Validity of the Yellow Flag Risk Form in People Treated for Low Back Pain with Mechanical Diagnosis and Therapy and the Pain Mechanism Classification System

Ronald Schenk[1*], Joseph Lorenzetti[2], Michael Ross[3], Ian Kulac[4], Kourtney Cox[3], Emily Gartenman[3], Anna Ruth Gsellman[3], Dylan Lodowski[3], Kristin Barnard[1], Joseph Baumgarden[1], Thomas Coleman[1], Michael Geraci[1], David May[1], Amy Murdock[1], Melissa Ramsay[1], Edyta Sullivan[1], Colleen Kashino[3], Joshua Cleland[2]

[1]Tufts University School of Medicine Doctor of Physical Therapy Program, USA

[2]Catholic Health System of Buffalo, USA

[3]Daemen College Department of Physical Therapy, USA

[4]Brigham Women's Hospital, USA

[*]**Corresponding author:** Ronald J Schenk, Tufts University School of Medicine, Doctor of Physical Therapy Program. 136 Harrison Ave, Boston, MA 02111, USA

Abstract

Background: Psychosocial aspects of pain are often associated with chronic low back pain, a condition for which the specific etiology is unknown. Psychosocial risk tools, such as the Yellow Flag Risk Form (YFRF) have been used to identify these factors and sub classify participants into clinically relevant subgroups, which are aligned with a specific intervention. The purpose of this research was to analyze patient outcomes in people with low back pain referred to physical therapists who utilize the YFRF Mechanical Diagnosis and Therapy (MDT), and Pain Mechanism Classification (PMCS) principles. **Methods:** One hundred seventy-nine people with Low Back Pain (LBP) were referred to a hospital-based physical therapy outpatient clinic in western New York State. Of the 179 participants, 26 met the exclusion criteria and 13 had incomplete data, resulting in an analysis of 140 participants. The participants were examined and classified based on MDT and the PMCS classifications by physical therapists trained in both systems. Participants were administered the YFRF, the Numerical Pain Rating Scale (NPRS), and the Focus on Therapeutic Outcomes (FOTO) tools at initial evaluation, at the 4th visit, and at discharge. **Results:** Of the 140 participants, 65% were experiencing chronic duration of symptoms and 60.7% of the sample scored greater than or equal to 50 on the YFRF. Among these participants, 92/140 (65.7% of the sample were classified as responders and 48/140 (34.3%) were classified as non-responders based on a statistically significant change score on either FOTO or the NPRS. A regression analysis of YFRF findings and outcome indicated that the model performed well in classifying patients as responders or non-responders. **Conclusion:** This research suggests that a high percentage of participants receiving musculoskeletal care may have symptom chronicity and psychosocial risk and still respond to physical therapy intervention. Further, MDT may be an effective musculoskeletal approach for participants classified with a derangement irrespective of psychosocial risk factors as indicated by change in YFRF score as a predictor of patient outcome.

Introduction

Chronic low back pain affects up to 80% of individuals within their lifetime, with current evidence suggesting that approximately 126.1 million adults in the United States experienced pain over the last 3 months, and approximately 25.3 million adults are suffering from chronic pain [1]. Chronic low back pain is a condition which may be described as having a duration of greater than 7 weeks, but which may also involve psychosocial constructs such as central sensitization [2,3]. Central sensitization has been recognized as a potential neurophysiological mechanism underlying a group of chronic pain conditions and may represent a sub-group of participants with chronic low back pain [3].

Psychosocial risk factors include depression, anxiety, fear avoidance, and pain catastrophizing. Tools assessing psychosocial risk can identify these factors and the condition may be further sub classified using the Pain Mechanism Classification System (PMCS). This method of sub classification showed preliminary evidence for effectively guiding treatment [4]. In the Kolski et al study [4], physical therapists trained in the PMCS were able to accurately classify participants into inflammatory, ischemic, peripheral neurogenic, central, and other categories when compared to classifications generated by a statistical model with cluster analysis based on patient signs and symptoms. The Yellow Flag Risk Form (YFRF) is a tool used in the PMCS to sub classify participants into intervention driven subgroups by categorizing participants based on the individual's psychological construct [4].

The concept of physical therapists classifying participants with musculoskeletal disorders was found to be efficacious in a study by Brennan et al. [5] and a systematic review of the literature by Cook et al [6] found that therapist directed exercise programs classified by patient response may produce favorable outcomes. Similarly, identification of psychosocial risk factors in participants with chronic pain may be critical for optimal recovery.

Various questionnaires are used to identify psychosocial risk factors in chronic pain participants including the previously validated STarT Back Tool (SBT) [7] and the Optimal Screening for Prediction of Referral and Outcome (OSPRO) [8]. The SBT can identify a patient's level of psychosocial risk by examining psychosocial and physical factors and categorizing a patient as high risk, medium risk, or low risk [7]. Although the SBT identifies the presence of psychosocial risk factors in participants with chronic pain, it does not classify participants into intervention driven subgroups. The OSPRO categorizes participants based on their "individual psychological construct", as described by Lentz et al [8] who described the need for a tool to allow therapists to concurrently measure a patient's level of psychosocial risk while determining their dominating pain mechanism. Moreover, the OSPRO specifically assesses the pain mechanism domains of fear avoidance, emotional coping, self-efficacy, sleep and presence of nerve related symptoms, pain intensity, and function [8] with functional as well as symptomatic improvement being the goal of therapeutic intervention.

A commonly used functional outcome measure is Focus on Therapeutic Outcomes (FOTO), a rehabilitation medical data management system that standardizes specific body part initial and discharge measurements [9]. Werneke, et al determined that there was a positive association between administering FOTO within the first 2 weeks of patient intake and patient outcomes when compared to participants who did not complete a FOTO assessment at any point during their episode of care. The authors hypothesized that administration of FOTO during the initial treatment sessions allows clinicians to adjust their management strategies, leading to more accurate treatment decisions and improved patient outcomes [9].

The SBT and the YFRF classify participants based on their level of psychosocial risk, which allows clinicians to direct treatments which vary from manual procedures to patient driven exercise programs such as those in the McKenzie Method of Mechanical Diagnosis and Therapy (MDT) system. The MDT approach uses repeated end range movements to classify participants into the categories of derangement, dysfunction, posture, or other. In a systematic review of participants with chronic low back pain, several pain measures showed that the McKenzie method may be efficacious in decreasing pain in the short term, while the disability measures determined that the McKenzie method is better in enhancing long-term function [10].

The MDT system of examination and intervention was also found to have an association with improving fear-avoidance beliefs, pain self-efficacy, depression and psychological distress in a systematic review by Kunhow, et al [11]. A primary objective of MDT is to determine a Directional Preference (DP), which is defined as the preferred direction of movement for the patient based on their response to repeated end range movements [12] and in that respect may empower the patient to self-manage the condition. A study by Werneke, et al [13] also provided preliminary evidence to support the use of MDT and the PMCS in physical therapy to screen for and manage chronic neuromusculoskeletal pain based on classification systems.

The primary purpose of this study was to analyze the ability of the YFRF to predict change in the NPRS and FOTO at the 4th visit and discharge.

Methods

Data were collected using the YFRF, FOTO, and the NPRS. Patient informed consent was required and obtained by the Catholic Health System of Buffalo Institutional Review Board and the Daemen College Human Subjects Research Review Committee. All eligible participants completed an informed consent form and agreed to have their de-identified data analyzed by the researchers. The plan for analysis was to determine if changes in psychosocial status was predictive of change in outcome as measured by the Focus on Therapeutic Outcomes (FOTO) and Numerical Pain Rating Scale (NPRS).

Participants

A total of 179 consecutive patients with LBP were admitted into the study via direct access or medical referral to a hospital-based outpatient physical therapy clinic in western New York State. Potential participants completed an intake form that screened for medical and surgical history and then completed an informed consent. Twenty-six participants were excluded due to meeting criteria that included red flag pathology or conditions such as cauda equina syndrome, spinal fracture, systemic infection, cancer, spondylolisthesis, spinal cord injury or joint laxity and were appropriately referred. Included, as participants were 153 patients who were from 18 to 85 years of age and who were willing to participate in physical therapy sessions. Following completion of the intake form and informed consent was a current history and physical examination conducted by a physical therapist trained in MDT and the PMCS. The physical examination included the testing of repeated end range spinal movements, which allowed for MDT classification into derangement or "other" categories. From the initial 179 participants 26 met the exclusion criteria and 13 were excluded from the analysis due to missing data (Figure 1). To be considered a "Responder" to intervention, the participants needed to demonstrate a favorable change in the FOTO of at least 8 and a favorable change in the NPRS of at least 4 at discharge. A favorable change of 8 in the FOTO meets the MCID of 8 for that functional measure and a favorable change of 4 in the NPRS is twice the MCID for that pain measure.
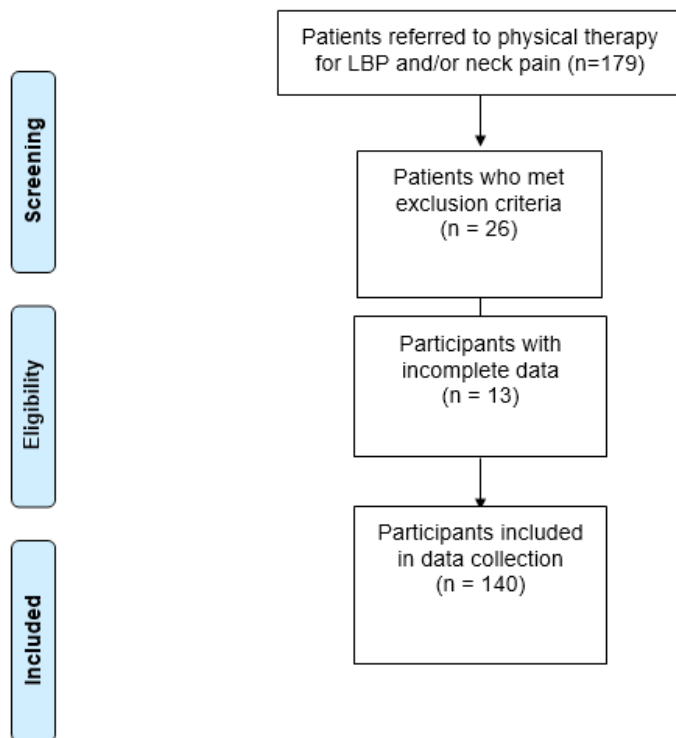

Figure 1: Participants included in the study.

Procedures

Participants received intervention directed by the examining physical therapist. These interventions included management according to MDT and PMCS principles, which included but were not limited to pain neuroscience education, particularly if the patient is found to be at moderate or high psychosocial risk. The MDT method involves testing of repeated end range movements to determine if a directional preference can be determined. If a DP is found, the patient is giving exercises consistent with their DP and may also receive manual procedures if the patient's symptoms are status quo or improving. Based on results at discharge on the FOTO and NPRS, participants were placed into the categories of Responders and Non-Responders. Ninety-two percent of the participants were categorized as Responders and 48 participants were categorized as Non-Responders. Of the patients described as Responders, 91.2% of were classified as derangement classification according to MDT. This is consistent with percentages classified as derangement according to Werneke [13]. The derangement classification is characterized by a directional preference, which is a preferred direction of movement and exercise based on the testing of repeated end range movements. The non-responders comprised 23/140 of the sample, or 16.4%, and were those who did not meet a change of at least 8 in the FOTO and at least 4 in the NPRS.

Examiners

Six physical therapists who were certified in MDT and trained in the PMCS performed examination and interventions for all participants who met the inclusion criteria and were deemed appropriate for management.

Analyses

All analyses were conducted using IBM SPSS statistical software version 27.0. Data were collected by clinicians in a hospital-based outpatient physical therapy clinic in western New York State in which all subjects were being seen for cervical or lower back pain.

The data set was reduced from 179 observations to 140 due to the exclusion criteria or missing data. A FOTO assessment score the YFRF and the NPRS were recorded at patient intake, fourth visit and discharge. If the patient was found to not be improving in either the YFRF, FOTO, or NPRS, they were referred to their physician. Other variables were recorded by clinicians included chronicity, initial classification, number of visits, and the duration of treatment days. Since the objective of the study was to determine if the YFRF is a predictor of patient outcome, the difference between the initial visit and the discharge visit was used for analysis and the change in YFRF was recorded.

Patient outcome was determined by the Minimal Clinically Important Difference (MCID). For FOTO the MCID is a change is at least eight points. A change of at least two points is needed to meet the MCID for the NPRS. Therefore, to be classified as a responder in this study, the patient needed to have a change in a

FOTO score that was greater than or equal to eight and a change in their NPRS score of at least four. These criteria were selected for patient outcome due the fact 85 of the 140 patients or 60.7% of the patients had an initial YFRF score of at least 50. This implies that about greater than 60% of patients were at minimum at moderate risk for psychosocial distress. Also, more than 53% of patients had an Initial NPRS of 5 or more.

Descriptive analyses and representativeness of sample

The patient demographics and classifications of the study participants is shown in Table 1 and the descriptive statistics are found in Table 2. Patient demographic baselines were summarized using means or medians, standard deviations, and ranges. Outcomes were compared between participants based on initial, interim, and discharge scores on FOTO, the YFRF, and the NPRS. Outcomes were analyzed using a logistic regression in order to determine change scores and statistical significance. Outcome measure data are demonstrated in Table 3.

Variable	N
Classification	
Chronicity (QTF):	
Total	140
Acute	30
Subacute	19
Chronic	91
Chronic %	65
Responders/Non responders	
Responders	92
Responder %	92/140 (65.7%)
Non-responder	48
Non-responder %	48/140 (34.3%)
Psychosocial	
YFRF score > 50	85/140
Moderate psychosocial risk %	60.7
MDT	
Derangement Classification	83
MDT Derangement %	83/140 (59.2%)
Other classification	57/140
Other classification %	40.7%
Sex	
Males	73
Females	67
Age	Mean 47
Duration of treatment (days)	47.49±32.27
Number of visits	8.35±6.13

Table 1: Patient Demographics.

Outcome Measure	Intake Score (Mean)	Discharge Score (Mean)	Change in score (Discharge – Initial)	p-value
FOTO	49.48	65.71	16.23	<0.001
YFRF	63.24	34.62	-28.62	<0.001
NPRS	4.55	1.22	-3.33	<0.001

Table 2: Descriptive Statistics.

	Intake YFRF Score (Mean)	**Discharge YFRF Score (Mean)**	**Change in score (Discharge – Initial)**	**p-value**
Responders	69.60	30.03	-39.57	<0.001
Non-responders	58.04	34.62	-23.42	<0.001

Table 3: Outcome Measure Data.

Results

Seventy-six percent of the participants were found to have chronic LBP based on chronicity (symptoms greater than 7 weeks duration [2], and this group comprised the majority of the study population (Table 2). Categorization also included numbers of participants with LE symptoms, average numbers of visits, average duration of treatment (Table 1). The average number of visits was 7.5 and the average duration of treatment was 44.6 days, despite the majority of these participants having chronic pain consistent with the Quebec Task Force (QTF) determination of chronicity [2].

A logistic regression was performed to determine the predictive value of the Yellow Flag Risk form as it related to patient outcome. The covariates of chronicity, number of visits, number of treatment days, and the change in the YFRF were included in the model as well as the two-way interactions between them. In the model, the number of responders was 92 of 140 patients 65.7%. Table 4 displays the model summary for the model. The change in YFRF covariate was significant with a p-value equal to 1e-07. Based upon the model output a one-point increase in the change of the YFRF increases the odds of a patient being a responder by $e^{0.06953}$% which is approximately 7.2%. These results are represented in Table 4 (Model Summary).

Model Summary				
Term	**Estimate**	**Std.error**	**Statistic**	**p.value**
Intercept	-2.17	0.4187	-5.19	2e-07
YFRF change	0.0624	0.0115771	5.389044	1e-07

Table 4: Model Summary.

The Receiver Operating Characteristic Curve (ROC curve) is frequently used as a visual representation for model performance in a classification model. The ROC curve plots the sensitivity versus (1-specificity) over all possible thresholds. Sensitivity was operationally defined in this study as the proportion of patients who are responders and are correctly classified by the model as a responder or equivalently the true positive rate. The proportion of patients who are identified by the model as non-responders and are in fact non-responders is known as the specificity. Therefore, (1-specificity) is the false positive rate. A threshold is a cut-off value for the predicted probably in which a patient is deemed a responder or non-responder. For example, if a threshold is set at 0.5 then, all subjects who have a predicted probability greater than 0.5 would be classified as responders and the others as non-responders. The ROC curve sequentially steps through a fine sequence of threshold values from zero to one. The predicted probabilities for the logistic model are shown in Figure 2 and the area under the curve (AUC) is a measure for model performance is shown in Figure 3 and demonstrates that the model outperforms chance.
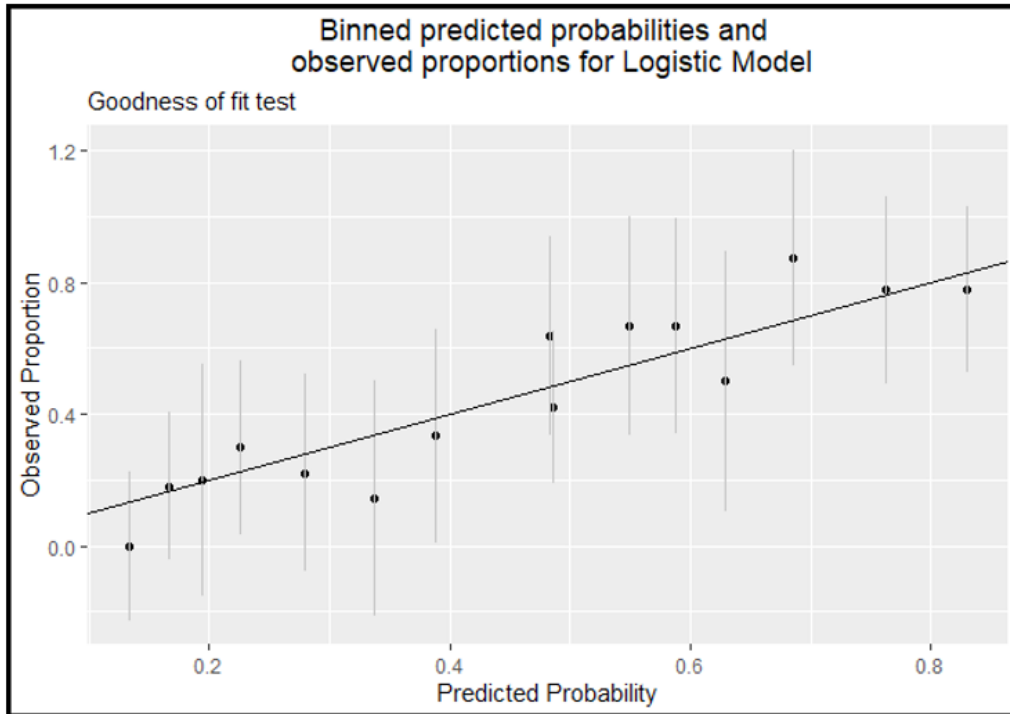
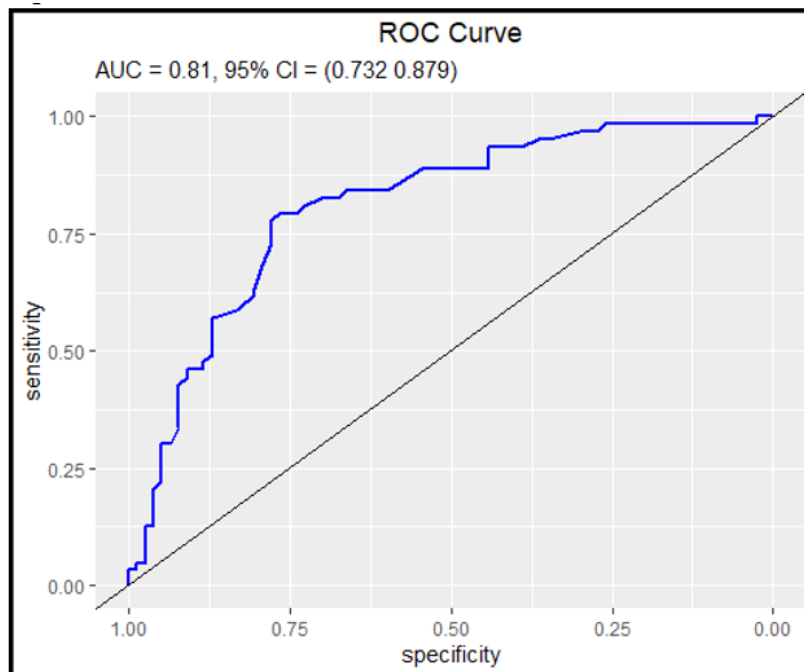**Figure 2:** Predicted Probabilities for the Logistic Model.

Figure 3: Area Under the Curve.

Discussion

The study suggests the validity of the YFRF in predicting a responder to MDT intervention in the management of patients with psychosocial risk factors. The majority of participants presenting with high scores on the YFRF along with the MDT classification of a derangement, and chronicity according the QTF [2], responded positively to treatment. Werneke et al [9] found that using MDT's classification of directional preference vs no-directional preference, there was a significant and clinically important difference in the change in functional status at discharge between participants with directional preference and the participants classified as having no directional preference. The present study also concluded that participants who had a derangement may have a directional preference and thus respond positively to intervention. Although our study did not directly use the SBT, we referenced studies that included this tool to format the design for our study. In a study by Fritz et al [7], SBT risk categories were associated closely with participants' progress; additionally, the SBT was able to quantify risk and predicted improvement. In a similar way, the YFRF was used in our study to identify participants who presented with psychosocial risk factors as well as categorize these participants into subcategories. The face validity of the YFRF is supported through its use at the Rehabilitation Institute of Chicago where the tool was administered to over 18,000 participants and was integral to a validation study of the PMCS [4]. In that study, physical therapists had good agreement with cluster analysis for categories of inflammatory pain, ischemic pain, peripheral pain and other. The YFRF might be useful in the group of participants who were non-responders. All participants scoring at least 50 on the YFRF were managed according to pain neuroscience education strategies utilized in the Kolski et al study validating the PMCS. These strategies included principles such as motivational interviewing, movement safe pain, graded exposure, and the traffic light guide. Although MDT has been shown to be effective in certain subgroups, particularly the derangement classification, it focuses on teaching patients to self-manage their condition, thereby becoming empowered and more independent from the clinician. Kolski et al's study found that trained therapists were able to classify participants using the PMCS into classifications.

Limitations

Limitations in this study include the use of only one healthcare system, which reduces the external validity of the study. A second limitation was that the treatment method was inclusive of only one treatment approach, the MDT method, whereas physical therapy practice typically includes more than one treatment. A third limitation to this study was a control group was not utilized in the study design, thus limiting the internal validity. Additionally, there was limited data from non-responders (n=12) which causes their data to be less accurate when compared to the large number of responders included in our study.

Conclusion

This research is consistent with other research that demonstrated the importance of classification and sub classification [2,5,6,13]. Utilization of the PMCS and MDT classification by trained clinicians allowed for optimal treatment within each sub classification leading to favorable outcomes. A high percentage of participants receiving musculoskeletal care has chronicity of symptoms in terms of duration but may still respond to mechanical treatment. Chronicity should also include consideration of psychosocial risk factors and in this investigation, MDT was found to be an effective treatment for the participants classified with a derangement irrespective of psychosocial risk. Further research should re-examine the term responder to include changes that meet the MCID in both the FOTO and NPRS. The regression model had a minimum change in FOTO score of 8 which is equal to the MCID. However, the minimum requirement for the NPRS was 4 which is twice the MCID. These criteria were selected for patient outcome due the fact 85 of the 140 patients or 60.7% of the patients had an initial YFRF score of at least 50 indicating minimum to moderate risk for psychosocial distress. Also, more than 53% of patients had an initial NPRS of 5 or more which further lends credence to the model. Based on the results of this study, the YFRF may be predictive out outcome in physical therapy patients experiencing chronic pain who are treated with MDT.

References

1. Louw A, Nijs J, Puentedura E (2017) A clinical perspective on a pain neuroscience education approach to manual therapy. J Man Manip Ther 25: 160-168.

2. Atlas S, Deyo R, Patrick D, Convery K, Keller R, et al. (1996) The Quebec task force classification for spinal disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. Spine (Phila Pa 1976) 21: 2885-2892.

3. Sanzarello I, Merlini L, Rosa MA, Perrone M, Frugiuele J, et al. (2016) Central sensitization in chronic low back pain: A narrative review. J Back Musculoskelet Rehabil 9: 625-633.

4. Kolski MC, O'Connor A, Van Der Laan K, Lee J, Kozlowski AJ, et al. (2016) Validation of a pain mechanism classification system (PMCS) in physical therapy practice. J Man Manip Ther 24: 192-199.

5. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, et al. (2006) Identifying subgroups of participants with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. Spine (Phila Pa 1976) 31: 623-631.

6. Cook C, Hegedus E, Ramey K (2005) Physical therapy exercise intervention based on classification using the patient response method: a systematic review of the literature. J Man Manip Ther 13: 152-162.

7. Fritz J, Beneciuk J, George S (2011) Relationship between categorization with the STarT back screening tool and prognosis for people receiving physical therapy for low back pain. Phys Ther 91: 722-732.

8. Lentz TA, Beneciuk JM, Bialosky JE, Zeppieri Jr G, Dai Y, et al. (2016) Development of a yellow flag assessment tool for orthopaedic physical therapists: results from the optimal screening for prediction of referral and outcome (OSPRO) cohort. J Orthop Sports Phys Ther 46: 327-343.

9. Werneke M, Deutscher D, Fritz J, Kallen MA, Cook KF, et al. (2020) Associations between interim patient-reported outcome measures and functional status at discharge from rehabilitation for non-specific lumbar impairments. Qual Life Res 29: 439-451.

10. Namnaqani FI, Mashabi AS, Yaseen KM, Alshehri MA (2019) The effectiveness of McKenzie method compared to manual therapy for treating chronic low back pain: a systematic review. J Musculoskelet Neuronal Interact 19: 492-499.

11. Kuhnow A, Kuhnow J, Ham D, Rosedale R (2020) The McKenzie Method and its association with psychosocial outcomes in low back pain: a systematic review. Physiother Theory Pract 37: 1283-1297.

12. Long A, Donelson R, Fung T (2004) Does it matter which exercise? a randomized control trial of exercise for low back pain. Spine 29: 2593-2602.

13. Werneke MW, Edmond S, Deutscher D, Ward J, Grigsby D, et al. (2016) Effect of adding McKenzie syndrome, centralization, directional preference, and psychosocial classification variables to a risk-adjusted model predicting functional status outcomes for participants with lumbar impairments. J Ortho Sports Phys Ther 46: 726-741.