Supplementary Material*

Fritz JM, Lane E, McFadden M, et al. Physical therapy referral from primary care for acute back pain with sciatica A randomized controlled trial. Ann Intern Med. 6 October 2020. [Epub ahead of print]. doi:10.7326/M20-4187

* This supplementary material was provided by the authors to give readers further details on their article. The material was reviewed but not copyedited.

 © 2020 American College of Physicians

Protocol Summary

Version: 10/27/2014

Management Strategies for Patients with Low Back Pain and Sciatica

Principal Investigator

Julie M. Fritz
Distinguished Professor, Department of Physical Therapy & Athletic Training
The University of Utah
520 Wakara Way
Salt Lake City, UT 84108
(801) 581-6287
julie.fritz@hsc.utah.edu

Co-Investigator(s)

Gerard Brennan
Senior Research Scientist and the Director for Clinical Quality and Outcomes Research
Intermountain Healthcare
5848 South 300 East
Murray, UT 84107
(801) 314-4100
Gerard.Brennan@imail.org

Tom Greene
Professor, Division of Clinical Epidemiology
The University of Utah
295 Chipeta Way
Salt Lake City, UT 84132
(801) 585-6667
Tom.Greene@hsc.utah.edu

Sponsor Contact Information

Agency for Health Research Quality
Kim Wittenberg
540 Gaither Road
Rockville, MD 20850
(301) 427-1321
kim.wittenberg@ahrq.hhs.gov

SPECIFIC AIMS

The research dissemination and demonstration project described in this proposal addresses an important evidence gap in the initial management of patients with a common condition; low back pain (LBP) with sciatica; a condition that can be disabling and costly for affected individuals and the healthcare system. Most patients with LBP and sciatica initially access healthcare in primary care settings. Practice guidelines[4,33] advocate initial management with first-line medications and patient education. Imaging (x-ray or MRI), opioid medications and invasive interventions such as injections or surgery are not indicated during the initial 4 weeks of management for the large majority of patients. Guidelines further advocate a shared decision-making process with respect to whether additional therapies should be used during the initial weeks of management, however there is currently very little high-quality evidence that can help inform this important decision. Uncertainty about the effectiveness of therapies that may be used early in the management of patients with LBP and sciatica is likely contributing to increasing utilization rates of MRI, opioids, injections, and surgery for these patients,[21,22,39,40] and rapidly growing costs without documented improvements in patient-centered outcomes.[41] More effective initial management strategies may help to curb rapidly escalating use of expensive and invasive tests and interventions. The window of opportunity to avoid these procedures may be in the initial management decisions emanating from the initial primary care encounter, but this hypothesis has not been adequately examined.

Physical therapy is often used at some point in the course of care for patients with LBP and sciatica, although its use during the initial 4-week management period is relatively uncommon and highly variable geographically. Our research team has conducted several preliminary studies indicating that the effectiveness of physical therapy in improving patient-centered outcomes and reducing risks for subsequent healthcare utilization for patients with LBP is enhanced if patients are referred early, instead of later, in the course of care.[47] Also, our team has conducted a series of studies to identify a physical therapy program specifically tailored to patients with LBP and sciatica that optimize clinical outcomes.[52,72,73] Based on this work we are proposing to evaluate the effectiveness of adding a standardized, evidence-based physical therapy program to primary care management early in the care for patients with LBP and sciatica. We hypothesize that the initial weeks of management provide a window of opportunity to avoid progression to persistent pain and disability and utilization of progressively more invasive and expensive management procedures if an evidence-based strategy is employed. The proposed study evaluates this overall hypothesis. If early physical therapy proves effective, broader dissemination into healthcare delivery could be evaluated. We will also seek to identify moderators of improvement with early physical therapy. Identifying moderators would provide information about the specific characteristics of patients most likely to benefit from early physical therapy (e.g, those with higher initial pain severity, etc.) which could inform shared decision-making for providers and patients in the future. This project fits the current AHRQ Comparative Effectiveness Research Portfolio. The project examines the impact of translation of new scientific information into clinical care, and develops evidence to inform clinical decision-making for an important priority condition.

We will conduct a randomized trial comparing the effectiveness and costs associated with adding physical therapy to primary care management during the first 4 weeks of care for patients with LBP and sciatica. Patients with LBP and sciatica with a new consultation in primary care will be randomized to receive guideline-based primary care management with or without early physical therapy. Patients will be followed for 1 year. Outcomes will include measures of disability, pain, satisfaction, healthcare utilization and costs.

Specific aims for the study are the following:

1. Compare the effectiveness of two initial management strategies for a subgroup of patients with LBP and sciatica. We hypothesize that early physical therapy will result in greater improvements in disability and pain over 1 year compared to a usual care strategy.
2. Compare utilization of specific healthcare procedures [and healthcare costs] associated with two initial management strategies for a subgroup of patients with LBP and sciatica. We hypothesize that early physical therapy will result in decreased utilization of MRI and injections over 1 year compared to a usual care strategy.
3. Evaluate moderators of improvement with early physical therapy as an initial management strategy. This aim is exploratory, however we hypothesize that pre-specified psychosocial and clinical factors will moderate improvement with early physical therapy.

clinics for recruitment. These assests will benefit the proposed project by permitting more efficient development of data management and recruitment procedures.

C.2 Overview of Approach

This study will be a pragmatic randomized trial comparing strategies for managing patients with a new episode of LBP and sciatica in primary care. One strategy is usual care (UC) directed within primary care during the initial 4 weeks. The other strategy is early physical therapy (EPT) with patients receiving 6-8 sessions in the initial 4 weeks. We will examine patient-centered outcomes and subsequent healthcare utilization [*and costs*] over a 1-year follow-up period. Based on our hypothesis that the first few weeks provide an opportunity to positively influence the clinical course of LBP with sciatica, the difference between strategies occurs in the 4 weeks after an initial primary care encounter. The study will use a pragmatic approach comparing these strategies under realistic clinical circumstances. Recruitment will occur from primary care clinics in the University of Utah Health Care system. Dr. Fritz has previously led successful clinical trials in this system. Patients visiting a primary care clinic with a chief complaint of LBP and sciatica will be potentially eligible. Eligible patients who provide informed consent will undergo baseline assessment followed by randomization to a treatment group (UC / EPT). Both groups will receive first-line medication as needed and advice and education emphasizing remaining active and the inappropriateness of early imaging or specialist referral as recommended by the APS/ACP guideline.[33] The UC group will be managed within primary care for the next 4 weeks with a stepped approach (ie, referral for additional care considered after 4 weeks if symptoms do not improve). The EPT group will be referred to physical therapy for 6-8 sessions in the first 4 weeks based on current best evidence supporting a centralizing approach. Follow-up evaluations will occur 4 weeks, 6 months and 1 year after enrollment.

C.3 Subjects

We will recruit 220 patients with LBP and sciatica making initial entry into the healthcare system in primary care. Initial entry is defined as not having treatment from any provider for LBP or sciatica in the past 6 months. Potential candidates must satisfy all inclusion criteria:

I. Symptoms of pain and/or numbness between the 12th rib and buttocks, which, in the opinion of the primary care provider, are originating from tissues of the lumbar region.
II. Symptoms of pain and/or numbness primarily into one leg that have extended below the knee in the last 72 hours, and correspond to a lower lumbar nerve root distribution (L4, L5, S1)
III. Current symptoms present for 90 days or fewer
IV. Age 18 - 60 years
V. Oswestry disability score $\geq$ 20%
VI. One or more of the following symptoms:
    1. Positive ipsilateral or contralateral straight leg raise test (reproduction of symptoms at <70^0)
    2. Reflex, sensory, or strength deficits in a pattern consistent with lower lumbar nerve root

These criteria will select patients meeting the operational definition of LBP and sciatica typically used in research and practice and are consistent with our preliminary work examining physical therapy treatments. Consistent with a pragmatic approach we will not require imaging studies because early imaging is not advocated by guidelines. Patients will be excluded if they meet any of the following exclusion criteria:

I. Any prior spine fusion surgery, or any surgery to the lumbosacral spine in the past year
II. Current pregnancy
III. Currently receiving treatment for LBP from another healthcare provider (e.g., chiropractic, massage therapy, injections, etc.) or any treatment for LBP in prior 6 months
IV. Judgment of primary care provider of “red flags” of a potentially serious condition including cauda equina, major or rapidly progressing neurologic deficit, fracture, cancer, infection or systemic disease

These criteria will exclude patients who do not fit the sub-group with LBP and sciatica. Patients with red flags may require early specialist referral. Patients who are pregnant comprise a separate sub-group requiring different management.[74,75] Lumbar traction and some exercises are contraindicated for recently post-surgical patients or those with a spinal fusion. [*Sciatica is very uncommon in adolescents under age 18.[130] Individuals over age 60 are increasingly likely to have lumbar spinal stenosis as an underlying cause of LBP and leg symptoms and practice guidelines differ for these patients.[12]*] Reasons for ineligibility will be tracked and eligibility and consent rates determined.

Subjects will be recruited into the study using procedures we are successfully employing in our ongoing LBP trial (*R18HS018672*). Recruitment will be conducted using the Utah Health Research Network (UHRN), a practice-based research network formed to promote collaboration between the University Health Sciences Center and the Community Physician Group and facilitate clinical research. Dr. Lisa Gren, UHRN Director, will assist in establishing the recruitment procedures through UHRN. Currently the UHRN includes 12 clinics, with approximately 80 primary care providers. The UHRN has an established recruitment process approved by the University Institutional Review Board (IRB). The process uses the electronic medical record to identify patients with specific characteristics based on demographic information and ICD-9 codes. These patients are informed of the study by clinic staff or by mail. Interested individuals are contacted within 24 hours by the study coordinator. For this project we will identify patients with a primary care visit (with no LBP visits in the past 6 months) with an ICD-9 code related to LBP and sciatica (722.1, 722.2, 722.52, 722.93, 724.3, 724.4) between the ages of 18-60. We performed a query of the EMR for a 12-month period (7/11 – 6/12) to gauge the number of potentially-eligible patients based on these criteria and identified a total of 1,886 individuals. This recruitment pool will be adequate to meet our accrual goals.

C.4 Informed Consent, Randomization and Blinding

When a potentially eligible patient speaks with the study coordinator, the coordinator will explain the project. If the patient meets preliminary eligibility criteria and has interest, the coordinator will meet with the patient to insure all eligibility criteria are met and have the patient sign an informed consent document approved by the University of Utah IRB. Patients who do not enroll will be instructed to follow-up with their primary care provider as needed. Once consent is obtained baseline examination procedures will be performed by a blinded research assistant. After the baseline examination the patient will again meet with the coordinator to receive the advice and education intervention consistent with current evidence (*details below*) then the coordinator will reveal the patient's treatment group assignment. All patients will be advised to return to their primary care provider as needed, consistent with a pragmatic design.

We will randomize individual patients instead of cluster-randomizing to avoid the loss of statistical power resulting from the latter.[77] Randomization will be conducted using a blocked procedure [*stratified by primary care clinic from which the patient is recruited*] and conducted using a random permuted block procedure. A random list of differing block sizes (2 or 4) will be generated prior to the study. Sequentially-numbered, sealed envelopes will be prepared containing the group assignment for each patient. The envelope will be opened by the coordinator after providing the advice and education intervention. Patients in the EPT group will be scheduled for physical therapy to begin within 3 days at [*a physical therapy clinic trained in the evidence-based physical therapy program. We selected two physical therapy clinics that will provide good geographic coverage of the region. Physical therapists in these clinics will be trained prior to recruitment with ongoing fidelity monitoring and follow-up training as outlined below.*]

Consistent with a pragmatic study we will not use placebos to blind patients or attempt to balance provider contact. At baseline the randomization assignment will not be revealed until the patient has completed procedures with both the primary care provider and study coordinator, reducing the potential for bias from the primary care provider offering advice or making medication decisions, or in the delivery of the advice and education by the coordinator. Follow-up assessments will be performed by a blinded research assistant. Any occurrences of un-blinding of the research assistant will be recorded. The primary care provider cannot remain blind to treatment received if a study patient returns for follow-up. If a patient returns to the primary care provider the provider will base care decisions on the patient's needs consistent with a pragmatic approach. [*It is possible that patients randomized to the UC group will eventually receive physical therapy, which could occur in a clinic trained in the evidence-based program. Our preliminary findings on the percentage of patients with LBP and sciatica who ever use physical therapy suggest this will not occur frequently. We will record the occurrence of patients in the UC group eventually receiving physical therapy and whether or not this occurs in a trained clinic. This will permit an exploration of the potential confounding impact should this occur with greater frequency than anticipated.*] This pragmatic approach will preserve appropriate patient-provider communication and permit an examination of effectiveness and future healthcare utilization under the most realistic clinical circumstances possible.

C.5 Baseline and Follow-Up Examination Procedures (*See Appendix A for paper copies of all forms)*

Baseline examination will consist of self-report questionnaires and a physical examination. The physical examination includes neurologic testing to confirm eligibility. The physical impairment index(PII)[78] composed of 7 measures of range of motion and strength, will be assessed. Each measure is graded positive (1) or

negative (0) for a total score of 0-7 with higher numbers indicating more impairment. Each measure has excellent reliability and validity.[78,79]

Study data will be collected using web-based data collection via REDCap (Research Electronic Data Capture). Data collection forms will require only modest modifications from our ongoing randomized trial (*R18-HS018672*). Our experience with web-based data collection is that patients find it easy to use and convenient, as reflected by compliance rates in our ongoing study. At baseline and each follow-up patients and researchers input data directly into REDCap. If a patient is unable to directly input data using a computer paper forms are used and data are uploaded at a later time.

Demographic Information

Patients will provide information including age, sex, ethnicity, race, employment status, and general medical and LBP history. Patients will be asked about expectations for treatments including physical therapy, injections, and surgery prior to their group assignment.[80] This information will be explored as a moderating factor and potential covariate in the analysis to mitigate the confounding effect of treatment preference.

Primary/Secondary Outcomes and Potential Moderators

Oswestry Disability Questionnaire (OSW),[81] a measure of LBP-related disability will be the primary outcome because functional ability is a primary concern for patients with LBP and sciatica.[82] The OSW is a 10-item scale scored from 0-100 and higher numbers indicating greater disability. We will use a modified version that replaces the sex life item with an employment/homemaking to improve compliance.[83,84] The OSW is widely used in research on non-operative management of patients with LBP with and without sciatica[85] with responsiveness equal or superior to other disability measures.[86,87] Our prior research has found this modified OSW to have excellent test-retest reliability (ICC = 0.90), validity, and responsiveness to change for patients with acute LBP, with a minimum clinically important difference (MCID) of 6 points.[83] A study of patients with LBP and sciatica reported an MCID of 8 points for the OSW.[86]

Numeric Pain Rating Scale (NPRS)*:* 0-10 NPRS ('0' = no pain, '10' = worst imaginable pain) will be used to assess pain intensity. We will use separate NPRS ratings of low back and leg pain. NPRS have excellent test-retest reliability.[84] Our prior research has found the NPRS responsive to change, with an MCID of 2 points among patients with acute LBP receiving physical therapy.[88] A 2-point MCID is also reported for patients with LBP and sciatica.[86] [*NPRS will be a secondary outcome and explored as a moderator.]*

EuroQol (EQ-5D): A generic quality of life instrument[91] will be used to assess health outcomes on a scale that may be referenced to other disease conditions.[92] The EQ-5D covers 5 domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each domain has 3 response categories: "no problems"; "some problems"; or "inability or extreme problems." Responses are combined to 5-digit health state classification. The EQ-5D yields a total of 243 possible health states. Valuation of each health state is available.[93] The EQ-5D is reliable and responsive to change in patients with LBP[94,95] and is commonly used in economic evaluation of interventions for LBP.

Fear Avoidance Beliefs Questionnaire (FABQ)*:* assesses patients' beliefs about how physical activity and work may affect their LBP and perceived risk for re-injury.[98] The FABQ has two subscales; a 7-item work subscale (FABQW), and 4-item physical activity subscale (FABQPA). Test-retest reliability of the subscales is high.[98,99] Validity is supported by associations with concurrent and future disability and work loss in patients with acute and chronic LBP.[100-103] [*FABQ will be explored as a potential moderator.]*

Pain Catastrophizing Scale (PCS)*:* a 13-item scale assessing the extent to which people catastrophize in response to pain.[106] Each item is scored from 0 - 4. The PCS is reported as a total score, and is composed of three sub-scales: rumination, magnification, and helplessness. Pain catastrophizing plays an important role in the transition from acute to chronic LBP.[102] [*PCS will be explored as a potential moderator.]*

Healthcare Utilization and Direct and Indirect Costs: [*We will collect both direct and indirect costs due to lost work productivity related to LBP over the 1-year follow-up.*] We will use a cost diary method integrated into the web-based data collection system to collect utilization and cost data. Cost diaries offer advantages in terms of feasibility and validity as compared with patient-reported questionnaires for collecting cost data and provide information at regular intervals instead of relying on patient recall over a long time period, minimizing recall error.[109] [*We will collect cost and utilization data monthly via REDCap using the input options outlined below.*] The cost diary will be modeled on paper-based methods with high compliance (85%) for economic analysis in studies of patients with LBP.[108,131] This compliance rate is consistent with our experience using this method in our ongoing study. Prior research reports no differences between cost

diaries completed for an entire year and extrapolation of data for limited time periods, indicating a robustness to impute scores for random missing data. Patient-reported cost diary data has been found to generally agree with insurance claims, supporting the validity of the self-report cost and utilization data.[108] [*Each monthly cost diary will ask patients if they have utilized healthcare resources in the past month specifically for LBP in 4 categories: 1) provider visits (traditional or complementary/alternative), 2) medications (prescription or over-the-counter), 3) interventions (injections, surgery, etc.), or testing (x-rays, MRI, etc.), and 4) lost time from work or decreased work productivity. Conditional logic will be used to reduce response burden for patients not seeking any care. Patients who are seeking care will be given follow-up questions to ascertain the number and nature of utilization in each category in the past month.*]

Follow-up examinations will be conducted 4 weeks, 6 months, and 1 year after baseline. The 4-week examination permits evaluation of the immediate effects of early physical therapy. 6-month and 1-year examinations evaluate intermediate and long-term effects of treatment received during the first 4 weeks. The 4-week follow-up will be conducted in-person by a blinded research assistant in order to re-assess the physical examination. 6-month and 1-year follow-ups will be conducted online by e-mailed links using REDCap. In addition to the self-reports described previously, follow-up assessments will also collect:

Satisfaction: Patient satisfaction with care for LBP will be measured at the 4-week follow-up using a 10-item instrument that has been validated and found capable of distinguishing different dimensions of satisfaction (caring, information and treatment effectiveness) among patients with LBP attending primary care.[111]

Global Rating of Change: will be completed at each follow-up using a 15-point scale[112] that asks the patient to rate the degree of change in his or her condition from enrollment to the present..

C.6 Procedures for Treatment (see *Appendix C for copies of treatment forms*)

Patients in both groups (UC or EPT) will continue to be managed by their primary care provider. Consistent with a pragmatic study, decisions about medication or additional referrals will be recorded but not controlled. All enrolled patients will be recommended by the coordinator to follow-up in primary care as-needed if unsatisfied with their progress. All patients will receive education on the favorable natural history of LBP with sciatica and reassurance that imaging and specialist referrals are not indicated early in the course of care. Patients will also receive advice to resume normal activity as soon as possible without bed rest. The advice and education will be provided by the study coordinator after the baseline assessment but prior to randomization to avoid bias. Coordinators will be licensed physical therapists and will receive additional training by the PI in provision of the advice and education intervention. All patients will be given a copy of the *Back Book* (*Appendix B*), a booklet developed to help modify beliefs and behavior of patients with LBP.[113] Messages are based on research demonstrating beneficial effects of remaining active. Research has found the *Back Book* to be well-accepted by patients, and capable of shifting beliefs about recovery and activity, particularly when delivered interactively.[113,114] The study coordinator will review the *Back Book* with the patient and answer any questions. Management common to all study patients therefore includes; 1) education and reassurance of the favorable prognosis of LBP with sciatica, the benefits of staying active, and the appropriateness of delaying imaging or referrals, 2) a copy of *The Back Book* with a review of its contents to reinforce key messages, and 3) ongoing primary care management as needed.

Usual Care (UC) Treatment Group

Patients in the UC group will be managed with a stepped care approach supported by current practice guidelines.[115] Initial management for the UC group will involve the education and assurance intervention as outlined above for the first 4 weeks following the primary care visit. Patients in the UC group will be recommended to follow-up with their primary care provider if unsatisfied with their progress after 4 weeks. At that time decisions on further treatments and/or referrals will be made by the primary care provider in consultation with the patient consistent with usual care.

Early Physical Therapy (EPT) Treatment Group:

Patients in the EPT group will receive the same education and advice intervention described above and will receive physical therapy during the initial 4 weeks following enrollment. Physical therapy treatment will be standardized based on evidence and our prior research evaluating a centralizing treatment program for the sub-group of patients with LBP and sciatica. The first physical therapy session will be scheduled within 3 days after enrollment and 6-8 sessions will be administered within the first 4 weeks [*in a clinic trained to provide the study treatment*] (2 weekly sessions in first two weeks and 1-2 sessions in weeks 3-4 based on therapist judgment). Each session will include a brief assessment and treatment with centralizing exercises

and spinal mobilizations. Mechanical traction is an optional component. Centralizing exercises will progress according to table 3, and consist of sustained and repeated movements to extend the spine as this most often produces centralization. Modifications can be made by the therapist to maximize centralization. Therapists will instruct patients to monitor symptoms and perform exercises in a range to maximize centralization and minimize discomfort. Patients will be provided handouts and instructed to perform assigned exercises at home every 4-5 hours on days between sessions. Spinal mobilizations will be performed with the patient prone or side-lying with varied amplitude and velocity to promote centralization of symptoms and spine extension. If mechanical traction is used it will be applied using a 3D ActiveTrac device (Empi, Minneapolis). We have used this device in prior studies and have access in our facilities. Static mechanical traction is applied for a maximum of 12 minutes with force parameters and patient positioning adjusted to maximize symptom centralization.

1. Lying on stomach	5. Prone press-up to extended arms with sag
2. Propping on elbows	6. Prone press-up to extended arms with overpressure
3. Prone press-up to elbows	7. Repeated extension in standing
4. Prone press-up to extended arms	
Goal: Tolerate for 5 minutes or 3 sets of 10 repetitions to maximize centralization	

Table 3. Progression of Centralizing Extension Exercises

Therapist Training and Fidelity Monitoring

[*We will train physical therapists in two outpatient clinics in the University of Utah Health Care system in the treatment program for EPT patients. These two clinics (University Orthopedic Center and South Jordan Health Center) provide good geographic coverage and have the necessary equipment to perform study treatments. There are 14 full-time physical therapists in these clinics that will be trained. New therapists will be trained on an ongoing basis. Dr. Fritz and the coordinator will lead the training. Dr. Fritz is experienced training therapists to successfully perform research-related treatments regardless of clinician experience.*[142] *A written procedure manual including study logistics and treatment techniques will be developed and an in-person training session held prior to enrollment. This session will explain the research process including issues related to human subjects, research ethics and adverse event reporting. Treatment procedures outlined above will be described, demonstrated and practiced during the session. Passing a written test related to study procedures will be required before a therapist will be able to treat study patients. We will conduct follow-up therapist meetings every 6 months to discuss progress and review procedures. More frequent sessions will be held if needed. Treatment fidelity* during the first 4 weeks will be monitored by identifying off-protocol interventions received at the 4-week follow-up. Fidelity for the EPT group will be monitored on two levels. Therapists will record patient's self-reported compliance with assigned exercises on the treatment *form from each session (appendix C). These forms are uploaded into the web-based data collection system. The study coordinator will audit at least 20% of forms to identify any concerns regarding treatment fidelity during the study.* Off-protocol events will be recorded and feedback given to therapists.

C.7 Subject Retention Procedures

[*We will use several strategies to enhance retention. At the time of obtaining consent we will clearly outline the time that will be required of participants. We will clarify for participants that study-related treatments will not be billed to them or their insurance. We will use reminder calls and e-mails for follow-up assessments. We will be in monthly contact with all participants via e-mail during the follow-up period to obtain cost and utilization data, helping to keep the participant connected to the study. The 4-week follow-up is done in person because this contact relatively soon after enrollment helps to maintain a sense of connection to the project for all participants, particularly those in the UC group. Long-term follow-ups are completed electronically to reduce the burden of an in person meeting. Our experience in our ongoing trial suggests this balance of in-person contact and the convenience of remote long-term follow-ups with frequent reminders and small gifts for completing assessments has been quite successful in retaining participants.*

C.8 Data Management and Quality Control Procedures

Data will be entered into REDCap, a software toolset and workflow methodology for electronic collection and management of research and clinical trial data. REDCap provides secure and easy data manipulation with audit trails for reporting, monitoring and querying records, and an automated export mechanism to common statistical packages (SPSS, SAS, Stata) facilitating quicker time to analyses. Periodically during the study data will be exported by the University of Utah Study Design and Biostatistics Center (SDBC) at the University's Center for Clinical and Translational Science (CCTS). The database will be maintained on a server supplied and maintained by University of Utah Health Sciences Center. The SDBC, directed by co-investigator Dr. Tom Greene, will provide additional analyses to monitor progress, including rates of recruitment, retention and adherence to study protocols.

Once a patient is enrolled, the Study Coordinator will create a patient profile in REDCap and the patient will receive a unique ID generated prior to beginning the study. The patient profile will be identified by the unique ID and will not contain any Personal Health Information in REDCap. Links between the ID and patients' identifying information will be maintained by the PI and will be available only to the Study Coordinator. After a patient profile is created the patient and investigators are able to input data directly into REDCap. Once data collection begins, the PI will be in at least weekly communication with the Study Coordinator to monitor overall progress, recruitment and identify any issues that could impact the safety or ethics of the study (e.g, adverse events, instances of unblinding, etc.).

C.9 Data Analysis

Data analyses will be carried out by the SDBC directed by Dr. Greene, who has extensive experience as a data coordinating statistician for clinical trials with recognized expertise in longitudinal data analysis. The SDBC employs biostatisticians with broad expertise, including Molly McFadden, who will carry out project analyses directed by Dr. Greene. The SDBC will internally review data distributions for extreme scores or inconsistent results and will notify the PI of problematic data. Because data quality and completeness will be monitored throughout, we expect any data discrepancies to be resolved and the study database closed within weeks of completing data collection. Following database closure the SDBC will provide data summaries with descriptive statistics including central tendency and dispersion computed for continuous data and frequency distributions for categorical data. Transformations will be sought for variables failing to meet distribution assumptions. We are aware that unadjusted between-group comparisons are valid if randomization is achieved, however pre-treatment characteristics of the groups will be compared to assess for chance imbalances. If differences are found these variables may be used as covariates in post-hoc sensitivity analyses. Intention-to-treat principles will be used with all patients analyzed in their randomized group regardless of compliance. We will compare compliance between groups and "per-protocol" secondary analyses may be considered if non-compliance is high or disproportionate between groups.

Multiple imputation (MI) will be used to address missing outcome data. MI incorporates baseline and follow-up factors beyond the variable being analyzed into imputation models to account for dependence of missing data on other factors. We will use the Markov Chain Monte Carlo method to generate imputed values[116] involving 4 steps: 1) preparing the dataset by identifying all outcome variables likely to be involved in later analyses and likely predictors of missingness, evaluating distributional assumptions with transformations as needed; 2) conducting the MI, with variance and covariance estimates based on observed data used to iteratively estimate maximum likelihood values for all subjects. Multiple replacement values for each missing score are drawn randomly from the posterior distribution and perturbed with error; 3) analyses are done identically on all versions of imputed data; and 4) parameter estimates and tests are combined and adjusted for between-imputation variance to yield final results. MI is valid if data are missing at random, meaning the probability of missingness is independent of the missing outcome values after accounting for predictor variables in the imputation model. Moderate violations of this assumption do not substantially bias MI results if the proportion of patients with missing data is low. We will evaluate the assumption with importance sampling to evaluate implications of violations of the missing at random assumption after completion of the basic MI procedure.[117] If this approach suggests strong dependence of the results on the missing at random assumption further sensitivity analyses will be performed using formal pattern mixture models.[118] Should sensitivity analyses indicate strong dependence of our results on untestable assumptions about missing data, this limitation to the interpretation of our results will be noted and conclusions appropriately qualified. Because this study involves procedures with minimal risk we will not plan interim analyzes to avoid inflating type I error rates. Treatment comparisons will be performed separately at each follow-up time to characterize treatment effects at critical junctures over the follow-up. The 4-week assessment will evaluate effects at completion of treatment; 6 month and 1 year assessments will evaluate intermediate and long-term effects respectively. We now lay out the analysis for each Specific Aim.

Specific Aim #1: "*Compare the effectiveness of two primary care management strategies for a subgroup of patients with LBP and sciatica.*"

For the primary (OSW) and secondary (NPRS, EQ-5D) outcomes we will use longitudinal mixed effects models to compare mean change in the outcome from baseline to each follow-up (4 weeks, 6 months, and 1 year) between the UC and EPT groups after controlling for the baseline level of the outcome [*and indicator variables to account for the different primary care clinics*.] An unstructured covariance matrix will be used to account for serial correlations in outcome measures within patients. Baseline adjustment

accounts for regression to the mean, increasing statistical power. [*Sensitivity analyses will be performed to examine the implications of possible clustering of outcomes by the physical therapist providing treatment in the EPT group by adding a random effect term for the physical therapist to the models*.] The primary analysis will compare mean OSW changes at 6 months. Comparisons between groups in mean OSW change at other time points and for other outcomes will be interpreted as secondary analyses. We will use a 2-sided alpha level of 0.05. Secondary analyses will be interpreted without adjustment for multiple comparisons. However we will employ a bootstrap procedure to obtain the probability of obtaining at least one p-value smaller than the minimum observed p-value from the secondary outcomes considered, based on observed correlations among the estimated treatment effects for each outcome.[119] This will provide an assessment of the probability that a given nominally significant result would have occurred by chance given the number of tests conducted and the observed association among the different outcome variables.

Specific Aim #2: "*Compare utilization of specific healthcare procedures and* [*healthcare costs*] *associated with two management strategies and evaluate the cost-effectiveness of the two strategies for a subgroup of patients with LBP and sciatica."*

Utilization outcomes evaluated during the year follow-up are: 1) advanced imaging (MRI or CT), 2) lumbar epidural injection, 3) emergency department visit, and 4) surgery (discectomy, decompression, laminectomy, nerve ablation, fusion). Utilization rates will be summarized on a patient basis (proportions of patients) and as rates over time (total utilization including multiple events per patient), and compared between groups with logistic regression and survival analysis methods for repeated events respectively.[127] We will compare numbers of patients missing work due to LBP between groups with similar procedures. [*We will examine costs from a societal perspective, collecting both direct and indirect costs using cost diaries. As recommended, we will value cost data with standard unit prices[132] based on University of Utah clinical research billing rates. Lost work productivity will be costed using methods described by Stewart.[28,133] Cost diaries document work absence or reduced performance (as a percentage of normal) due to losing concentration, repeating tasks, fatigue, etc, at work due to LBP. Lost productive time will be calculated as the number of days of work absence ("absenteeism") and the number of hours of reduced productivity ("presenteeism") multiplied by the patient's self-reported salary.[28] Patients not employed outside the home (homemakers, students) will report absenteeism and presenteeism. Lost productive time will be based on average salary for age- and sex-matched individuals according to the Bureau of Labor Statistics. Sensitivity analyses will be used to examine the impact of differing methods of valuing direct and indirect costs.*]

[*We will use parametric and non-parametric methods to compare costs between treatments. As the mean is the most useful statistic to evaluate costs related to treatment implementation,[134] we will calculate means and standard deviations for direct, indirect and total cost by group. We will explore cost data distributions graphically and statistically. With univariate and multivariate techniques we will examine the relationship between treatment and total cost. As cost data are typically skewed we will use nonparametric bootstrapping methods with 2000 pair-wise replications to compare mean costs and avoid distributional assumptions.[135] Confidence intervals around the mean cost difference will be calculated with bias corrected and accelerated methods.[136] Regression analyses will also be used to allow costs to be adjusted for patient factors such as socioeconomic status or co-morbidities. We will perform separate analyses for patients with complete cost data and those with imputed costs. Missing cost data will be handled using MI as described.[137] MI is preferred for missing cost data over alternatives such as last score forward or mean imputation.[139] Short- and long-term cost-effectiveness (CE) analyses will be conducted. The short-term analysis will use direct and indirect costs over the first 4 weeks. The effectiveness of these costs will be based on OSW and EQ-5D scores at 4 weeks. Long-term CE analysis will use cost and effect data measured at 1-year. Costs and effects will be measured in the same year; eliminating the need for inflation adjustment or discounts. Once 3 parameters are computed: 1) mean differences in cost and effect between treatments, 2) variances for differences in costs and effects, and 3) covariance between effectiveness and cost difference, the incremental cost-effectiveness ratio or incremental net benefits summarizing the monetary value of the intervention will be calculated.[140] A CE acceptability curve will be used to quantify and graphically depict uncertainty in the analysis.[142] To consider uncertainty in parameters, probabilistic sensitivity analysis will be conducted. To reflect the uncertainties in costs and in effects, a gamma distribution will be adopted.*]

Specific Aim #3: *"Evaluate moderators of improvement with early physical therapy."*

Moderators are baseline variables that interacts with the type treatment to predict outcomes.[126] We will pre-specify potential moderators of improvement with early physical therapy based on theory or preliminary

findings suggesting a moderating effect including; 1) symptom centralization at baseline (yes/no), 2) severe baseline disability (baseline OSW <40 or ≥40), 3) leg pain rated higher than LBP at baseline (leg NPRS > LBP NPRS or leg NPRS ≤ LBP NPRS). Mixed effects models described above will be extended by adding interaction terms for moderating variables using separate models for each outcome and each variable. The extended models will be used to provide separate estimates of treatment effects within each level of the moderator and to provide statistical tests of the dependence of treatment effects on the moderator. To mitigate risk of low power in subgroup analyses, the NPRS, for which the study has high power to detect important differences, will serve as the primary outcome for evaluating moderators. We may explore different follow-up points or dependent variables, and may consider additional putative moderators if evidence emerges suggesting a potentially important variable that may provide information to patients and providers about types of patients for whom early physical therapy may be particularly beneficial (or unhelpful).

C.10 Proposed Time Line

The timeline is based on our preliminary estimate of recruiting 8 patients per month. At this rate we need 28 months to recruit 220 patients. We allow 3 months at the beginning of the study to establish procedures and 4 months at the end to complete data analysis, manuscript preparation, etc.

<table>
<tr><th></th><th>Month 1</th><th>Month 2</th><th>Month 3</th><th>Month 4</th><th>Month 5</th><th>Month 6</th><th>Month 7</th><th>Month 8</th><th>Month 9</th><th>Month 10</th><th>Month 11</th><th>Month 12</th></tr>
<tr><td>Year 1</td><td colspan="3">Establish study procedures</td><td colspan="8">Recruit approximately 8 patients per month</td></tr>
<tr><td>Year 2</td><td colspan="11">Recruit 8 patients per month</td></tr>
<tr><td>Year 3</td><td colspan="8">Recruit 8 patients per month</td><td colspan="3">Complete follow-up evaluations</td></tr>
<tr><td>Year 4</td><td colspan="8">Complete follow-up evaluations</td><td colspan="3">Data analysis, report writing, etc.</td></tr>
</table>

C.11 Sample Size Justification

Assuming at least 90% of patients complete the OSW at 6 months, 110 subjects per group (total = 220) provides at least 86% power to detect a difference of 7 points on the OSW to 6 months, assuming a standard deviation of 16 points (treatment effect = 43.8% of 1 standard deviation). MCID for the OSW has been estimated at 6-8 points.[83,86] This sample size also provides at least 82% power to detect a treatment effect on the 1-year OSW change assuming the same standard deviation and 80% follow-up. In addition, this sample size provides > 99% power to detect a clinically important difference of 2.0 on the NPRS assuming a standard deviation of 2.4 and at least 80% follow-up. Our prior work with patients with LBP and sciatica indicate these estimates of effect are realistic.[52,72] Our power calculations are slightly conservative as they assume no information will be obtained from patients with missing outcomes; we can expect slightly smaller minimum detectable effects (i.e. greater power) with MI.

C.12 AHRQ Priority Populations

This project will include subjects from AHRQ Priority Populations including women, children (age 18-21), and minority groups as they are represented in the population from which study subjects will be recruited. Because this project is the first to examine the effectiveness and subsequent healthcare utilization resulting from the management strategies described in this proposal, we are taking a pragmatic approach to recruitment instead of targeting the recruitment specifically towards priority populations. Once more information is known about the effectiveness of these strategies, studies specifically examining translation into priority populations such as individuals in rural regions, low income or minority groups will be possible.

C.13 Limitations and Contingencies

As with most clinical studies, a primary concern is adequate recruitment. While we believe our recruitment accrual estimate is realistic and obtainable, we will maintain a contingency plan in case accrual falls short of our estimates. The UHRN practice-based research network includes primary care clinics that are not in our recruitment estimate because they are greater than 20 miles from Salt Lake City, but could be added if necessary. Another potential limitation would be subjects' ability to access physical therapy clinical sites for treatment. We are able to use space in physical therapy clinics affiliated with the University of Utah system located throughout Salt Lake City and could train additional clinics to provide the study treatment in order to increase the likelihood that a subject will be able to access a clinic for treatment.

LITERATURE CITED

1. *Initial National Priorities for Comparative Effectiveness Research.* Washington, DC: The Institute of Medicine; 2009.
2. Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol.* 2009;62:781-796.
3. Hayden JA, Dunn KM, van der Windt DA, Shaw WS. What is the prognosis of back pain? *Best Pract Res Clin Rheumatol.* 2010;24:167-179.
4. Chou R, Huffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med.* 2007;147:492-504.
5. Patel AT, Ogle AA. Diagnosis and management of acute low back pain. *Am Fam Physician.* 2000;61:1779-1790.
6. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine.* 2006;31:2724-2727.
7. Deyo RA, Phillips WR. Low back pain. A primary care challenge. *Spine.* 1996;21:2826-2832.
8. Cypress BK. Characteristics of physician visits for back symptoms: a national perspective. 1983;73:389-395.
9. Licciardone JC. The epidemiology and medical management of low back pain during ambulatory medical care visits in the United States. *Osteopath Med Prim Care.* 2008;2:11.
10. Towards Optimizing Practice. Guideline for the Evidence-Informed Primary Care Management of Low Back Pain. 2nd Ed. 2011. Institute for Health Economics, available at: http://www.topalbertadoctors.org/cpgs.php?sid=65&cpg_cats=90
11. Armon C, Argoff CE, Samuels J, Backonja MM. Assessment: use of epidural steroid injections to treat radicular lumbosacral pain: report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology.* 2007;68:723-729.
12. Chou R, Huffman LH. Clinical Guideline for the Evaluation and Management of Low Back Pain. Glenview, IL: American Pain Society; 2009.
13. Grotle M, Foster NE, Dunn KM, Croft P. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain.* 2010;151:790-797.
14. Hill JC, Konstantinou K, Egbewale BE, Dunn KM, Lewis M, van der Windt D. Clinical outcomes among low back pain consulters with referred leg pain in primary care. *Spine* 2011;36:2168-2175.
15. Selim AJ, Ren XS, Fincke G, et al. The importance of radiating leg pain in assessing health outcomes among patients with low back pain. Results from the Veterans' Health Study. *Spine.* 1998;23:470-474.
16. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum.* 2008;59:632-641.
17. Turner JA, Franklin G, Fulton-Kehoe D, et al. ISSLS prize winner: early predictors of chronic work disability: a prospective, population-based study of workers with back injuries. *Spine.* 2008;33:2809-2818.
18. Stover BD, Turner JA, Franklin G, et al. Factors associated with early opioid prescription among workers with low back injuries. *J Pain.* 2006;7:718-725.
19. Webster BS, Courtney TK, Huang YH, Matz S, Christiani DC. Physicians' initial management of acute low back pain versus evidence-based guidelines. Influence of sciatica. *J Gen Intern Med.* 2005;20:1132-1135.
20. Arts MP, Peul WC. Timing and minimal access surgery for sciatica: a summary of two randomized trials. *Acta Neurochir* 2011;153:967-974.
21. Deyo RA, Mirza SK, Turner JA, Martin BI. Overtreating chronic back pain: time to back off? *J Am Board Fam Med.* 2009;22:62-68.

22. Friedly J, Chan L, Deyo R. Increases in lumbosacral injections in the Medicare population: 1994 to 2001. *Spine.* 2007;32:1754-1760.
23. Martin BI, Deyo RA. Expenditures and health status among adults with back and neck problems. *JAMA.* 2008;299:656-664.
24. Martin BI, Turner JA, Mirza SK, Lee MJ, Comstock BA, Deyo RA. Trends in health care expenditures, utilization, and health status among US adults with spine problems, 1997-2006. *Spine.* 2009;34:2077-2084.
25. Yelin E. Cost of musculoskeletal diseases: impact of work disability and functional decline. *J Rheumatol.* 2003;68 (suppl):8-11.
26. Cohen JW, Monheit AC, Beauregard KM, et al. The Medical Expenditure Panel Survey: a national health information resource. *Inquiry.* 1996;33:373-389.
27. Druss BG, Marcus SC, Olfson M, Pincus HA. The most expensive medical conditions in America. *Health Affairs.* 2002;21:105-111.
28. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time and cost due to common pain conditions in the US workforce. *JAMA.* 2003;290:2443-2454.
29. Itz CJ, Geurts JW, van Kleef M, Nelemans P. Clinical course of non-specific low back pain: A systematic review of prospective cohort studies set in primary care. *Eur J Pain.* May 28 2012;e-pub ahead of print.
30. Menezes Costa Lda C, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LO. The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ.* Aug 7 2012;184(11):E613-624.
31. Vroomen PC, de Krom MC, Knottnerus JA. Predicting the outcome of sciatica at shortterm follow-up. *Br J Gen Pract.* 2002;52:119-123.
32. Luijsterburg PA, Verhagen AP, Ostelo RW, van Os TA, Peul WC, Koes BW. Effectiveness of conservative treatments for the lumbosacral radicular syndrome: a systematic review. *Eur Spine J.* 2007;16:881-899.
33. Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med.* 2007;147:478-491.
34. Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am.* 2012;50:569-585.
35. Chou R. Pharmacological management of low back pain. *Drugs.* 2010;70:387-402.
36. Weiner DK, Kim YS, Bonino P, Wang T. Low back pain in older adults: are we utilizing healthcare resources wisely? *Pain Med.* 2006;7:143-150.
37. Deyo RA, Mirza SK. Trends and variations in the use of spine surgery. *Clin Orthop Relat Res.* 2006;443:139-146.
38. Gray DT, Deyo RA, Kreuter W, et al. Population-based trends in volumes and rates of ambulatory lumbar spine surgery. *Spine.* 2006;31:1957-1963.
39. Mitchell JM. Utilization trends for advanced imaging procedures: evidence from individuals with private insurance coverage in California. *Med Care.* 2008;46:460-466.
40. Luo X, Pietrobon R, Hey L. Patterns and trends in opioid use among individuals with back pain in the United States. *Spine.* 2004;29:884-890.
41. Freburger JK, Holmes GM, Agans RP, al e. The rising prevalence of chronic low back pain. *Arch Intern Med.* 2009;169:251-258.
42. Wasiak R, Kim J, Pransky G. Work disability and costs caused by recurrence of low back pain: longer and more costly than in first episodes. *Spine.* 2006;31:219-225.
43. Webster BS, Courtney TK, Huang YH, Matz S, Christiani DC. Survey of acute low back pain management by specialty group and practice experience. *J Occup Environ Med.* 2006;48:723-732.

44. Freburger JK, Carey TS, Holmes GM. Physician referrals to physical therapists for the treatment of spine disorders. *Spine J.* 2005;5:530-541.
45. Cummins J, Lurie JD, Tosteson TD, et al. Descriptive epidemiology and prior healthcare utilization of patients in the Spine Patient Outcomes Research Trial's (SPORT) three observational cohorts: disc herniation, spinal stenosis, and degenerative spondylolisthesis. *Spine.* 2006;31:806-814.
46. Gelhorn AC, Chan L, Martin BI, Friedly J. Management patterns in acute low back pain: the role of physical therapy. *Spine.* 2012;37:775-782.
47. Fritz JM, Childs JD, Wainner RS, Flynn TW. Primary care referral of patients with a new consultation for low back pain to physical therapy: the impact of the timing and content of care on future healthcare utilization and costs. *Spine.* 2012;37:2114-2121.
48. Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. *Phys Ther.* 2001;81:1018-1028.
49. Poitras S, Blais R, Swaine B, Rossignol M. Management of work-related low back pain: a population-based survey of physical therapists *Phys Ther.* 2005;85:1168-1181.
50. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying sub-groups of patients with "non-specific" low back pain: results of a randomized clinical trial. *Spine.* 2006;31:623-631.
51. Fritz JM, Delitto A, Erhard RE. Comparison of a classification-based approach to physical therapy and therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine.* 2003;28:1363-1372.
52. Fritz JM, Lindsay W, Matheson JW, et al. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction?: Results of a randomized clinical trial and subgrouping analysis. *Spine.* 2007;32:E793-E800.
53. Long AL, Donelson R. Does it matter which exercise? A randomized trial of exercise for low back pain. *Spine.* 2004;29:2593-2602.
54. Kent P, Mjosund HL, Petersen DH. Does targeting manual therapy and/or exercise improve patient outcomes in nonspecific low back pain? A systematic review. *BMC Med.* 2010;8:22.
55. Coxhead CE, Inskip H, Meade TW, North.W.R., Troup JD. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet* 1981;1:1065-1068.
56. Lidström A, Zachrisson M. Physical therapy on low back pain and sciatica. An attempt at evaluation. *Scand J Rehabil Med.* 1970;2:37-42.
57. Hofstee DJ, Gijtenbeek JM, Hoogland PH, et al. Westeinde sciatica trial: randomized controlled study of bed rest and physiotherapy for acute sciatica. *J Neurosurg.* 2002;96:45-49.
58. Luijsterburg PA, Verhagen AP, Ostelo RW, et al. Physical therapy plus general practitioners' care versus general practitioners' care alone for sciatica: a randomised clinical trial with a 12-month follow-up. *Eur Spine J.* 2008;17:509-517.
59. Luijsterburg PA, Lamers LM, Verhagen AP, et al. Cost-effectiveness of physical therapy and general practitioner care for sciatica. *Spine* 2007;32:1942-1948.
60. Albert HB, Manniche C. The efficacy of systematic active conservative treatment for patients with severe sciatica. *Spine.* 2012;37:531-542.
61. Cherkin DC, Deyo RA, Loeser JD, Bush T, Waddell G. An international comparison of back surgery rates. *Spine.* 1994;19:1201-1206.
62. Squires DA. The U.S. health system in perspective: a comparison of twelve industrialized nations. In: Fund TC, ed. Vol 1532 New York: The Commonwealth Fund; 2011.
63. Gudavalli MR, Cambron JA, McGregor M, et al. A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for patients with chronic low back pain. *Eur Spine J.* 2006;15:1070-1082.
64. Klaber-Moffett JA, Carr J, Howarth E. High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. *Spine.* 2004;29:1167-1172.

65. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *J Personal Social Psychol.* 1986;51:1173-1182.
66. Miles CL, Pincus T, Carnes D, et al. Can we identify how programmes aimed at promoting self-management in musculoskeletal pain work and who benefits? A systematic review of sub-group analysis within RCTs. *Eur J Pain.* 2011;15: e711-755.
67. Bigos S, Bowyer O, Braen G, et al. *Acute Low Back Problems in Adults.* AHCPR Publication 95-0642 ed. Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, US Department of Health and Human Services; 1994.
68. Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. *J Orthop Sports Phys Ther.* 2007;37:290-302.
69. McKenzie RA, May S. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. 2nd ed. Minneapolis, MN: Orthopedic Physical Therapy Products; 2003.
70. Skytte L, May S, Petersen P. Centralization: its prognostic value in patients with referred symptoms and sciatica. *Spine.* 2005;30:E293-299.
71. Albert HB, Hauge E, Manniche C. Centralization in patients with sciatica: are pain responses to repeated movement and positioning associated with outcome or type of disc lesion? *Eur Spine J.* 2012;21:630-636.
72. Browder DA, Childs JD, Cleland JA, Fritz JM. Effectiveness of an extension oriented treatment approach in a subgroup of patients with low back pain: a randomized clinical trial. *Phys Ther.* 2007;87:1608-1618.
73. Fritz JM, Thackeray A, Childs JD, Brennan GP. A randomized clinical trial of the effectiveness of mechanical traction for sub-groups of patients with low back pain: study methods and rationale. *BMC Musculoskelet Disord.* 2010;11:81.
74. van Tulder MW, Koes B, Seitsalo S, Malmivaara A. Outcome of invasive treatment modalities on back pain and sciatica: an evidence-based review. *Eur Spine J.* 2006;15(suppl):S82-92.
75. Young G, Jewell D. Interventions for preventing and treating pelvic and back pain in pregnancy. *Cochrane Database of Systematic Reviews.* 2002;1:CD001139.
76. Fritz JM, Cleland JA, Speckman M, Brennan GP, Hunter SJ. Physical therapy for acute low back pain: associations with subsequent health care costs. *Spine.* 2008;33:1800-1805.
77. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. New York: Oxford University Press; 2000.
78. Waddell G. Clinical assessment of lumbar impairment. *Clin Orthop.* 1987;221:110-120.
79. Fritz JM, Piva SR. Physical impairment index: reliability, validity, and responsiveness in patients with acute low back pain. *Spine.* 2003;28:1189-1194.
80. Torgerson DJ, Klaber M, J A., Russell IT. Patient preferences in randomised trials: threat or opportunity? *J Health Serv Res Policy.* 1996;1:194-197.
81. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy.* 1980;66:271-273.
82. Ong BN, Konstantinou K, Corbett M, Hay E. Patients' own accounts of sciatica: a qualitative study. *Spine.* 2011;36:1251-1256.
83. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry disability questionnaire and the Quebec back pain disability scale. *Phys Ther.* 2001;81:776-788.
84. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain.* 1994;58:387-392.
85. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine.* 1998;23:2003-2013.
86. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskelet Disord.* 2006;7:82.

87. Walsh TL, Hanscom B, Lurie JD, Weinstein JN. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. *Spine.* 2003;28:607-615.
88. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine.* 2005;30:1331-1335.
89. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine.* 1995;20:1899-1908.
90. Grøvle L, Haugen AJ, Keller A, Natvig B, Brox JI, Grotle M. Reliability, validity, and responsiveness of the Norwegian versions of the Maine-Seattle Back Questionnaire and the Sciatica Bothersomeness and Frequency Indices. *Spine.* 2008;33:2347-2353.
91. EuroQol Group. EuroQol–A new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16:199-208.
92. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. *Med Desc Making.* 2006;26:410-420.
93. Dolan P, Gudex C, Kind P, et al. The time trade-off method: results from a general population survey. *Health Econ.* 1996;5:141-154.
94. Campbell H, Rivero-Arias O, Johnston K, Gray A, Fairbank JC, Frost H. Responsiveness of objective, disease-specific, and generic outcome measures in patients with chronic low back pain: an assessment for improving, stable, and deteriorating patients. *Spine.* 2006;31:815-822.
95. Solberg TK, Olsen JA, Ingebrigtsen T, Hofoss D, Nygaard OP. Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *Eur Spine J.* 2005;14:1000-1007.
96. Chan CW, Goldman S, Ilstrup DM, Kunselman AR, O'Neill PI. The pain drawing and Waddell's nonorganic physical signs in chronic low-back pain. *Spine.* 1993;18:1717-1722.
97. Beattie PF, Meyers SP, Stratford P, LMillard RW, Hollenberg GM. Associations between patient report of symptoms and anatomic impairment visible on lumbar magnetic resonance imaging. *Spine.* 2000;25:819-828.
98. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain.* 1993;52:157-168.
99. Jacob T, Baras M, Zeev A, Epstein L. Low back pain: reliability of a set of pain measurement tools. *Arch Phys Med Rehabil.* 2001;82:735-742.
100. Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than fear itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain.* 1999;80:329-339.
101. Fritz JM, George SZ, Delitto A. The role of fear avoidance beliefs in acute low back pain: relationships with current and future disability and work status. *Pain.* 2001;94:7-15.
102. Buer N, Linton SJ. Fear-avoidance beliefs and catastrophizing: occurrence and risk factor in back pain and ADL in the general population. *Pain.* 2002;99:485-491.
103. Fritz JM, George SZ. Identifying specific psychosocial factors in patients with acute, work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther.* 2002;82:973-983.
104. Arnow BA, Blasey CM, Lee J, et al. Relationships among depression, chronic pain, chronic disabling pain, and medical costs. *Psychiatr Serv.* 2009;60:344-350.
105. Zigmont AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; 67:361-370.
106. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess.* 1995;7:524-532.
107. Goossens ME, Rutten Van Molken MP, Kole Snijders AM, Vlaeyen JW, Van Breukelen G, Leidl R. Health economic assessment of behavioural rehabilitation in chronic low back pain: a randomised clinical trial. *Health Econ.* 1998;7:39-51.

108. Goossens ME, Rutten-van Molken MP, Vlaeyen JW, van der Linden SM. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol.* 2000;53 688-695.

109. Freer CB. Health diaries: a method of collecting health information. *J R Coll Gen Pract.* 1980;30:279-282.

110. Goossens ME, Rutten Van Molken MP, Leidl R, Bos SG, Vlaeyen JW, et al. Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. Economic evaluation. *J Rheumatol.* 1996;23:1246-1254.

111. Cherkin DC, Deyo RA, Berg AO. Evaluation of a physician education intervention to improve primary care for low back pain. Impact on patients. *Spine.* 1991;16:1173-1178.

112. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10:407-415.

113. Burton AK, Waddell G, Tillotson KM, Summerton N. Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine.* 1999;24:2484-2491.

114. George SZ, Teyhen DS, Wu SS, et al. Psychosocial education improves low back pain beliefs: results from a cluster randomized clinical trial (NCT00373009) in a primary prevention setting. *Eur Spine J.* 2009;18:1050-1058.

115. Van Tulder MW, Becker A, Bekkering T, et al. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J.* 2006;15 (suppl):S169-191.

116. Schafer JL. Multiple imputation: A primer. *Stat Methods Med Res.* 1999;8:3-15.

117. Carpenter J, Kenward M. Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Stat Methods Med Res.* 2007;16:259-275.

118. Daniels M, Hogan J. Reparameterizing the pattern mixture model for sensitivity analyses under informative dropout. *Biometrics.* 2000;56:1241-1248.

119. Westfall PJ, Young SS. *Resampling-Based Multiple Testing*. New York: John Wiley & Sons; 1993.

120. van der Roer N, Boos N, van Tulder MW. Economic evaluations: a new avenue of outcome assessment in spinal disorders. *Eur Spine J.* 2006;15 Suppl 1:S109-117.

121. Barber JA, Thompson SG. Analysis and interpretation of cost data in randomised controlled trials: review of published studies. *BMJ.* 1998;317:1195-2000.

122. Barber JA, Thompson SG. How should cost data in pragmatic trials be analysed? *BMJ.* 2000; 320:1197-1200.

123. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the nonparametric bootstrap. *Stat Med.* 2000;19:3219-3226.

124. Oostenbrink JB, Al MJ, Rutten Van Molken MP. Methods to analyse cost data of patients who withdraw in a clinical trial setting. *Pharmacoecon.* 2003;21:1103-1112.

125. Oostenbrink JB, Al MJ. The analysis of incomplete cost data due to dropout. *Health Econ.* 2005;14:763-776.

126. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatr* 2002;59:877-883

127. Therneau T and Grambsch P. Modeling Survival Data: Extending the Cox Model, Chapter 8, Multiple Events per Subject, pp 169-230, Springer, New York, 2000

128. Fritz JM, Brennan GP, Hunter SJ, Magel JS. Initial management decisions after a new consultation for low back pain: implications for the usage of physical therapy for subsequent health care costs and utilization. *Arch Phys Med Rehabil* 2013;94:808-816.

129. Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? *Med Care* 2007;45:973-980.

130. Stafford MA, Peng P, Hill DA. Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. *Br J Anaesth* 2007;99:461-473.

131. Goossens ME, Rutten Van Molken MP, Kole Snijders AM, et al. Health economic assessment of behavioural rehabilitation in chronic low back pain: a randomised clinical trial. *Health Econ* 1998;7:39-51.

132. van der Roer N, Boos N, van Tulder MW. Economic evaluations: a new avenue of outcome assessment in spinal disorders. *Eur Spine J* 2006;15 Suppl 1:S109-17.

133. Stewart WF, Ricci JA, Chee E, et al. Lost productive work time costs from health conditions in the Uunited States: results from the American Productivity Audit. *J Occup Environ Med* 2003;45:1234-46.

134. Barber JA, Thompson SG. Analysis and interpretation of cost data in randomised controlled trials: review of published studies. *BMJ* 1998;317:1195-2000.

135. Barber JA, Thompson SG. How should cost data in pragmatic trials be analysed? *BMJ* 2000;320:1197-200.

136. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the nonparametric bootstrap. *Stat Med* 2000;19:3219-26.

137. Oostenbrink JB, Al MJ, Rutten Van Molken MP. Methods to analyse cost data of patients who withdraw in a clinical trial setting. *Pharmacoecon* 2003;21:1103-12.

138. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996;91:1913-25.

139. Oostenbrink JB, Al MJ. The analysis of incomplete cost data due to dropout. *Health Econ* 2005;14:763-76.

140. Willan AR and Briggs AH. Statistical analysis of cost-effectiveness data. West Sussex, England. John Wiley. 2006.

141. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves-facts, fallacies and frequently asked questions. Health Economics. 2004; 13: 405-415.

142. Whitman JM, Fritz JM, Childs JD. The influence of experience and specialty certifications on clinical outcomes for patients with low back pain treated within a standardized physical therapy management program. *J Orthop Sports Phys Ther* 2004;34:662-72.

Statistical Supplement

1. *Scope of Statistical Analyses for the current manuscript*. This manuscript presents our principal statistical analyses for the effect of the randomized treatment interventions on the primary and secondary patient reported outcomes and on utilization outcomes. Additional analyses which are detailed in the protocol, particularly those related to cost-effectiveness and moderation of treatment effects by baseline factors are planned for subsequent publications. Our final analyses for the first manuscript also implemented the following modifications from those specified in the original protocol: a) We present an as-treated analysis in place of a per-protocol analysis as a sensitivity analysis to address patients who did not receive their randomly assigned therapy, b) analyses involving repeated events for utilization outcomes, clustering by physical therapist and primary care clinics, and bootstrap assessments of studywise type-1 error were omitted for simplicity.

2. *Analysis of primary and secondary numeric outcomes*. We used longitudinal mixed effects analysis of covariance models to estimate the effects of the randomized treatments on the mean changes from baseline to each follow-up assessment of the primary OSW outcome and numeric outcomes while controlling for the baseline levels of each outcome variable. For each outcome, the mixed model used unstructured covariance matrices to model the serial correlation in responses across the three follow-up visits at 4 weeks, 6 months and 1 year. The model included fixed effects for follow-up time treated as a categorical variable, the baseline level of the outcome, randomized treatment group, and interactions of follow-up time with the randomized treatment and the baseline level. Model parameters were estimated using restricted maximum likelihood estimation.

In sensitivity analyses, the mixed effects analyses of the primary and secondary numeric outcomes were repeated using an as-treated strategy in which 10 of the 110 EPT subjects who did actually receive the intervention were assigned to the usual care group and 5 of the 110 UC subjects who did receive the EPT intervention were assigned to the EPT group.

3. *Multiple imputation* analysis of primary and secondary outcomes. The results of the mixed effects analyses assume that missing data follow a missing data follow a missing at random (MAR) pattern in which the probability of missingness may depend on other observed outcome values in the model, but are not related to the unobserved values of missing responses themselves. To assess the robustness of our results to bias from missing data, we repeated the mixed effects analyses of the primary and secondary numeric outcomes after multiply imputing missing data using fully sequential imputation under an imputation model that incorporated additional auxiliary variables that were selected using subject matter considerations as potential predictors of the outcomes or the probability of missingness. The multiple imputation models included all variables in the analytic model plus the baseline assessments of gender, age, baseline anxiety, depression, concurrent neck or upper back pain, BMI, length in days of current episode of LBP, and the pain catastrophizing scale, as well as indicator variables for surgery or injections prior to the 4 week and 6 month assessments, and nonmissing longitudinal assessments of the OSW, FABQ physical activity and work subscales, EQ-5D quality of life index, EQ-5D self-rating, and the ratings of low back pain and leg pain in the preceding 24 hours. A total of 25 imputed data sets were created. Results were pooled across these 25 data sets using Rubin's formulae to adjust standard errors to account for variation in results between the imputed data sets in order to account for uncertainty resulting from imputation of missing data.

The sensitivity analyses based on multiple imputation retain a MAR assumption, but the assumption is relaxed to allow dependence of missing outcomes and probability of missingness of the auxiliary variables in the imputation model.

4. *Analyses of utilization outcomes and patient-reported success*. Our main analyses of each of the utilization outcomes – surgery, lumbar epidural injection, advanced imaging and emergency department visits - were performed by calculating relative risk statistics and 95% confidence intervals.

We used relative risk rather than Cox regression in our main analyses of each utilization outcome because these outcomes generally occurred early in follow-up and we did not view the timing of the outcome events as highly clinically relevant. In order to assess the dependence of the results to early dropout and missing assessments, we repeated our analyses of each of the utilization outcomes using Cox regression analysis to relate the outcome to the randomized treatment assignment, with follow-up censored at the final monthly utilization assessment. The Cox regression analyses were performed using a counting process style input in which intermittently missing monthly assessments were excluded from the risk set.

We applied separate analyses of relative risk to relate patient-reported success at the 4 week, 6 month and 1 year assessments to the randomized treatment assignment.

5. *Analyses of missed work days*. We used a generalized estimating equations (GEE) analysis with robust standard errors under a negative binomial outcome model with logarithmic link function to compare the number of missed work days due to LBP per month over the 1 year follow-up period between the randomized groups. The GEE analysis was performed using a compound symmetry working covariance matrix to account for serial correlation over the monthly assessments during the follow-up period. The model included randomized treatment assignment and follow-up month coded as a categorical factor as predictor variables.

A separate GEE analysis used a Bernoulli outcome model with logarithmic link to compare the proportion of patients missing at least one work day in a given month between the randomized treatment groups. This analysis also used a compound symmetry working covariance model and robust standard errors for statistical inference.

Primary and Secondary Outcomes for As Treated Sensitivity Analysis
Repeated Measures Mixed Models Analysis

For the “as-treated” analysis; 10 participants randomized to EPT group who did not receive PT during the 4-week treatment period were analyzed with the UC group. Five participants randomized to UC group who did receive PT during the 4-week treatment period were analyzed with the EPT group

Supplement Table 1. As Treated Sensitivity Analysis

Outcome	Visit	Usual Care (n=115) (no PT sessions in the 4 week treatment period)		Early Physical Therapy (n=105) (≥ 1 PT sessions in the 4 week treatment period)		**Difference Between Groups[+]** Relative Difference (95% CI)
		Mean (95% CI)	Mean change from baseline (95% CI)	Mean (95% CI)	Mean change from baseline (95% CI)	
Oswestry Disability Index Score	Baseline	35.3 (32.5, 38.1)		39.5 (36.8, 42.2)		
	4 week	27.4 (24.7, 30.1)	-9.5 (-12.2, -6.7)	20.4 (17.6, 23.2)	-16.5 (-19.3, -13.7)	-7.0 (-10.9, -3.1)
	6 month	20 (17.2, 22.9)	-16.9 (-19.7, -14.0)	14.3 (11.4, 17.2)	-22.6 (-25.5, -19.7)	-5.7 (-9.8, -1.6)
	1 Year	18.6 (15.7, 21.5)	-18.3 (-21.2, -15.4)	15.1 (12.2, 18.0)	-21.8 (-24.7, -18.9)	-3.5 (-7.7, 0.6)
Numeric Pain Rating (back)	Baseline	4.9 (4.6, 5.3)		5.0 (4.6, 5.3)		
	4 week	3.8 (3.4, 4.2)	-1.1 (-1.5, -0.7)	2.5 (2.1, 2.9)	-2.4 (-2.8, -2.0)	-1.2 (-1.8, -0.7)
	6 month	3.3 (2.9, 3.7)	-1.6 (-2.0, -1.2)	2.6 (2.2, 3)	-2.3 (-2.7, -1.8)	-0.7 (-1.3, -0.1)
	1 Year	3.2 (2.8, 3.6)	-1.7 (-2.1, -1.2)	2.4 (1.9, 2.8)	-2.5 (-2.9, -2.1)	-0.9 (-1.5, -0.3)
Numeric Pain Rating (leg)	Baseline	4.0 (3.5, 4.4)		4.1 (3.7, 4.6)		
	4 week	2.9 (2.5, 3.3)	-1.1 (-1.5, -0.7)	2.2 (1.8, 2.7)	-1.8 (-2.2, -1.3)	-0.7 (-1.3, -0.1)
	6 month	2.1 (1.7, 2.6)	-1.9 (-2.3, -1.4)	2.2 (1.8, 2.7)	-1.8 (-2.2, -1.3)	0.1 (-0.6, 0.7)
	1 Year	2.1 (1.7, 2.6)	-1.9 (-2.3, -1.4)	1.9 (1.4, 2.3)	-2.1 (-2.6, -1.7)	-0.3 (-0.9, 0.4)
Fear Avoidance Beliefs – Physical Activity	Baseline	14.1 (13.0, 15.2)		15.1 (14, 16.1)		
	4 week	11.0 (9.8, 12.2)	-3.4 (-4.6, -2.2)	9.7 (8.5, 11)	-4.7 (-5.9, -3.5)	-1.3 (-3, 0.4)
	6 month	9.4 (8.1, 10.6)	-5.1 (-6.3, -3.8)	9.1 (7.8, 10.3)	-5.4 (-6.7, -4.1)	-0.3 (-2.1, 1.5)
	1 Year	10.1 (8.9, 11.4)	-4.3 (-5.6, -3.0)	7.8 (6.5, 9.0)	-6.7 (-8.0, -5.4)	-2.4 (-4.2, -0.6)

Supplement Table 1. As Treated Sensitivity Analysis						
		Usual Care (n=115) (no PT sessions in the 4 week treatment period)		Early Physical Therapy (n=105) (≥ 1 PT sessions in the 4 week treatment period)		**Difference Between Groups+**
Outcome	Visit	Mean (95% CI)	Mean change from baseline (95% CI)	Mean (95% CI)	Mean change from baseline (95% CI)	Relative Difference (95% CI)
Fear Avoidance Beliefs – Work	Baseline	13.9 (11.7, 16.1)		15.4 (13, 17.7)		
	4 week	12.4 (10.8, 14.1)	-1.7 (-3.4, 0)	11.4 (9.7, 13.1)	-2.7 (-4.5, -1)	-1.0 (-3.4, 1.4)
	6 month	10.8 (9.1, 12.6)	-3.3 (-5.1, -1.6)	9.3 (7.5, 11.1)	-4.9 (-6.6, -3.1)	-1.5 (-4, 1)
	1 Year	10.6 (8.8, 12.3)	-3.6 (-5.3, -1.8)	9.0 (7.2, 10.8)	-5.1 (-7, -3.3)	-1.6 (-4.1, 0.9)
EQ-5D Quality of Life☒	Baseline	0.64 (0.6, 0.67)		0.64 (0.6, 0.68)		
	4 week	0.71 (0.67, 0.74)	0.06 (0.03, 0.09)	0.75 (0.72, 0.79)	0.10 (0.07, 0.14)	0.04 (0, 0.09)
	6 month	0.79 (0.76, 0.83)	0.15 (0.11, 0.18)	0.79 (0.76, 0.83)	0.15 (0.11, 0.18)	0.00 (-0.05, 0.05)
	1 Year	0.79 (0.76, 0.83)	0.14 (0.11, 0.18)	0.81 (0.77, 0.85)	0.16 (0.13, 0.2)	0.02 (-0.03, 0.07)
EQ-5D Overall Health Self-Rating	Baseline	57.2 (53.5, 60.8)		55.3 (51.4, 59.2)		
	4 week	62.5 (59.2, 65.8)	5.9 (2.6, 9.2)	69.8 (66.4, 73.2)	13.2 (9.7, 16.6)	7.3 (2.5, 12.1)
	6 month	68.5 (65, 72)	11.9 (8.4, 15.4)	68.5 (64.9, 72.2)	11.9 (8.3, 15.5)	0.0 (-5, 5)
	1 Year	72.0 (68.4, 75.7)	15.4 (11.8, 19)	76.0 (72.3, 79.7)	19.4 (15.7, 23.1)	4.0 (-1.2, 9.2)
Pain Catastrophizing Score	Baseline	19.6 (17.4, 21.9)		20.3 (17.9, 22.7)		
	4 week	15.5 (13.7, 17.3)	-3.7 (-5.6, -1.9)	13.7 (11.8, 15.6)	-5.6 (-7.5, -3.7)	-1.8 (-4.5, 0.8)
	6 month	12.4 (10.4, 14.3)	-6.9 (-8.8, -4.9)	11.4 (9.4, 13.3)	-7.9 (-9.8, -5.9)	-1.0 (-3.7, 1.7)
	1 Year	11.1 (9.1, 13.0)	-8.2 (-10.1, -6.3)	10.7 (8.6, 12.7)	-8.6 (-10.6, -6.5)	-0.4 (-3.2, 2.4)

☒ EQ5D Quality of life score based on U.S. valuations ranging from 0.0 – 1.0 with higher scores indicating greater quality of life. EQ5D overall health self-rating ranges from 0 – 100 with higher numbers indicating greater self-rated overall health

+ Analyses were conducted with unstructured covariance matrix. Differences between groups for each outcome controlled for the baseline level of the outcome. A positive difference favors Early Physical Therapy for the EQ-5D Quality of Life and Self-Rated Health outcomes. A negative difference favors Early Physical Therapy for the Oswestry, numeric pain rating for back and leg pain, fear-avoidance and pain catastrophizing outcomes.

Primary and Secondary Outcomes for Multiple Imputation Sensitivity Analysis Repeated Measures Mixed Models Analysis

For the multiple imputation analysis we used fully sequential multiple imputation with an imputation model that included auxiliary variables (listed in statistical supplement) to reduce the risk of bias due to loss-to-follow-up and missing follow-up measurements. The imputation model generated multiple replacement values for each missing score which were drawn randomly from the posterior distribution and perturbed with error; then analyses were done identically on all versions of imputed data. Parameter estimates and tests were combined and adjusted for between-imputation variance to yield final results.

Supplement Table 2. Multiple Imputation Sensitivity Analysis

Outcome	Visit	Usual Care (n=110)		Early Physical Therapy (n=110)		**Difference Between Groups+**
		Mean (95%CI)	Mean Change from Baseline (95%CI)	Mean (95%CI)	Mean Change from Baseline (95%CI)	Relative Difference (95% CI)
Oswestry Disability Index	Baseline	35.8 (34.1, 37.5)		38.9 (37.4, 40.3)		- .
	4 weeks	28.3 (25.4, 31.1)	-9.1 (-11.9, -6.2)	20.6 (17.7, 23.4)	-16.8 (-19.6, -13.9)	-7.7 (-11.8, -3.6)
	6 months	20.1 (17.4, 22.8)	-17.2 (-20, -14.5)	15.4 (12.7, 18.1)	-21.9 (-24.6, -19.2)	-4.7 (-8.6, -0.8)
	1 year	20 (17.1, 22.9)	-17.3 (-20.2, -14.4)	15.2 (12.2, 18.2)	-22.1 (-25.1, -19.1)	-4.8 (-8.9, -0.6)
Numeric Pain Rating (back pain)	Baseline	4.8 (4.6, 5)		5.1 (4.9, 5.3)		
	4 weeks	3.9 (3.5, 4.3)	-1.0 (-1.4, -0.7)	2.5 (2.1, 2.9)	-2.4 (-2.8, -2.1)	-1.4 (-2, -0.9)
	6 months	3.4 (3.0, 3.8)	-1.6 (-2.0, -1.1)	2.8 (2.3, 3.2)	-2.2 (-2.6, -1.8)	-0.7 (-1.3, -0.1)
	1 year	3.4 (3.0, 3.8)	-1.6 (-2.0, -1.1)	2.5 (2.0, 2.9)	-2.5 (-3.0, -2.1)	-1.0 (-1.6, -0.3)
Numeric Pain Rating (leg pain)	Baseline	3.8 (3.6, 4.1)		4.2 (4.0, 4.5)		.
	4 weeks	3.0 (2.6, 3.5)	-1.0 (-1.4, -0.6)	2.2 (1.8, 2.6)	-1.8 (-2.2, -1.4)	-0.8 (-1.4, -0.2)
	6 months	2.2 (1.8, 2.7)	-1.8 (-2.3, -1.4)	2.3 (1.9, 2.8)	-1.7 (-2.2, -1.2)	0.1 (-0.5, 0.8)
	1 year	2.3 (1.9, 2.8)	-1.7 (-2.2, -1.3)	2.0 (1.5, 2.4)	-2.1 (-2.6, -1.6)	-0.4 (-1.0, 0.3)
Fear Avoidance Beliefs – Physical Activity	Baseline	14.0 (13.3, 14.6)		15.2 (14.5, 15.8)		
	4 weeks	11.3 (10.2, 12.5)	-3.2 (-4.4, -2.0)	9.7 (8.5, 10.8)	-4.9 (-6.1, -3.7)	-1.7 (-3.4, 0.0)
	6 months	9.7 (8.5, 11.0)	-4.8 (-6.1, -3.6)	9.1 (7.8, 10.5)	-5.4 (-6.7, -4.1)	-0.6 (-2.4, 1.2)
	1 year	10.7 (9.4, 12.0)	-3.9 (-5.2, -2.6)	7.6 (6.2, 9.0)	-7.0 (-8.4, -5.6)	-3.1 (-5, -1.2)
	Baseline	14.0 (12.7, 15.2)		15.7 (14.4, 17.1)		

Supplement Table 2. Multiple Imputation Sensitivity Analysis						
		Usual Care (n=110)		Early Physical Therapy (n=110)		**Difference Between Groups+**
Outcome	Visit	Mean (95%CI)	Mean Change from Baseline (95%CI)	Mean (95%CI)	Mean Change from Baseline (95%CI)	Relative Difference (95% CI)
Fear Avoidance Beliefs – Work	4 weeks	13.2 (11.5, 14.9)	-1.6 (-3.3, 0.0)	11.9 (10.2, 13.7)	-2.9 (-4.7, -1.2)	-1.3 (-3.7, 1.1)
	6 months	11.3 (9.6, 13.0)	-3.5 (-5.3, -1.8)	9.8 (8.0, 11.6)	-5.0 (-6.8, -3.3)	-1.5 (-4.0, 0.9)
	1 year	12.2 (10.2, 14.1)	-2.7 (-4.7, -0.7)	9.1 (7.0, 11.2)	-5.7 (-7.8, -3.6)	-3.1 (-6.0, -0.1)
EQ-5D Quality of Life ☒	Baseline	0.64 (0.62, 0.67)		0.63 (0.61, 0.65)		
	4 weeks	0.69 (0.66, 0.73)	0.05 (0.02, 0.09)	0.76 (0.72, 0.79)	0.12 (0.09, 0.15)	0.07 (0.02, 0.12)
	6 months	0.78 (0.74, 0.81)	0.14 (0.10, 0.18)	0.79 (0.75, 0.82)	0.15 (0.11, 0.18)	0.01 (-0.04, 0.06)
	1 year	0.77 (0.73, 0.81)	0.13 (0.10, 0.17)	0.81 (0.77, 0.85)	0.17 (0.13, 0.21)	0.04 (-0.01, 0.09)
EQ-5D Self-Rated Health	Baseline	57.8 (55.6, 59.9)		54.4 (52, 56.8)		
	4 weeks	61.6 (58.1, 65.1)	5.5 (2.0, 9.0)	69.7 (66.3, 73.2)	13.7 (10.1, 17.2)	8.1 (3.1, 13.1)
	6 months	67.3 (63.3, 71.3)	11.2 (7.1, 15.3)	67.8 (63.9, 71.7)	11.7 (7.7, 15.7)	0.5 (-5.1, 6.2)
	1 year	70.4 (66.9, 74.0)	14.4 (10.8, 17.9)	73.6 (70.2, 77.0)	17.5 (14.0, 21.1)	3.2 (-1.9, 8.2)

☒ EQ5D Quality of life score based on U.S. valuations ranging from 0.0 – 1.0 with higher scores indicating greater quality of life. EQ5D overall health self-rating ranges from 0 – 100 with higher numbers indicating greater self-rated overall health

+ Analyses were conducted with unstructured covariance matrix. Differences between groups for each outcome controlled for the baseline level of the outcome. A positive difference favors Early Physical Therapy for the EQ-5D Quality of Life and Self-Rated Health outcomes. A negative difference favors Early Physical Therapy for the Oswestry, numeric pain rating for back and leg pain, fear-avoidance and pain catastrophizing outcomes.

Sensitivity Analyses Outcomes Using Cox Regression Analyses

Cox Regression analyses were conducted for the time from randomization to the first instance of the utilization events of injections, surgery (discectomy, laminectomy, fusion), or advanced imaging: MRI or CT. Follow-up months for which subjects did not provide utilization data were excluded from the risk set. Subjects were censored at the last month in which the utilization questionnaire was completed.

Supplement Table 3. Cox Regression for First Event						
	Summary of Number of Event and Censored Values				Treatment Group Comparison	
Outcome	Total	Event	Censored	% Censored	Hazard Ratio	95% CI
Injection	2135	29	2106	98.3%	1.12	(0.54 – 2.32)
Surgery (discectomy, laminectomy, or fusion)	2216	16	2200	99.3%	1.34	(0.50 – 3.61)
Advanced Imaging: MRI or CT	1946	55	1891	97.2%	0.76	(0.45 – 1.30)

Analysis of Missed Work Days

The mean number of workdays missed each month was modeled using generalized estimating equations for a negative binomial outcome with a working compound symmetry covariance model to account for serial correlation within subjects. Statistical inference was based on robust standard errors. The analysis included treatment group and follow-up month as predictor variables.

.

Supplement Table 4. Analysis of the Mean Number of Work Days Missed per Month

	Adjusted Mean Days Missed (95% CI)			Ratio of Mean Missed Workdays Between Treatment Groups	95% Confidence Interval for Ratio
Treatment Group	EPT	0.7	(0.5, 1.1)	0.60	(0.33, 1.11)
	UC	1.2	(0.8, 2.0)		
Follow-up Month	1	2.1	(1.5, 3.1)		
	2	1.7	(1.1, 2.6)		
	3	2	(1.3, 3.0)		
	4	0.8	(0.5, 1.2)		
	5	0.8	(0.5, 1.2)		
	6	0.8	(0.5, 1.2)		
	7	0.6	(0.4, 1.0)		
	8	0.5	(0.3, 0.8)		
	9	0.9	(0.6, 1.4)		
	10	0.9	(0.5, 1.6)		
	11	0.8	(0.5, 1.4)		
	12	0.7	(0.4, 1.1)		

The probability of missing at least 1 workday in a given month was modeled using generalized estimating equations for a Bernoulli outcome with a logarithmic link function and working compound symmetry covariance model to account for serial correlation within subjects. Statistical inference was based on robust standard errors. The analysis included treatment group and follow-up month as predictor variables.

Supplement Table 5. Analysis of the Proportion who Missed at Least 1 Workday Per Month

	Adjusted % who Missed ≥ 1 Workday (95% CI)			Ratio of % Who Missed at Least 1 Workday Between Treatment Groups	95% Confidence Interval for Ratio
Treatment Group	EPT	14.6	(11.0, 19.4)	0.897	(0.61, 1.31)
	UC	16.3	(12.3, 21.6)		
Month reported	1	25.6	(20.4, 32.1)		
	2	17.4	(13.0, 23.4)		
	3	20.4	(15.7, 26.6)		
	4	17.1	(12.8, 22.9)		
	5	15.3	(11.2, 20.9)		
	6	14.5	(10.5, 20.0)		
	7	11.0	(7.5, 16.1)		
	8	13.3	(9.4, 18.7)		
	9	15.3	(11.1, 21)		
	10	12.8	(8.9, 18.3)		
	11	14.2	(10.1, 19.8)		
	12	13.4	(9.5, 19.0)		