an interaction term to limit this variable to patients with IA), and baseline scores on painDETECT (total score), Brief Fatigue Inventory (total score), and Beck Depression Inventory-II (total score) 2. Mixed model, where the raw score (x) at any time is the dependent variable, and independent fixed effect variables include current treatment, time since enrollment, all of the covariates listed above for linear regression, and the random effect is the patient The primary and secondary analytical approaches for the secondary outcomes will be the same as for the co-primary outcome measures for all continuous variables. Outcomes based on classification will use analysis of proportions of patients transitioning from one class to another.
Safety and Monitoring	Standardized definitions and timelines will be used for the reporting of serious and non-serious adverse events. A VA-assigned Data Monitoring Committee (DMC) and the Institutional Review Board of the VA Boston Healthcare System will monitor safety and all other aspects of the study.

2. Study Abstract (summary for lay-persons)

Naltrexone is an FDA approved drug (for alcoholism) that has found widespread use “off-label” to treat pain and fatigue at much lower doses than are used for the approved indication. There are a few scientific studies in three conditions (fibromyalgia, Crohn's disease, and multiple sclerosis) that suggest that this drug has benefit and is safe. However, considering the extent of use in other conditions, and uncertainty about the mechanism of action (purely a pain reliever? other benefits on brain chemistry? anti-inflammatory?), study is needed in diverse diseases. The current study is intended to generate preliminary data in several rheumatologic conditions (osteoarthritis and multiple forms of inflammatory arthritis) in order to select such conditions for future study in larger clinical trials. Although it is a pilot study, a placebo-controlled component is used because of the prominent placebo group effect seen in studies in which self-reported pain is the main outcome. A "blinded cross-over" design is used so that patients will not know when they might be transitioning between placebo and naltrexone.

3. Study Endpoints

a. Primary Outcome

- Average interference of pain with general activity (question 9 on the Brief Pain Inventory, an average of 7 sub-questions, each 0-10) will be compared during naltrexone treatment and during placebo treatment.

b. Secondary Outcomes

The secondary outcome measures include:

- Brief Pain Inventory [other individual questions than those used for the primary outcome, particularly question 5 (average pain severity)]
- painDETECT (continuous measure 0-38, or classified as nociceptive/unclear/neuropathic per the questionnaire guidelines)
- Brief Fatigue Inventory, specifically question 2 (usual level in past 24 hours, 0-10) and question 4 (interference in the past 24 hours, average of 6 questions 0-10 each)
- PROMIS-29 (total score, continuous measure, 28-150)
- Beck Depression Inventory-II (continuous measure 0-63, or classified as minimal/mild/moderate/severe per the questionnaire guidelines)
- Use of “as-needed” analgesic medications (weekly, expressed as % of maximum prescription dose)
- Patient global assessment of improvement or worsening on a 7-point scale [Clinical Global Impression of Severity (CGI-S) and Improvement (CGI-I) Scales]
- CRP
- Adverse events, collected using the IRB’s and DMC’s standard forms

Secondary outcome measures for specific diseases:

- Rheumatoid arthritis: DAS28, as a continuous measure or classified as good/moderate/no response per EULAR criteria
- Spondyloarthritis: BASDAI, as a continuous measure

4. Background and Rationale

a. Background

Chronic Pain and Arthritis

According to an Institute of Medicine report, 116 million Americans are affected by chronic pain, at an overall annual cost of $635 billion (1). Veterans are expected to be disproportionately affected for multiple reasons, including service-related injuries and psychiatric comorbidities such as post-traumatic stress disorder and depression. Treatment for chronic pain has been and continues to be a priority research area for VA CSR&D.

The two most common causes of chronic pain are osteoarthritis (OA) and back pain (2). OA is defined by cartilage loss that is usually attributable in large part to mechanical causes, but abnormalities are numerous and the relationships between them are complex (3, 4). OA is estimated to affect 12.1 million Americans, at an annual cost of $89 billion in medical expenses and additional costs related to reduced productivity (2). Important risk factors for OA in different locations include obesity, advancing age, and prior injury, all of which are highly relevant to the population served by the VA.

Inflammatory arthritis (IA) includes multiple diseases, but the most common and destructive are rheumatoid arthritis (RA) and spondyloarthritis (SpA), a term that encompasses psoriatic arthritis, ankylosing spondylitis, reactive arthritis, and arthritis associated with inflammatory bowel disease. Together, RA and SpA affect 2-3% of persons and usually require treatment with immune-suppressive drugs in order to prevent severe joint damage. Most patients achieve good control of disease, but there is often evidence of some ongoing inflammation; e.g. in one large cohort, only 25% met criteria for remission one year after enrollment, and 72% of those patients had already been in remission at enrollment (5). SpA is the one form of arthritis that is more common in men than women and often has onset in early adulthood, thus is enriched in the VA population; an important registry in SpA (Program to Understand the Long-term Outcomes in Spondyloarthritis, PULSAR) was established in the VA in 2007 and is ongoing.

Mechanisms of Pain in Arthritis

Mechanisms of pain include peripheral and central sources (2, 6). In turn, although the neurons that conduct painful sensations are the same regardless of the stimulus, peripheral sources can be subdivided into those that are purely nociceptive and those that include an inflammatory component. In all persons experiencing pain, central nervous system mechanisms are involved, and these mechanisms are variably active in different persons.

"Central sensitization" is a characteristic of chronic pain of various origins, and in fibromyalgia – the quintessential syndrome of chronic widespread pain – central sensitization may explain the entire syndrome, since there is no evidence of damage in most areas where pain is experienced.

Pain in arthritis results from a combination of these sources. Although one might expect that pain in OA would be attributable to purely non-inflammatory, peripheral, nociceptive input, the other sources play a role. Inflammation is apparent in OA both pathologically and by measurement of inflammatory cytokines, and the degree of synovial inflammation seen on MRI is one factor associated with pain in persons with OA (4, 7). Self-reported pain attributable to OA is more severe in persons with evidence of central sensitization (8). In turn, pain in the inflammatory arthritides is not entirely attributable to inflammation. Joint damage leads to non-inflammatory peripheral pain, and patients with RA and fibromyalgia report higher levels of joint pain than patients with RA who do not also have fibromyalgia (1, 9). For unclear reasons, the prevalence of fibromyalgia is higher in persons with RA (10) and in several other autoimmune diseases than in the general population. Despite great advances in treatment of inflammation in RA, 40-65% of all RA patients have reported inadequate control of pain despite modern therapies (6), and 12% of patients meeting criteria for remission reported chronic pain of clinically significant severity in one study (5).

Thus, approaches to pain management in the arthritides could include reduction of inflammation, nociception, and central sensitization.

Treatment of Arthritic Pain

Unfortunately, management of pain in OA, as in management of chronic pain more generally, remains a challenge despite great effort. Acetaminophen produces little if any benefit above placebo (11). Non-steroidal anti-inflammatory drugs (NSAIDs) are clearly superior to placebo or acetaminophen, with effect sizes 0.3 – 0.5 ("small" in common interpretation (12)) in meta-analyses of multiple trials (11, 13), but this benefit corresponds to median absolute reductions in pain, relative to placebo, of less than 20%, a common standard for the minimum clinically significant difference (12, 14, 15). However, excess cardiovascular risk has now been attributed to NSAIDs as an entire class (16), not just to Cox-2 selective drugs, and NSAIDs were already known to confer important risks of peptic ulcer disease and kidney damage. Tramadol, stronger opioids, and a variety of non-pharmacologic interventions perform no better than NSAIDs based on effect sizes and improvement in pain reduction versus placebo (13) (and see numerous Cochrane reviews), and an "epidemic" of abuse of prescription opioids has received great attention in both the medical literature and popular media. FDA approval of duloxetine to treat pain in OA, on the basis of an absolute difference in pain relief of 10% compared to placebo (17), is arguably the biggest advance in pain management in OA in the past decade. Intra-articular injections of corticosteroids (CS) or hyaluronic acid (HA) derivatives have higher absolute improvements than oral drugs, but since the placebo effect of intra-articular injection is also high, the improvement after CS or HA injection relative to placebo remains modest (18). Combination pharmacologic therapy also has been disappointing (19). Total joint

replacement (TJR) is the only treatment with gratifying improvements in pain in most patients, but many patients are not good surgical candidates due to young age, old age, or comorbidities, and TJR brings multiple short-term risks plus the long-term risk of prosthesis failure particularly in young persons.

Pain management per se has not been studied as extensively in the inflammatory arthritides, since the focus has been on control of inflammation and prevention of structural damage. However, persistent pain despite apparent control of inflammation is common, as noted above, and use of NSAIDs and opioids has the same problems with effectiveness and safety in IA as in OA.

Considering the mediocre benefit of individual modalities in controlling chronic pain, combination therapies involving collaboration across disciplines have been studied – including VA-based studies published in high-profile journals (20-22) – and are widely advocated, as evident in the National Pain Strategy (23). However, the benefit of such approaches is still less than clinically significant in many patients, likely not even additive for the benefits of individual interventions. The pharmacologic component of multi-modality pain management is clearly a weak point in such interventions and arguably the most amenable to improvement.

Low-Dose Naltrexone

Naltrexone is an opioid *antagonist* that is used in high doses for emergency treatment of opioid overdose and is FDA approved in an oral daily dose of 50 mg to prevent recidivism in alcoholics (see Prescribing Information). At much lower doses of 4 – 4.5 mg daily, however, it has been shown in small, blinded, randomized trials to improve pain in fibromyalgia (24), gastrointestinal symptoms in Crohn's disease (25, 26), and quality of life in multiple sclerosis (27). In addition, case reports indicating considerable benefit in patients with complex regional pain syndrome (28), low back pain (29), and scleroderma (30) have been published, although publication bias for case reports is presumed to be high. The mechanism/s is/are unclear, with proposals including not only modulation of central pain-processing pathways but also mitigation of inflammatory roles of microglia (31). In addition to blocking opioid receptors, naltrexone blocks TLR-4, and some studies suggest that the anti-inflammatory effects of low dose naltrexone are independent of opioid receptors (31). Even if effects are mediated entirely by mu, kappa, and delta opioid receptors, the relationship of pharmacokinetics to downstream effects is likely to be complex, by analogy to the finding that opioid agonists can produce analgesia or hyperalgesia at different doses (31).

There is also a lot of enthusiastic discussion about low-dose naltrexone (LDN) on the internet (see results of a Google search of "low dose naltrexone arthritis"), based mostly on testimonials from patients and the recommendations of web-based MDs and other self-appointed experts, most of whom recommend a range of other unproven remedies as well. The testimonials are essentially case reports using patient-reported outcomes (as a class, the type of outcome used in all studies of pain management), but of course they must

be viewed with skepticism because of massive selection bias and uncertainty about the patients' diagnoses. Considering the modest benefit of other options, use of LDN will continue or increase in the absence of studies that support its effectiveness and systematically gather data on safety. Since the drug can be prescribed but only via compounding pharmacies, its use produces significant cost to patients (~$40/month).

What makes LDN of greater interest than most unproven treatments for arthritic pain is a combination of pharmacologic plausibility, the afore mentioned trials in other painful conditions, and striking results of an internet-based survey (http://curetogether.com/Arthritis/ig/treatment-effectiveness-vs-popularity). The survey itself was unbiased in asking persons to rate a large number of approaches they might have used for arthritic pain, but there was no way to validate what type of arthritis the participants had, and the population of responders (who sign up as members based on self-reported conditions) was undoubtedly biased in multiple ways. That being acknowledged, *LDN was ranked as high as any therapy, and most notably, the other therapies that did as well are known to be highly effective in IA or OA* (Enbrel, oral or injectable corticosteroids, TJR), whereas the numerous approaches that did not perform as well include all of the pharmacologic and non-pharmacologic approaches known to provide mediocre benefit based on randomized trials (Fig. 1). On closer inspection, the survey is not nearly as large as it appears: although the group has 1554 members, fewer than 200 produced the 2127 total evaluations of all therapies, and only 17 persons commented on LDN.

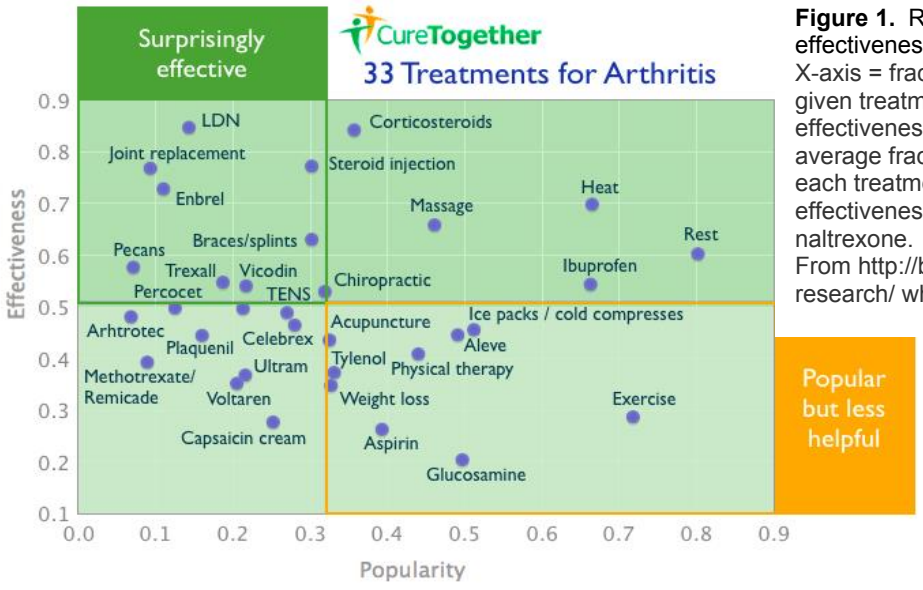

Figure 1. Results of an online survey about effectiveness of 33 treatments for arthritis pain. X-axis = fraction of respondents who tried a given treatment. Y-axis = average rated effectiveness of a given treatment. Vertical line = average fraction of respondents who had tried each treatment. Horizontal line = average rated effectiveness of all treatments. LDN = low dose naltrexone.
From http://blog.23andme .com/23andme-research/ what-patients-say-works-for-arthritis/

Intriguingly, 23 members of a group reporting RA (presumably overlapping with the "arthritis" group) also placed LDN with TJR and corticosteroids and seemingly above DMARDs, biologics, and non-pharmacologic therapies. In a group of patients reporting OA, only 5 commented on LDN, and only 2 of those reported benefit. In fibromyalgia, 72 respondents again gave LDN the highest rating among pharmacologic therapies, with the only approaches giving comparable relief being rest/sleep, application of heat, and stress

reduction. Scanning across other disease states was both encouraging and concerning for the survey reaching a biased population: LDN was at or near the top of therapies reported in patients with multiple sclerosis (n=128), fatigue (n=112), Crohn's disease (n=33), neuropathy (v=22), Sjogren's syndrome (v=9), Hashimoto's (v=9), psoriasis (v=7), hepatitis C (v=6), sarcoidosis (v=4), symptom relief in breast cancer (v=4) or non-small-cell lung cancer (v=4), psoriatic arthritis (v=3), ankylosing spondylitis (v=3), polymyositis (v=3), scleroderma (v=3), HIV (v=2), lupus (v=1), and complex partial seizures (v=1). Reassuringly with regard to bias based on belief systems, LDN was *not* reported as being reliably helpful in 4-20 patients each with irritable bowel syndrome, anxiety, depression, asthma, dandruff, Raynaud's, night cramps, Lyme disease, low back pain, or prostate cancer. Reassuringly with regard to possible bias by the survey designers, LDN was not suggested as an option for diabetes, COPD, gout, eczema, congestive heart failure, or many other conditions.

These data provide further support that LDN may provide an important advance in management of pain and other symptoms – particularly in inflammatory diseases – but do little to alleviate the concern that hype has exceeded evidence.

Other attractive features of LDN are its safety and tolerability, based on the known risks at the much higher FDA-approved doses and the side effects reported in clinical trials (see Section 9).

Considering the diversity of conditions proposed to benefit from LDN; mechanism(s) of action that may or may not include modulation of inflammatory, nociceptive, and central pain pathways; and the unequivocal need for better therapies in these conditions; *high-quality trials are needed in both inflammatory and non-inflammatory conditions in which central pathways play variable roles*. However, funding agencies appropriately expect high-quality preliminary data before investing in definitive trials. This study is designed to obtain such data in a small but placebo-controlled study, powered to detect an effect size as small as that seen with NSAIDs or the most beneficial non-pharmacologic approaches.

b. Rationale for the Clinical Trial Design

The relatively rapid onset of action of naltrexone means that cross-over designs are feasible, increasing power to detect significant benefit. Indeed, the afore mentioned studies in fibromyalgia and Crohn's used cross-over designs. One disadvantage of a cross-over design using subjective outcomes is that placebo/nocebo effects are likely to be amplified during the second treatment period if the cross-over time is unblinded, because the patient's experience during the first period will change his/her expectation for the second period. Therefore, the cross-over time will be blinded. The PI has been advised by an expert in the field to design the study so that each patient is taking LDN for at least 8 weeks and ideally longer (J. Younger, personal communication), and to keep the placebo period at least 6 weeks in the LDN-first group so as to minimize carry-over of benefit.

The most important outcomes will be patient-reported. The number of in-person visits is moderate, chosen to ensure collection of sufficient patient-reported data and distribution of study drug without being a strong disincentive to participation.

5. Study Design and Methods

a. Overview

This randomized, double-blinded, placebo-controlled, cross-over trial will enroll 60 patients (a minimum of 30 with osteoarthritis and a minimum of 10 with inflammatory arthritis) with persistent pain despite standard management approaches. Patients will receive low-dose naltrexone for 8 weeks and placebo for 8 weeks, according to either of two schedules in which the patient is blinded to the durations of treatment and the possible cross-over times. Medical conditions and treatment (pharmacologic and non-pharmacologic) relevant to pain must be stable for at least 8 weeks prior to enrollment, and patients will only be enrolled if they and the investigator do not plan to make other treatment changes (pharmacologic or non-pharmacologic) over the next 16 weeks.

Assessments for the primary outcome and most of the secondary outcomes will be by patient-reported outcomes on a weekly basis. In patients with IA, assessment of disease activity will also occur at the 3 study visits over the course of 16 weeks.

The conditions to be studied are common. We anticipate seeing at least 4 eligible patients per week and enrolling 25% of them, thus completing enrollment within 18 months. The only eligibility criteria that are likely to limit enrollment are pain severity and absence of recent change in other medications.

The study design meets the great majority of the 25 recommendations for design of clinical trials in knee OA made by the Osteoarthritis Research Society International (OARSI)(14): effective randomization strategy; pre-specified stratification and subset analyses; adequate blinding procedures; specification of who is blinded; indistinguishable drug and placebo; minimization of confounding by changes in concomitant medications; assessment of potential participants' consistency in reporting; explicit exclusion based on specific comorbidities; characterization of pain subphenotypes; minimum baseline pain ≥ 4 on a 0-10 scale; use of validated patient-reported outcome measures; definition of primary and secondary outcome measures a priori; and assessment of three core clinical measures (pain, physical function, and patient global assessment). Seven additional recommendations apply only to studies that are unblinded or assess structural changes in OA, and three others relate to registering and reporting (and will also be followed). This study only falls short of these recommendations on two counts: by including a heterogeneous group of conditions (deliberately), and by not including a set of objective measures of physical function.

The study also includes the four outcome domains identified by consensus in the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT): pain intensity, physical functioning, emotional functioning, and global rating of improvement (12).

b. Identification of Patients

i. Inclusion Criteria

Patients must meet all of the following criteria in order to be eligible for enrollment:

- One or more of the following chronic conditions: osteoarthritis, rheumatoid arthritis, peripheral spondyloarthritis (which may include shoulder or hip involvement)
- Average daily pain interference with function (average of the 7 parts of question 9 on the BPI) rated at least 4 on a scale of 0-10, and no higher than 9
- No increase in medication in the past 8 weeks made with the expectation of improving pain, including acetaminophen, any NSAID, any opioid, tramadol, gabapentin, pregabalin, cyclobenzaprine, any tricyclic antidepressant, duloxetine, systemic or injectable corticosteroids, or injectable viscosupplements
- No plan to start another medication or a non-pharmacologic treatment regimen likely to affect pain during the next 16 weeks
- Age at least 18
- Capable of informed consent, and willingness to comply with study procedures, including receipt of weekly phone calls from the study coordinator

ii. Exclusion Criteria

- Use of opioids, including tramadol, in the past 7 days (determined by asking participant if they have used any opioid containing medications in the past 7 days or plan to use any opioid containing medications)
- Pregnant, breast feeding, or unwilling to engage in contraceptive practices if sexually active and capable of conceiving (determined by asking participant if they are pregnant or breast feeding or plan to become pregnant during the duration of the study)
- Schizophrenia, bipolar disorder, or poorly controlled depression or anxiety
- Previous use of naltrexone for more than 8 weeks or in the past 2 weeks, at a low dose or FDA approved dose
- Back pain described by the patient as greater in severity than arthritic pain in all peripheral locations
- Significant kidney disease, defined as glomerular filtration rate < 30 ml/min
- Liver cirrhosis. There is no specific screening procedure to exclude cirrhosis.
- Peripheral neuropathy described by the patient as greater in severity than arthritic pain. There is no specific screening procedure.
- Plan to have surgery during the next 16 weeks

- Other qualitative circumstances that the investigator feels would make the patient a poor candidate for this clinical trial, such as an unstable social situation or unreliable transportation

iii. Notes on Inclusion and Exclusion Criteria

Patients will be regarded as having rheumatoid arthritis or spondyloarthritis based on their assessment by a rheumatologist (including the PI) in the VA Boston Healthcare System. Although criteria exist for diagnosis of RA and psoriatic arthritis, these criteria were created primarily to have high specificity for enrollment of patients in clinical trials in which disease activity and structural damage (rather than pain) are the outcomes of interest. Applying such criteria in the current study may be too restrictive.

Average daily pain of at least 4 on a scale of 0-10 is widely recommended and used in trials in OA (14, 32), since lower levels of pain may not induce a patient to seek treatment (i.e. are not clinically significant), and inclusion of patients with lower levels of pain may make it more difficult to see differences between treatment arms.

Patients with severe psychiatric illnesses will be excluded in part because of uncertain safety of the treatment and in part because of concerns about reliability of data. Current DSM-5 diagnoses of bipolar and related disorders or schizophrenia spectrum and other psychotic disorders (APA, 2013) will be grounds for exclusion. Patients with severe depression will be identified and excluded on the basis of a total score ≥ 29, or a score of 2 or 3 on the question of suicidality, on the Beck Depression Inventory-II administered at screening.

Previous extended use of naltrexone is an exclusion so as to avoid enrolling patients who have previously reported benefit at a low dose, or who had side effects at the FDA-approved dose.

Patients with moderate to severe chronic kidney or liver disease are excluded due to delayed drug elimination and metabolism, in accordance with the Prescribing Information for the (much larger) FDA-approved 50 mg dose.

Patients with peripheral neuropathy as the major source of pain are excluded only in order to prevent sources of pain from being too heterogeneous, not for safety reasons.

c. Study Medications

An investigational product, also known as investigational medicinal product in some regions, is defined as a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical study, including products already with a marketing authorization but used or assembled (formulated or packaged) differently than the authorized form, or

used for an unauthorized indication, or when used to gain further information about the authorized form. The investigational product will be stored in a secure area according to local regulations. It is the responsibility of the investigator to ensure that investigational product is only dispensed to study participants. The investigational product will be dispensed only from official study sites by authorized personnel according to local regulations.

In this protocol, the investigational products are:

- Naltrexone for oral use, 4.5 mg per capsule, to be taken daily
- Placebo to match naltrexone for oral use

These products will be manufactured as capsules by the VA Boston Investigational Drug Service. Naltrexone powder will be purchased from Letco Medical (Decatur, AL):
Naltrexone HCL, USP (Dihydrate), NDC No. 62991-1243-02

Capsules will be colored so that naltrexone and placebo are indistinguishable in appearance. They will be stored in the pharmacy until the day of distribution and will be distributed in bottles sufficient for 28 days of use. Patients will be instructed to use the bottles in the numbered order. Bottles 1 and 2 will be distributed at enrollment or mailed to the participant. Bottles 3 and 4 will be distributed at the in-person visit at week 8 or mailed to the participant. Bottles 1 and 3 will each contain 28 capsules, and bottles 2 and 4 will each contain 10 extra capsules in case the patient has a delay in an in-person visit beyond exactly 8 weeks. It will therefore be evident to the patient that s/he will be receiving the same drug for a given 4-week period, but it will not be evident to the patient that treatment will *not* change in the transitions from bottle 1 and 2 and bottle 3 to 4, and will *always* change in the transition from bottle 2 to 3.

i. Randomized Treatment Periods: Methods and Doses

At enrollment, eligible patients will undergo randomization to receive oral naltrexone and placebo according to one of two schemes (Table 1). Randomization will be performed by the research pharmacy. Patients will be informed that they will receive naltrexone for 8 weeks and placebo for 8 weeks but will not be given the details of the possible treatment timelines.

Table 1. Treatment groups.

Group	Weeks 1-8	Weeks 9-16
1	Low-Dose Naltrexone	Placebo
2	Placebo	Low-Dose Naltrexone

Low-dose naltrexone = 4.5 mg

To ensure that equal numbers of patients with OA and IA are incorporated into the two groups, stratification by these two diagnostic groups will be used.

ii. Prohibited Medications

Use of any opioid agonist, including tramadol, is an exclusion criterion for enrollment, and these drugs must also be avoided during the trial. If a patient must use an opioid agonist for more than 2 days during the trial, that patient must be removed from the trial, and data will be censored at the time of that treatment.

Medications to treat pain must not have been increased for 8 weeks prior to enrollment. Opioids, including tramadol, may not have been used for at least 7 days prior to enrollment. If a patient must have a change in treatment (pharmacologic or non-pharmacologic) related to pain or arthritis, that patient must be removed from the trial, and data will be censored at the time of that treatment change. Stable treatment with IV medications that require pre-treatment with other medications to prevent infusion reactions (e.g., rituximab, infliximab, IVIG) will not be regarded as a change in treatment. Corticosteroid or viscosupplementation injection intended to improve musculoskeletal pain must have been given at least 8 weeks prior to enrollment.

It is expected that many patients will be using acetaminophen or NSAIDs as needed before enrollment. Rather than requiring patients to move to a fixed schedule, data on use of such drugs will be collected and used in analyses.

d. Study Procedures

i. Recruitment

Recruitment will occur through the clinical practices of the rheumatologists at the three main campuses of the VA Boston Healthcare System. Primary care providers will be notified about the trial and will be encouraged to contact the PI about potential participants. Pre-screening by discussion with the referring physician and review of medical records will greatly reduce the number of screening failures. Referring physicians will be asked to ask prospective enrollees to rate their average daily pain severity and pain interference (0-10) for the past two weeks, but will be asked *not* to tell patients that the criterion for eligibility is ≥4. In this way, the time between identification of a potential enrollee and the screening/baseline visit will serve as a run-in period for the purposes of severity and stability of daily pain. An IRB-approved brochure will be available to inform potential patients about the study. If permitted by the staff in the primary care clinics or pain clinic, the brochure will be displayed in common areas (waiting rooms or hallways) for patients to pick up themselves. Providers who are not investigators will not be asked to distribute brochures. Brochures will not be mailed. This recruitment technique will not require additional study staff.

The PI plans to include five other rheumatologists in the VA Boston Healthcare System as co-investigators. However, other colleagues may become credentialed for research. If necessary and practical, these rheumatologists will be added as co-investigators for

the current study. The assessment of disease activity in RA and SpA is easily derived from standard clinical practice.

ii. Screening

Screening will be performed by the PI or co-investigator at an in-person visit. Patients who meet eligibility criteria may enroll at the same visit. They may receive study drug either at enrollment or later by mail or in-person, but they must begin study drug within 14 days of screening.

iii. Enrollment

Details of the goals of the research and the risks and benefits of the protocol will be reviewed with each patient who meets eligibility criteria. Consent will be obtained by the investigator, a sub-investigator, a research assistant, or a study coordinator. Patients who choose to enroll will be randomized to one of two treatment groups as above, in a double-blinded manner. Study drug (sufficient for 66 days, in case a visit cannot be scheduled at exactly 56 days) will be distributed in person or by mail. The IRB of the VA Boston Healthcare System will approve the protocol and the consent document prior to enrollment of any patient.

iv. Visit Frequency/ Visit Schedule

There will be 3 in-person study visits: screening / enrollment and 8 and 16 weeks after enrollment. Study drug will be distributed at (in person) or shortly after (by mail or in person) the first two of these visits. Patients who are unable to attend the second and/or third in-person visits may remain in the study if completed questionnaires, new questionnaires, and study drug are delivered by mail.

v. Study Assessments

1). Study Assessment

Patient-reported outcomes will be reported weekly on paper forms. Subjects may receive reminders by phone. The outcome measures used will be:

Primary outcome measure (all patients):

- Brief Pain Inventory-short form (BPI) (33): The BPI is a 9-item self-report questionnaire that allows patients to rate the severity of their pain and the degree to which their pain interferes with common dimensions of feeling and function. For the purpose of the proposed study, we will be particularly interested in pain “interference”. A recent consensus panel recommended that the two domains measured by the BPI – pain intensity (severity) and the impact of pain on functioning (interference) – be included as outcomes in all chronic-pain clinical

trials (IMMPACT, (34)). The IMMPACT panel (www.immpact.org) specifically identified the interference items of the BPI, rated on a 0–10 scale, as one of the two scales recommended for assessment of pain-related functional impairment (35). It has excellent reliability and validity and it has been used widely in OA research (36). (weekly questionnaire)

Secondary outcome measures:

- Brief Pain Inventory short form, other questions than those used for the primary outcome, particularly question 5, average pain severity (weekly)
- painDETECT (for neuropathic component of pain) (37, 38) (only at weeks 0, 4, 8, 12, and 16)
- PROMIS-29 (survey of quality of life in multiple domains) (40) (only at weeks 0, 4, 8, 12, and 16)
- Brief Fatigue Inventory, specifically question 2 (usual level in past 24 hours, 0-10) and question 4 (interference in the past 24 hours, average of 6 questions 0-10 each)
- Beck Depression Inventory-II (purchase pending), a widely used 21-question assessment of the severity of depression (only at weeks 0, 4, 8, 12, 16)
- Clinical Global Impression of Severity (CGI-S) and Improvement (CGI-I) Scale. The CGI-S/I is the most widely used clinician-rated measure of treatment related functioning and response. The CGI-S score rates severity of illness on a 7-point scale, ranging from 1 (“normal”) to 7 (“among the most severely ill patients”). The CGI-I score rates clinical improvement on a 7-point scale ranging from 1 (“very much improved” to 7 (“very much worse”) (weekly)
- Use of “as-needed” analgesic medications, using a form that will contain the names of the specific medications (only at weeks 0, 4, 8, 12, 16)
- C reactive protein (CRP), a widely used clinical measure of inflammation.
- Rheumatoid arthritis: DAS28 (only at in-person visits)
- Spondyloarthritis: BASDAI (only at in-person visits)

Assessment of blinding (all patients):

- Question asking whether the patient thinks s/he is taking LDN, or placebo, or is not sure (weekly)

The investigator and coordinator will also inquire about adverse events and medication use, including whether there has been interruption in use of study drug, at the 2 in-person visits following enrollment. At the weekly phone-calls to remind patients to complete questionnaires, they will also be asked if they have experienced adverse events and will be reminded to inform the study team of serious adverse events as soon as possible after they occur.

Although efforts will be made to schedule patients exactly 8 weeks and 16 weeks after they start study drug, the in-person cross-over day may occur up to 7 days before or up

to 10 days after week 8, i.e., between days 49 and 66. If the "week 8" cross-over visit does not occur at day 56, the clock will be re-set for weeks 9-16. Thus, the final "week 16" visit may occur any time between 49 and 66 days after the actual "week 8" cross-over date. Sufficient drug is distributed at enrollment and week 8 to allow 66 days of treatment. If a participant is not able to come to an in-person visit within the window, an out of window visit may occur and will be noted as a protocol deviation. The study consists of two consecutive 8-week treatment periods. An interruption in use of the study drug may occur for up to 4 weeks when it occurs *between* the first and second study periods. Allowing this extended time frame off study drug between treatment periods is appropriate; inclusion of a protocolized "wash-out" period is a common design in cross-over studies. Because there is no defined wash-out period in this study, the analysis plan already included omission of data from the first two weeks after cross-over (see Section 7).

2). Study Schedule Table

	Screen / Enroll	Wk 0	Wks 1-3	Wk 4	Wks 5-7	Wk 8	Wks 9-11	Wk 12	Wks 13-15	Wk 16
Consent	X									
Eligibility	X									
Phone reminders		X	X	X	X		X	X	X	
Complete weekly questionnaires	X	X	X	X	X	X	X	X	X	X
Complete every-4-week questionnaires	X			X		X		X		X
Return questionnaires						X				X
Distribute questionnaires	X					X				
Distribute study drug	X					X				
Collect unused study drug						X				X
Report AEs			X	X	X	X	X	X	X	X
Assess disease activity	X					X				X
Assess medication use	X		X	X	X	X	X	X	X	X
Laboratory tests *	X					X				X

In-person visits are at screening / enrollment and weeks (Wks) 8 and 16. Wk 0 (Baseline) may occur in-person at enrollment or within 14 days at home. Phone reminders to complete questionnaires weekly may occur but are not essential to the protocol. Non-serious adverse events (AEs) will be collected at in-person visits; subjects will be instructed to report potential serious adverse events (SAEs) as soon as possible and will be asked about non-serious AEs during the weekly phone conversations when those occur.
* Complete blood count with differential, electrolytes, BUN, creatinine, liver function tests, ESR, CRP. Inability to provide these samples will be regarded simply as missing data (as with an incomplete questionnaire) rather than as a protocol deviation.

e. Common Closing Date - Study Duration

The study will close 16 weeks after enrollment of the last subject. Data will not be collected from subjects after they have completed the 16 weeks of the study.

f. Criteria for Withdrawal of Study Medication – Early Termination

- Use of any opioid agonist, including tramadol, for more than two days
- Rise in AST or ALT to a level 3-fold above the upper limit of normal, or more than 2-fold above the patient's level at baseline if that baseline was abnormal.
- Decline in kidney function (drop in GFR to < 30 ml/min)
- Pregnancy or nursing
- Addition or discontinuation of any medication given on a scheduled basis to treat pain. For medications taken as-needed, changes in frequency will not be grounds for withdrawal; rather, patients will be asked how frequently they have been using such medications.

g. Outcome Definitions

See Section 7, Data Analysis

6. Safety Monitoring and Adverse Event Reporting

a. Nature of Study

In determining what type of adverse events will be reported, several facts about the drug being tested and the nature of the underlying disease need to be considered.

Naltrexone is an FDA approved drug at 50 mg, a dose 11-fold higher than is to be used in this study. The side effects with this higher dose – as well as a 300 mg dose that is no longer used – are described in the Prescribing Information. Side effects of low-dose

naltrexone have been described in several published trials, but those encompass only about 170 patients followed for a few months.

b. Study Oversight

The Principal Investigator has primary oversight responsibility of this clinical trial. The IRB of the VA Boston Healthcare System has oversight responsibility for this clinical trial. A Data Monitoring Committee (DMC) will be assigned by VA Central Office. The DMC and IRB will review accrual, patterns and frequencies of all adverse events and protocol compliance every 4 months. The DMC and IRB make recommendations to the Principal Investigator regarding the continuation status of the protocol.

The Principal Investigator and the research team are responsible for identifying adverse events. Adverse events and protocol compliance will be reviewed once a month by the Principal Investigator.

c. Definitions

This section defines the types of adverse events and outlines a process for the appropriate collecting, grading, and reporting procedures. The information in this section complies with ICH Guidelines E2A: Clinical Safety Data Management: Definitions and Standards for Expedited Reporting of the International Conference of Harmonization (ICH) Guideline for Good Clinical Practice and applies the Standards set forth in the National Cancer Institute (NCI), Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

Adverse event is defined as: "…an unfavorable and unintended sign, symptom or disease associated with a patient's participation in a study."

Serious adverse events (SAE) will be defined according to 21CFR 312.32.

An SAE is any untoward medical occurrence that at any dose:

- Results in death
- Is life-threatening
- Requires hospitalization or prolongs an existing hospitalization
- Results in persistent or significant disability/incapacity
- Is a congenital anomaly/birth defect
- An important medical event may be considered an SAE when, based upon appropriate medical judgment, it may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed above.

In addition, other events that will be reported as SAEs in conjunction with this trial include:

- Pregnancy
- Cancer

- Overdose of the study drug
- Suspected transmission of an infectious agent (eg, pathogenic or nonpathogenic) via the study drug

An unexpected adverse event is defined as any adverse experience, the specificity or severity of which is not consistent with the risks of information described in the protocol. Therefore, expected adverse events are those that are identified in the research protocol, package insert, or investigational brochure as having been previously associated with the study agents or are known consequences of a person's medical condition and thus having the potential to arise as a consequence of participation in the study.

d. Toxicity Grading of Adverse Events

Toxicity grades will be assessed according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) (http://ctep.cancer.gov/reporting/ctc.html). The purpose of using the CTCAE system is to provide a standard language to describe toxicities, to facilitate tabulation and analysis of the data and to facilitate the assessment of the clinical significance of all adverse events. The CTCAE provides the following grades and descriptions in the CTCAE manual (v4.0). Adverse events should be recorded and graded 1 to 5 according to the CTCAE grade provided below:

- Grade 1 = Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.
- Grade 2 = Moderate; minimal, local or noninvasive intervention indicated; limiting age appropriate instrumental activities of daily living
- Grade 3 = Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care activities of daily living
- Grade 4 = Life-threatening consequences; urgent intervention indicated
- Grade 5 = Death related to adverse event

As only naltrexone and placebo will be considered to be study drugs in this trial, only adverse events possibly, probably, or definitely related to naltrexone will be considered reportable for this study.

e. Relation to Study Therapy

The relation or attribution of an adverse event to an investigational product is determined by the investigator and then recorded on the appropriate case report form and/or SAE reporting form. The CTCAE provides the following descriptors and definitions for assigning an attribution to each adverse event.

Code Descriptor Definition

"**Unrelated" Category Code**

1	Unrelated	The adverse event is clearly not related to the investigational product

"Related" Category Codes

2	Unlikely	The adverse event is doubtfully related to the investigational product
3	Possible	The adverse event may be related to the investigational product
4	Probable	The adverse event is likely related to the investigational product
5	Definite	The adverse event is clearly related to the investigational product

f. Standard Elements

A set of standard elements for adverse event data will be collected. These elements include: patient ID, dates for event/event reported/date resolved, the event itself, event severity, whether it was expected and/or serious (as defined above), patient status, place of adverse event treatment (to further determine serious events), causality, and subsequent changes to protocol or consent form. Additionally, the reporter may write a more detailed description of the event and any other pertinent information.

g. Expected / Known Risks and Adverse Events Associated with Study Intervention and Procedures

i. Study Drug/Intervention: For known risks of study intervention, see Section 9.
ii. Study Procedures: For risks of study procedures, see Section 9.

h. Reporting Timeline

- The investigator will fulfill all reporting requirements of the IRB of the VA Boston Healthcare System.
- Within **24 hours** (of learning of the event), co-investigators must report to the PI any SAE whether related or unrelated to study drug. The PI in turn must review and report the SAE to the IRB within 24 hours of learning of the event, and to the DMC within 2 working days of learning of the event.
- Expected or unexpected adverse events that are grade 1 will only be collected or reported at the discretion of the PI, for example, neuropsychiatric side effects that have been reported in previous studies
- All other expected or unexpected reportable adverse events must be reported within **20 working days** of the notification of the event or of the site becoming aware of the event.

i. Investigational New Drug Application (IND)

The FDA has made a formal ruling that this study is exempt from needing an IND.

j. Planned Interim Analysis

There is no plan for an interim analysis. Progress reports will be submitted to the DMC every 4 months.

k. Criteria to Suspend Enrollment

There are no pre-specified criteria to suspend enrollment.

7. Data Analysis and Statistical Considerations

a. Primary Outcome

Interference of pain with general activity (question 9 on the Brief Pain Inventory, an average of 7 sub-questions, each 0-10) is the primary outcome measure. Pain severity is the primary outcome reported in studies of pain in OA. However, some patients will choose to increase activity at the expense of a level of pain that they have learned to tolerate, so "pain interference" is of at least equal interest (12, 23).

i. Primary Analytic Approach to the Primary Outcome

The goal is to determine the difference in pain interference during treatment with LDN versus placebo. We will take advantage of the multiple data points obtained from each patient to improve the precision of that estimate, reducing the risk of type II error. The summary of each patient's response to LDN compared to placebo will be made using linear contrasts. Thus, for each patient:

$$d = a_{n1}x_{n1} + a_{n2}x_{n2} \ldots + a_{n8}x_{n8} - a_{p1}x_{p1} - a_{p2}x_{p2} \ldots - a_{p8}x_{p8}$$

where each "a" indicates a weight assigned to that observation on the basis of fitting the pre-post crossover data from all patients with a cubic spline or two linear splines; n indicates treatment with naltrexone; p indicates treatment with placebo; x indicates pain severity or pain interference; and numbers indicate weeks on the treatment. The first observation following cross-over will be dropped in each patient. Thus, in group 1, weights a_{p1} and a_{p2} will be 0, and in group 2, weights a_{n1} and a_{n2} will be 0. Weights will be adjusted in individual patients in the event of missing data, and data obtained during an active adverse event will not be included; the weight given to the summary value of each patient's data (d) in the full analysis, however, will be equal. In each patient, the sum of all weights will be set to 0.

The distribution of d among all patients will then be compared to a null distribution by t-test.

ii. Secondary Analytic Approaches to the Primary Outcome

Effects of covariates on response to treatment will be analyzed in two ways:

1. Linear regression, where response (d) is the dependent variable, and independent variables include age, sex, diagnosis of OA or IA, group assignment (1 or 2), baseline pain severity, baseline CRP (with or without an interaction term to limit this variable to patients with IA), and baseline scores on painDETECT (total score), Brief Fatigue Inventory (total score), and Beck Depression Inventory-II (total score)

2. Mixed model, where the raw score (x) at any time is the dependent variable, and independent fixed effect variables include current treatment, time since enrollment, all of the covariates listed above for linear regression, and the random effect is the patient.

The ability to detect a clinically significant difference is also considered essential for trials of pain in OA and other conditions. The proportions of patients experiencing improvement or worsening of at least 2 (on 0-10 scales) on pain interference and/or pain severity relative to the scores at enrollment after 8 weeks on LDN or placebo will be compared. This standard meets or exceeds standards proposed by various organizations (11, 12, 14). This analysis will be between-patient and limited to the first 8 weeks in the study, since carry-over and time-dependent effects will make data after cross-over difficult to interpret. An overall assessment of significant improvement will be made on the basis of improvement of at least 2 points in *either* pain severity or pain interference, without worsening of 2 or more on the other scale.

iii. Sample Size Calculations

The primary study end point will be the ability of LDN (denoted as 'n' below) to reduce pain interference to a greater extent than placebo (p).

The minimum effect size (ES) that characterizes treatments that are regarded as being effective for reducing pain in OA is about 0.3 compared to oral placebo (18). Estimating that the response during placebo will have a standard deviation σ = 2.0 on a scale of 0-10 (32), the minimum absolute mean difference between LDN and placebo that we wish to detect is δ = ES * σ = 0.3 * 2 = 0.6.

The two crossover patterns are [pppppppp_nnnnnnnn] and [nnnnnnnn_pppppppp]. Removing the two observations just after cross-over, the patterns are either [8p and 6n] or [8n and 6p]. With 30 subjects per pattern the overall mean difference, d = [D(8p, 6n) + D(8n,6p)]/60 where D denotes the sum of the 30 'n-p' differences in each pattern. We assume an autocorrelation model, that r is the correlation for adjacent times, and that r^j is the correlation between times j units apart. Then, the variance of d roughly doubles, compared to assuming that r = 0. Exact calculations show that with 26 to 30 subjects per pattern group and $0.5 \leq r \leq 0.7$ that the variance of d ranges from 0.014 to 0.018. Hence, we assume Var(d) = 0.016. For this form of the t-test, it follows that:

The design has 80% power to detect a mean difference between treatments of 0.6 or larger with a two-sided type I error of 5%. This allows for up to 4 of 30 patients in each group to drop out before 12 weeks.

b. Secondary Outcomes

The secondary outcome measures include:

- Brief Pain Inventory (other individual questions than those used for the primary outcome, especially average pain severity = question 5)
- painDETECT (continuous measure 0-38, or classified as nociceptive/unclear/neuropathic per the questionnaire guidelines)
- Brief Fatigue Inventory, specifically question 2 (usual level in past 24 hours, 0-10) and question 4 (interference in the past 24 hours, average of 6 questions 0-10 each)
- PROMIS-29 (total score, continuous measure, 28-150)
- Beck Depression Inventory-II (continuous measure 0-63, or classified as minimal/mild/moderate/severe per the questionnaire guidelines)
- Use of "as-needed" analgesic medications (expressed as % of maximum prescription dose)
- Patient global assessment of improvement or worsening on a 7-point scale [Clinical Global Impression of Severity (CGI-S) and Improvement (CGI-I) Scales]
- CRP
- Adverse events, including pre-specified minor adverse events (vivid dreams, headache, dizziness, insomnia, fatigue, nausea) and categories of severe adverse events (cardiovascular, neuropsychiatric, malignancy, serious infection)

Secondary outcome measures for specific diseases:

- Rheumatoid arthritis: DAS28, as a continuous measure or classified as good/moderate/no response per EULAR criteria
- Spondyloarthritis: BASDAI, as a continuous measure

The primary and secondary analytical approaches for the secondary outcomes will be the same as for the primary outcome measures for all continuous variables. Weightings will change for outcomes measured every 4 or 8 weeks rather than weekly. Outcomes based on classification will use analysis of proportions of patients transitioning from one class to another. Proportions of patients experiencing categories of adverse events while receiving LDN or placebo will be compared by Fisher's exact tests.

8. Data Management

a. Registration

At enrollment, each patient will be assigned an ID number beginning with the type of arthritis (OA or IA) followed by order of enrollment among patients with OA or IA (01-60). This ID number will be seen by the investigator and study participant and used on all questionnaires as well as the case report forms used for the additional data obtained at in-person visits. A file that links the patient's name and chronological ID number to the random study ID number will be maintained by the study coordinator on a VA network drive. This process will minimize the chance that the PI will be able to identify individual subjects if he is participating in data analysis, particularly important for any exploratory analyses beyond those that are pre-specified.

b. Randomization

The research pharmacy will perform the randomization using a varying block design provided by the statistician, and the study staff will remain blinded to group assignment. The pharmacist will log treatment assignments. The patient and investigator will remain blinded to the treatment assignments until completion of the last visit of the last patient enrolled. At that time, the blind will be broken and the PI will be informed about treatment assignments and will notify the patients and their MDs which treatments they received. Patients will be notified of their treatment assignments by the sending of an IRB approved letter within one month after the last patient's last visit or within one month after IRB approval of the letter, whichever is later. The participants' PCPs and/or rheumatologists will be notified by being copied on a CPRS note containing the approximate dates of study treatment and the treatment assignments within one month after the last patient's last visit. Unblinding of the treatment assignment during the study will be performed only in rare instances where knowledge of the treatment assignment is felt to be necessary to protect patient safety, as determined by the IRB of the VA Boston Healthcare System. The two research pharmacists will be unblinded to treatment and can be reached 24/7 by phone and are able to access pharmacy records off-site. Request for unblinding will be made simultaneously in writing. The IRB and DMC will be notified of the treatment assignment revealed by unblinding according to the timeline of SAE reporting, since it is expected that any unblinding will be associated with an SAE.

c. Data Entry

All study data will be recorded on paper forms by the patient or transcribed from the patient's oral report by the investigator or study coordinator, then collected and copied into electronic files by the coordinator or a research assistant. Although direct electronic data entry by patients would be preferred by some patients and would avoid the problem of questionnaires being lost, it is felt that relying on electronic forms would exclude many potential participants. Data will be copied from paper forms to similarly formatted electronic versions.

Paper forms produced by this study including signed consent forms will be stored in a locked office or other secure facility for a minimum defined time period per VA and FDA regulations. Electronic files including study data, the file linking patient identifiers to coded data, IRB communications, and other regulatory documents will be stored on a secure SharePoint site set up specifically for this study behind the VA firewall. Electronic copying of study data from paper forms will occur directly on the SharePoint site. Research records will be kept indefinitely or until the law allows their destruction in accordance with the VA Record Control Schedule. Paper records will be shredded, and electronic records will be destroyed in a manner in which they cannot be retrieved.

d. Data Quality Control

As an assessment of compliance with treatment, patients will be asked to return the previous containers of study drug at the next in-person visit, and asked not to throw out any extra doses that might remain because they forgot to take them.

At the second and third in-person study visits, it will be clear to the study coordinator which patients are completing study questionnaires and which are not. One of the reasons for planning to average the results of patient-reported outcome measures at multiple time points in a weighted manner is so that data will still be usable if a patient only completes the forms at in-person visits. At in-person visits, patients will complete questionnaires without the participation of the investigator of coordinator, since their presence might influence reporting compared to data recorded when alone.

Two steps will be taken to maximize the numbers of patients who complete all 3 in-person visits. First, they will be given a financial incentive at a level that is not expected to raise concerns about coercion: $20 for the screening visit, $20 for visit 2, $20 for visit 3, and an additional $20 at visit 3 if all 3 visits were completed. Second, the study coordinator and PI will arrange schedules such that patients can be seen at at least 2 of the 3 campuses of the VA Boston Healthcare System.

9. Protection of Human Subjects

a. GCP Statement

This clinical trial will be conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki, and that are consistent with Good Clinical Practice and all applicable regulatory requirements.

b. Benefits and Risks

The potential benefits of participating in this study are improvement in pain and/or improvement in function as limited by pain. The potential risks of this study are side effects

of naltrexone, delay in seeking other approaches to pain management, and potential loss of confidentiality.

i. Risks of Low-Dose Naltrexone (LDN)

In studies of low-dose naltrexone (LDN) that included about 170 patients, no liver toxicity or severe depression has been reported. No other dangerous side effects have been reported, either, but the number of patients in these studies is too small to be confident that there are no dangers that might occur rarely (24-27). The only common side effect repeatedly reported for LDN has been vivid dreams (sometimes but not usually described as nightmares). Headaches, dizziness, insomnia, fatigue and nausea have been reported in some studies but not others (24-27).

At doses 11 times higher than those used in this study (50 mg rather than 4.5 mg), naltrexone use has been associated with the following:

- Liver toxicity: usually this was mild, only apparent on blood testing, and without symptoms, but some cases of hepatitis (abdominal pain, loss of appetite, and jaundice) have been reported. Serious liver injury has only been reported with much higher doses (300 mg daily) that are no longer used.
- Depression and suicidal thoughts
- Symptoms of opioid withdrawal in patients also taking opioid pain relievers

All of these risks have been used to create exclusion criteria for this study.
The 50 mg dose has only been used to treat patients with alcoholism or opioid abuse problems, who are already at high risk for liver disease, depression, suicide, and use of opioid medications without telling their physicians. The full package insert is available on-line and is considered to be the Investigator's Brochure for this study.

Even in the event that a patient is using opioids surreptitiously or via prescription outside the VA, naltrexone is unlikely to cause severe opioid withdrawal symptoms at the low dose used in this study. However, it is possible that it would make such drugs less effective as pain relievers. Opioid agonists have been formulated with low-dose antagonists in an effort to reduce risk of addiction or development of tolerance, so there is substantial evidence of safety of such combinations.

ii. Other Risks

The main procedures being performed for research purposes are questionnaires. Additional physical examination and laboratory testing will be performed at the 4 in-person visits but the disease activity scales generated by including such information are advocated for routine clinical use and therefore are consistent with standard of care practice. All non-medication study procedures therefore confer minimal risk.

The risk associated with potential delay in seeking other approaches to pain management is also considered to be minimal. Only patients who report stable pain for the previous 2 months and stable use of medications to control pain, and who report no intention of making other efforts to control pain during the coming 16 weeks, will be approached for participation in this study. Standard approaches to preserving confidentiality of research data are noted in Section 8.

10. References

1. Lee YC. Effect and treatment of chronic pain in inflammatory arthritis. Current rheumatology reports. 2013 Jan;15(1):300. PubMed PMID: 23292816. Pubmed Central PMCID: 3552517.
2. Cohen E, Lee YC. A Mechanism-Based Approach to the Management of Osteoarthritis Pain. Current osteoporosis reports. 2015 Dec;13(6):399-406. PubMed PMID: 26419467. Pubmed Central PMCID: 4623875.
3. Ding C, Zhang Y, Hunter D. Use of imaging techniques to predict progression in osteoarthritis. Curr Opin Rheumatol. 2013 Jan;25(1):127-35. PubMed PMID: 23080226.
4. Torres L, Dunlop DD, Peterfy C, Guermazi A, Prasad P, Hayes KW, et al. The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. Osteoarthritis Cartilage. 2006 Oct;14(10):1033-40. PubMed PMID: 16713310.
5. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. Arthritis Res Ther. 2011;13(3):R83. PubMed PMID: 21651807. Pubmed Central PMCID: 3218896.
6. American College of Rheumatology Pain Management Task F. Report of the American College of Rheumatology Pain Management Task Force. Arthritis Care Res (Hoboken). 2010 May;62(5):590-9. PubMed PMID: 20461782.
7. Neogi T, Guermazi A, Roemer F, Nevitt MC, Scholz J, Arendt-Nielsen L, et al. Association of Joint Inflammation With Pain Sensitization in Knee Osteoarthritis: The Multicenter Osteoarthritis Study. Arthritis & rheumatology. 2016 Mar;68(3):654-61. PubMed PMID: 26554395.
8. Neogi T, Frey-Law L, Scholz J, Niu J, Arendt-Nielsen L, Woolf C, et al. Sensitivity and sensitisation in relation to pain severity in knee osteoarthritis: trait or state? Ann Rheum Dis. 2015 Apr;74(4):682-8. PubMed PMID: 24351516. Pubmed Central PMCID: 4062615.
9. Ranzolin A, Brenol JC, Bredemeier M, Guarienti J, Rizzatti M, Feldman D, et al. Association of concomitant fibromyalgia with worse disease activity score in 28 joints, health assessment questionnaire, and short form 36 scores in patients with rheumatoid arthritis. Arthritis Rheum. 2009 Jun 15;61(6):794-800. PubMed PMID: 19479706.
10. Wolfe F, Hauser W, Hassett AL, Katz RS, Walitt BT. The development of fibromyalgia--I: examination of rates and predictors in patients with rheumatoid arthritis (RA). Pain. 2011 Feb;152(2):291-9. PubMed PMID: 20961687.
11. Bannuru RR, Schmid CH, Kent DM, Vaysbrot EE, Wong JB, McAlindon TE. Comparative effectiveness of pharmacologic interventions for knee osteoarthritis: a systematic review and network meta-analysis. Ann Intern Med. 2015 Jan 6;162(1):46-54. PubMed PMID: 25560713.
12. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, 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 Feb;9(2):105-21. PubMed PMID: 18055266.

13. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage. 2010 Apr;18(4):476-99. PubMed PMID: 20170770.
14. McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, et al. OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis Cartilage. 2015 May;23(5):747-60. PubMed PMID: 25952346.
15. Wolfe F, Michaud K. Assessment of pain in rheumatoid arthritis: minimal clinically significant difference, predictors, and the effect of anti-tumor necrosis factor therapy. J Rheumatol. 2007 Aug;34(8):1674-83. PubMed PMID: 17611989.
16. Coxib, traditional NTC, Bhala N, Emberson J, Merhi A, Abramson S, et al. Vascular and upper gastrointestinal effects of non-steroidal anti-inflammatory drugs: meta-analyses of individual participant data from randomised trials. Lancet. 2013 Aug 31;382(9894):769-79. PubMed PMID: 23726390. Pubmed Central PMCID: 3778977.
17. Brown JP, Boulay LJ. Clinical experience with duloxetine in the management of chronic musculoskeletal pain. A focus on osteoarthritis of the knee. Therapeutic advances in musculoskeletal disease. 2013 Dec;5(6):291-304. PubMed PMID: 24294303. Pubmed Central PMCID: 3836379.
18. Bannuru RR, McAlindon TE, Sullivan MC, Wong JB, Kent DM, Schmid CH. Effectiveness and Implications of Alternative Placebo Treatments: A Systematic Review and Network Meta-analysis of Osteoarthritis Trials. Ann Intern Med. 2015 Sep 1;163(5):365-72. PubMed PMID: 26215539.
19. Ramiro S, Radner H, van der Heijde DM, Buchbinder R, Aletaha D, Landewe RB. Combination therapy for pain management in inflammatory arthritis: a Cochrane systematic review. The Journal of rheumatology Supplement. 2012 Sep;90:47-55. PubMed PMID: 22942329.
20. Dobscha SK, Corson K, Perrin NA, Hanson GC, Leibowitz RQ, Doak MN, et al. Collaborative care for chronic pain in primary care: a cluster randomized trial. JAMA. 2009 Mar 25;301(12):1242-52. PubMed PMID: 19318652.
21. Bair MJ, Ang D, Wu J, Outcalt SD, Sargent C, Kempf C, et al. Evaluation of Stepped Care for Chronic Pain (ESCAPE) in Veterans of the Iraq and Afghanistan Conflicts: A Randomized Clinical Trial. JAMA internal medicine. 2015 May;175(5):682-9. PubMed PMID: 25751701.
22. 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 Jul 16;312(3):240-8. PubMed PMID: 25027139.
23. Ballantyne JC, Sullivan MD. Intensity of Chronic Pain--The Wrong Metric? N Engl J Med. 2015 Nov 26;373(22):2098-9. PubMed PMID: 26605926.
24. Younger J, Noor N, McCue R, Mackey S. Low-dose naltrexone for the treatment of fibromyalgia: findings of a small, randomized, double-blind, placebo-controlled, counterbalanced, crossover trial assessing daily pain levels. Arthritis Rheum. 2013 Feb;65(2):529-38. PubMed PMID: 23359310.
25. Smith JP, Bingaman SI, Ruggiero F, Mauger DT, Mukherjee A, McGovern CO, et al. Therapy with the opioid antagonist naltrexone promotes mucosal healing in active Crohn's disease: a randomized placebo-controlled trial. Dig Dis Sci. 2011 Jul;56(7):2088-97. PubMed PMID: 21380937. Pubmed Central PMCID: 3381945.
26. Smith JP, Field D, Bingaman SI, Evans R, Mauger DT. Safety and tolerability of low-dose naltrexone therapy in children with moderate to severe Crohn's disease: a pilot study. Journal of clinical gastroenterology. 2013 Apr;47(4):339-45. PubMed PMID: 23188075. Pubmed Central PMCID: 3586944.

27. Cree BA, Kornyeyeva E, Goodin DS. Pilot trial of low-dose naltrexone and quality of life in multiple sclerosis. Annals of neurology. 2010 Aug;68(2):145-50. PubMed PMID: 20695007.
28. Chopra P, Cooper MS. Treatment of Complex Regional Pain Syndrome (CRPS) using low dose naltrexone (LDN). Journal of neuroimmune pharmacology : the official journal of the Society on NeuroImmune Pharmacology. 2013 Jun;8(3):470-6. PubMed PMID: 23546884. Pubmed Central PMCID: 3661907.
29. Ghai B, Bansal D, Hota D, Shah CS. Off-label, low-dose naltrexone for refractory chronic low back pain. Pain medicine. 2014 May;15(5):883-4. PubMed PMID: 24967470.
30. Frech T, Novak K, Revelo MP, Murtaugh M, Markewitz B, Hatton N, et al. Low-dose naltrexone for pruritus in systemic sclerosis. Int J Rheumatol. 2011;2011:804296. PubMed PMID: 21918649. Pubmed Central PMCID: 3171757.
31. Younger J, Parkitny L, McLain D. The use of low-dose naltrexone (LDN) as a novel anti-inflammatory treatment for chronic pain. Clin Rheumatol. 2014 Apr;33(4):451-9. PubMed PMID: 24526250. Pubmed Central PMCID: 3962576.
32. Chappell AS, Desaiah D, Liu-Seifert H, Zhang S, Skljarevski V, Belenkov Y, et al. A double-blind, randomized, placebo-controlled study of the efficacy and safety of duloxetine for the treatment of chronic pain due to osteoarthritis of the knee. Pain practice : the official journal of World Institute of Pain. 2011 Jan-Feb;11(1):33-41. PubMed PMID: 20602715.
33. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Annals of the Academy of Medicine, Singapore. 1994 Mar;23(2):129-38. PubMed PMID: 8080219.
34. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003 Dec;106(3):337-45. PubMed PMID: 14659516.
35. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005 Jan;113(1-2):9-19. PubMed PMID: 15621359.
36. Mendoza T, Mayne T, Rublee D, Cleeland C. Reliability and validity of a modified Brief Pain Inventory short form in patients with osteoarthritis. European journal of pain. 2006 May;10(4):353-61. PubMed PMID: 16051509.
37. Freynhagen R, Baron R, Gockel U, Tolle TR. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Current medical research and opinion. 2006 Oct;22(10):1911-20. PubMed PMID: 17022849.
38. Freynhagen R, Tolle TR, Gockel U, Baron R. The painDETECT project - far more than a screening tool on neuropathic pain. Current medical research and opinion. 2016 Feb 23:1-109. PubMed PMID: 26907456.
39. Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, et al. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. Cancer. 1999 Mar 1;85(5):1186-96. PubMed PMID: 10091805.
40. Hinchcliff M, Beaumont JL, Thavarajah K, Varga J, Chung A, Podlusky S, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-Reported Outcomes Measurement Information System 29-item Health Profile and Functional Assessment of Chronic Illness Therapy-Dyspnea short form. Arthritis Care Res (Hoboken). 2011 Nov;63(11):1620-8. PubMed PMID: 22034123. Pubmed Central PMCID: 3205420.