American Pain Society
RESEARCH
EDUCATION
TREATMENT
ADVOCACY
PUBLISHED BY ELSEVIER

The Journal of Pain, Vol 16, No 1 (January), 2015: pp 19-30
Available online at www.jpain.org and www.sciencedirect.com

Subgrouping for Patients With Low Back Pain: A Multidimensional Approach Incorporating Cluster Analysis and the STarT Back Screening Tool

Jason M. Beneciuk,[*,†] Michael E. Robinson,[‡,§] and Steven Z. George[*,†,§]

*Departments of *Physical Therapy and [‡]Clinical and Health Psychology, University of Florida, Gainesville, Florida.*
[†]Brooks Rehabilitation–University of Florida College of Public Health and Health Professions Research Collaboration, Jacksonville, Florida.
[§]Center for Pain Research and Behavioral Health, University of Florida, Gainesville, Florida.

Abstract: **Early screening for psychological distress has been suggested to improve patient management for individuals experiencing low back pain. This study compared 2 approaches to psychological screening (ie, multidimensional and unidimensional) so that preliminary recommendations on which approach may be appropriate for use in clinical settings other than primary care could be provided. Specifically, this study investigated aspects of the STarT Back Screening Tool (SBT): 1) discriminant validity by evaluating its relationship with unidimensional psychological measures and 2) construct validity by evaluating how SBT risk categories compared to empirically derived subgroups using unidimensional psychological and disability measures. Patients (N = 146) receiving physical therapy for LBP were administered the SBT and a battery of unidimensional psychological measures at initial evaluation. Clinical measures consisted of pain intensity and self-reported disability. Several SBT risk–dependent relationships (ie, SBT low < medium < high risk) were identified for unidimensional psychological measure scores, with depressive symptom scores associated with the strongest influence on SBT risk categorization. Empirically derived subgroups indicated that there was no evidence of distinctive patterns among psychological or disability measures other than high or low profiles; therefore, 2 groups may provide a clearer representation of the level of pain-associated psychological distress, maladaptive coping, and disability in this setting compared with 3 groups as suggested when using the SBT in primary care settings.**

Perspective: ***This study suggests that the SBT can replace administering several unidimensional psychological measures as a first-line screening measure for psychological distress. However, clinicians need to be aware of the potential for misclassification with SBT results when compared to unidimensional measures. This study also suggests that a modified SBT risk stratification scheme based on empirically derived subgroups could potentially assist in identifying elevated levels of pain-associated psychological distress, maladaptive coping, and disability in practice settings outside of primary care. Patients identified with elevated levels of pain-associated distress and maladaptive coping may be indicated for additional assessment using construct-specific questionnaires.***

Key words: *Psychological screening, psychological subgrouping, STarT Back Screening Tool, low back pain, physical therapy.*

Received November 27, 2013; Revised October 2, 2014; Accepted October 8, 2014.

This study was funded by a 2009 award from the Brooks Health System to the University of Florida Foundation. J.M.B. was supported by a National Institutes of Health T32 Neuromuscular Plasticity Research Training Fellowship grant (T32 HD043730). This manuscript was written while J.M.B. received support from the National Institutes of Health Rehabilitation Research Career Development Program (K12-HD055929).

The authors declare no financial interest in the results of this research.

Address reprint requests to Jason M. Beneciuk, PT, PhD, Department of Physical Therapy, University of Florida, Gainesville, FL 32610-0154. E-mail: beneciuk@phhp.ufl.edu

1526-5900/$36.00

http://dx.doi.org/10.1016/j.jpain.2014.10.004

Early risk factor screening for poor clinical outcomes has been identified as a potential method to improve the efficiency and effectiveness of care.[9,27,50] Findings from recent reviews[10,47] indicate consistent relationships between elevated levels of psychological factors measured during initial assessment and poor future clinical outcomes. Therefore, routine and early identification of psychological risk factors (ie, screening) has been suggested as a method to improve patient management strategies, with the ultimate goal being the prevention of future low back pain (LBP)-associated

activity limitations.[9,38,43] Despite these suggestions, potential problems exist when attempting to translate research findings from one clinical setting to another. For example, differences in psychological profiles have been reported for patients seeking care for LBP in primary care settings when compared to secondary care settings.[46] Moreover, decision-making dilemmas may be more common for secondary care setting providers who are commonly referred patients from primary care physicians, as previous studies have suggested that primary care physicians may not adequately screen for psychological distress.[32,36] Therefore, there is the potential that some patients referred to physical therapy may be more appropriate for referral to mental health providers, which may have negative implications in regard to initial management strategies and subsequent patient outcomes for those provided by secondary care providers (eg, physical therapists). As a result, measures or tools that were developed and intended to be used in primary care settings by physicians or physical therapists providing care in primary care settings require additional testing before being implemented in secondary care settings (eg, outpatient physical therapy).

Two methods of screening for psychological factors include the use of unidimensional or multidimensional self-report questionnaires. Unidimensional measures consist of several items, with each item representing the same psychological construct, whereas multidimensional measures commonly consist of several items, with each item representing a different psychological construct and possibly also consisting of items representing other domains (eg, physical impairment). There are strengths and limitations to using multidimensional measures to screen for psychological factors.[27,47,63,64] For example, multidimensional measures can provide information related to general psychological distress and require less time to administer in comparison to using several unidimensional measures that provide more detailed information about specific psychological factors. Alternatively, multidimensional measures may not provide as robust estimates of construct reliability in comparison to unidimensional measures consisting of multiple items related to the same construct.

The STarT Back Tool (SBT) is an example of a multidimensional screening measure consisting of 9 items related to physical and psychological factors. Based on SBT overall and psychosocial subscale scoring, individuals are categorized as SBT low, medium, or high risk for persistent disabling LBP in primary care settings.[25] SBT low- and medium-risk categories are primarily distinguished by SBT overall scoring, whereas SBT medium- and high-risk categories are primarily distinguished by SBT psychosocial scale scoring. Relevant to the purpose of this current study is the methodology used during the development phase to generate SBT categorization cutoff scores. Specifically, receiver operating characteristic (ROC) curves and area under the curve (AUC) analyses were used to test overall and psychosocial scale scores against dichotomized reference standard (eg, disability) scores.[25] Although findings from previous studies[1,63] support the clinical utility of the SBT compared with unidimensional psychological measures, there is a need for more direct comparisons between multidimensional and unidimensional screening approaches to provide more definitive clinical recommendations. For example, the identification of patient subgroups has been implicated as a high priority for future LBP-related research,[16,22,50] with particular concern for the influence that psychological factors have on LBP outcomes being emphasized[43,49,60]; however, there is no clear direction for clinicians as to the most appropriate method to screen for psychological distress and to identify psychological subgroups.[36] In addition, patients receiving initial consultation for LBP may be associated with different psychosocial profiles based on the type of clinical setting[46]; therefore, additional data are needed to evaluate the SBT's validity and determine its applicability in other health care settings.

The purpose of this study was to compare multidimensional and unidimensional approaches to screening for pain-associated psychological distress, maladaptive coping, and disability. First, we assessed SBT discriminant validity by evaluating relationships between SBT risk categories with full-length psychological measures and a self-report LBP-related disability measure. Second, we investigated SBT construct validity by evaluating how SBT categories compare to empirically derived subgroups using cluster analysis from unidimensional psychological and LBP-related disability questionnaires. Then we compared the SBT categories and empirically derived subgroups to each other and with clinical measures of pain intensity and self-report disability. The overall goal of conducting these analyses was to provide preliminary future research suggestions for assessment of these psychological constructs in clinical settings other than primary care (eg, outpatient physical therapy) because the SBT was specifically designed to support first contact care decision making.[17]

Methods

Data for this cross-sectional study were obtained from a larger observational, prospective cohort study that assessed the predictive validity of the SBT in comparison to unidimensional psychological measures for 6-month clinical outcomes.[1] Data were collected between December 14, 2009, and February 5, 2012, from 4 outpatient physical therapy clinics of Brooks Rehabilitation located in Jacksonville, Florida, and 2 outpatient physical therapy clinics of Shands Rehabilitation located in Gainesville, Florida.

Participants

Consecutive patients seeking treatment for LBP at 6 participating outpatient physical therapy clinics were screened for study eligibility by a physical therapist. All patients were referred for physical therapy by a physician; therefore, this setting was considered secondary care. Potential study participants met both of the following criteria before being enrolled into this study:

1) adults between the ages of 18 and 65 years seeking physical therapy for LBP (defined as having symptoms at T12 or lower, including radiating pain into the buttocks and lower extremity) and 2) the ability to read and speak the English language. We included patients with a full range of LBP based on self-reported current symptom duration (ie, acute [≤14 days], subacute [15–90 days], and chronic [≥91 days]). These broad inclusion criteria were to allow for a cohort that was applicable to clinical practice. Potential study participants were ineligible to participate in this study if any of the following criteria were met: 1) the presence of systemic involvement related to metastatic or visceral disease; 2) recent spinal fracture; 3) osteoporosis; or 4) pregnancy. Physical therapists provided all patients that met study eligibility criteria with a brief explanation of the study and a study advertisement with primary investigator contact information. Clinicians emphasized to patients that participating in this study would not dictate the treatment they received for their LBP, and that if they elected not to participate, they would receive the same treatment. This study was approved by the University of Florida institutional review board, and informed consent was obtained from each study participant.

Demographic and Historical Variables

Study participants were asked to complete a standardized self-report questionnaire consisting of demographic items related to age, sex, race, and employment status. Additionally, information involving LBP clinical characteristics (ie, prior surgery, symptom duration, symptom onset, symptom location, work-related LBP) was obtained.

SBT

The SBT is a 9-item multidimensional screening measure used to identify subgroups of patients with LBP in primary care settings based on the presence of modifiable prognostic factors that may be useful in matching patients with targeted interventions.[24,25] The SBT contains items related to physical and psychosocial factors that have been identified as strong independent predictors for persistent disabling LBP. SBT overall scores (ranging from 0 to 9) are determined by summing all positive responses, and SBT psychosocial subscale scores (ranging from 0 to 5) are determined by summing items related to bothersomeness, fear, catastrophizing, anxiety, and depression. Based on patient responses, the SBT categorizes patients as "high risk" (psychosocial subscale scores ≥4) in whom high levels of psychosocial prognostic factors are present with or without physical factors present; "medium risk" (overall score >3; psychosocial subscale score <4) in whom physical and psychosocial factors are present but not a high level of psychosocial factors; or "low risk" (overall score 0–3) in whom few prognostic factors are present.[25] Acceptable test-retest reliability and internal consistency have been reported for SBT overall and psychosocial subscale scores.[25] SBT overall scores have demonstrated acceptable to outstanding discriminant validity for physical reference standards (eg, disability and referred leg pain), whereas SBT psychosocial subscale continuous scores best discriminated psychosocial reference standards (eg, catastrophizing, fear, and depression) in primary care settings.[25] The SBT has demonstrated good concurrent validity in comparison to a similar screening instrument.[26]

Psychological Measures

Our selection of unidimensional psychological measures was primarily based on SBT psychosocial scale items and constructs related to maladaptive coping strategies and psychological distress in response to pain. Psychological measures were collected at baseline and are described in more detail below.

Maladaptive Coping

Fear-Avoidance Beliefs Questionnaire (FABQ). Fear-avoidance beliefs specific to LBP were assessed with the FABQ.[61] The FABQ consists of a 4-item FABQ physical activity scale (FABQ-PA, potentially ranging from 0 to 24) and a 7-item FABQ work scale (FABQ-W, potentially ranging from 0 to 42), with higher scores indicating higher levels of fear-avoidance beliefs for both FABQ scales. Both FABQ scales have been found to have acceptable reliability[31,55,61] and have demonstrated internal consistency.[55,57,58,61] The FABQ-W has demonstrated predictive validity for disability and work loss in patients with LBP.[18,19,21,61]

Tampa Scale of Kinesiophobia (TSK-11). The TSK-11 was used to assess the degree of fear of movement and injury or reinjury in individuals with LBP.[65] The TSK-11 is an 11-item questionnaire with a potential range of 11 to 44, with higher scores indicating greater fear of movement and injury or reinjury due to pain. The TSK-11 has been found to have good test-retest reliability and internal consistency.[65] Predictive and concurrent validity have also been reported for the TSK-11.[65]

Pain Catastrophizing Scale (PCS). The PCS was used to assess the degree of catastrophic cognitions due to painful experiences.[56] Pain catastrophizing has been broadly defined as an exaggerated negative orientation toward actual or anticipated pain experiences.[56] The PCS is a 13-item questionnaire with a potential range of 0 to 52, with higher scores indicating higher levels of pain catastrophizing. Previous studies have supported the PCS as a reliable and valid measure.[11,12,48,56]

Psychological Distress

Patient Health Questionnaire (PHQ-9). The PHQ-9 was used to assess the degree to which depressive symptoms are present in patients with LBP. The PHQ-9 is a 9-item questionnaire with a potential range of 0 to 27, with higher scores indicating elevated depressive symptoms. The PHQ-9 has demonstrated various types of validity in different healthcare settings[29,40] and has been used in studies involving patients with LBP.[13]

State-Trait Anxiety Inventory (STAI)

The trait portion of the STAI (STAI-T) was used to assess the degree that dispositional anxiety has on patients with LBP.[54] The STAI-T is a 20-item questionnaire with a

potential range of 20 to 80, with higher scores indicating elevated levels of anxiety. The STAI-T has been found to be reliable and valid.[4,35] We reported the trait portion of the STAI as this construct is considered to be relatively stable over time and this measure was only assessed at intake.

Clinical Measures

Clinical measures were collected at baseline and are described in more detail below.

Numerical Pain Rating Scale (NPRS)

Pain intensity was rated using an NPRS ranging from 0 (no pain) to 10 (worst pain imaginable).[7,8,33] Participants were asked to rate their current pain intensity as well as their best and worst level of pain intensity over the past 24 hours. These 3 pain ratings were averaged and used as the NPRS variable in this study.[34]

Oswestry Disability Questionnaire (ODQ)

LBP-related disability was assessed with the ODQ, which has 10 items that assess how LBP affects common daily activities.[20,30] The ODQ has a range of 0% (no disability due to LBP) to 100% (completely disabled due to LBP), with higher scores indicating higher LBP-related disability. The ODQ has been found to have high levels of test-retest reliability, internal consistency, validity, and responsiveness.[15,20,52]

Roland-Morris Disability Questionnaire (RMDQ)

LBP-related disability was also assessed with the RMDQ, which has 24 items that assess the functional status over the past 24 hours in patients with LBP.[53] The RMDQ has a range of 0 (no disability due to LBP) to 24 (maximum disability due to LBP), with higher scores indicating higher LBP-related disability. The RMDQ has been found to have high levels of test-retest reliability, internal consistency, validity, and responsiveness.[52,53]

Data Analysis

All data analyses were performed using SPSS, version 20.0 (IBM Corp, Armonk, NY). Intake descriptive statistics were calculated for demographic variables and psychological and clinical measures. Raw scores for each full-length specific psychological measure and the RMDQ were transformed to z scores to provide standardized scores for subsequent cluster analysis techniques; however, raw scores are reported for descriptive purposes because they are more clinically interpretable.

SBT Discriminant Validity

We assessed SBT discriminant validity by evaluating relationships between SBT categories and differences in full-length psychological measure and the RMDQ scores. First, relationships between full-length psychological measure and RMDQ scores by SBT categorization were evaluated using 1-way analysis of variance with Bonferroni post hoc testing as appropriate. Next, discriminant function analysis (DFA) with cross-validated jackknifed classification was performed on the same sample as a follow-up to interpret 1) how the psychological measure and RMDQ scores differentiated SBT categorization and 2) the accuracy in SBT categorization using psychological measure and RMDQ scores. Briefly, DFA is a multivariate statistical procedure used to determine if a set of variables (ie, psychological measure and RMDQ scores) can predict group membership (ie, SBT categorization).[59] Eigenvalues were reported as a measure of variance, indicating how well the discriminant function discriminated between SBT categories, with higher eigenvalues indicating greater discrimination. Canonical correlations were reported as a measure of the relationship between SBT categorization and the discriminant function, with chi-square tests used to determine the significance of the relationship. A summary of classification results from the DFA was generated to evaluate for accuracy in SBT categorization. Jackknifed (ie, 1 case at a time deleted) classification was used because it estimates the ability of predictors to separate groups and has been suggested to reduce bias in classification.[59] Finally, 2 separate linear regression models were used to provide a test of linear effect across the 3 SBT categories, with intake NPRS and ODQ scores serving as dependent variables.

SBT Construct Validity

We investigated SBT construct validity by evaluating how SBT categories compare to empirically derived subgroups and to determine which method is more strongly associated with clinical measures of pain intensity (NPRS) and self-report disability (ODQ) at intake. A cluster analysis using unidimensional psychological measure and RMDQ scores was used to generate empirically derived subgroups without considering SBT categorization. Specifically, an exploratory hierarchical agglomerative cluster analysis was performed using Ward's clustering method with squared Euclidean distances as the similarity measure to create homogeneous cluster profiles among unidimensional psychological measures and the RMDQ. Agglomeration coefficients were inspected and plotted to establish the most optimal cluster solution based on the percent change between adjacent cluster solutions[45] and plot characteristics (ie, elbow criterion).[37] Then, to identify potential cluster group differences in demographic and clinical measure scores at intake, independent t-tests were used. Next, we compared the distribution of SBT categories by our empirically derived subgroups using chi-square analysis. Finally, SBT categories and cluster solutions were independently assessed for relationships with NPRS and ODQ clinical measures using 1-way analysis of variance and independent t-testing, respectively.

SBT Cutoff Scores Associated With Empirically Derived Psychological Subgroups

ROC curve analysis was used to determine SBT overall scores that predicted empirically derived subgroup categorization. AUC was interpreted as the probability of correctly identifying individuals categorized as "elevated

pain-associated psychological distress, maladaptive coping and disability." The range of AUC scores is from .5 (no better than chance identification) to 1.0 (perfect identification). Cutoff scores were then calculated for SBT overall scores (to match current SBT categories) for predicting empirically derived psychological subgroup categorization. Cutoff scores were determined by calculating sensitivity, specificity, and likelihood ratios for each potential SBT overall score.

Results

During this study period, 275 patients were screened for eligibility criteria. Of these patients, 123 were excluded from study participation, with the most common reason being that they were older than 65 years (n = 47). The remaining 152 patients provided informed consent and were enrolled into the study. Of these patients, 6 were not able to complete the study for personal reasons. Therefore, intake data were obtained from 146 patients. Patient demographic and clinical characteristics and psychological scores at intake by SBT categorization are displayed in Table 1. The SBT categorized 53 (36.3%) patients as low risk, 55 (37.7%) as medium risk, and 38 (26.0%) as high risk. There were no differences in demographic or clinical characteristics among SBT subgroups ($P > .05$).

SBT Discriminant Validity

Comparisons among all SBT risk subgroups indicated an SBT risk–dependent relationship (ie, SBT low < medium < high risk) with several psychological measures and the RMDQ. Specifically, patients categorized as SBT low risk had lower intake FABQ-PA, PCS, TSK-11, PHQ-9, and RMDQ scores compared to those categorized as being medium or high risk ($P < .05$). Comparisons among SBT low- and high-risk subgroups indicated that patients initially categorized as SBT low risk had lower intake FABQ-W and STAI-T scores compared with those categorized as high risk ($P < .05$).

DFA run with simultaneous entry method with 1 self-report disability (RMDQ: Wilks $\lambda = .61$, $P < .001$) and 6 psychological predictors (FABQ-PA: Wilks $\lambda = .80$, $P < .001$; FABQ-W: Wilks $\lambda = .95$, $P = .032$; PCS: Wilks $\lambda = .78$, $P < .001$; TSK-11: Wilks $\lambda = .81$, $P < .001$; PHQ-9: Wilks

Table 1. Descriptive Statistics for Study Sample

Variable	Total Sample (N = 146)	STarT Low Risk (n = 53)	STarT Medium Risk (n = 55)	STarT High Risk (n = 38)	P Value
Demographic and clinical					
Age (y)	41.1 (13.5)	38.8 (14.0)	44.6 (12.8)	40.3 (13.0)	.069
Sex, n female	89 (61.0%)	28 (52.8%)	37 (67.3%)	24 (63.2%)	.291
Race					.635
Caucasian/white	110 (75.9%)	39 (75.0%)	41 (74.5%)	30 (78.9%)	
African American/black	25 (17.2%)	12 (23.1%)	9 (16.4%)	4 (10.5%)	
Other	11 (6.9%)	2 (1.9%)	5 (9.1%)	4 (10.5%)	
Employment					.219
Employed	94 (65.7%)	35 (67.3%)	39 (72.2%)	20 (54.0%)	
Unemployed	39 (27.3%)	15 (28.9%)	9 (16.7%)	15 (40.5%)	
Retired	10 (7.0%)	2 (3.8%)	6 (11.1%)	2 (5.4%)	
Previous low back surgery (yes)	26 (17.9%)	8 (15.1%)	11 (20.4%)	7 (18.4%)	.773
Symptom duration					.079
Acute (≤14 d)	17 (11.8%)	3 (5.8%)	11 (20.8%)	3 (7.9%)	
Subacute (15–90 d)	56 (39.2%)	20 (38.5%)	17 (32.1%)	19 (50.0%)	
Chronic (≥91 d)	70 (49.0%)	29 (55.8%)	25 (47.2%)	16 (42.1%)	
Symptom location					.059
LBP only	49 (33.6)	25 (47.2%)	15 (27.3%)	9 (23.7%)	
LBP and buttock or thigh	72 (49.3%)	23 (43.4%)	30 (54.5%)	19 (50.0%)	
LBP and lower leg	25 (17.1%)	5 (9.4%)	10 (18.2%)	10 (26.3%)	
Work-related LBP (yes)	19 (13.0%)	8 (15.1%)	7 (12.7%)	4 (10.5%)	.813
STarT measures					
STarT overall score (range: 0–9)	4.5 (2.5)	1.8 (.9)	5.1 (1.1)	7.4 (1.1)	<.001
STarT psychosocial score (range: 0–5)	2.3 (1.6)	.9 (.8)	2.3 (.9)	4.5 (.5)	<.001
Psychological measures					
FABQ-PA (range: 0–24)	14.5 (5.7)	11.8 (5.5)	14.6 (5.3)	18.4 (4.4)	<.001
FABQ-W (range: 0–42)	12.5 (10.9)	9.6 (9.6)	13.7 (10.7)	15.6 (12.6)	.023
PCS (range: 0–52)	16.8 (12.1)	10.7 (9.6)	17.1 (10.5)	25.3 (13.0)	<.001
TSK-11 (range: 11–44)	25.1 (6.8)	21.9 (6.4)	25.2 (5.4)	29.7 (6.9)	<.001
PHQ-9 (range: 0–27)	7.2 (6.1)	3.7 (3.8)	7.5 (5.5)	12.4 (6.0)	<.001
STAI-T (range: 20–80)	36.1 (9.1)	33.4 (7.6)	36.2 (9.9)	39.5 (9.4)	.008
Clinical Outcome Measures					
NPRS (range: 0–10)	5.3 (2.0)	4.4 (1.9)	5.6 (1.9)	6.6 (1.5)	<.001
ODQ (range: 0–100%)	32.4 (16.7)	19.9 (12.7)	37.0 (14.1)	43.5 (14.2)	<.001
RMDQ (range: 0–24)	11.2 (5.9)	6.7 (4.5)	12.1 (5.1)	16.2 (4.1)	<.001

λ = .69, $P < .001$; STAI-T: Wilks λ = .94, $P = .012$) suggested that each predictor contributed uniquely to SBT categorization and resulted in 2 discriminant functions, which is expected with 3 SBT risk categories. The overall test of the 2 functions was significant ($\chi^2[14] = 102.84$, Wilks λ = .47, $P < .001$), indicating that predictor scores were able to discriminate among the 3 SBT categories. The test for function 2 alone was not significant ($\chi^2[6] = 2.00$, Wilks λ = .98, $P = .919$); however, function 1 accounted for 51.8% (canonical R = .72) of the total relationship between predictors and SBT categories. The pooled within-group correlations between discriminating variables and standardized canonical discriminant functions, as well as the standardized canonical discriminant function coefficients (analogous to multiple regression beta weights), are provided in Table 2. RMDQ and PHQ-9 scores demonstrated the strongest positive relationships with the discriminant function, whereas FABQ-PA and PCS scores demonstrated moderate positive relationships and TSK-11 scores demonstrated the weakest positive relationship. FABQ-W and STAI-T scores demonstrated weak negative relationships with the discriminant function. The overall accuracy for classification using the discriminant function was 59.4% using the cross-validated jackknifing technique. The percentages classified correctly were 78.8% for SBT low-risk, 37.7% for SBT medium-risk, and 63.2% for SBT high-risk. Based on sample distribution, the prior probabilities for chance assignment were 36.3% for SBT low-risk, 37.7% for SBT medium-risk, and 26.0% for SBT high-risk. Classification by the discriminant function exceeded chance classification for SBT low- and high-risk categories and was identical for SBT medium-risk when the cross-validated values were compared.

Results from 2 separate linear regression models indicated that SBT categorization accounted for 19.8% and 32.3% of the variance in intake NPRS (β = .44, $P < .01$) and ODQ (β = .57, $P < .01$) scores, respectively. SBT risk categorization–dependent relationships with intake NPRS and ODQ scores are provided in Figs 1A and 1B.

Table 2. Coefficients of Unidimensional Psychological and Disability Measures of the Discriminant Function

Unidimensional Psychological and Disability Measures	Discriminant Function 1		Discriminant Function 2	
	Standardized Coefficients*	Correlation Coefficients†	Standardized Coefficients*	Correlation Coefficients†
FABQ-PA	.288	.475	.540	.475
FABQ-W	−.290	.213	−.302	−.262
PCS	.284	.509	−.031	.138
TSK-11	.084	.465	.325	.403
PHQ-9	.413	.645	.454	.217
STAI-T	−.072	.246	−.133	.042
RMDQ	.640	.770	−.764	−.582

*Standardized canonical discriminant function coefficients.
†Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions.

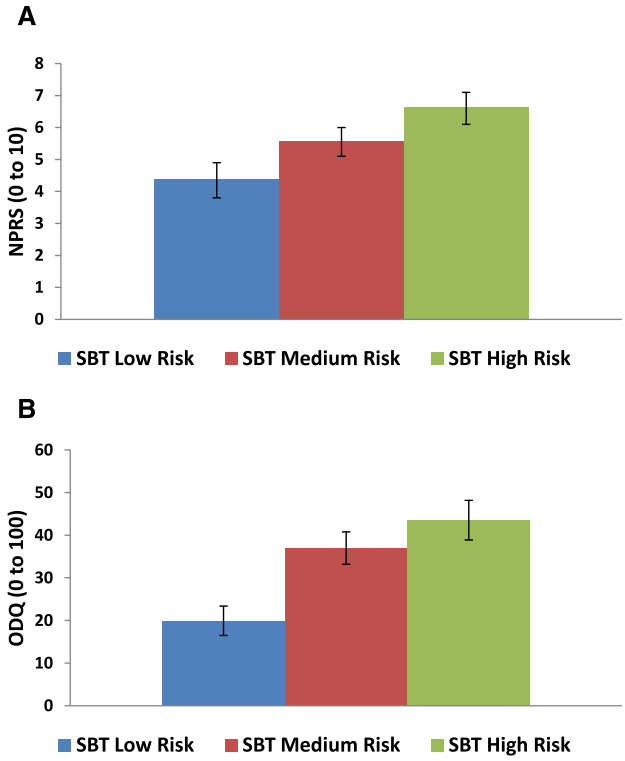

Figure 1. (A) Intake NPRS scores by SBT categorization. **(B)**. Intake ODQ scores by SBT categorization.

Empirically Derived Subgroups

Inspection of all predictor z-scores indicated that absolute values did not exceed 4.0 (range = −2.5 to 3.1), suggesting that the data did not contain extreme outliers.[44,59] Inspection of agglomeration coefficients from a hierarchical agglomerative cluster analysis of 6 psychological measures and a single self-report disability measure revealed that the percent change was moderate (41.8%) between the 2- and 1-cluster solutions, with relatively smaller changes in preceding steps, suggesting that a 2-cluster solution is appropriate, which was further confirmed by visual inspection of plotted agglomeration coefficients.[37,45] The cluster profiles are shown in Fig 2. Cluster 1 was labeled "Low Pain-Associated Psychological Distress, Maladaptive Coping, and Disability" (n = 78, 54.5%) and was composed of

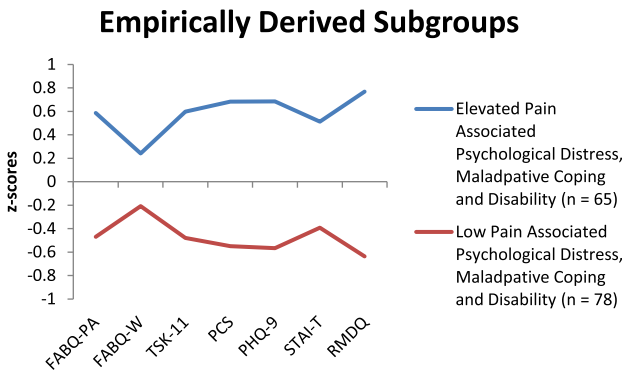

Figure 2. Empirically derived psychological subgroups.

individuals who were associated with lower psychological measure and RMDQ scores when compared to Cluster 2, which was labeled "Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability" (n = 65, 45.5%) and was composed of individuals who were associated with higher psychological measure and RMDQ scores. There were no significant differences between the clusters in demographic or LBP clinical characteristic variables ($P > .05$). Descriptive statistics for psychological measure and RMDQ raw scores for each cluster are shown in Table 3 to make interpretation easier.

The distribution of cluster profiles by SBT categorization is provided in Fig 3. In the Low Pain-Associated Psychological Distress, Maladaptive Coping, and Disability cluster, 60.3% of participants were categorized as SBT low risk, compared to SBT medium and high risk (35.9% and 3.8%, respectively), and in the Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability cluster, 53.8% of participants were categorized as SBT high risk, compared to SBT medium and low risk (38.5% and 7.7%, respectively) ($\chi^2 = 60.36$, $P < .001$). Inspection of standardized residuals indicated that the proportion of those categorized as SBT low risk was less than expected (standardized residual = −3.8) and those categorized as SBT high risk was greater than expected (standardized residual = 4.3) in the Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability cluster.

Using SBT categorization to evaluate for relationships among clinical measures, an SBT risk–dependent relationship (ie, SBT low < medium < high risk) was identified for NPRS scores ($P < .01$; Cohen's d estimates: low vs medium, .63; low vs high, 1.29; medium vs high, .59). Patients categorized as SBT low risk were associated with lower ODQ scores in comparison to those categorized as SBT medium (Cohen's d = 1.28) and high risk (Cohen's d = 1.75) ($P < .01$); however, the ODQ scores for SBT medium and high risk were similar ($P > .05$). Using our 2 empirically derived subgroups to evaluate for similar relationships, we found that patients allocated to the Low Pain-Associated Psychological Distress, Maladaptive Coping, and Disability cluster were associated with lower NPRS ($P < .01$; Cohen's d = 1.14) and ODQ ($P < .01$; Cohen's d = 1.52) scores in comparison to those allocated to the Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability cluster (Table 3).

Table 3. Psychological and Clinical Measure Scores Across Empirically Based Psychological Subgroups

	Low Pain-Associated Psychological Distress, Maladaptive Coping, and Disability Cluster (n = 78)	*Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability Cluster (n = 65)*	P *Value*	*Effect Size*
Psychological measures				
FABQ-PA	11.9 (5.1)	18.0 (4.4)	<.01	1.28
FABQ-W	10.4 (9.6)	15.5 (11.9)	<.01	.47
PCS	10.3 (8.6)	25.2 (11.1)	<.01	1.50
TSK-11	21.9 (5.8)	29.3 (5.7)	<.01	1.29
PHQ-9	4.0 (3.5)	11.6 (5.9)	<.01	1.57
STAI-T	32.5 (6.8)	40.7 (9.6)	<.01	.99
Clinical measures				
NPRS	4.5 (1.9)	6.4 (1.4)	<.01	1.14
OSW	23.3 (13.4)	43.6 (13.3)	<.01	1.52
RMDQ	7.4 (4.6)	15.8 (3.9)	<.01	1.97

Values are mean (SD).

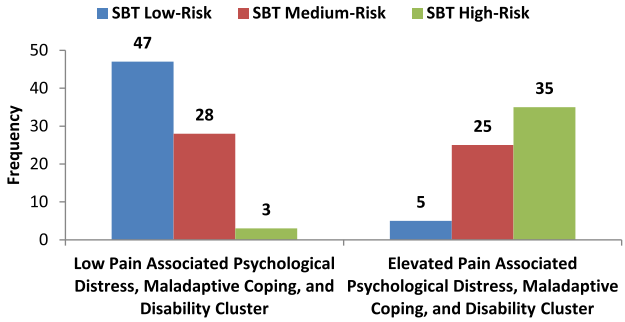

Figure 3. STarT category distribution by cluster profile.

SBT Cutoff Scores Associated With Empirically Derived Psychological Subgroups

The AUC for SBT overall scoring was .907 (95% confidence interval: .857–.957). Upon visual inspection, the cutoff point nearest to the upper-left hand corner of the ROC curve was 4.5 points for the SBT overall score (sensitivity = .892; specificity = .821; +LR = 4.98; −LR = .13), indicating that patients with SBT total scores ≥5 would have increased odds to be categorized as Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability, and those scoring ≤4 would have decreased odds to be categorized as Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability.

Discussion

The purpose of this study was to provide a direct comparison between the SBT with several psychological measures and a self-report LBP-related disability measure. The overall intent of this study was to provide preliminary suggestions for measurement of key psychological constructs in settings other than primary care. The SBT has been successfully applied in primary care[25] and chiropractic[39] settings; therefore, additional data are needed to evaluate the SBT's validity and determine the promising applicability of this screening measure in other health care settings because the SBT was specifically designed to support first-contact care decision making.[17] Our primary findings indicated that several SBT risk–dependent relationships (ie, SBT low < medium < high risk) exist for full-length unidimensional psychological

measure scores and that with the exception of self-reported disability, depressive symptoms may have the strongest influence on risk categorization by the SBT. Our empirically derived subgroups indicated that there was no evidence of distinctive patterns among psychological or self-report disability measures other than high or low profiles in this setting compared with 3 groups as suggested when using the SBT in primary care settings. Therefore, the SBT has potential for use as a first-line screening measure for identifying pain-associated psychological distress, maladaptive coping, and disability in physical therapy settings; however, future prospective studies are required to evaluate potential modification of SBT risk category cutoff scores that may provide a more accurate representation of the subgroups in this particular clinical setting.

The identification of SBT risk–dependent relationships across several unidimensional psychological measure scores provide indications that administering the SBT can ably replace administering multiple unidimensional psychological measures at initial assessment as a first-line screening measure for psychological distress. Previous studies conducted in physical therapy[1] and primary care[63] settings have indicated similar suggestions. This parsimonious measurement option may be advantageous for clinicians practicing in busy outpatient settings where time is becoming more limited and utilization of brief validated screening tools has been suggested as a method to improve clinical decision-making efficiency.[26,27,47] However, clinicians using the SBT as a first-line screening measure need to be aware of the potential for misclassification because the SBT provides an indication of overall disability risk. The SBT does not provide detailed information about specific psychological factors; therefore, certain patients may require additional assessment for specific psychological factors. For example, our findings indicated that when only considering psychological measures, PHQ-9 scores were associated with the strongest positive relationship with SBT categorization, potentially suggesting that further detailed assessment for depressive symptoms may be appropriate for those patients categorized as SBT high-risk. This additional information could then be implemented into clinical decision making to determine if these patients are potential candidates for referral to mental health professionals.

We used cluster analysis to generate 2 empirical-based subgroups, whereas the SBT categorizes patients into 3 subgroups. The SBT was developed to incorporate baseline prognostic information to distinguish patients categorized as SBT low-risk from those categorized as medium- and high-risk; however, the expected patient response to treatment is primarily used to distinguish SBT medium risk from high risk and would require analysis of prospective data.[17] Therefore, we acknowledge that when comparing these 2 different subgrouping approaches our findings may have been a reflection of the cross-sectional design and analysis. Future prospective studies are necessary to determine if the 2 empirically derived subgroups identified in our cluster analysis provide predictive capabilities for clinical outcomes as well as the 3 SBT subgroups already described.[1] Specifically, a prospective comparison of the 2 and 3 SBT subgroup options would be able to determine if incorporating only 2 subgroups is sensitive enough to distinguish outcome differences from SBT medium- and high-risk groups. This is an important issue to resolve in future studies because a 2-subgroup option may be more pragmatic for screening purposes; however, if consolidation of risk subgroups limits how the SBT can be used to guide initial treatment decisions, then the 2-subgroup option may not be acceptable for clinical adoption.

Despite the methodologic differences in generating subgroups, visual inspection of the distribution of SBT scores generated from our sample did not reflect a trimodal distribution (Figs 4A and 4B); rather, it was more closely indicative of a bimodal distribution that parallels the number of our empirically derived subgroups. These findings are consistent with previous suggestions that current 3 SBT categorical designations may not represent underlying distributions of SBT scores for patients experiencing LBP in all clinical settings.[51] Therefore, a modified risk stratification scoring system using the SBT has the potential to be more suitable in this clinical setting. Nevertheless, future studies are required prior to providing any preliminary recommendations. In our analyses, the ROC-generated SBT cutoff scores that can be used to categorize patients as Elevated Pain-Associated Psychological Distress, Maladaptive Coping, and Disability were similar, compared to a SBT psychosocial scale score equal to or greater than 4 points for the same categorization in the primary care setting. Our empirically derived psychological subgroups were also

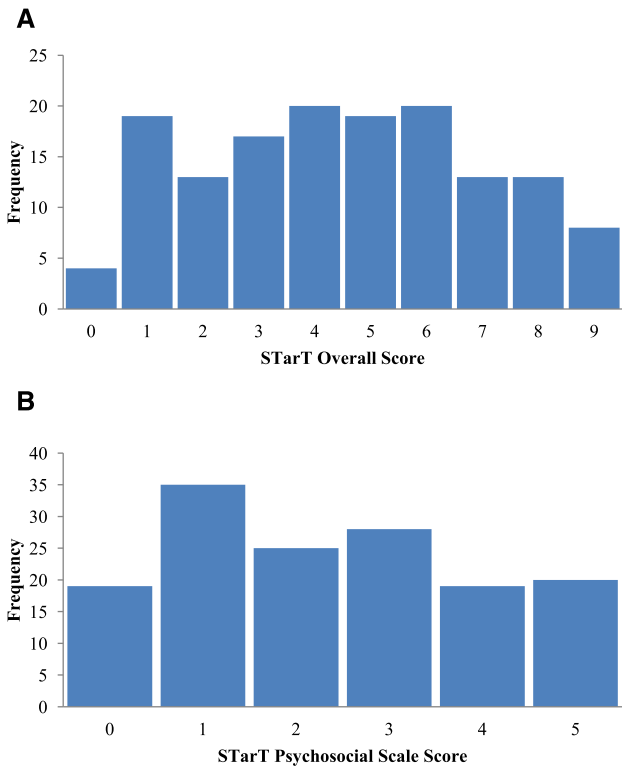

Figure 4. **(A)**. Distribution of STarT overall scores. **(B)**. Distribution of STarT psychosocial scale scores.

associated with low and high levels for multiple clinical measure scores (ie, NPRS and ODQ), each representing a unique clinical outcome domain (ie, pain intensity and self-reported disability, respectively), which can have future research implications. Future longitudinal studies are required to determine if our 2-cluster solution provides an accurate prediction of pain and disability outcomes.

Related to our cluster analysis findings, the lack of specificity among psychological measures to generate patterns other than high or low pain-associated psychological distress and maladaptive coping potentially suggests that initial screening for specific psychological domains through the use of multiple measures may not be necessary for all patients. Rather, first-line psychological risk factor screening with a single multidimensional measure may be used to identify patients requiring additional psychological assessment. There is preliminary evidence that the ability to identify greater than 2 psychological subgroups may be dependent on the type of psychological distress assessed.[2,5,6,14,62] For example, Beneciuk et al[2] identified 3 different subgroups by incorporating a measure of patient-specific pain-related fear in addition to measures of general pain-related fear and catastrophizing. Beneciuk et al[2] suggested that specific fear should then be accounted for separately with individuals experiencing LBP and that multidimensional measures (eg, the STarT Back Tool) may not account for a potentially important patient subgroup because they do not include items associated with specific fears. Alternatively, there is support that the cumulative relationship among different elevated psychological factors may have an additive adverse effect on patient prognosis and clinical outcomes.[3,42,64] For example, Bergbom et al[3] identified 4 subgroups with varying psychological profiles based on pain catastrophizing and depressive symptom scores, where patients categorized as high risk were associated with the greatest levels of self-report disability and pain intensity when compared to other subgroups. Collectively, the above-referenced studies provide insight to the ongoing debate regarding identification of the appropriate number of subgroups when screening for psychological distress and how that information can be used for prognostic purposes. Furthermore, the number of subgroups identified may be not only related to the clinical setting but also dependent upon the psychological measures that are administered. Future studies that incorporate consistent measures will help to resolve the issue of determining an appropriate number of psychological subgroups in clinical settings.

Findings from this cross-sectional study add to the literature involving the SBT as a screening measure for LBP-associated psychological distress; however, this study does not provide information related to treatment. Results from a recent clinical trial in primary care settings indicated that initial treatment decisions based on SBT categorization on 3 levels (ie, SBT low, medium, and high risk) were associated with greater improvements in clinical outcomes and cost savings when compared to usual care for patients with LBP.[28] Specifically, future studies should evaluate the SBT to guide risk-stratified treatment in secondary care settings and determine if use of 2 categories is potentially more appropriate for decision making and providing effective treatment. Future studies should also consider investigating the SBT in work-related injury settings to determine how risk stratification may differ for those patient samples.

Several limitations should be considered when interpreting the results of this study. First, because the results were based on cluster analysis, we employed various "stopping rules" combined with practical judgment and theoretical foundations in determining our final number of cluster solutions.[23] Nevertheless, in our opinion, this subgrouping methodology should also be considered a strength of this study as our empirically derived subgroups reflected the underlying distribution of unidimensional psychological and disability measure scores and were not based on arbitrary cutoff scores. Future studies should consider establishing an a priori optimal number of cluster solutions with an empirical basis, and then determine if agglomeration coefficients and plot characteristics confirm or refute their hypothesis (eg, confirmatory cluster analyses). Second, we acknowledge that our generated subgroups were based on the measures used in this study that were primarily aligned with the fear-avoidance model of musculoskeletal pain[41] and that including other potentially important psychological factors (eg, self-efficacy) may have influenced our findings. Third, although we cannot confirm that patients in our study sample were or were not prescreened by primary care physicians before being referred for physical therapy, patients in this study sample were associated with increased severity evidenced by higher mean pain intensity ratings (5.3 vs 3.2 points) and self-reported disability scores (OSW scores: 16.2 points vs Roland-Morris Disability Questionnaire scores: 6.7 points) when compared to primary care patients enrolled in the SBT development cohort.[25] Finally, this study did not incorporate any predictive analyses with the empirically derived subgroups as that was beyond the scope of our study. Previous study findings[1] suggest that the SBT may be valuable as a prognostic indicator for self-reported disability outcomes at 6 months compared with unidimensional psychological measures, thereby providing evidence for the SBT's predictive validity in physical therapy settings.

Conclusion

The SBT provides a viable option for use as an initial screening measure to identify elevated levels of pain-associated psychological distress and maladaptive coping as it adequately distinguishes among commonly used single-construct, unidimensional measures. Our findings do provide suggestions for future research related to evaluation of a modified risk stratification scheme using different cutoff scores to identify low and high pain-associated psychologically distressed and maladaptive coping LBP patients referred for treatment in physical therapy and other secondary care settings.

Acknowledgments

This study was conducted as part of J.M.B.'s doctoral dissertation; therefore, he would like to acknowledge other dissertation committee members (ie, Mark D. Bishop, PT, PhD; Julie M. Fritz, PT, PhD; Nabih R. Asal, PhD, FACE). We would also like to acknowledge the following clinicians and personnel from Brooks Rehabilitation in Jacksonville, Florida (M. Brian Hagist, Tim Shreve, Jason Kral, Matthew Stafford, Ryan Reed, Michael Spigel, Holly Morris, Flo Singletary, Amanda Osborne, and Robert Rowe) and the University of Florida Orthopaedics and Sports Medicine Institute and Shands Rehabilitation Center at Magnolia Parke in Gainesville, Florida (Giorgio Zeppieri, Josh Barabas, Debi Jones, Derek Miles, Zack Sutton, Dalton Reed, Michael Hodges, Shannon Long, Tim Shay, Amy Borut, and Yvette Silvey). Dr. Robinson's lab (Anne Nisenzon, Calia Torres, and Laura Wandner) assisted with data collection and management.

References

1. Beneciuk JM, Bishop MD, Fritz JM, Robinson ME, Asal NR, Nisenzon AN, George SZ: The STarTBack Screening Tool and individual psychological measures: Evaluation of prognostic capabilities for low back pain clinical outcomes in outpatient physical therapy settings. Phys Ther 93: 321-333, 2013

2. Beneciuk JM, Robinson ME, George SZ: Low back pain subgroups using fear-avoidance model measures: Results of a cluster analysis. Clin J Pain 28:658-666, 2012

3. Bergbom S, Boersma K, Overmeer T, Linton SJ: Relationship among pain catastrophizing, depressed mood, and outcomes across physical therapy treatments. Phys Ther 91: 754-764, 2011

4. Bieling PJ, Antony MM, Swinson RP: The State-Trait Anxiety Inventory, Trait version: Structure and content reexamined. Behav Res Ther 36:777-788, 1998

5. Boersma K, Linton SJ: Screening to identify patients at risk: Profiles of psychological risk factors for early intervention. Clin J Pain 21:38-43, 2005. discussion 69-72

6. Boersma K, Linton SJ: Psychological processes underlying the development of a chronic pain problem: A prospective study of the relationship between profiles of psychological variables in the fear-avoidance model and disability. Clin J Pain 22:160-166, 2006

7. Bolton JE: Accuracy of recall of usual pain intensity in back pain patients. Pain 83:533-539, 1999

8. Childs JD, Piva SR, Fritz JM: Responsiveness of the numeric pain rating scale in patients with low back pain. Spine 30: 1331-1334, 2005

9. Chou R, Qaseem A, Snow V, Casey D, Cross JT Jr, Shekelle P, Owens DK, Clinical Efficacy Assessment Subcommittee of the American College of Physicians, American College of Physicians, American Pain Society Low Back Pain Guidelines Panel: Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 147: 478-491, 2007

10. Chou R, Shekelle P: Will this patient develop persistent disabling low back pain? J Am Med Assoc 303:1295-1302, 2010

11. Crombez G, Eccleston C, Baeyens F, Eelen P: When somatic information threatens, catastrophic thinking enhances attentional interference. Pain 75:187-198, 1998

12. Crombez G, Vlaeyen JW, Heuts PH, Lysens R: Pain-related fear is more disabling than pain itself: Evidence on the role of pain-related fear in chronic back pain disability. Pain 80: 329-339, 1999

13. Damush TM, Wu J, Bair MJ, Sutherland JM, Kroenke K: Self-management practices among primary care patients with musculoskeletal pain and depression. J Behav Med 31:301-307, 2008

14. Denison E, Asenlof P, Sandborgh M, Lindberg P: Musculoskeletal pain in primary health care: Subgroups based on pain intensity, disability, self-efficacy, and fear-avoidance variables. J Pain 8:67-74, 2007

15. Fairbank JC, Pynsent PB: The Oswestry Disability Index. Spine (Phila Pa 1976) 25:2940-2952, 2000. discussion 2952

16. Foster NE, Dziedzic KS, van der Windt DA, Fritz JM, Hay EM: Research priorities for non-pharmacological therapies for common musculoskeletal problems: Nationally and internationally agreed recommendations. BMC Musculoskelet Disord 10:3, 2009

17. Foster NE, Hill JC, O'Sullivan P, Hancock M: Stratified models of care. Best Pract Res Clin Rheumatol 27:649-661, 2013

18. Fritz JM, George SZ: Identifying psychosocial variables in patients with acute work-related low back pain: The importance of fear-avoidance beliefs. Phys Ther 82:973-983, 2002

19. Fritz JM, George SZ, Delitto A: The role of fear-avoidance beliefs in acute low back pain: Relationships with current and future disability and work status. Pain 94: 7-15, 2001

20. Fritz JM, Irrgang JJ: A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. Phys Ther 81:776-788, 2001

21. George SZ, Fritz JM, Bialosky JE, Donald DA: The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: Results of a randomized clinical trial. Spine (Phila Pa 1976) 28:2551-2560, 2003

22. Goldstein MS, Scalzitti DA, Craik RL, Dunn SL, Irion JM, Irrgang J, Kolobe TH, McDonough CM, Shields RK: The revised research agenda for physical therapy. Phys Ther 91: 165-174, 2011

23. Hair JE, Anderson RE, Tatham RL, Black WC: Cluster analysis, in Boyd J (ed): Multivariate Data Analysis. Upper Saddle River, NJ, Prentice Hall, 1998, pp 469-515

24. Hay EM, Dunn KM, Hill JC, Lewis M, Mason EE, Konstantinou K, Sowden G, Somerville S, Vohora K, Whitehurst D, Main CJ: A randomised clinical trial of subgrouping and targeted treatment for low back pain compared with best current care. The STarT Back Trial Study Protocol. BMC Musculoskelet Disord 9:58, 2008

25. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM: A primary care back pain screening tool: Identifying patient subgroups for initial treatment. Arthritis Rheum 59: 632-641, 2008

26. Hill JC, Dunn KM, Main CJ, Hay EM: Subgrouping low back pain: A comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. Eur J Pain 14:83-89, 2010

27. Hill JC, Fritz JM: Psychosocial influences on low back pain, disability, and response to treatment. Phys Ther 91: 712-721, 2011

28. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, Konstantinou K, Main CJ, Mason E, Somerville S, Sowden G, Vohora K, Hay EM: Comparison of stratified primary care management for low back pain with current best practice (STarT Back): A randomised controlled trial. Lancet 378:1560-1571, 2011

29. Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL: Using the Patient Health Questionnaire–9 to measure depression among racially and ethnically diverse primary care patients. J Gen Intern Med 21: 547-552, 2006

30. Hudson-Cook N, Tomes-Nicholson K, Breen A: A revised Oswestry disability questionnaire, in Roland MO, Jenner JR (eds): Back Pain: New Approaches to Rehabilitation and Education. New York, NY, Manchester University Press, 1989, pp 187-204

31. Jacob T, Baras M, Zeev A, Epstein L: Low back pain: Reliability of a set of pain measurement tools. Arch Phys Med Rehabil 82:735-742, 2001

32. Jellema P, van der Windt DA, van der Horst HE, Blankenstein AH, Bouter LM, Stalman WA: Why is a treatment aimed at psychosocial factors not effective in patients with (sub)acute low back pain? Pain 118:350-359, 2005

33. Jensen MP, Turner JA, Romano JM, Fisher LD: Comparative reliability and validity of chronic pain intensity measures. Pain 83:157-162, 1999

34. Jensen MP, Turner LR, Turner JA, Romano JM: The use of multiple-item scales for pain intensity measurement in chronic pain patients. Pain 67:35-40, 1996

35. Kabacoff RI, Segal DL, Hersen M, Van Hasselt VB: Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. J Anxiety Disord 11:33-47, 1997

36. Kent PM, Keating JL, Taylor NF: Primary care clinicians use variable methods to assess acute nonspecific low back pain and usually focus on impairments. Man Ther 14: 88-100, 2009

37. Ketchen DJ, Shook CL: The application of cluster analysis in strategic management research: An analysis and critique. Strategic Manage J 17:441-458, 1996

38. Koes BW, van Tulder MW, Thomas S: Diagnosis and treatment of low back pain. BMJ 332:1430-1434, 2006

39. Kongsted A, Johannesen E, Leboeuf-Yde C: Feasibility of the STarT Back Screening Tool in chiropractic clinics: A cross-sectional study of patients with low back pain. Chiropr Man Therap 19:10, 2011

40. Kroenke K, Spitzer RL, Williams JB: The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16: 606-613, 2001

41. Leeuw M, Goossens ME, Linton SJ, Crombez G, Boersma K, Vlaeyen JW: The fear-avoidance model of musculoskeletal pain: Current state of scientific evidence. J Behav Med 30:77-94, 2007

42. Linton SJ, Nicholas MK, MacDonald S, Boersma K, Bergbom S, Maher C, Refshauge K: The role of depression and catastrophizing in musculoskeletal pain. Eur J Pain 15: 416-422, 2011

43. Main CJ, George SZ: Psychologically informed practice for management of low back pain: Future directions in practice and research. Phys Ther 91:820-824, 2011

44. Meyers LW, Gamst G, Guarino AJ: Applied Multivariate Research: Design and Interpretation. Thousand Oaks, CA, Sage Publications, 2006

45. Milligan GA, Cooper MC: An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159-179, 1985

46. Morso L, Kent P, Albert HB, Manniche C: Is the psychosocial profile of people with low back pain seeking care in Danish primary care different from those in secondary care? Man Ther 18:54-59, 2013

47. Nicholas MK, Linton SJ, Watson PJ, Main CJ, "Decade of the Flags" Working Group: Early identification and management of psychological risk factors ("yellow flags") in patients with low back pain: A reappraisal. Phys Ther 91:737-753, 2011

48. Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L: The Pain Catastrophizing Scale: Further psychometric evaluation with adult samples. J Behav Med 23:351-365, 2000

49. Pincus T, Smeets RJ, Simmonds MJ, Sullivan MJ: The fear avoidance model disentangled: improving the clinical utility of the fear avoidance model. Clin J Pain 26:739-746, 2010

50. Pransky G, Borkan JM, Young AE, Cherkin DC: Are we making progress? The Tenth International Forum for Primary Care Research on Low Back Pain. Spine (Phila Pa 1976) 36:1608-1614, 2011

51. Robinson ME, George SZ: Screening for problematic low back pain: STarT. Pain 153:2159-2160, 2012

52. Roland M, Fairbank J: The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine (Phila Pa 1976) 25:3115-3124, 2000

53. Roland M, Morris R: A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976) 8:141-144, 1983

54. Spielberger CD, Gorsuch RL, Lushene RE, Vagg PR, Jacobs GA: Manual for the State and Trait Anxiety Inventory (Form Y). Palo Alto, CA, Consulting Psychologists Press, 1983

55. Staerkle R, Mannion AF, Elfering A, Junge A, Semmer NK, Jacobshagen N, Grob D, Dvorak J, Boos N: Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in a Swiss-German sample of low back pain patients. Eur Spine J 13:332-340, 2004

56. Sullivan M, Bishop S, Pivik J: The Pain Catastrophizing Scale: Development and validation. Psychol Assess 7:524-532, 1995

57. Swinkels-Meewisse IE, Roelofs J, Schouten EG, Verbeek AL, Oostendorp RA, Vlaeyen JW: Fear of movement/(re)injury predicting chronic disabling low back pain: A prospective inception cohort study. Spine (Phila Pa 1976) 31:658-664, 2006

58. Swinkels-Meewisse IE, Roelofs J, Verbeek AL, Oostendorp RA, Vlaeyen JW: Fear-avoidance beliefs,

disability, and participation in workers and non-workers with acute low back pain. Clin J Pain 22:45-54, 2006

59. Tabachnick BG, Fidell LS: Using Multivariate Statistics, 5th ed. Boston, MA, Pearson/Allyn and Bacon, 2007

60. van der Windt D, Hay E, Jellema P, Main C: Psychosocial interventions for low back pain in primary care: Lessons learned from recent trials. Spine (Phila Pa 1976) 33:81-89, 2008

61. Waddell G, Newton M, Henderson I, Somerville D, Main CJ: A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 52:157-168, 1993

62. Westman AE, Boersma K, Leppert J, Linton SJ: Fear-avoidance beliefs, catastrophizing, and distress: A longitudinal subgroup analysis on patients with musculoskeletal pain. Clin J Pain 27:567-577, 2011

63. 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 153:2182-2191, 2012

64. Wideman TH, Sullivan MJ: Development of a cumulative psychosocial factor index for problematic recovery following work-related musculoskeletal injuries. Phys Ther 92:58-68, 2012

65. Woby SR, Roach NK, Urmston M, Watson PJ: Psychometric properties of the TSK-11: A shortened version of the Tampa Scale for Kinesiophobia. Pain 117:137-144, 2005