HHS Public Access

Author manuscript

J Orthop Sports Phys Ther. Author manuscript; available in PMC 2018 December 01.

Published in final edited form as:

J Orthop Sports Phys Ther. 2018 June ; 48(6): 460–475. doi:10.2519/jospt.2018.7811.

Optimal Screening for Prediction of Referral and Outcome (OSPRO) for Musculoskeletal Pain Conditions: Results from the Validation Cohort

Steven Z. George, PT, PhD, FAPTA[1], **Jason M. Beneciuk, PT, PhD, MPH**[2], **Trevor A. Lentz, PT, MPH**[3], **Samuel S. Wu, PhD**[4], **Yunfeng Dai, MS**[4], **Joel E. Bialosky, PT, PhD**[2], and **Giorgio Zeppieri Jr, PT, MPT, SCS**[5]

[1]Duke Clinical Research Institute and Department of Orthopaedic Surgery, Duke University, Durham NC

[2]Department of Physical Therapy, Brooks - PHHP Research Collaboration. University of Florida, Gainesville, Florida

[3]Rehabilitation Science, Department of Physical Therapy, University of Florida, Gainesville, FL

[4]Department of Biostatistics, University of Florida, Gainesville, FL

[5]University of Florida Orthopaedic and Sports Medicine Institute, University of Florida, Gainesville, FL

Abstract

Study Design—Observational, prospective cohort

Background—Musculoskeletal pain is a common reason to seek healthcare and earlier non-pharmacological treatment and enhancement of personalized care options are two high priority areas. Validating concise assessment tools is an important step in establishing better care pathways.

Objectives—To determine the predictive validity of Optimal Screening for Prediction of Referral and Outcome (OSPRO) tools for individuals with neck, low back, shoulder, or knee pain.

Corresponding Author. Steven Z. George PT, PhD, 2400 Pratt Street Durham, NC 27715, steven.george@duke.edu.

Author Contributions

SZG secured funding, provided overall design, gave input on the analysis plan, and approved final version of the manuscript. JMB provided input on design and analysis plan, and approved final version of the manuscript. TAL provided input on design, assisted with coordinating collection of outcome measures and approved final version of the manuscript. SSW and YD led the analysis plan and approved final version of the manuscript. All other authors (JEB, GZ) assisted with collection of outcome measures and approved final versions of the manuscript.

Institutional Approval

This study was approved by the University of Florida Gainesville Health Science Center Institutional Review Board (IRB-01). All patients provided informed consent before participating in this study.

Trial Registry

N/A

Financial Disclosure and Conflict of Interest

The authors have no financial, conflicting, or competing interests to report.

Methods—A convenience sample (n = 440) was gathered by Orthopaedic Physical Therapy-Investigator Network clinics (n = 9). Participants completed questionnaires for demographic, clinical, comorbidity, and the OSPRO tools and were followed for 12-month outcomes in pain intensity, region-specific disability, quality of life, and comorbidity change. Analyses predicted these 12-month outcomes with models that included the OSPRO review of systems and yellow flag tools and planned covariates (accounting for comorbidities and established demographic and clinical factors).

Results—The 10 item OSPRO yellow flag tool (baseline and 4 week change score) consistently added to predictive models for 12-month pain intensity, region-specific disability, and quality of life. The 10 item OSPRO review of system tool added to a predictive model for quality of life (mental summary score) and 13 additional items of the OSPRO review of system+ tool added to prediction of 12-month comorbidity change. Other consistent predictors included age, race, income, previous episode of pain in same region, comorbidity number, and baseline measure for the outcome of interest.

Conclusion—The OSPRO review of system and yellow flag tools statistically improved prediction of multiple 12 month outcomes. The additional variance explained was small and future research is necessary to determine if these tools can be used as measurement adjuncts to improve management of musculoskeletal pain.

Introduction

Musculoskeletal pain is a common reason to seek healthcare and a national initiative has provided guidance on priorities for improving management of this costly and disabling condition.[33] Two elements stressed in progressive pain management strategies are earlier non-pharmacological treatment[13, 44] and enhancement of personalized/tailored care options.[33] One way physical therapists can meet the demands of these initiatives is to develop concise assessment tools that aid clinical decision making for these elements.[27] In musculoskeletal pain management two important components of almost every patient encounter are identification of symptoms that may indicate co-existing systemic pathology[15] and consideration of pain-associated distress and coping styles.[49] These components are important to consider because their results could alter a care episode by indicating the need for additional diagnostic testing before starting traditional non-pharmacological treatment[14, 28] alone or supplemented with principles of psychologically informed practice.[5, 31, 42]

The Orthopaedic Physical Therapy-Investigator Network (OPT-IN) was formed to develop and validate concise assessment tools for individuals with a primary complaint of neck, shoulder, low back, or knee pain. OPT-IN provided the clinical infrastructure necessary to recruit for the Optimal Screening for Prediction of Referral and Outcome (OSPRO) development and validation cohorts. The OSPRO cohort studies occurred in sequence, with the development cohort (a cross-sectional study for tool development) and the separately recruited validation cohort (a longitudinal study to test the predictive validity of the newly developed tools). The instruments were directly aligned with assessment of examination components that could influence a care episode. A Review of System (OSPRO-ROS) tool was developed for assessing symptoms of systemic pathology[19] and a Yellow Flag (OSPRO-

YF) tool was developed for assessing psychosocial aspects of pain vulnerability and resilience.[39] Details on the OSPRO-ROS and YF tools have been previously reported,[19, 39] and they will be described in more detail in the Methods.

Development of the OSPRO-ROS and YF tools was encouraging, but all prior work was done in a cross-sectional manner.[19, 39] Longitudinal studies provide a more optimal design to test the capabilities of these tools and determine predictive validity for outcomes of relevance in clinical decision making. Therefore, the purpose of the current paper is to report the primary analyses for the OSPRO validation cohort of individuals with primary complaint of neck, low back, shoulder, or knee pain. These analyses involved prediction of 12-month pain, quality of life, region-specific disability, and comorbidity outcomes. Conceptually our predictive models were built to determine the OSPRO tools contribution to 12 month outcomes after demographic, clinical, and baseline variables were already considered. In addition, we considered interaction between anatomical region and the OSPRO tools to determine if tool performance varied based on primary site of pain. This approach provided a relatively high bar to determine the predictive validity of the new tools because models included previously established predictive factors and anatomical region as planned covariates. Based on prior studies showing that change in psychological factors may improve outcome prediction for low back pain[4, 26, 55, 57] we also entered 4 week change in the OSPRO-YF tool into the last step of the prediction models. Our over-riding hypotheses were that the OSPRO-YF tool would improve prediction of pain and disability outcomes, while the OSPRO-ROS tool would improve prediction of quality of life and comorbidity outcomes.

Methods

Overview

The OSPRO validation cohort study was approved by the University of Florida Human Subjects Institutional Review Board and all participants provided consent to participate in the study. A convenience sample was gathered from December 2014 and December 2015 by participating OPT-IN clinics (n = 9). The OPT-IN clinics that participated in data collection represented 5 of 8 geographic regions for the United States including the Mideast, Southeast, Great Lakes, Rocky Mountain States, and Far West. The majority of the patients (275/440, 62.5%) were recruited from clinics in the Southeast region. The New England, Plains, and Southwest regions were not represented. An attempt was made to balance between urban and rural settings over the entire OPT-IN network, though for pragmatic reasons that balance was not provided within each geographic region. Methodological details for the OSPRO validation cohort have been previously reported in a cohort profile paper.[21] In the current paper we present an abbreviated version of the methods that allows for interpretation of primary analyses.

Participants

Physical therapists determined participant eligibility at initial evaluation with matching criteria from the development cohort[19, 39]:

Inclusion criteria—Patients between the ages of 18 and 65 years of age were eligible to participate in this study if they: 1) were seeking outpatient physical therapy treatment for musculoskeletal pain, 2) had primary complaints involving the cervical spine, lumbar spine, shoulder or knee, and 3) were able to read and comprehend English language (this criterion was necessary due to the large number of self-report forms).

Exclusion criteria—Patients were excluded from study participation for any diagnosis indicative of 1) widespread chronic pain syndrome (e.g. fibromyalgia or irritable bowel syndrome), 2) neuropathic pain syndrome (e.g. complex regional pain syndrome or diabetic neuropathy), 3) psychiatric history (currently in care of mental health care provider or taking ≥ 2 prescription psychiatric medications), 4) cancer (currently receiving treatment for active cancer), 5) neurological disorder (e.g. stroke, spinal cord injury, or traumatic brain injury).

Baseline and follow-up data collection occurred online at the clinic or at home (based on individual preference), with participants completing all survey assessments on the study website. Eligible participants were directed to a secure, University of Florida hosted website for the informed consent process and baseline assessment. All assessments were self-report and completed electronically by the participant in a de-identified manner.

Follow up time points were at 4 weeks, 6 months, and 12 months and participants were notified of a pending assessment by an email that directed them back to the study website to complete their follow-up assessment. If participants did not complete their follow-up assessment within 1 week of the first email notification, an additional email reminder was sent each week for up to 3 weeks. Participants who were not responsive to any of these email reminders were contacted by telephone. Only 12-month data were reported in this paper, and there are no plans to report the 6-month data separately.

Predictive Measures

Demographic and clinical information—Participants completed a standard intake form previously used in our clinical studies;[3, 23] this form captures information including: age, sex, race, income, employment, education, insurance, geographic region, pain location, pain duration, pain onset type, previous episode in same location, and history of surgery. Historical data included anatomical location of the pain, onset of symptoms, duration of symptoms, previous episodes in same anatomical region, and previous treatments.

Comorbidities—Health history was determined with the Charlson and Functional Comorbidity Indices.[9, 24] For analysis purposes a comorbidity count was derived by adding unique number of comorbidities reported (i.e. similar comorbidities reported in both indices were only counted once). The number of comorbidities reported at baseline was used as a covariate.

OSPRO Tools

Review of systems—The OSPRO-ROS tool includes standard symptom descriptors previously used to aid with screening for potential systemic involvement.[19] It includes questions related to symptoms of the cardiovascular, gastrointestinal, endocrine, nervous,

integumentary, pulmonary, and musculoskeletal systems that were identified based on their ability to predict any one positive response to a larger item bank. The 10- and 23-item versions of the OSPRO-ROS tool had differing accuracy in predicting positive response to the larger item bank.[19] Therefore, these versions were considered separately in predictive analyses. The OSPRO-ROS tool was scored by summing the positive responses, providing a potential range of 0–23 if all 23 items are used. Higher OSPRO-ROS indicate higher levels of red flag symptom complaints. In this analysis we separated the 23 items to determine if they uniquely contributed to outcomes of interest. Therefore, OSPRO-ROS refers to the first 10 items of the tool and OSPRO-ROS+ refers to the additional 13 items.

Yellow flags—The OSPRO-YF tool includes items from pain vulnerability domains (negative affect and fear-avoidance) and pain resilience domains (positive affect and self-efficacy) to aid with efficient identification of pain associated psychological distress and coping.[39] The OSPRO-YF tool estimates scores for full-length parent questionnaires with increased accuracy based on 10- and 17-item versions of the tool. The OSPRO-YF tool was considered in predictive analyses by testing the 10-item version and additional 7 items separately.[39] The OSPRO-YF tool was scored by summing all item responses from the original parent questionnaires on the original scale, with pain resilience items reverse scored, providing a potential range of 6–89 if all 17 items are used. Higher OSPRO-YF scores indicate higher psychological distress as evidence by higher pain vulnerability and lower pain resilience.

Outcome Measures

Outcome measures were captured at baseline and at 12 month follow up. The baseline value of a given measure was included in the corresponding prediction model for 12 month outcomes. Pain intensity was assessed with the 0–10 numeric rating scale (NRS) with 0= no pain at all and 10= worst pain imaginable and participants rated their current pain intensity, as well as their best (lowest) and worst (highest) pain intensity over the past 24 hours.[6, 10, 35] The average of these three ratings were used to represent pain intensity in these analyses.

Region specific disability was assessed by participants completing one of the following questionnaires that matched the primary site of pain complaint; 1) the Neck Disability Index (NDI),[53] Oswestry Disability Questionnaire (ODQ),[17] Quick Disability of Arm Shoulder and Hand (DASH),[2] or International Knee Documentation Committee (IKDC) Subjective Knee Form.[34] The individual region specific measures were included in the analysis as z-scores because of different scaling and so that they could all be included in the same predictive models, consistent with how this was done in analyses from the OSPRO development cohort.[7]

The Medical Outcomes Study Short-Form Health Survey (SF-8) was collected as a general quality of life measure and reported as the corresponding Physical Component Summary (PCS) and Mental Component Summary (MCS) scores.[18, 38] Comorbidity at 12 months was included as an outcome measure to determine change in disease burden and it was assessed in the same manner as was described in the Predictive Measures section.

Data Analysis

Our primary analyses assessed the accuracy of predicting 12-month outcomes by the OSPRO-ROS and OSPRO-YF tools. We fit separate general linear models for the continuous outcomes measures of 12-month pain intensity, region specific disability, quality of life and comorbidity, using the OSPRO tools as planned fixed effects. In each prediction model we had planned covariates to enter into the model before OSPRO tools were considered, and these included (as examples - full set reported in Table 1) age, sex, race, income, employment, education, type of insurance, geographic region, pain location, pain duration, pain onset type, previous episode in same location, history of surgery, comorbidities, and the corresponding outcome measure at baseline. This modeling approach is consistent with other reports in the literature[3, 16, 25] and resulted in the following model structure consistently applied for each outcome of interest:

- Block 1) Demographic, clinical, and comorbidity;
- Block 2) Baseline dependent variable;
- Block 3) OSPRO tools, short version of tools (10 item versions for the ROS and YF);
- Block 4) OSPRO tools, longer version (13 additional items for the ROS and 7 additional items for the YF); and
- Block 5) OSPRO-YF 4-week change score.

After block 5, interaction terms for OSPRO-YF and ROS tools by anatomical region were included in the models to investigate specificity of use based on primary site of pain complaint. Predictive analyses were first conducted with completed cases in full (all covariates) and parsimonious (backward selection) models. Then, missing 12-month outcomes were accounted for using regression imputed methods.[47] In addition, considering that the data may be not missing at random, we also performed regressions inversely weighted by inclusion (non-missing) probability, which was estimated based on logistic regressions with logit link and the following predictors: age, education, type of insurance, pain onset type and baseline dependent variable.[8] Therefore, this paper reported results from the following models 1) completed cases; 2) regression imputed; and 3) inverse probability weighted. Presenting the results in this manner remained true to the original analysis plan, while also presenting models that appropriately accounted for loss of follow-up.

Power Analysis

There are no uniform standards for determining sample size in cohort studies. For the OSPRO studies sample size estimates were based on precision for the assessment tools. The sample size was calculated so that 95% confidence intervals for the accuracy of predicting 23-item versions of the OSPRO-ROS tool from the abbreviated 10-item version would have a width of at most ±5%. Specifically, we required that sample size N satisfies sqrt(p*(1−p)/N)*1.96<0.05, where p is the prediction accuracy. This calculation yielded 385 patients with neck, shoulder, low back, or knee pain. A liberal estimate of 20% loss to follow up at 1 year results in a required total sample size of 462, or approximately 115 patients for each anatomical region.

Results

Recruitment and Follow Up Summary

A descriptive summary of the OSPRO validation cohort is reported in Table 1 and additional data is available in the cohort profile paper.[21] A total of 440 participants completed baseline measures with primary complaints of neck (n = 98, 22.3%), shoulder (n = 107, 24.3%), low back (n = 118, 26.8%), or knee (n = 117, 26.6%) pain. A total of 279 (63.4%) participants completed the 12 month follow up with primary complaints of neck (n = 59, 21.1%), shoulder (n = 66, 23.7%), low back (n = 72, 25.8%), or knee (n = 82, 29.4%). While there were no differences in follow up rates by anatomical region, there were several differences between those that completed 12 month follow up and those that didn't complete the follow up (Table 1).

Those that completed the follow up were more likely to be younger, have higher income and completed higher levels of education. In addition, there were differences in insurance type, clinic site, and onset of symptoms based on 12 month follow up. Finally, those that completed follow up had lower scores on OSPRO-YF, neck disability, pain intensity, and composite z-score for region specific disability. Therefore, those that did not complete follow up for this study were more likely to have lower income and education levels, be uninsured or on disability, covered by Medicaid or worker compensation, and be experiencing higher pain and pain associate distress. All variables reported in Table 1 had already been planned as covariates, so no additional covariates were added to the prediction models based on differences in follow up rates.

Overall Model Performance

Overall performance for completed cases, regression imputed, and inverse probability weighted models is summarized in Table 2 and individual predictors across these models are summarized in Table 3. There was a consistent pattern for prediction of 12-month pain, disability, and quality of life. The baseline value of the outcome of interest explained the most additional variance after accounting for demographic and clinical variables. Then, the 10 item version of the OSPRO-YF tool explained variance beyond baseline scores, but the 10 item version of the OSPRO-ROS tool only explained additional variance in mental component summary scores. Overall the additional amount of variance explained at baseline by OSPRO tools was small (increment range from 0.01 to 0.07). When the 4-week change in the 10 item OSPRO-YF tool was added to prediction models it explained additional variance in 12-month pain, disability, and quality of life outcomes. Again the overall amount of variance added was small from change scores (increment range from 0.04 – 0.07).

The pattern for predicting 12-month comorbidity change differed from the other outcome measures. Baseline number of comorbidities still explained the most additional variance after accounting for demographic and clinical variables. However, only the 13 additional items from the OSPRO-ROS+ tool explained variance in the models predicting 12-month comorbidity change.

Individual Predictors of Outcome

Parsimonious models were used to identify individual predictors because they provided a conservative estimate of the overall model's predictive ability (i.e. parsimonious models had lowest total variance explained - Table 2). Also parsimonious models were deemed appropriate for identifying individual predictors due to our lower than anticipated follow up rate at 12 months, meaning that the full models had potential to be over-fit if all covariates were included. Finally, we valued efficiency in identifying individual predictors (i.e. reporting fewest number) and wanted to avoid over-reporting individual predictors. Fewer individual predictors make future risk model building easier by better prioritizing clinical data collection. Results from completed cases and regression imputed models were reported to allow for direct comparisons of model stability.

Model parameters for individual predictors are provided in detail in Table 4 (Pain Intensity and Region Specific Disability), Table 5 (Quality of Life), and Table 6 (Comorbidity). The estimates provided in the Tables 4–6 represent how much the outcome variable would be expected to change per one unit change in a given predictor variable. For example, in Table 4 for 12-month pain intensity outcome the estimate for 'previous episode' as a categorical predictor is 0.83 (completed cases model). This means that a "yes" response to 'previous episode' would increase the expected 12-month pain intensity score by an additional 0.83 points on the 0–10 scale. As another example from Table 4, 'baseline pain intensity' is a continuous predictor with estimate of 0.41 (completed cases model). This means that a 'baseline pain intensity' score of 6 would be predicted to be 2.5 (6 * 0.41) for 12-month pain intensity on the 0–10 scale. A brief summary of individual predictors for each outcome is provided below:

12 Month Pain Intensity—Previous episode in same region, baseline pain intensity, and the OSPRO-YF tool (10 item and 4-week change) were predictors in the inverse probability weighted model (Table 4). These predictors matched the completed cases model, while the regression imputed model also included educational level.

12 Month Region Specific Disability—The inverse probability weighted model included gender, race, comorbidity, baseline score, and OSPRO YF tool (10 item and 4-week change) as individual predictors (Table 4). These predictors matched the completed case model. The regression imputed model included different demographic factors for predictors (e.g. age and previous episode) and also had consideration of anatomical region. The nature of the interaction indicated that prediction of disability outcomes for the shoulder region differed from those at the knee. Otherwise its individual predictors matched the other models.

12 Month Physical Component Summary—Race, comorbidity, baseline PCS scorer, and the OSPRO-YF tool (10 item and 4-week change) were predictors in the inverse probability weighted model (Table 5). These predictors matched the completed cases model, and the regression imputed model differed by including age (instead of race) as an individual predictor.

12 Month Mental Component Summary—The inverse probability weighted model included age, baseline MCS score, 10 item OSPRO-ROS, and OSPRO-YF tool (10 item and 4-week change) as individual predictors (Table 5). These predictors matched the completed cases model, and the regression imputed model included income as an additional individual predictor.

12 Month Comorbidity—Education, baseline number of comorbidities, and 13 additional items of the OSPRO-ROS+ were individual predictors in the inverse probability weighted model (Table 6). The completed cases model only included the baseline number of comorbidities and 13 additional items of the OSPRO-ROS+, and the regression imputed model included age, the baseline number of comorbidities and 13 additional items of the OSPRO-ROS+.

Discussion

Analyses from the OSPRO validation cohort provided additional information on use of concise assessment tools for prediction of musculoskeletal pain outcomes. The 10 item OSPRO-YF added statistically to the prediction of 12-month pain intensity, disability, and quality of life (physical and mental); a finding consistent with other concise tools for pain-associated distress (i.e. Orebro Musculoskeletal Pain Questionnaire[32, 41] and STarT Back Tool[29, 30]). The 10 item OSPRO-ROS tool added statistically to the prediction of 12-month quality of life (mental), while the 13 item OSPRO-ROS+ tool added statistically to the prediction of comorbidity status. The OSPRO-ROS and OSPRO-ROS+ findings were novel as there are no other tools available we are aware of for direct comparison and these data provided preliminary support for this tool's predictive validity. All predictive models included demographic, clinical, and baseline variables as planned covariates consistent with previous modelling strategies.[3, 20, 25] The OSPRO tools added relatively small amounts of variance to models containing covariates (i.e. demographic and clinical factors, comorbidity, and baseline outcome scores), similar to another report focused on psychological measures.[16] Therefore, the OSPRO tools may have limited potential to enhance clinical decision making when considered in conjunction with demographic variables and baseline outcomes scores. The OSPRO tools are intentionally concise and consistently contributed to outcome prediction across a variety of domains in the parsimonious prediction models. Therefore it is our assertion that these tools could still be useful measurement adjuncts for health systems developing clinical pathways that determine appropriateness of non-pharmacological pain management,[44] facilitate delivery of tailored psychologically informed treatment options,[42] and/or consider the impact of disease burden on patient management strategies.[48] However we acknowledge that the individual clinical relevance (if any) of OSPRO tools will need to be determined in follow up studies from additional cohorts.

The 10 item OSPRO-YF tool consistently contributed small amounts of additional variance to predictive models for 12-month pain intensity, region specific disability, and quality of life (mental and physical) outcomes. This finding is similar to the predictive abilities of the aforementioned assessment tools (e.g. Orebro Musculoskeletal Pain Questionnaire[32, 41] and STarT Back Tool[29, 30]) and a recent study suggests it is unlikely that any one screening tool will be superior for prediction when compared head to head to other screening tools.[37]

Therefore, a few caveats deserve mention in interpreting results from this cohort. First, the OSPRO-YF tool was predictive of multiple outcome domains, while the other assessment tools tend to have stronger predictive capabilities for functional outcomes.[36] Second, the OSPRO-YF tool included items for pain resilience, a dimension not captured in the other tools but may be relevant for predicting pain-related outcomes. Third, the OSPRO-YF tool can be used as total score (e.g. these analyses) or to estimate scores of 11 different full length parent questionnaires for negative mood, fear-avoidance, and positive coping style (e.g. in development paper[39]). However, we acknowledge that OSPRO predictive contributions to outcomes were small in magnitude and additional research must be completed before informed recommendations for clinical use are provided.

The contribution of the 10 item OSPRO-YF 4-week change score to outcome prediction expands the concept of treatment monitoring for individuals with neck, shoulder, and knee pain. That is, considering an immediate change in pain-associated psychological distress may improve prediction of longer term clinical outcomes. Treatment monitoring via change in psychological measures has been established for patients with low back pain.[26, 43, 50, 55, 56] In this cohort we considered the OSPRO-YF tool for its treatment monitoring capacity across several other musculoskeletal pain conditions. Consistently the 4-week change in the 10 item OSPRO-YF tool contributed small amounts of additional variance to the prediction of 12-month outcomes for pain intensity, region-specific disability, and quality of life. The OSPRO-YF change score contribution, while small in magnitude, was of equal weight as the baseline score for a given prediction model. This finding suggests that to enhance outcome prediction via treatment monitoring psychological assessments should be structured to capture baseline status and a follow-up measure since they both equally contributed to the outcome of interest.

The 10 item OSPRO-ROS tool was narrower in its predictive scope by being specific to 12-month quality of life (mental). The finding for mental component summary scores suggests that the OSPRO-ROS (short version) can be used in tandem with the OSPRO-YF tool for better accuracy on mental health outcomes. The 10 item OSPRO-ROS tool correlated with depressive symptoms in the cross-sectional development cohort,[19] and this was a corroborative finding in the longitudinal validation cohort. Collectively these findings suggest that even though the items on the OSPRO-ROS are focused on red flag symptomology there is a link between these symptoms and overall mental health status, even after other psychological factors are considered (i.e. by the OSPRO-YF tool in these analyses).

The additional 13 items from the OSPRO-ROS+ contributed small amounts of additional variance to the prediction of 12-month comorbidity change. Traditionally red flag symptom assessment has been geared towards determining existing pathology, but this strategy has been questioned due to low accuracy.[14, 52] An alternate approach to red flag assessment is determining association with change in medical, health, or disease status.[19, 46] In these analyses we focused on whether the 10 item OSPRO-ROS tool or the additional 13 items from the OSPRO-ROS+ was predictive of 12-month comorbidity change. Comorbidity status was selected because musculoskeletal pain burden may be exacerbated by the presence of multiple comorbid conditions, which can independently influence the

trajectories of perceived health status, functional impairment, and disability.[45, 51, 54] As a result, there is surging interest in the implications multiple comorbidities (i.e. multimorbidity) has for individual patient care and decision-making.[1] In order to better understand the impact of multimorbidity and more clearly define who is at risk for poor outcomes, physical therapists and other healthcare providers will need assessment tools that provide a reasonable estimate of future disease burden. Information on future disease burden can then be combined with other existing methods for predicting clinical outcomes resulting in an approach that generates care pathways addressing issues specific to multimorbidity. The additional 13 items of the OSPRO-ROS+ consistently predicted 12-month comorbidity change, adding to models that already included the baseline number of comorbidities. This is an encouraging finding that could aid future clinical decision making for value based care in musculoskeletal pain[40, 48] but it will need to be investigated in additional studies for replication.

The OSPRO tools added statistically to the prediction of outcomes after considering baseline outcome scores, but contributions may have limited clinical relevance. For example, the baseline 10 item OSPRO-YF score (range = 3 – 53) would have to vary by 30 points in order to correspond with a 2-point difference in 12-month pain intensity outcome. This likely means that the OSPRO-YF would be used to refine a prediction after an initial trajectory is determined by baseline pain intensity score. Similarly large differences in baseline OSPRO-YF or OSPRO-ROS scores are needed to predict clinically relevant differences for other outcomes. In the case of the OSPRO-YF tool the 4-week change score can be used to further refine outcome prediction which may enhance its utility, but this comes with the burden of an additional measurement point. Future utility of the OSPRO tools can only be determined in subsequent studies that directly link tool use to clinical decision making.

These findings did indicate that the OSPRO tools can be used broadly across individuals with neck, shoulder, low back, and knee pain. There was very little evidence of influence of anatomical region for the OSPRO-YF and the OSPRO-ROS tools. This finding was similar to our previous work in depressive symptoms,[20] fear-avoidance beliefs,[22] and pain-associated distress.[7] The influence of psychological symptoms on clinical outcomes as not being region dependent has been reported from other cohorts too.[11, 12] However, the regression imputed analyses did indicate a potential for differences in tool used based on disability measures that are specific to an anatomical region. For example, our analyses indicate slightly higher 12-month disability scores would be expected for shoulder pain compared to knee pain, given the same baseline OSPRO-YF score. This finding of needing to consider anatomical specificity with yellow flag assessment converges with the initial validation of a modified STarT Back Tool.[29] The reasons for these contrasting findings from the regression imputed models cannot be determined or resolved within this cohort. However, they do provide focus for future study in this area by determining whether OSPRO tool interpretation needs to be adjusted based on anatomical region if the outcome of interest is region specific disability.

Primary limitations of the OSPRO validation cohort have been previously described in the cohort profile paper and these include convenience sampling and lack of individual treatment parameters.[21] Another primary limitation is that we did not include specific

medical diagnoses and severity of injury in the predictive models. Therefore these predictive models may not have application when a specific medical diagnosis is a strong predictor of clinical outcomes. An additional limitation of this analysis was that the 63.4% follow up rate was lower than was anticipated. Furthermore, there were multiple differences from those that completed followed up and those that did not. Specifically, these predictive models may need to be adjusted for those of non-white race, with lower income and education levels, those that are uninsured, receiving Medicaid or worker's compensation, and those with higher pain and pain associated distress. In order to account for this lower than anticipated and differential follow up rate we were transparent in interpreting results from parsimonious models (to avoid reporting over fit models) that accounted for the missing data (to avoid loss to follow up bias). Most of the time the completed cases and imputed models showed very good convergence, but these analyses indicated that the prediction of comorbidity outcomes was most affected by the loss to follow up. Another limitation is that all outcomes for these analysis were self-reported. Future studies should consider incorporating a corresponding physical performance measure and medical record verification of the 12-month comorbidity status. Finally, in our analysis we did not weight the OSPRO-YF tool based on its different components (e.g. negative affect, fear-avoidance, and positive coping). Therefore, a limitation in interpreting the OSPRO-YF tool score is we don't know which individual components may be better targets for intervention approaches or if there are dominant components of the OSPRO-YF tool for predicting outcomes.

The OSPRO validation cohort generated several areas of future research. First, the musculoskeletal conditions recruited in this cohort were selected because they were highly prevalent and commonly treated by physical therapists in outpatient settings. Future study of the OSPRO tools in less prevalent patient groups are necessary to determine refinements to the existing tools. Second and specific to the OSPRO-ROS tool, there may be an interest in determining if the tool can be used to identify the need for additional diagnostic testing. Although this direction was not our intent in the validation cohort, the OSPRO-ROS tool certainly could be investigated in appropriately designed future studies for improving accuracy in identifying systemic pathology (e.g. would have to include diagnostic standards). Third and specific to the OSPRO-YF tool, investigating whether relevant domains not originally included in tool development (e.g. perceived injustice and optimism) improve predictive performance of the tool is another area of future study. The original OSPRO cohort studies were designed to be proof of concept study for tool development and initial validation. Finally, the original OSPRO tool development did not include item response theory and using such an analytical approach could generate different tools to compare performance in future predictive testing.

Future work will determine if or how OSPRO tools improve clinical decision making for musculoskeletal pain. The OSPRO tools could be used to direct tailored treatment options for higher pain associated psychological distress linked to poor outcomes or for symptom reports indicating increased disease burden. The current study was predictive but future studies can investigate whether these tools can be used to identify responders via treatment effect modification or to verify their use as treatment monitoring tools via mediation analyses. Another area of future work is to incorporate the OSPRO tools into existing electronic health records and/or patient registries. The OSPRO tools provide a concise way

to capture relevant risk adjustment parameters often missing from large scale datasets for musculoskeletal pain. Pragmatic use of these tools will allow for more precise estimates of their predictive capabilities for clinical outcomes and exploration of their ability to predict future healthcare utilization. For example, these tools may be used to identify patients that start in a non-pharmaceutical care pathway but then transition to higher risk options like opioids, injections, or surgery. Earlier identification of these patients may allow for additional tailored strategies to be explored for preventing unwarranted utilization of high risk, low benefit treatments for musculoskeletal pain.

Conclusion

The primary analyses from the OSPRO validation cohort demonstrated how the OSPRO tools added statistically to the prediction of 12-month outcomes for common musculoskeletal pain conditions. Specifically, the 10 item OSPRO-YF tool (which assesses negative mood, fear-avoidance, and positive coping styles) improved prediction of 12-month pain intensity, region specific disability, and quality of life (physical and mental). The 10 item OSPRO-ROS tool (which assesses red flag symptomology) improved prediction of 12-month quality of life (mental) and the additional 13 items from the OSPROROS+ improved prediction of 12-month comorbidity status. OSPRO tools contributed small amounts of variance to prediction models that included demographic and clinical factors, comorbidity, and baseline scores. The OSPRO validation cohort was not designed to be a definitive study, so future research is needed to determine if these tools have a role in improving clinical decision making for better management of musculoskeletal pain.

Acknowledgments

Brooks Rehabilitation provided resources for travel related to this study (e.g. clinic training and presentation of preliminary findings).

The author team graciously acknowledges those individuals assisting with data collection at participating OPT-IN clinical sites. Their selfless efforts were vital to the success of this project.

OPT-IN Network Participants included: UF Health: Daniel Broome, Marty Huegel, Debi Jones, Steve Emery, Mike Hodges, Derek Miles, Jodi Davis, Charlene Stubbington, Mike Darcy; ATI Physical Therapy: Ellen Shanley, Thomas Denninger, Jenna Bartsokas, Elise Harris, Jordan Floyd, Wade Harrell; University of Southern California: Lori Michener, Amy Pomrantz, Brooks Rehabilitation: Raine Osborne, Nata Salvatori, John Leschitz, Brian Hagist, Laura Langer, Tim Shreve, Nando Malaman, Michael Bourassa, Justin Zych, Tasha Mouton Shanklin; University of Illinois at Chicago: Aaron Keil, Brad Myers, Deb Davey, Justin Payette, Adam Wielechowski, Richard Severin, Erik Martinez; Indiana State University: Ryan Hanigan, Carolina Valencia, Danielle Jena, Nicole Woodard; Arcadia University: Angela Tate; Life's Work Physical Therapy: Sandra Stryker, Aaron Leonard, Erin Courtney, Brandon Little, Kathryn Jankord, Brad Simpson, Charleen Hall, Paige Nixon, Julia Neufeld; University of Colorado, Denver: Paul Mintken, Virginia Arnette, Andrea Barsch.

Finally, we thank the Orthopaedic Section's Clinical Research Network Advisory Board for their feedback and guidance during the funding period.

Funding Statement

All authors on this project were supported by the 2013 Clinical Research Network grant from the Orthopaedic Section, American Physical Therapy Association. SZG and JMB received additional support from Brooks Rehabilitation while designing this study and writing this protocol paper. JMB received support from the American National Institutes of Health (NIH) Rehabilitation Research Career Development Program (K12-HD055929). TAL received support from the Foundation for Physical Therapy with a Promotion of Doctoral Studies I and II (PODS I and II) Award.

References

1. Atun R. Transitioning health systems for multimorbidity. Lancet (London, England). 2015; 386:721–722.
2. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three item-reduction approaches. The Journal of bone and joint surgery. American volume. 2005; 87:1038–1046. [PubMed: 15866967]
3. Beneciuk JM, Bishop MD, Fritz JM, et al. The STarT back screening tool and individual psychological measures: evaluation of prognostic capabilities for low back pain clinical outcomes in outpatient physical therapy settings. Physical therapy. 2013; 93:321–333. [PubMed: 23125279]
4. Beneciuk JM, Fritz JM, George SZ. The STarT Back Screening Tool for prediction of 6-month clinical outcomes: relevance of change patterns in outpatient physical therapy settings. The Journal of orthopaedic and sports physical therapy. 2014; 44:656–664. [PubMed: 25098194]
5. Beneciuk JM, George SZ. Pragmatic Implementation of a Stratified Primary Care Model for Low Back Pain Management in Outpatient Physical Therapy Settings: Two-Phase, Sequential Preliminary Study. Physical therapy. 2015; 95:1120–1134. [PubMed: 25858972]
6. Bolton JE. Accuracy of recall of usual pain intensity in back pain patients. Pain. 1999; 83:533–539. [PubMed: 10568862]
7. Butera KA, Lentz TA, Beneciuk JM, George SZ. Preliminary Evaluation of a Modified STarT Back Screening Tool Across Different Musculoskeletal Pain Conditions. Physical therapy. 2016; 96:1251–1261. [PubMed: 26847006]
8. Cao W, Tsiatis AA, Davidian M. Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. Biometrika. 2009; 96:723–734. [PubMed: 20161511]
9. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987; 40:373–383. [PubMed: 3558716]
10. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. Spine. 2005; 30:1331–1334. [PubMed: 15928561]
11. Christensen JO, Johansen S, Knardahl S. Psychological predictors of change in the number of musculoskeletal pain sites among Norwegian employees: a prospective study. BMC musculoskeletal disorders. 2017; 18:140. [PubMed: 28376786]
12. de Vos Andersen NB, Kent P, Hjort J, Christiansen DH. Clinical course and prognosis of musculoskeletal pain in patients referred for physiotherapy: does pain site matter? BMC musculoskeletal disorders. 2017; 18:130. [PubMed: 28356140]
13. Dowell D, Haegerich TM, Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain--United States, 2016. Jama. 2016; 315:1624–1645. [PubMed: 26977696]
14. Downie A, Williams CM, Henschke N, et al. Red flags to screen for malignancy and fracture in patients with low back pain: systematic review. BMJ (Clinical research ed.). 2013; 347:f7095.
15. Ferguson FC, Morison S, Ryan CG. Physiotherapists' understanding of red flags for back pain. Musculoskeletal care. 2015; 13:42–50. [PubMed: 24965065]
16. Foster NE, Thomas E, Bishop A, Dunn KM, Main CJ. Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. Pain. 2010; 148:398–406. [PubMed: 20022697]
17. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. Physical therapy. 2001; 81:776–788. [PubMed: 11175676]
18. Gandhi SK, Salmon JW, Zhao SZ, Lambert BL, Gore PR, Conrad K. Psychometric evaluation of the 12-item short-form health survey (SF-12) in osteoarthritis and rheumatoid arthritis clinical trials. Clinical therapeutics. 2001; 23:1080–1098. [PubMed: 11519772]
19. George SZ, Beneciuk JM, Bialosky JE, et al. Development of a Review-of-Systems Screening Tool for Orthopaedic Physical Therapists: Results From the Optimal Screening for Prediction of Referral and Outcome (OSPRO) Cohort. The Journal of orthopaedic and sports physical therapy. 2015; 45:512–526. [PubMed: 25996361]

20. George SZ, Coronado RA, Beneciuk JM, Valencia C, Werneke MW, Hart DL. Depressive symptoms, anatomical region, and clinical outcomes for patients seeking outpatient physical therapy for musculoskeletal pain. Physical therapy. 2011; 91:358–372. [PubMed: 21233305]
21. George SZ, Lentz TA, Beneciuk JM, Wu SS. The Optimal Screening for Prediction of Referral and Outcome (OSPRO) in patients with musculoskeletal pain conditions: a longitudinal validation cohort from the USA. BMJ open. 2017
22. George SZ, Stryker SE. Fear-avoidance beliefs and clinical outcomes for patients seeking outpatient physical therapy for musculoskeletal pain conditions. The Journal of orthopaedic and sports physical therapy. 2011; 41:249–259. [PubMed: 21335927]
23. George SZ, Zeppieri G Jr, Cere AL, et al. A randomized trial of behavioral physical therapy interventions for acute and sub-acute low back pain (NCT00373867). Pain. 2008; 140:145–157. [PubMed: 18786762]
24. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. Journal of clinical epidemiology. 2005; 58:595–602. [PubMed: 15878473]
25. Hart DL, Werneke MW, Deutscher D, George SZ, Stratford PW. Effect of fear-avoidance beliefs of physical activities on a model that predicts risk-adjusted functional status outcomes in patients treated for a lumbar spine dysfunction. The Journal of orthopaedic and sports physical therapy. 2011; 41:336–345. [PubMed: 21471649]
26. Hart DL, Werneke MW, Deutscher D, George SZ, Stratford PW, Mioduski JE. Using intake and change in multiple psychosocial measures to predict functional status outcomes in people with lumbar spine syndromes: a preliminary analysis. Physical therapy. 2011; 91:1812–1825. [PubMed: 22003164]
27. Hart DL, Werneke MW, George SZ, Deutscher D. Single-item screens identified patients with elevated levels of depressive and somatization symptoms in outpatient physical therapy. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2012; 21:257–268.
28. Henschke N, Maher CG, Refshauge KM, et al. Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. Arthritis and rheumatism. 2009; 60:3072–3080. [PubMed: 19790051]
29. Hill JC, Afolabi EK, Lewis M, et al. Does a modified STarT Back Tool predict outcome with a broader group of musculoskeletal patients than back pain? A secondary analysis of cohort data. BMJ open. 2016; 6:e012445.
30. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis and rheumatism. 2008; 59:632–641. [PubMed: 18438893]
31. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet (London, England). 2011; 378:1560–1571.
32. Hockings RL, McAuley JH, Maher CG. A systematic review of the predictive ability of the Orebro Musculoskeletal Pain Questionnaire. Spine. 2008; 33:E494–500. [PubMed: 18594447]
33. Institute of Medicine Committee on Advancing Pain Research C, EducationRelieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research Washington (DC): National Academies Press (US) National Academy of Sciences; 2011The National Academies Collection: Reports funded by National Institutes of Health.
34. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the international knee documentation committee subjective knee form. The American journal of sports medicine. 2001; 29:600–613. [PubMed: 11573919]
35. Jensen MP, Turner JA, Romano JM, Fisher LD. Comparative reliability and validity of chronic pain intensity measures. Pain. 1999; 83:157–162. [PubMed: 10534586]
36. Karran EL, McAuley JH, Traeger AC, et al. Can screening instruments accurately determine poor outcome risk in adults with recent onset low back pain? A systematic review and meta-analysis. BMC medicine. 2017; 15:13. [PubMed: 28100231]

37. Karran EL, Traeger AC, McAuley JH, Hillier SL, Yau YH, Moseley GL. The Value of Prognostic Screening for Patients With Low Back Pain in Secondary Care. The journal of pain : official journal of the American Pain Society. 2017; 18:673–686. [PubMed: 28153805]
38. Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in Orthopaedics: A Brief Guide. The Journal of bone and joint surgery. American volume. 2015; 97:1628–1634. [PubMed: 26446970]
39. Lentz TA, Beneciuk JM, Bialosky JE, et al. Development of a Yellow Flag Assessment Tool for Orthopaedic Physical Therapists: Results From the Optimal Screening for Prediction of Referral and Outcome (OSPRO) Cohort. The Journal of orthopaedic and sports physical therapy. 2016; 46:327–343. [PubMed: 26999408]
40. Lentz TA, Harman JS, Marlow NM, George SZ. Application of a Value Model for the Prevention and Management of Chronic Musculoskeletal Pain by Physical Therapists. Physical therapy. 2017
41. Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire. The Clinical journal of pain. 2003; 19:80–86. [PubMed: 12616177]
42. Main CJ, George SZ. Psychologically informed practice for management of low back pain: future directions in practice and research. Physical therapy. 2011; 91:820–824. [PubMed: 21451091]
43. Mansell G, Hill JC, Main C, Vowles KE, van der Windt D. Exploring What Factors Mediate Treatment Effect: Example of the STarT Back Study High-Risk Intervention. The journal of pain : official journal of the American Pain Society. 2016; 17:1237–1245. [PubMed: 27565304]
44. Qaseem A, Wilt TJ, McLean RM, Forciea MA. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. Annals of internal medicine. 2017
45. Ritzwoller DP, Crounse L, Shetterly S, Rublee D. The association of comorbidities, utilization and costs for patients identified with low back pain. BMC musculoskeletal disorders. 2006; 7:72. [PubMed: 16982001]
46. Roach KE, Brown M, Ricker E, Altenburger P, Tompkins J. The use of patient symptoms to screen for serious back problems. The Journal of orthopaedic and sports physical therapy. 1995; 21:2–6. [PubMed: 7889028]
47. Robins JM, Rotnitzky A, Zhao LP. Estimation of Regression Coefficients When Some Regressors are not Always Observed. Journal of the American Statistical Association. 1994; 89:846–866.
48. Rundell SD, Gold LS, Hansen RN, Bresnahan BW. Impact of co-morbidities on resource use and adherence to guidelines among commercially insured adults with new visits for back pain. Journal of evaluation in clinical practice. 2017
49. Singla M, Jones M, Edwards I, Kumar S. Physiotherapists' assessment of patients' psychosocial status: are we standing on thin ice? A qualitative descriptive study. Manual therapy. 2015; 20:328–334. [PubMed: 25454686]
50. Smeets RJ, Vlaeyen JW, Kester AD, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. The journal of pain : official journal of the American Pain Society. 2006; 7:261–271. [PubMed: 16618470]
51. van Dijk GM, Veenhof C, Schellevis F, et al. Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. BMC musculoskeletal disorders. 2008; 9:95. [PubMed: 18582362]
52. Verhagen AP, Downie A, Popal N, Maher C, Koes BW. Red flags presented in current low back pain guidelines: a review. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2016; 25:2788–2802.
53. Vernon H. The Neck Disability Index: state-of-the-art, 1991–2008. Journal of manipulative and physiological therapeutics. 2008; 31:491–502. [PubMed: 18803999]
54. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. Journal of general internal medicine. 2007; 22(Suppl 3):391–395. [PubMed: 18026807]
55. Werneke MW, Hart DL, George SZ, Deutscher D, Stratford PW. Change in psychosocial distress associated with pain and functional status outcomes in patients with lumbar impairments referred

to physical therapy services. The Journal of orthopaedic and sports physical therapy. 2011; 41:969–980. [PubMed: 22146493]

56. Whittle R, Mansell G, Jellema P, van der Windt D. Applying causal mediation methods to clinical trial data: What can we learn about why our interventions (don't) work? European journal of pain (London, England). 2017; 21:614–622.
57. Wideman TH, Hill JC, Main CJ, Lewis M, Sullivan MJ, Hay EM. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. Pain. 2012; 153:2182–2191. [PubMed: 22800410]

Key Points

Findings

- The baseline score and 4-week change in the 10 item OSPRO-YF tool improved statistically the prediction of 12-month pain intensity, region-specific disability, and quality of life (physical and mental) outcomes.
- The baseline score of the 10 item OSPRO-ROS tool improved statistically the prediction of 12-month quality of life (mental) and the additional 13 items in the OSPRO-ROS+ improved statistically the prediction of comorbidity change.

Implications

- The OSPRO tools can be used for baseline assessment and treatment monitoring for commonly occurring musculoskeletal conditions (i.e. neck, shoulder, low back, or knee pain)
- The OSPRO tools contributed to outcome prediction, but their contribution to the models was small in magnitude.
- It is our assertion the OSPRO tools may be used for directing care in pain management pathways that deliver early non-pharmacological treatments, psychologically informed approaches, or want to consider the impact of multimorbidity.

Caution

- This sample was not recruited consecutively and there was high loss to follow up. Therefore these results may not be entirely representative of patient populations.
- There is some evidence of anatomical specificity in these tools use for predicting region specific outcomes, and that issue will need to be considered in future studies.
- Additional studies are needed to determine the utility of the OSPRO tools for clinical decision making.

Table 1

Descriptive Summary of OSPRO Validation Cohort

Variable		Label	Overall Sample (n = 440)	Missing 12-Month (n =161)	Completed 12-Month (n = 279)	p-value
Demographics	*Age*	*Mean±SD*	*45.2±15.8*	*43.1±15.3*	*46.5±16.0*	*0.024*
		Median (Min, Max)	*45 [18, 75]*	*42 [18, 74]*	*47 [18, 75]*	
	Gender	Male	164 (37.3%)	69 (42.9%)	95 (34.1%)	0.145
		Female	275 (62.5%)	92 (57.1%)	183 (65.6%)	
		Prefer not to answer	1 (0.2%)		1 (0.4%)	
	Race	American Indian/Alaska Native	3 (0.7%)	1 (0.6%)	2 (0.7%)	0.121
		Asian	25 (5.7%)	8 (5.0%)	17 (6.1%)	
		Black or African American	62 (14.1%)	32 (19.9%)	30 (10.8%)	
		White	343 (78.0%)	117 (72.7%)	226 (81.0%)	
		Don't know / Prefer not to answer	7 (1.6%)	3 (1.9%)	4 (1.4%)	
	Ethnicity	Hispanic or Latino	31 (7.0%)	10 (6.2%)	21 (7.5%)	0.310
		Not Hispanic or Latino	376 (85.5%)	135 (83.9%)	241 (86.4%)	
		Don't know / Prefer not to answer	33 (7.5%)	16 (9.9%)	17 (6.1%)	
	Income	*Less than $20,0000*	*59 (13.4%)*	*29 (18.0%)*	*30 (10.8%)*	*0.004*
		$20,000 to $35,000	*53 (12.0%)*	*14 (8.7%)*	*39 (14.0%)*	
		$35,001 to $50,000	*50 (11.4%)*	*17 (10.6%)*	*33 (11.8%)*	
		$50,001 to $70,000	*56 (12.7%)*	*24 (14.9%)*	*32 (11.5%)*	
		Greater than $70,000	*156 (35.5%)*	*44 (27.3%)*	*112 (40.1%)*	
		Don't know / Prefer not to answer	*66 (15.0%)*	*33 (20.5%)*	*33 (11.8%)*	
	Employment	Full-time employed	237 (53.9%)	88 (54.7%)	149 (53.4%)	0.007
		Part-time employed	62 (14.1%)	21 (13.0%)	41 (14.7%)	
		Unemployed	61 (13.9%)	31 (19.3%)	30 (10.8%)	
		Retired	58 (13.2%)	11 (6.8%)	47 (16.8%)	
		Prefer not to answer	22 (5.0%)	10 (6.2%)	12 (4.3%)	
	Education	*Less than high school*	*11 (2.5%)*	*10 (6.2%)*	*1 (0.4%)*	*<0.0001*
		Graduated from high school	*38 (8.6%)*	*21 (13.0%)*	*17 (6.1%)*	

Variable		Label	Overall Sample (n = 440)	Missing 12-Month (n =161)	Completed 12-Month (n = 279)	p-value
		Some college	*112 (25.5%)*	*49 (30.4%)*	*63 (22.6%)*	
		Graduated from college	*120 (27.3%)*	*41 (25.5%)*	*79 (28.3%)*	
		Some post-graduate course work	*56 (12.7%)*	*15 (9.3%)*	*41 (14.7%)*	
		Completed post graduate degree	*97 (22.0%)*	*23 (14.3%)*	*74 (26.5%)*	
		Prefer not to answer	*6 (1.4%)*	*2 (1.2%)*	*4 (1.4%)*	
	Insurance	*Private*	*273 (62.0%)*	*93 (57.8%)*	*180 (64.5%)*	*0.009*
		Medicare	*52 (11.8%)*	*13 (8.1%)*	*39 (14.0%)*	
		Medicaid	*19 (4.3%)*	*9 (5.6%)*	*10 (3.6%)*	
		Worker's compensation	*14 (3.2%)*	*8 (5.0%)*	*6 (2.2%)*	
		Disability	*4 (0.9%)*	*3 (1.9%)*	*1 (0.4%)*	
		Uninsured	*7 (1.6%)*	*5 (3.1%)*	*2 (0.7%)*	
		Other	*45 (10.2%)*	*15 (9.3%)*	*30 (10.8%)*	
		Unknown / Prefer not to answer	*26 (5.9%)*	*15 (9.3%)*	*11 (3.9%)*	
	Clinic Site	*Portland, OR*	*20 (4.5%)*	*4 (2.5%)*	*16 (5.7%)*	*0.025*
		Los Angeles, CA	*61 (13.9%)*	*18 (11.2%)*	*43 (15.4%)*	
		Greenville, SC	*86 (19.5%)*	*28 (17.4%)*	*58 (20.8%)*	
		Boulder, CO	*17 (3.9%)*	*1 (0.6%)*	*16 (5.7%)*	
		Jacksonville, FL	*50 (11.4%)*	*21 (13.0%)*	*29 (10.4%)*	
		Gainesville, FL	*139 (31.6%)*	*59 (36.6%)*	*80 (28.7%)*	
		Philadelphia, PA	*20 (4.5%)*	*7 (4.3%)*	*13 (4.7%)*	
		Chicago, IL	*24 (5.5%)*	*13 (8.1%)*	*11 (3.9%)*	
		Terra Haute, IN	*23 (5.2%)*	*10 (6.2%)*	*13 (4.7%)*	
Anatomical Region		Neck	98 (22.3%)	39 (24.2%)	59 (21.1%)	0.375
		Low back	118 (26.8%)	46 (28.6%)	72 (25.8%)	
		Shoulder	107 (24.3%)	41 (25.5%)	66 (23.7%)	
		Knee	117 (26.6%)	35 (21.7%)	82 (29.4%)	
Pain Duration (# days)		Mean±SD	413.8±1757.6	506.5±1454.1	360.2±1911.4	0.955
		Median (Min, Max)	90 [0, 29565]	90 [0, 10000]	90 [1,29565]	

Variable		Label	Overall Sample (n = 440)	Missing 12-Month (n =161)	Completed 12-Month (n = 279)	p-value
Onset of Symptoms		*Gradual*	*239 (54.3%)*	*74 (46.0%)*	*165 (59.1%)*	*0.028*
		Sudden	*138 (31.4%)*	*60 (37.3%)*	*78 (28.0%)*	
		Traumatic	*63 (14.3%)*	*27 (16.8%)*	*36 (12.9%)*	
Previous Episodes (in past year)		Yes	224 (50.9%)	81 (50.3%)	143 (51.3%)	0.187
		No	185 (42.0%)	64 (39.8%)	121 (43.4%)	
		Do not remember	31 (7.0%)	16 (9.9%)	15 (5.4%)	
Surgery for Primary Complaint		Yes	83 (18.9%)	37 (23.0%)	46 (16.5%)	0.094
		No	357 (81.1%)	124 (77.0%)	233 (83.5%)	
Unique Number of Comorbidities		Mean±SD	2.1±2.3	1.9±2.1	2.2±2.3	0.158
		Median (Min, Max)	1.5 [0, 13]	1 [0, 13]	2 [0, 11]	
Distribution of Comorbidities		0	134 (30.6%)	55 (34.2%)	79 (28.5%)	0.437
		1	85 (19.4%)	31 (19.3%)	54 (19.5%)	
		2 +	219 (50.0%)	75 (46.6%)	144 (52.0%)	
OSPRO-ROS	10 item score (range: 0–10)	Mean±SD	2.7±2.4	2.8±2.5	2.6±2.3	0.515
		Median (Min, Max)	2 [0, 10]	2 [0, 10]	2 [0, 10]	
	13 item score (range: 0–13)	Mean±SD	1.2±1.8	1.4±2.1	1.2±1.6	0.924
		Median (Min, Max)	1 [0, 12]	1 [0, 12]	1 [0, 9]	
	23 item score (range: 0–23)	Mean±SD	3.9±3.8	4.2±4.3	3.8±3.5	0.631
		Median (Min, Max)	3 [0, 21]	3 [0, 21]	3 [0, 17]	
OSPRO-YF	*10 item score (range: 3–53)*	*Mean±SD*	*17.4±6.7*	*18.4±7.2*	*16.9±6.3*	*0.033*
		Median (Min, Max)	*17 [4, 47]*	*18 [4, 47]*	*16 [4, 40]*	
	7 item Score (range: 3–46)	*Mean±SD*	*14.9±5.5*	*15.9±6.3*	*14.3±4.9*	*0.005*
		Median (Min, Max)	*15 [3, 34]*	*16 [3, 34]*	*15 [3, 28]*	
	17 item score (range: 6–89)	*Mean±SD*	*32.4±11.2*	*34.3±12.5*	*31.2±10.3*	*0.009*
		Median (Min, Max)	*32 [8, 81]*	*33 [8, 81]*	*31 [9, 68]*	
Average Pain Intensity (range: 0–10)		*Mean±SD*	*4.2±2.0*	*4.8±2.3*	*3.9±1.7*	*<.0001*
		Median (Min, Max)	*4 [0, 9.7]*	*4.7[0, 9.7]*	*3.7 [0, 8]*	
Quality of Life (range: 0–100)	Physical Summary Score	Mean±SD	42.7±8.5	41.9±9.1	43.2±8.1	0.138

Variable		Label	Overall Sample (n = 440)	Missing 12-Month (n =161)	Completed 12-Month (n = 279)	p-value
		Median (Min, Max)	43.7 [22.4, 59]	42.1 [22.4, 58.6]	44.4 [24.9, 59]	
	Mental Summary Score	Mean±SD	50.9±9.1	49.9±9.3	51.5±9.0	0.070
		Median (Min, Max)	53 [22.6, 68.8]	51.8 [26, 65.7]	53.7 [22.6, 68.8]	
Neck Disability Index (range: 0–100)		*Mean±SD*	*28.6±16.1*	*31.8±15.6*	*26.5±16.3*	*0.045*
		Median (Min, Max)	*24 [2, 76]*	*30 [2, 62]*	*22 [2, 76]*	
Revised Oswestry Questionnaire (range:0–100)		Mean±SD	28.7±18.2	30.7±18.8	27.4±17.9	0.279
		Median (Min, Max)	26 [0, 86]	31 [0, 86]	24 [2, 82]	
QuickDASH (range: 0–100)		Mean±SD	38.8±20.1	43.6±22.5	35.8±17.9	0.071
		Median (Min, Max)	34.1 [2.3, 97.7]	43.2 [2.3, 97.7]	31.8 [6.8, 77.3]	
IKDC Total Score (range: 0–100)		Mean±SD	39.6±15.7	39.6±17.6	39.6±14.9	0.908
		Median (Min, Max)	39.2 [7.2, 77.3]	38.1 [15.5, 73.2]	39.2 [7.2, 77.3]	
Region Specific Disability (units in z score)		*Mean±SD*	*0.4±0.9*	*0.6±1.0*	*0.4±0.9*	*0.032*
		Median (Min, Max)	*0.3 [−1.6, 3.6]*	*0.5 [−1.3, 3.6]*	*0.2 [−1.6, 3.4]*	

Table Key:

- Italic font indicates variable differs based on follow up status
- OSPRO-YF = Optimal Screening for Prediction of Referral and Outcome Yellow Flag Tool; OSPRO-ROS = Optimal Screening for Prediction of Referral and Outcome Review of Systems Tool; DASH = Disability of Arm, Shoulder, and Hand; IKDC = International Knee Documentation Committee

Table 2

Model Performance for OSPRO Tools Predicting 12 Month Clinical Outcomes

Dependent Variable (12 months)	Pain Intensity	Region Specific Disability	Physical Component Summary	Mental Component Summary	Comorbidity
Results presented in following order for predictive models: Completed cases; Regression imputed; Inverse probability weighted					
Block 1	.276#	.368#	.314#	.291#	.677#
Demographic, Clinical and Comorbidity	.305#	.352#	.291#	.311#	.519#
	.306#	.397#	.299#	.312#	.463#
Block 2	.374*	.516*	.402*	.435*	--------
Baseline Dependent Variable	.407*	.485*	.376*	.433*	
	.406*	.549*	.385*	.425*	
Block 3	.406*	.544	.467*	.483*	.705*
OSPRO-YF (10 items)	.422*	.501*	.413*	.453*	.530*
OSPRO-ROS (10 items)	.442*	.580*	.452*	.473*	.485*
Block 4	.427	.559	.497	.491	.730
OSPRO-YF+ (7 items)	.436	.504	.429	.460	.541*
OSPRO-ROS+ (13 items)	.461	.592	.478	.478	.502
Block 5	.456	.587*	.531*	.522*	.732*
OSPRO-YF (4-week change)	.458*	.536*	.469*	.501*	.544*
	.507	.628*	.550*	.554*	.700
Parsimonious Model (Backward Selection)	.317#	.458#	.417#	.412#	.626#
	.369#	.467#	.353#	.424#	.444#

J Orthop Sports Phys Ther. Author manuscript; available in PMC 2018 December 01.

Dependent Variable (12 months)	Pain Intensity	Region Specific Disability	Physical Component Summary	Mental Component Summary	Comorbidity
	.350#	.491#	.433#	.430#	.613#

Table Key:

- Pain intensity was from 0–10 numeric rating scale; region specific disability was from z-scores derived from validated questionnaires for neck, shoulder, back, and knee disability; Physical and mental component summary was from Medical Outcomes Survey-Short Form; Comorbidity change was from the Charlson and Functional Comorbidity Indices
- OSPRO-YF = Optimal Screening for Prediction of Referral and Outcome Yellow Flag Tool; OSPRO-ROS = Optimal Screening for Prediction of Referral and Outcome Review of Systems Tool; + = additional items in longer versions of the corresponding tool
- # = r-square $p < 0.05$ (first step of model); * = r-square increment $p < 0.05$ (compared to previous step)

Table 3

Predictors of 12 Month Clinical Outcomes from Completed Cases and Regression Imputed Approaches

Model Type	Demographic, Clinical, and Comorbidity	Baseline Dependent	Baseline OSPRO Short	Baseline OSPRO Long	4-Week Change OSPRO-YF	Region Interaction
Pain Intensity						
Completed Cases	Previous Episodes	Yes	YF-10		Yes	
Regression Imputed	Income, Previous Episodes	Yes	YF-10		Yes	
Inverse Probability Weighted	Race, Previous Episodes	Yes	YF-10			ROS-10
Region Specific Disability						
Completed Cases	Gender, Race, Comorbidity	Yes	YF-10		Yes	
Regression Imputed	Age, Previous Episodes, Anatomical Region, Comorbidity	Yes	YF-10		Yes	YF-10
Inverse Probability Weighted	Race, Income, Previous Episodes	Yes	YF-10		Yes	
Physical Component Summary						
Completed Cases	Race, Comorbidity	Yes	YF-10		Yes	
Regression Imputed	Age, Comorbidity	Yes	YF-10		Yes	
Inverse Probability Weighted	Race, Income, Anatomical Region, Comorbidity	Yes	YF-10		Yes	
Mental Component Summary						
Completed Cases	Age	Yes	ROS-10 YF-10		Yes	
Regression Imputed	Age, Income	Yes	ROS-10 YF-10		Yes	
Inverse Probability Weighted	Age, Income, Education	Yes	YF-10		Yes	
Comorbidity						
Completed Cases		Yes		ROS+		
Regression Imputed	Age	Yes		ROS+		
Inverse Probability Weighted	Education	Yes		ROS+		

Table Key:

- Pain intensity was from 0–10 numeric rating scale; region specific disability was from z-scores derived from validated questionnaires for neck, shoulder, back, and knee disability; Physical and mental component summary was from Medical Outcomes Survey-Short Form; Comorbidity was from the Charlson and Functional Comorbidity Indices
- OSPRO-YF = Optimal Screening for Prediction of Referral and Outcome Yellow Flag Tool; OSPRO-ROS = Optimal Screening for Prediction of Referral and Outcome Review of Systems Tool; YF-10 or ROS-10 = 10 item version of the tool; ROS+ = additional 13 items of the OSPRO-ROS.

Table 4

Parameters for Individual Predictors of Pain Intensity and Region Specific Disability

12 Month Pain Intensity (potential range = 0 – 10)					
Completed Cases	*Estimate*	*Regression Imputed*	*Estimate*	*Inverse Probability Weighted*	*Estimate*
Intercept	−0.91	Intercept	−0.34		−1.12
		Education (reference – High School or lower)			
		Some college	0.84		
		Graduated from college	−0.01		
		Some post-graduate courses	0.05		
		Completed post-graduate	−0.22		
		Prefer not to answer	−0.31		
Previous Episode (reference – No)		Previous Episode (reference – No)		Previous Episode (reference – No)	
Yes	0.83	Yes	0.61	Yes	0.89
Baseline Pain Intensity	0.41	Baseline Pain Intensity	0.35	Baseline Pain Intensity	0.43
OSPRO-YF (10 items)	0.07	OSPRO-YF (10 items)	0.05	OSPRO-YF (10 items)	0.08
OSPRO-YF (change)	0.08	OSPRO-YF (change)	0.06	OSPRO-YF (change)	0.08
12 Month Region Specific Disability (potential range = units in standard deviation)					
Completed Cases	*Estimate*	*Regression Imputed*	*Estimate*	*Inverse Probability Weighted*	*Estimate*
Intercept	−1.22	Intercept	−1.41	Intercept	−1.30
Gender (reference – Female)		Age (years)	0.01	Gender (reference – Female)	
Male	−0.29			Male	−0.31
Race (reference – White)		Previous Episode# (reference – No)		Race (reference – White)	
American Indian/Alaska Native	1.06	Yes	0.28	American Indian/Alaska Native	0.97
Asian	−0.16			Asian	−0.16
Black or African American	0.26			Black or African American	0.32
Prefer not to answer	2.22			Prefer not to answer	2.48

		Anatomical Region (reference – Knee) Neck Low Back Shoulder	 −0.01 −0.13 −0.76		
Comorbidity (# baseline)	0.08	Comorbidity (# baseline)	0.06	Comorbidity (# baseline)	0.08
Baseline Z Score	0.35	Baseline Z Score	0.34	Baseline Z Score	0.36
OSPRO-YF (10 items)	0.03	OSPRO-YF (10 items)	0.01	OSPRO-YF (10 items)	0.04
OSPRO-YF (change)	0.04	OSPRO-YF (4-week change)	0.04	OSPRO-YF (change)	0.04
		Anatomical Region × YF (reference – Knee) Neck Low Back Shoulder	 0.03 0.02 0.06		

Table Key:

- Pain intensity was from 0–10 numeric rating scale; region specific disability was from z-scores derived from validated questionnaires for neck, shoulder, back, and knee disability;
- OSPRO-YF = Optimal Screening for Prediction of Referral and Outcome Yellow Flag Tool; OSPRO-ROS = Optimal Screening for Prediction of Referral and Outcome Review of Systems Tool; Comorbidity included unique number from Charlson and Functional Comorbidity Indices

Table 5

Parameters for Individual Predictors of Quality of Life

12 Month Physical Component Summary (potential range = 0 – 100)					
Completed Cases	*Estimate*	*Regression Imputed*	*Estimate*	*Inverse Probability Weighted*	*Estimate*
Intercept	45.33	Intercept	48.14	Intercept	47.1
Race (reference – White)		Age (years)	−0.07		Race (reference – White)
American Indian/Alaska Native	−4.67			American Indian/Alaska Native	−5.12
Asian	3.26			Asian	3.51
Black or African American	−3.16			Black or African American	−2.94
Prefer not to answer	−18.76			Prefer not to answer	−19.58
Comorbidity (# baseline)	−0.83	Comorbidity (# baseline)	−0.54	Comorbidity (# baseline)	−0.92
Baseline PCS score	0.27	Baseline PCS score	0.23	Baseline PCS score	0.24
OSPRO-YF (10 items)	−0.37	OSPRO-YF (10 items)	−0.33	OSPRO-YF (10 items)	−0.40
OSPRO-YF (change)	−0.40	OSPRO-YF (change)	−0.38	OSPRO-YF (change)	−0.41
12 Month Mental Component Summary (potential range = 0 – 100)					
Completed Cases	*Estimate*	*Regression Imputed*	*Estimate*	*Inverse Probability Weighted*	*Estimate*
Intercept	38.76	Intercept	34.15	Intercept	38.93
Age (in years)	0.09	Age (in years)	0.07	Age (in years)	0.08
		Income# (reference < $20K)			
		$20K – $35K	4.71		
		$35+K – $50K	6.40		
		$50+K – $75K	4.88		
		$75+K	6.33		
		Prefer not to answer	5.29		
Baseline MCS Score	0.34	Baseline MCS Score	0.28	Baseline MCS Score	0.35

OSPRO-ROS (10 items)	−0.58	OSPRO-ROS (10 items)	−0.34	OSPRO-ROS (10 items)	−0.65
OSPRO-YF (10 items)	−0.42	OSPRO-YF (10 items)	−0.28	OSPRO-YF (10 items)	−0.43
OSPRO-YF (change)	−0.41	OSPRO-YF (change)	−0.38	OSPRO-YF (change)	−0.41

Table Key:

- Physical and mental component summary scores were from the Medical Outcomes Survey-Short Form
- OSPRO-YF = Optimal Screening for Prediction of Referral and Outcome Yellow Flag Tool; OSPRO-ROS = Optimal Screening for Prediction of Referral and Outcome Review of Systems Tool; Comorbidity included unique number from Charlson and Functional Comorbidity Indices

Table 6

Parameters for Individual Predictors of Comorbidity Status

12 Month Comorbidity Change					
Completed Cases	***Estimate***	***Regression Imputed***	***Estimate***	***Inverse Probability Weighted***	***Estimate***
Intercept	−0.16	Intercept	−0.34	Intercept	−0.38
		Age (years)	0.01		
Comorbidity (# baseline)	−0.79	Comorbidity (# baseline)	−0.87	Comorbidity (# baseline)	−0.79
OSPRO-ROS+ (13 items)	0.19	OSPRO-ROS+ (13 items)	0.14	OSPRO-ROS+ (13 items)	0.25
				Education (reference – High school or lower)	
				Completed post graduate degree	−0.02
				Some post-graduate course work	−0.03
				Graduated from college	0.17
				Some college	0.72
				Prefer not to answer	4.32

Table Key:

- OSPRO-ROS+ = Additional 13 items from the long version of the Optimal Screening for Prediction of Referral and Outcome Review of Systems Tool; Comorbidity change was from Charlson and Functional Comorbidity Indices