Design of the Spine Pain Intervention to Enhance Care Quality And Reduce Expenditure Trial (SPINE CARE) study: Methods and lessons from a multi-site pragmatic cluster randomized controlled trial


Center for Healthcare Delivery Sciences (C4HDS), Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
Clinical Excellence Research Center, Stanford University School of Medicine, Palo Alto, CA, USA
Department of Orthopaedic Surgery, Center for Musculoskeletal Research, Vanderbilt University Medical Center, Nashville, TN, USA
Department of Physical Medicine and Rehabilitation, Osher Center for Integrative Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
Center for Health System Improvement, University of Tennessee Health Science Center, Memphis, TN, USA
HonorHealth Clinical Research Institute, Scottsdale, Arizona, USA
Department of Internal Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
Department of Physical Medicine & Rehabilitation, Vanderbilt University Medical Center, Nashville, TN, USA

ABSTRACT

Background: Low back and neck pain (together, spine pain) are among the leading causes of medical visits, lost productivity, and disability. For most people, episodes of spine pain are self-limited; nevertheless, healthcare spending for this condition is extremely high. Focusing care on individuals at high-risk of progressing from acute to chronic pain may improve efficiency. Alternatively, postural therapies, which are frequently used by patients, may prevent the overuse of high-cost interventions while delivering equivalent outcomes.

Methods: The SPINE CARE (Spine Pain Intervention to Enhance Care Quality And Reduce Expenditure) trial is a cluster-randomized multi-center pragmatic clinical trial designed to evaluate the clinical effectiveness and healthcare utilization of two interventions for primary care patients with acute and subacute spine pain. The study was conducted at 33 primary care clinics in geographically distinct regions of the United States. Individuals ≥18 years presenting to primary care with neck and/or back pain of ≤3 months’ duration were randomized at the clinic-level to 1) usual care, 2) a risk-stratified, multidisciplinary approach called the Identify, Coordinate, and Enhance (ICE) care model, or 3) Individualized Postural Therapy (IPT), a standardized postural therapy method of care. The trial’s two primary outcomes are change in function at 3 months and spine-related spending at one year. 2971 individuals were enrolled between June 2017 and March 2020. Follow-up was completed on March 31, 2021.

Discussion: The SPINE CARE trial will determine the impact on clinical outcomes and healthcare costs of two interventions for patients with spine pain presenting to primary care.

Trial registration number: NCT03083886

1. Background

Neck and back pain (together, spine pain) have a life-time prevalence of 48%–84% and are the leading global cause of functional limitation and absence from work [1,2]. Spine pain ranks as the second most common reason for seeking health care globally and, in 2016, accounted...
for $134.5 billion in health spending in the U.S., which was the highest amount for any health condition [3,4].

Practice guidelines recommend non-pharmacologic options or medications such as nonsteroidal anti-inflammatory drugs and skeletal muscle relaxants as first line treatments for patients presenting with spine pain [5–7]. Guidelines also recommend against the routine use of imaging. Despite this, low-value spine care continues to burden the health system. For example, a quarter of patients with acute low back pain undergo some form of imaging in the first six weeks of diagnosis, despite numerous randomized clinical trials demonstrating no improvement in clinical outcomes [8–12]. More than 50% of patients with back pain are prescribed opioids before more conservative options [11–13]. Similar trends are observed in the use of invasive procedures and spine surgery [14].

Interventions to reduce the use of low-value care for patients with spine pain have been tested but many have proven ineffective and the studies evaluating them have had methodological shortcomings [15]. In addition, while the majority of patients with spine pain initially present to their primary care providers, interventions in this setting have received limited attention. There have also been few direct comparisons of different approaches.

Two promising strategies may overcome the limitations of the existing evidence base. First, the Identify, Coordinate, and Enhance (ICE) care model uses risk stratification to determine the intensity of a multidisciplinary intervention that consists of physical therapy, counseling from health coaches to address biopsychosocial barriers and consultative guidance to a patient’s primary care provider from a specialist physician. The approach is motivated by the observation that most patients with spine pain have improvements in their symptoms within weeks of symptom onset [16–18]; thus, health quality and efficiency may be improved by focusing care on individuals at high-risk of progressing from acute to chronic pain [19,20]. The ICE model has been estimated to reduce spine-related spending by up to 25% [21] but has not been subject to rigorous evaluation.

Second, postural therapy attempts to address pain by re-aligning and rebalancing spinal muscles. While several specific techniques have been shown to have durable benefits for patients with chronic back pain [22–24], the Egoscue method uses individual postural alignment therapy (IPT) for patients with chronic as well as more acute spine care [25–27]. It is currently delivered in a standardized fashion at numerous centers throughout the U.S. and internationally [25–27] but has not been rigorously tested.

We designed and launched the Spine Pain Intervention to Enhance Care Quality And Reduce Expenditure (SPINE CARE) trial to test the impact of adding the ICE model or IPT to usual primary care. We hypothesized that ICE and IPT will decrease spine related healthcare spending while improving physical function and quality of life.

2. Methods

2.1. Study design

SPINE CARE is a three-arm pragmatic trial [28] funded by unrestricted philanthropic gifts to Stanford University and approved by the institutional review boards at Brigham and Women’s Hospital, Stanford University, and all institutions enrolling patients. It is registered with clinicaltrials.gov (NCT03083886). The authors are responsible for the design and conduct of this study and all meet International Committee of Medical Journal Editors (ICMJE) criteria.

2.2. Study setting

This trial was conducted at 33 U.S. primary care clinics in three geographically distinct integrated delivery networks. Sites included 12 practices affiliated with an academic medical center (Vanderbilt Medical Center; Nashville, Tennessee), 15 practices in a community-based integrated delivery network (Honor Health; Phoenix, Arizona) and 6 privately-owned primary care practices in Southern California (Laguna Hills, North Hollywood, and Oxnard) and Houston, Texas belonging to a larger network (United BioSource Corporation, LLC [UBC]). Eligible practices had to provide primary care to adult participants with acute back and neck pain and be located within 30-minute driving distance of an Egoscue IPT clinic as well as a physical therapy clinic that could be trained in the ICE protocol. Clinics that had an existing comprehensive spine care practice model were excluded.

Study enrollment began in June 2017. Virtually all participants had completed study-specific interventions (described below) prior to workplace closures resulting from the COVID-19 pandemic; 2% of patients were enrolled after March 1, 2020. Additionally, less than 5% of ICE patients were still undergoing physical therapy after March 1, 2020 and were transitioned to virtual visits. In contrast, patients in the IPT arm had the option of utilizing virtual visits throughout the study and so no changes were made in response to the pandemic. Follow-up of all trial participants ended on March 31, 2021.

2.3. Randomization

Practices were randomized in a 1:1:1 ratio to ICE, IPT or usual care such that all primary care providers and their patients within a practice were assigned to the same treatment arm. While our primary hypotheses relate to patient-level changes in health spending and physical functioning, we chose cluster randomization at the practice level as the intervention involved clinic-level workflow and to avoid contamination between patients cared for by the same provider and providers within the same clinic.

Practices were stratified and randomized within the resultant strata. Given the pragmatic nature of the trial, the specific approach differed by delivery network to reflect their unique characteristics and number of participating clinics, and therefore, the stratification variables felt to be most important to achieve balance across the treatment arms. In all cases, each stratum in each network contained 3 clinics of which one was randomized to each of the three treatment arms. In specific, the 15 Honor Health clinics were stratified into 5 groups based on their quintile of potentially eligible. The 12 Vanderbilt clinics were stratified into 4 groups based upon whether they provided walk-in care (yes/no) and whether they were above or below the median distance to the local IPT clinic. Because the UBC sites were from two distinct geographies (California and Texas), the 6 sites were divided into 2 groups based on state.

2.4. Eligible subjects and enrollment

The study population included individuals ≥18 years presenting to primary care with neck and/or back pain of ≤3 months’ duration. Complete inclusion and exclusion criteria are summarized in Table 1.

Screening occurred in 3 stages. First, electronic medical record schedules were reviewed daily to identify individuals with an upcoming primary care appointment for a spinal complaint, defined as back or neck pain, lumbago, sciatica, hip pain, or back/neck sprain and strain. Patients were excluded at this stage if their medical record indicated that they were pregnant, were actively receiving treatment for cancer, were receiving worker’s compensation, had undergone spine surgery or invasive procedure within the prior 6 months, were currently participating in another spine-related clinical trial or had received physical therapy in the previous 3 months. Second, when patients presented for their in-person visits, they were approached by research staff, informed about the study and asked to complete a screening form that confirmed the duration of pain and repeated screening for exclusion criteria. Participants with spine pain who did not undergo pre-screening (for example, those presenting for same-day urgent appointments or whose documented visit reason was for a different complaint), were also identified and screened at this stage. Third, the primary care providers of potentially eligible patients were asked to confirm that the patients did
unable to complete baseline questionnaires on the day of their in-person
electronically through a REDCap database \[[32]\]. Some patients were
tient demographic form, the EQ-5D questionnaire \[[29]\], the Lorig self-
ies. Consenting patients completed baseline surveys including a pa
study. Those patients who expressed interest were then contacted by the
staff member contact them about the potential of participating in a
Vanderbilt sites, primary care providers asked patients who were unable
were adapted to clinic workflows, staff availability, and operations of
- not have any clinical “red flags” (specifically, fever related to spine
pathology, night sweats, unintentional weight loss, bowel and bladder
dysfunction, neurologic weakness, or current intravenous drug use).

not have any clinical “red flags” (specifically, fever related to spine
pathology, night sweats, unintentional weight loss, bowel and bladder
dysfunction, neurologic weakness or current intravenous drug use) or
additional clinical indications for exclusion from the study such as
physical limitations or a major mental illness diagnosis.

Because of the pragmatic nature of the study, enrollment procedures
were adapted to clinic workflows, staff availability, and operations of
the specific clinics participating in the trial. The six UBC sites did not
perform routine pre-screening but instead conducted these activities
when patients presented for care. In some cases, after examining pa-
tients who had not been identified earlier in the process, primary care
providers at all sites referred eligible patients for study inclusion. At the
Vanderbilt sites, primary care providers asked patients who were unable
to be approached in person about their willingness to have a research
staff member contact them about the potential of participating in a
study. Those patients who expressed interest were then contacted by the
research staff over the phone; screening, consent, and baseline ques-
tionnaires were then completed within two days of the patient being
seen by their provider in clinic.

Patients who met all eligibility criteria were asked to provide
informed consent. The eighteen Vanderbilt and UBC clinics obtained
written informed consent; the fifteen Honor Health clinics initially used
written consent at the request of the IRB but then transitioned to verbal
informed consent in January 2020 based on updated institutional pol-
 gets. Consenting patients completed baseline surveys including a pa-
tient demographic form, the EQ-5D questionnaire \[[29]\], the Lorig self-
efficacy scale \[[30]\], the Oswestry Disability Index (ODI) \[[31]\], and the
STaRT Back Tool \[[19]\]. All baseline surveys were completed on paper or
electronically through a REDCap database \[[32]\]. Some patients were
unable to complete baseline questionnaires on the day of their in-person
visit and instead were given up to 2 days to complete these forms
electronically. All surveys had to be completed prior to the start of study
specific interventions for patients to remain eligible. All patients
received up to a $100 honorarium for completing patient-reported
outcomes over a one-year period.

Study participants, study staff interacting with patients, and treating
physicians were not blinded to group assignment, as knowledge of the
treatment protocols was essential to the study interventions. That said,
these individuals were not aware of the details of treatment arms to
which they had not been assigned. Study investigators and data analysts
remained blinded until all follow-up data were obtained and the primary
analytic strategies were finalized.

### 2.5. Intervention arms

#### 2.5.1. Identify, Coordinate, and Enhance (ICE) care model

The ICE care model was developed for adults with acute and sub-
acute neck or back pain not using high-dose opioid medications or
receiving spine-related long-term disability payments \[[33]\]. The eight-
week care plan was provided by a care team consisting of a physical
therapist, who served as the rehabilitation lead, a “Spine Coach”, who
served as a behavioral counselor trained in motivational interviewing
\[[34,35]\], and an “ICE MD”, a physician specialized in physiatry or pain
management who provided specialist consultation to primary care
physicians for high-risk patients who did not improve after an initial
course of therapy. All providers used the electronic medical record,
email, fax, or telephone to communicate a summary of their visits with
the other members of the care team through a standardized communi-
 cation process (see Study Protocol Appendix K).

Patients enrolled in the ICE arm of the study were risk stratified using
the STaRT Back Tool. This 9-item self-administered survey was designed
to identify modifiable biomedical, psychological, and social risk factors
in order to classify patients based on their risk of developing chronic
pain into low, medium, and high-risk prognostic subgroups \[[19]\]. For
this study, we combined medium- and high-risk groups so that bio-
 psychosocial risk factors could be addressed for as many people as
possible early in the course of care. In specific, participants with STaRT
 Back scores of \( \leq 3 \) and \( > 4 \) were classified as low and medium or high-
risk, respectively.

As described in Table 2, patients deemed to be low risk received one
Spine Coach call and one physical therapy visit. Medium and high risk
patients received three Spine Coach call and three physical therapy
visits. The Spine Coach calls were scheduled within three business
days of the patients’ index primary care visit and a physical therapy visit
within seven business days of their primary care visit. The appointments
for physical therapy visits and Spine Coach calls were made by the
research staff after completion of the baseline questionnaire, generally
prior to the patient leaving their primary care provider’s office after
enrollment.

The Spine Coach call focused on self-management and coping stra-
tegies to manage pain and adherence to exercises prescribed by the
physical therapists. As part of these visits, patients were also given
written educational materials \[[36]\]. Patients deemed to be medium or
high risk had additional emphasis on addressing physical function,
normalizing their experience, and building self-management skills. The
physical therapists were asked to follow their standard practice guide-
lines for management of patients with spine pain with a focus on
providing patients with a customized exercise program and education on
activity modification. Physical therapists were free to provide treatment
based on their evaluation and professional judgment.

At the completion of the ICE treatment course, the ICE MD conducted
an “e-consultation” for all medium or high-risk patients using a struc-
tured consultation form. For this study, the ICE MD’s role was fulfilled
by two physical medicine and rehabilitation specialists and two pain
specialists. Their consultations included a medical record review and the
 provision of specialized recommendations to the primary care provider
delivered through the electronic medical record. In those cases where
the ICE MD did not have access to a patient’s medical records, the
consultation was done by telephone with the patient’s primary care

### Table 1

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults ( \geq 18 ) years willing and able to provide informed consent</td>
<td>Severe, active psychosis, major depression, or physical disability limiting ability to participate in treatment</td>
</tr>
<tr>
<td>Back or neck pain of ( \leq 3 ) months’ duration</td>
<td>Presence of ‘red flags’ or other medical contraindications as assessed by primary care providers</td>
</tr>
</tbody>
</table>

#### Table 2

<table>
<thead>
<tr>
<th>Subject risk</th>
<th>Physical therapy</th>
<th>Spine coach call</th>
<th>ICE MD consultation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td>1 session</td>
<td>1 session</td>
<td>None*</td>
</tr>
<tr>
<td>High or medium risk</td>
<td>3 sessions</td>
<td>3 sessions</td>
<td>All patients</td>
</tr>
</tbody>
</table>

* Patients who do not improve after 6 weeks will be referred back to their primary care providers.
2.5.2. Individualized Postural Therapy (IPT)

Patients enrolled in this study arm received IPT delivered according to the Egoscue Method [25,26]. This approach involves a standardized evaluation of an individual’s posture while standing and moving to identify alignment deviations and a personalized corrective exercise program aiming to improve alignment, muscle balance, coordination, and postural control [26]. The Egoscue method was created in 1971 and is currently provided in 32 clinics worldwide [25,27].

Patients enrolled in the IPT arm of the study were scheduled for an initial IPT visit within seven days of enrollment prior to leaving their primary care provider’s office. The initial consultation was scheduled for up to 90 minutes, during which a personalized exercise plan was developed for each patient. Patients were encouraged to perform their exercises daily, which typically take 15–45 minutes. Follow-up IPT visits could take place in person or via video conference (i.e., Skype, FaceTime or Zoom) over the course of the subsequent eight weeks.

Upon completion of the IPT sessions, a summary report of patients’ progress was sent to the primary care provider. Patients in the IPT arm did not have any out-of-pocket costs for study recommended visits.

2.5.3. Usual care

Patients enrolled in this arm received usual care as directed by their PCP. Treatment modalities included medications, specialty referrals, imaging, procedures such as injections, or physical or psychological therapy, as determined by their PCPs. Patients randomized to the usual care arm of the study paid standard copayments as determined by their insurance carrier for all services recommended by their PCPs.

2.6. Outcomes

The trial’s two primary outcomes are spine-related cost of care at one year and change in pain-related disability from baseline to 3 months [31] (Table 3).

The schedule of outcome data collection is described in Table 4.

Costs will be estimated by applying unit costs incurred by patients and third-party payers to patient-reported resource utilization. Resource utilization included health care provider visits, emergency room encounters, hospital admissions, imaging studies (including MR and CT), interventions and medications [11] and were captured using checklists based on those employed by Fritz et al [37], originally developed by Goosens et al [38]. The overall instructions and each specific question in the checklist ask subjects to report resource utilization directly related to their back or neck pain.

Unit cost of these different services will be estimated using publicly available data sources. In specific, for outpatient visits to PCPs, specialists, physical therapists and other providers and for tests and procedures, we will use the Centers for Medicare & Medicaid Services Physician Fee Schedule Look-Up Tool [39]. For hospitalizations, we will use the hospital adjusted expenses per inpatient day value calculated by the Kaiser Family Foundation [40]. For emergency department visits, we will use estimates from the Medical Expenditure Panel Survey [41]. To calculate the average retail price of medications, we will use average retail prices from GoodRx.com [42]. A 12-month period was chosen for costs to adequately capture care that may have been delayed but not entirely prevented during the active intervention phases of ICE and IPT, and because health insurers and employers conventionally consider 1-year time horizons when making coverage decisions. Cost data will be converted to monthly estimates prior to performing imputation, as resource use information was collected at irregularly-spaced intervals (although patients were always asked about care received since the last survey time point).

Change in pain-related function was assessed using the Oswestry Disability Index 2.1a (ODI) [31,43–45]. The ODI contains 1 item on pain and 9 items on activities of daily living (i.e., personal care, lifting, walking, sitting, standing, sleeping, sex life, social life and travel). Each item is measured on a 0–5 scale with 5 representing the greatest disability. The scores for all sections are added together, giving a possible score of 50. The total score is then doubled and expressed as a percentage. A change in ODI of 6 points (out of the 100-point score) is considered to be clinically meaningful for patients with spine pain [37]. We will evaluate a change in ODI from baseline to 3 months to capture the short-term clinical impact of the interventions.

Secondary outcomes include change in ODI from baseline to 12 months (to mirror the time period used to assess costs), quality of life as assessed using EQ-5D-5L and a self-efficacy scale adapted from work by Lorig [30] at 12 months. EQ-5D-5L is a standardized instrument for measuring general health status. The instrument consists of a descriptive section and a visual analogue scale. The descriptive section measures 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) with each dimension having 5 levels ranging from “no problem” to “extreme problem”. The visual analogue scale is a self-rated horizontal health scale with end points labelled as “the best health you can imagine” and “the worst health you can imagine”. The raw scores are converted into index scores based on US norms [46]. Lorig et al.’s self-efficacy scale for individuals with musculoskeletal complaints has three subscales that focus on pain, functioning, and other symptoms [30]. Because spine-specific pain is already captured with ODI, we used the “functioning” and “other symptoms” subscales to measure self-efficacy.

We also asked patients about their employment status. If they stated they were currently employed, we asked about their income and how many days of work they had missed.

Finally, we will measure implementation-related outcomes to assess intervention fidelity. These will include the number of protocol-specified appointments (i.e., Spine Coach, physical therapist, or the Egoscue practitioner) scheduled and attended, the timing of visits and the length of the treatment course, which we originally expected to last 6 to 8 weeks.

Patients completed these surveys online using REDCap or over the phone. Clinical research coordinators at all sites were trained on phone outreach procedures and were provided standardized scripts for communication and collection of patient-reported outcomes. To ensure high rates of follow-up survey completion, each site used additional strategies to engage patients including using survey reminder completion postcards, text messaging, survey mailings, and increased frequency of follow-up calls.

2.7. Analytic plan

We will report means and frequencies of pre-randomization variables separately for each of the three arms. Comparisons of baseline values for each of the intervention arms to the usual care arm will be performed using generalized estimating equations with Proc GENMOD in SAS with link functions and error distributions appropriate to the outcomes, adjusting for the cluster and stratified randomized design.

The trial outcomes will be evaluated using intention-to-treat
principles among all randomized patients based on the study arm to which a subject was assigned at study enrollment. Our primary analyses will compare each of the two treatment arms separately with usual care (i.e., ICE compared with usual care and IPT with usual care). Because we are evaluating two primary outcomes, we will use a Bonferroni-corrected two-tailed type I error of 0.025 [47,48]. We will not adjust for multiple testing between treatment arms for several reasons. First, while the chance of finding at least one false positive among several tests is greater than 5%, in this case a Bonferroni correction will be too conservative because the multiple comparisons among the treatment arms share the same exposure groups [49,50]. Secondly, a systematic review of multi-arm trials found that more than half of all randomized trials with multiple unrelated exposure groups do not adjust for multiple comparisons, on the basis that if each exposure was compared with control in a separate trial, no adjustment would be necessary [51].

Our primary analyses will be performed after imputing missing data with multiple imputation. We will perform 20 imputations using Proc MI in SAS with a fully conditional specification. This procedure assumes that data are missing at random and generates imputed values sequentially by specifying a regression model for each variable given the other variables, rather than drawing the imputations from a pre-specified joint distribution [52]. This approach achieves in-range values and has high relative efficiency [53]. Each imputed dataset will be analyzed separately, and the results will be pooled using Proc MIANALYZE. The overall estimate’s variance is a function of the variance within and between each of the 20 imputed datasets. The following variables will be used to perform the imputation: age, study arm, study site, clinic, sex, race/ethnicity, BMI, exercise frequency at baseline, education level, employment status, smoking status, other medical conditions at baseline, number of medications used for spine pain at baseline, length of pain at baseline, number of previous pain episodes, STarT back score, baseline ODI, baseline self-efficacy, baseline EQ-5D, and scores for patient-reported outcomes at every follow-up timepoints (ODI, cost, Lorig self-efficacy scale, and EQ-5D).

For spine-related cost of care, costs will be summed after imputation to generate estimates of 12-month spine-related spending. We will use generalized estimating equations (Proc GENMOD in SAS with a log link and Poisson-distributed errors) to evaluate outcomes for each imputed dataset and Rubin’s rules (Proc MIANALYZE in SAS) to combine the parameter estimates from the individual analyses [54]. If many subjects have no spending in follow-up (which we do not expect, as the eligible patients are those presenting to primary care with spine pain) or if we observe overdispersion, we will use a zero-inflated negative binomial model executed in Proc GENMOD [55]. Because cost data in healthcare is not normally distributed with zero values for participants with no resource utilization during follow-up, we will winsorize estimated costs at the 95th percentile in secondary analyses [56]. We will report relative spending differences and absolute treatment differences, as appropriate.

Analyses of the change in ODI from baseline to three months, as well as secondary outcomes, will be performed identically to those for spending, except using an identity link function and normally distributed errors. For these analyses we report absolute differences with confidence intervals and two-tailed p values. All models will adjust for age and sex, account for the cluster and block randomized study design and include fixed effects for the 3 different delivery networks that were used for recruitment. Secondary analyses will additionally adjust for covariates that differ statistically between groups at baseline despite randomization.

As a sensitivity analysis, we will conduct a complete-case analysis on all non-missing values. The complete-case spine-related cost of care and change in ODI models will be performed with Proc GENMOD in SAS with a log link function and Poisson-distributed errors for spine-related health spending and an identity link function and normally distributed errors for other outcomes. As resource utilization and ODI were collected at multiple time points, we will also evaluate non-missing cost and ODI data from each time point using a repeated measures design. In specific, we will use generalized estimating equations, with link functions and error distributions appropriate to the study outcomes, to assess changes in outcomes over the 12 months of follow-up, while accounting for correlations in the repeated measurements. For spine-related health spending, these analyses will be conducted using the time intervals at which raw data were collected as well as after converting estimates to evenly-spaced months. We will also conduct exploratory analyses comparing the ICE and IPT treatment arms against each other, though this analysis likely to be underpowered. As with our primary analyses, all of our sensitivity analyses will account for the cluster and block randomized design, include fixed effects for delivery network and adjust for age and sex. Similarly, we will report relative spending differences and absolute treatment differences, as appropriate.

Finally, we will evaluate the effect of treatment within pre-specified subgroups including age, sex, STarT Back risk group, location of pain, pain duration, and whether or not this was the subject’s first pain episode.

2.8. Sample size considerations

We powered our study to detect clinically and economically-meaningful differences between each of the treatment arms and usual care. For ODI, assuming a standard deviation of 16-point based on published literature [57], an ICC of 0.01, 10% loss to follow-up, we estimated that we would need to enroll 699 patients (233 per arm) in order to achieve a power of 80% to detect a clinically meaningful difference of 6 points (out of the 100-point score) [37]. For spine-related spending, we estimated that ICE would result in a 20% relative reduction in spending compared to usual care based on the existing literature [58–60]. Because of the lack of peer-reviewed literature evaluating the economic impact of IPT, we assumed the same effect size for this arm. We further assumed spine-related spending in the usual care of $894 based upon prior work by Fritz et al. [61], a standard deviation 1.25 times the mean, an intra-cluster correlation (ICC) of 0.01, a 10% loss to follow-up. On this basis, we estimated that we would need to enroll 3096 patients (1032 per arm) in order to achieve a power of 80% to detect our assumed differences in spending between either treatment arm and usual care. Given the substantially larger sample size required for our economic analyses, we powered our trial on the basis of this outcome, which would also give us more than 98% power to detect a 6-point change in ODI between each intervention arm and the usual care arm.

Because of slower than anticipated enrollment, funding limitations in the context of the COVID-19 pandemic required us to stop enrollment by March 31, 2020 (after 2 years and 10 months of enrollment) by
time we had recruited 2919 patients (i.e. 94% of planned enrollment). Based upon our initial assumptions (rather than any interim analyses of trial outcome data), we estimated that this sample size would provide us 78% power to detect our assumed differences in health spending. Conversely, given the inherent uncertainty about the assumptions we used for our initial power calculations, we alternatively estimated that 80% power could be achieved with our reduced sample size despite the observed 15% loss to follow-up if the standard deviation of health spending was equal to the mean (instead of 1.25 times the mean) and the ICC was 0.033 (instead of 0.01). The revised sample size would still give us more than 98% power to detect a 6-point change in ODI between each intervention arm and the usual care arm.

2.9. Baseline characteristics

A total of 25249 patients were screened for eligibility (Fig. 1). The final sample consisted of 2971 patients cared for by 328 physicians. The sample was mostly female (60.3%), white (71.5%), had a college education or above (56.5%), and was employed part- or full-time (66.9%) (Table 5). In addition, 1613 patients (54.5%) reported at least one other health condition and 2280 patients (76.9%) reported exercising 1 or more hours per week. The mean BMI of the sample was 30.2.

The majority of participants (62.1%) had pain lasting less than 1 month and 68.5% had previously experienced spine pain. They were prescribed an average of 2.5 prescription medications at the time of enrollment; 1347 patients (58.2%) reported NSAID use while 129 (5.6%) reported being prescribed an opioid.

3. Discussion

The prevalence of spine pain has increased dramatically in the US and worldwide [62,63] with resultant increases in health spending and spine-related disability [64]. As a result, it is imperative to develop models of care for patients with this condition that are evidence-based, practical, and can be used early in the course of disease.
Table 5
Baseline characteristics by study arm.

<table>
<thead>
<tr>
<th></th>
<th>ICE (N = 829)</th>
<th>IPT (N = 1150)</th>
<th>Usual Care (N = 992)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>50.9</td>
<td>51.2 (16.0)</td>
<td>51.2 (16.0)</td>
</tr>
<tr>
<td>Male Sex, N(%)</td>
<td>514 (38.0%)</td>
<td>462 (40.2%)</td>
<td>402 (40.6%)</td>
</tr>
<tr>
<td>Race/ethnicity*, N(%)</td>
<td>603 (73.0%)</td>
<td>651 (75.5%)</td>
<td>120 (65.8%)</td>
</tr>
<tr>
<td>White</td>
<td>80 (9.7%)</td>
<td>207 (20.9%)</td>
<td>29 (4.5%)</td>
</tr>
<tr>
<td>Black or African-American</td>
<td>7 (0.9%)</td>
<td>11 (1.1%)</td>
<td>61 (7.9%)</td>
</tr>
<tr>
<td>Asian</td>
<td>19 (2.3%)</td>
<td>22 (2.2%)</td>
<td>39 (4.3%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>15 (17.5%)</td>
<td>108 (10.9%)</td>
<td>61 (9.5%)</td>
</tr>
<tr>
<td>Other</td>
<td>60 (7.3%)</td>
<td>42 (4.2%)</td>
<td>39 (4.3%)</td>
</tr>
<tr>
<td>Employed, N(%)</td>
<td>580 (70.2%)</td>
<td>771 (82.6%)</td>
<td>632 (63.9%)</td>
</tr>
<tr>
<td>Education, N(%)</td>
<td>28 (3.4%)</td>
<td>45 (4.6%)</td>
<td>45 (4.6%)</td>
</tr>
<tr>
<td>Did not finish high school</td>
<td>302 (36.6%)</td>
<td>468 (47.3%)</td>
<td>477 (48.2%)</td>
</tr>
<tr>
<td>Completed high school</td>
<td>142 (16.4%)</td>
<td>142 (14.3%)</td>
<td>142 (14.3%)</td>
</tr>
<tr>
<td>College degree or higher</td>
<td>60 (60.0%)</td>
<td>477 (48.2%)</td>
<td>477 (48.2%)</td>
</tr>
<tr>
<td>Chief Complaint, N(%)</td>
<td>100 (12.1%)</td>
<td>139 (14.0%)</td>
<td>139 (14.0%)</td>
</tr>
<tr>
<td>Neck Pain only</td>
<td>685 (82.6%)</td>
<td>840 (84.7%)</td>
<td>840 (84.7%)</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
<td>29.9 (8.7)</td>
<td>30.0 (8.1)</td>
<td>30.8 (12.2)</td>
</tr>
<tr>
<td>Exercise, N(%)</td>
<td>175 (21.2%)</td>
<td>274 (27.7%)</td>
<td>274 (27.7%)</td>
</tr>
<tr>
<td>Less than 1 h/week</td>
<td>178 (21.6%)</td>
<td>202 (20.4%)</td>
<td>202 (20.4%)</td>
</tr>
<tr>
<td>1–2 h/week</td>
<td>135 (16.4%)</td>
<td>142 (14.3%)</td>
<td>142 (14.3%)</td>
</tr>
<tr>
<td>2.5–5 h/week</td>
<td>337 (40.9%)</td>
<td>372 (37.6%)</td>
<td>372 (37.6%)</td>
</tr>
<tr>
<td>Medical History*, N(%)</td>
<td>175 (21.2%)</td>
<td>193 (19.5%)</td>
<td>193 (19.5%)</td>
</tr>
<tr>
<td>Heart attack</td>
<td>14 (1.7%)</td>
<td>23 (2.3%)</td>
<td>23 (2.3%)</td>
</tr>
<tr>
<td>Stroke</td>
<td>18 (2.2%)</td>
<td>15 (1.5%)</td>
<td>15 (1.5%)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>78 (9.5%)</td>
<td>119 (12.0%)</td>
<td>119 (12.0%)</td>
</tr>
<tr>
<td>High cholesterol</td>
<td>183 (22.2%)</td>
<td>260 (26.3%)</td>
<td>260 (26.3%)</td>
</tr>
<tr>
<td>Depression</td>
<td>142 (17.3%)</td>
<td>165 (16.7%)</td>
<td>165 (16.7%)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>191 (23.2%)</td>
<td>193 (19.5%)</td>
<td>193 (19.5%)</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>98 (11.9%)</td>
<td>101 (10.2%)</td>
<td>101 (10.2%)</td>
</tr>
<tr>
<td>Rheumatoid Arthritis</td>
<td>14 (1.7%)</td>
<td>31 (3.1%)</td>
<td>31 (3.1%)</td>
</tr>
<tr>
<td>None of the above</td>
<td>384 (46.7%)</td>
<td>455 (46.0%)</td>
<td>455 (46.0%)</td>
</tr>
<tr>
<td>Current Smoker, N(%)</td>
<td>61 (7.4%)</td>
<td>120 (12.2%)</td>
<td>120 (12.2%)</td>
</tr>
<tr>
<td>Number of prescription medications, mean (SD)</td>
<td>2.6 (2.6)</td>
<td>2.6 (2.6)</td>
<td>2.4 (2.4)</td>
</tr>
<tr>
<td>Medications taken for spine pain*, N(%)</td>
<td>366</td>
<td>489</td>
<td>489</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>168 (57.2%)</td>
<td>492 (63.5%)</td>
<td>492 (63.5%)</td>
</tr>
<tr>
<td>Muscle relaxers</td>
<td>168 (26.3%)</td>
<td>185 (23.9%)</td>
<td>185 (23.9%)</td>
</tr>
<tr>
<td>Opioids</td>
<td>29 (4.5%)</td>
<td>61 (7.9%)</td>
<td>61 (7.9%)</td>
</tr>
<tr>
<td>Steroids</td>
<td>61 (9.5%)</td>
<td>57 (7.4%)</td>
<td>57 (7.4%)</td>
</tr>
<tr>
<td>Other</td>
<td>58 (9.1%)</td>
<td>84 (10.8%)</td>
<td>84 (10.8%)</td>
</tr>
<tr>
<td>When current episode began, N(%)</td>
<td>541</td>
<td>647</td>
<td>647</td>
</tr>
<tr>
<td>1 month ago or less</td>
<td>114 (14.1%)</td>
<td>630 (64.4%)</td>
<td>630 (64.4%)</td>
</tr>
<tr>
<td>1–2 months ago</td>
<td>66 (67.7%)</td>
<td>212 (21.2%)</td>
<td>212 (21.2%)</td>
</tr>
</tbody>
</table>

Table 5 (continued)

<table>
<thead>
<tr>
<th></th>
<th>ICE (N = 829)</th>
<th>IPT (N = 1150)</th>
<th>Usual Care (N = 992)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of previous episodes, N(%)</td>
<td>257 (31.7%)</td>
<td>325 (33.2%)</td>
<td>325 (33.2%)</td>
</tr>
<tr>
<td>None</td>
<td>188 (22.2%)</td>
<td>226 (23.1%)</td>
<td>226 (23.1%)</td>
</tr>
<tr>
<td>1–2</td>
<td>115 (14.2%)</td>
<td>127 (13.0%)</td>
<td>127 (13.0%)</td>
</tr>
<tr>
<td>3–4</td>
<td>61 (7.9%)</td>
<td>70 (7.2%)</td>
<td>70 (7.2%)</td>
</tr>
<tr>
<td>5–6</td>
<td>196 (24.2%)</td>
<td>230 (23.5%)</td>
<td>230 (23.5%)</td>
</tr>
<tr>
<td>More than 6 episodes</td>
<td>5.0 (2.2%)</td>
<td>4.6 (2.2%)</td>
<td>4.5 (2.5%)</td>
</tr>
<tr>
<td>Baseline STarTBack Score, mean (SD)</td>
<td>31.2</td>
<td>28.9 (17.4)</td>
<td>28.9 (17.4)</td>
</tr>
<tr>
<td>Baseline ODI Score, mean (SD)</td>
<td>0.69</td>
<td>0.70 (0.14)</td>
<td>0.70 (0.14)</td>
</tr>
<tr>
<td>Baseline EQ-SD Score, mean (SD)</td>
<td>69.7</td>
<td>69.5 (19.2)</td>
<td>69.5 (19.2)</td>
</tr>
<tr>
<td>Baseline Self-efficacy Score (function subscale), mean (SD)</td>
<td>82.9</td>
<td>82.6 (18.5)</td>
<td>82.6 (18.5)</td>
</tr>
<tr>
<td>Baseline Self-efficacy Score (other symptoms subscale), mean (SD)</td>
<td>72.7</td>
<td>74.2 (21.5)</td>
<td>74.2 (21.5)</td>
</tr>
</tbody>
</table>

Abbreviations: ICE: Identify, Coordinate, and Enhance, IPT: Individualized Postural Therapy, BMI: Body Mass Index, ODI: Oswestry Disability Index. * Participants were allowed to select multiple categories. As such, the total percentages for each category do not add up to 100%.

Third, while an equivalent number of screened patients were eligible in all three treatment arms, a lower proportion of individuals allocated ICE and IPT agreed to participate. We hypothesize that this was the result of the additional protocol-specified care required for subjects in these arms (for example, travel and time to receive treatment even though travel expenses were reimbursed).

Fourth, although we designed our study to detect economically and clinically relevant benefits from the interventions under study and included a larger number of clustering units than most published cluster-randomized controlled trials [73], an even larger number of smaller clinics would have provided additional power and further enhanced generalizability. For example, based upon our original trial assumptions, doubling the number of clusters would have reduced our overall sample size to 2040 and the average number of patients per cluster to 31. That said, even large delivery networks like those included in our trial, have a finite number of practice sites. Thus, doubling the number of clusters would have required also doubling the number of delivery networks. Appropriate research-capable environments are challenging to identify, and in the case of our study, would likely have added at least another year to site identification. Further, each additional site has substantial fixed costs to initiate and maintain study procedures which, in some cases, may have more than offset the savings from reduced patient volumes.

Finally, should the ICE intervention prove effective, new provider roles and new processes for coordination of care will need to be developed. Similarly, should the study show benefit from the IPT model, integration of these models into routine care for spine pain will require
Table 6
Site-specific recommendations for improved recruitment and retention.

<table>
<thead>
<tr>
<th></th>
<th>HonorHealth</th>
<th>Vanderbilt</th>
<th>UBC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timeline</strong></td>
<td>Recruitment</td>
<td>Strategy</td>
<td>Strategy</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>March 2019</td>
<td>Enroll patients on weekends from selected clinics that have higher volumes</td>
<td>September 2018</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>April 2019</td>
<td>Implemented verbal consent in addition to written consent</td>
<td>April 2019</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>February 2019</td>
<td>Provided feedback and used social norming strategies through bimonthly clinic-specific dashboards highlighting the highest enrolling clinics and providers</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>April 2019</td>
<td>Incentive program for clinics ($10 giftcards) to any clinic provider who goes out of their way to help enroll patients in the trial.</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>February 2019</td>
<td>Cover the cost of transportation for patients enrolled in the ICE and IPT arms.</td>
<td>April 2019</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>June 2017</td>
<td>Identified PCP champions throughout the health system</td>
<td>April 2019</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>June 2017</td>
<td>Send automated reminder emails, make phone calls to non-responders, and send reminder postcards</td>
<td>August 2018</td>
</tr>
<tr>
<td></td>
<td>September 2019</td>
<td>Send reminder text messages</td>
<td></td>
</tr>
</tbody>
</table>

In conclusion, this cluster randomized trial will compare usual care with two novel models of care for the management of spine pain in the primary care setting and determine their impact on healthcare utilization and pain-related disability.

Declaration of Competing Interest

Kristin R. Archer is a current member of the NeuroSpinal Innovation, Inc. advisory board and a past consultant for Pacira and NeuroPoint Alliance. Susan Butterworth owns a consulting firm that provides motivational interviewing training.

All other authors (Niteesh Choudhry, Constance Fontanet, Roya Ghazinouri, Sheila Fifer, Nancy Haff, Harvinder Deogun, Shannon Block, Angelina Cooper, Ellen Sears, Parul Goyal, Rogelio Coronado, Byron Schneider, Eugene Hsu, and Arnold Milstein) declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Patients who participated, study staff and contributing team members especially: Jennifer Scott, Sonam Mehta, Claude Pinnock, Keith Marton, Jessica Lee, Kyle Moraswi, Mufaddal Mahersi, Mehdi Najafzadeh, Roberta Aham, Claudia Davidson, Pete Miles, Layle Gurrieri, Traci Stewart-Solomon, Claude Edward Shackelford, Sabrina Finney, Matthew Abbate, Lynn Holliday, Karin Moolman, Shanita Coleman-Dockery, Ilaben Patel, Frederica Angel, Jennifer Green, Kevin Mitchell, Sandra Ermini, and James McDonald.

Funding

This work was supported by unrestricted philanthropic gifts to Stanford University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cct.2021.106602.

References


