TITLE: Adding Physical Impairment to Risk Stratification Improved Outcome Prediction in Low Back Pain

RUNNING HEAD: Physical Impairment and Risk Stratification for LBP

TOC CATEGORY: Musculoskeletal

ARTICLE TYPE: Original Research

AUTHOR BYLINE: Jason M. Beneciuk, Steven Z. George

AUTHOR INFORMATION:

J.M. Beneciuk, PT, DPT, PhD, MPH Department of Physical Therapy, University of Florida, Gainesville, Florida, and Brooks Rehabilitation Clinical Research Center, 3901 University Boulevard South, Suite 103, Jacksonville, FL 32216 (USA). Address all correspondence to Dr Beneciuk at: beneciuk@phhp.ufl.edu. @JBeneciuk https://orcid.org/0000-0002-5888-3630

S.Z. George, PT, PhD, FAPTA, Department of Orthopaedic Surgery and Duke Clinical Research Institute, Duke University, Durham, North Carolina. @DukeMSK https://orcid.org/0000-0003-4988-9421

KEYWORDS: Prognosis, Low Back Pain, Subgrouping, Cluster Analysis

ACCEPTED: August 16, 2020

SUBMITTED: February 24, 2020

Objective. Identifying subgroups of low back pain (LBP) has the potential to improve prediction of clinical outcomes. Risk stratification is one such strategy that identifies similar characteristics indicative of a common clinical outcome trajectory. The purpose of this study was to determine if an empirically derived subgrouping approach based on physical impairment measures improves information provided from the STarT Back Tool (SBT).

Methods. At baseline in this secondary analysis of a cohort study, patients (N = 144) receiving physical therapy for LBP completed the SBT and tests (active lumbar flexion, extension, lateral bending, and passive straight-leg-raise) from a validated physical impairment index. Clinical outcomes were assessed at 4 weeks and included the Numerical Pain Rating Scale (NPRS) and Oswestry Disability Index (ODI). Exploratory hierarchical agglomerative cluster analysis identified empirically derived subgroups based on physical impairment measures. Independent samples *t* testing and chi-square analysis assessed baseline subgroup differences in demographic and clinical measures. Spearman rho correlation coefficient was used to assess baseline SBT risk and impairment subgroup relationships, and a 3-way mixed-model ANOVA was used to assessed SBT risk and impairment subgroup relationships with clinical outcomes at 4 weeks.

Results. Two physical impairment-based subgroups emerged from cluster analysis: (1) Low-Risk Impairment (n = 119, 81.5%), characterized by greater lumbar mobility and (2) High-Risk Impairment (n = 25, 17.1%), characterized by less lumbar mobility. A weak, positive relationship

was observed between baseline SBT risk and impairment subgroups (r_s = .170). An impairment-by-SBT risk-by-time interaction effect was observed for ODI scores but not for NPRS scores at 4 weeks.

Conclusions. Physical impairment subgroups were not redundant with SBT risk categories and could improve prediction of 4-week LBP disability outcomes. Physical impairment subgroups did not improve the prediction of 4-week pain intensity scores.

Impact. Subgroups based on physical impairment and psychosocial risk could lead to better prediction of LBP disability outcomes and eventually allow for treatment options tailored to physical and psychosocial risk.

Strategies that minimize heterogeneity for common musculoskeletal pain conditions and optimize prognostic decision making have been recently emphasized.[1–6] Risk stratification is one such strategy that identifies subgroups of patients with similar characteristics that are indicative of a common risk trajectory for clinical outcomes.

Identifying subgroups predictive of clinical outcomes is an important component to risk stratification. This can be especially true if the prognostic indicators have potential to be modified by treatment. For example, patients identified as being high psychosocial risk for persistent low back pain (LBP) related disability using the STarT Back Tool (SBT) may benefit from receiving psychologically informed physical therapy.[7–9] Impairments of body function have also been used to classify patients with specific physical impairments (eg, range of motion, movement coordination, spinal segmental mobility) as modifiable targets for treatment.[10,11]

Patient heterogeneity beyond pain-associated distress has not been widely considered. Yadollahpour and colleagues[12] recently highlighted this issue in an analysis that considered subgroup composition based on intake psychological, disability, pain, and physical impairment measures. That study compared the empirically derived groups with SBT classification (primarily psychosocial risk), finding that the empirically derived clusters clearly differentiated between SBT low and high-risk groups. However, less agreement was observed with the SBT medium risk profile because of differences in physical impairment measures. These findings support the notion that patient heterogeneity extends past pain associated distress indicating need for LBP subgrouping that considers psychosocial and physical impairment measures.[13]

Therefore, the purpose of this analysis was to determine if an empirically derived subgrouping approach based on an established physical impairment measure adds to predicting LBP clinical outcomes when combined with the STarT Back Tool (SBT). Physical impairment was selected because it is a modifiable factor, and addressing physical impairment is an established part of physical therapy practice.[10,11] The SBT was selected as the reference prognostic tool for this analysis because it has well established prognostic capabilities,[14–18] and we wanted to determine if adding physical impairment measures had the potential to improve prediction for this already accurate risk stratification tool.

[H1] METHODS

This was a secondary analysis of an observational, cohort study involving patients receiving physical therapy for LBP.[15]

[H2] Participants

Consecutive patients seeking treatment for LBP at six participating community-based outpatient physical therapy clinics in Jacksonville, Florida as part of a single health system were screened for study eligibility by a physical therapist. 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. 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 if they elected not to participate they would receive the same treatment. Informed consent was obtained in compliance with the University of Florida Institutional Review Board.

[H2] Demographic Measures

Study participants completed a standardized self-report questionnaire consisting of items related to age, sex, race, ethnicity, education, household income, marital and employment status. Information involving LBP clinical characteristics (ie, prior surgery, symptom duration, symptom onset, symptom location, work-related LBP) was also obtained.

[H2] Clinical Measures

[H3]STarT Back Tool (SBT). The SBT is a 9-item screening measure used to identify subgroups of patients with LBP based on the presence of modifiable prognostic factors which may be useful in matching patients with targeted interventions.[14] SBT overall scores (ranging from 0 to 9) and SBT psychosocial subscale scores (ranging from 0 to 5) are determined by summing all positive responses. The SBT then categorizes patients as 'high-risk' (psychosocial subscale scores ≥4) in which high levels of psychosocial prognostic factors are present with or without physical factors present, 'medium-risk' (overall score >3; psychosocial subscale score <4) in which physical and psychosocial factors are present, but not a high level of psychosocial factors, or 'low-risk' (overall score 0-3) in which few prognostic factors are present.

[H3]Physical Impairment Index (PII). Several physical examination tests (ie, active lumbar flexion, extension, lateral bending, passive straight-leg-raise, active bilateral straight-leg-raise, active sit-up, and spinal tenderness) were administered to quantify impairment due to LBP as components of the Physical Impairment Index (PII).[19] The PII has been validated in previous studies investigating clinical outcomes.[20,21] Active lumbar flexion, extension, lateral bending, passive straight-leg-raise measures were reported in degrees of range of motion obtained using a handheld inclinometer.[21] For the purpose of this study, we only included PII tests (ie, active lumbar flexion, extension, lateral bending, and passive straight-leg-raise range of motion) that had consistent numerical scaling. Other PII tests (ie, active bilateral straight-leg-raise, active sit-up, and spinal tenderness) were not included in our cluster analysis because they were measured using a binary scale (ie, positive or negative).

[H2] Outcome Measures

Outcome measures were administered at intake and four weeks later and described in more detail below.

[H3]Numerical Pain Rating Scale (NPRS). Pain intensity was rated using a NPRS, ranging from "0" (no pain) to "10" (worst pain imaginable). 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 three pain ratings were averaged and used as the NPRS variable in this study.[22] The NPRS has been found to have sound psychometric properties[22–24] with a minimal clinically important difference reported to be 2 points.[25]

[H3]Oswestry Disability Index (ODI). LBP-related disability was assessed with the ODI, which has 10 items that assesses how LBP affects common daily activities.[20,26] The ODI has a range of

0% “no disability due to LBP” to 100% “completely disabled due to LBP”, with higher scores indicating higher disability from LBP. The ODI has been found to have sound psychometric properties[20,27] with a minimal clinically important difference reported to be 10 percentage points.[28]

[H2] Data Analysis

All statistical analyses were performed using SPSS version 25.0 (SPSS Inc., Chicago, Illinois). Descriptive statistics were used to provide a summary of continuous (with means and standard deviations) and categorical (frequency counts and percentages) data. The data analysis was comprised of different approaches for 1) subgroup derivation, 2) cross-sectional subgroup validation, and 3) longitudinal subgroup validation. These are described in more detail below.

[H3] *Subgroup Derivation*

Kent and colleagues[29] describe ‘supervised’ (using longitudinal clinical outcome data) and ‘unsupervised’ (using cross-sectional baseline data) analytical approaches for identifying patient subgroups. For subgroup derivation we used cluster analysis as an ‘unsupervised’ approach to identify latent relationships that would allow for subgroups to be derived for later validation.[30,31]

PII measures (ie, active lumbar flexion, extension, lateral bending, and passive straight-leg-raise range of motion) were used in the cluster analysis without considering demographic characteristics, SBT risk, or intake outcome measure scores. PII measures were transformed to z-scores to provide standardized scores for subsequent cluster analysis techniques as this approach would allow for generation of subgroups based on physical impairment measures only. 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 physical impairment based measures. Agglomeration coefficient schedules were inspected and plotted using a dendrogram (supplementary figure) to establish the most optimal cluster solution based on the percent change between adjacent cluster solutions and scree plot characteristics (ie, elbow criterion).[32–34]

[H3] *Cross-Sectional Subgroup Validity*

The empirically derived subgroups were further evaluated by investigating differences in intake PII measures. Discriminant function analysis (DFA) with cross-validated jackknifed classification was performed to determine how PII measures differentiated impairment based risk.[35] Eigenvalues from the DFA were reported as a measure of variance, indicating the degree of discrimination between impairment based risk subgroups with higher eigenvalues indicating greater discrimination. Canonical correlations were reported as a measure of the relationship between individual PII measures and the discriminant function. Jackknifed (ie, one case deleted at a time) classification evaluated accuracy in impairment based risk subgrouping.[35]

Then, independent samples t-testing and chi-square analysis were used to identify cluster group differences in intake demographic and clinical measure scores. Finally, relationships between cluster solutions and SBT risk status were assessed with Spearman's rho correlation coefficients.

[H3] *Longitudinal Subgroup Validity*

Whether the newly derived physical impairment risk subgroups added to the SBT risk stratification was investigated by group by time interactions for 4-week clinical outcomes for pain intensity (NPRS ratings) and disability (ODI scores). These interactions were assessed with separate 3-way, mixed-model ANOVA with time (baseline and 4 weeks) as the within-subject factor and impairment based cluster and SBT risk group as the between-subject factors for the 2

clinical outcomes. Interactions ($P \leq .05$) involving the physical impairment subgroups were decomposed using Bonferroni post hoc procedures.

[H2] Role of the Funding Source

The funders played no role in the design, conduct, or reporting of this study.

[H1] RESULTS

During the 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 greater than 65 years of age (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 due to personal reasons and 2 were not included in this analysis because PII data was missing. Therefore, baseline data was obtained from 144 patients and 4-week follow-up data was obtained from 128 patients (88.9%). Descriptive characteristics of the study sample (n = 144) and the resultant impairment based cluster solutions are provided in Table 1.

[H2] Subgroup Derivation - Cluster Analysis

Inspection of all predictor z-scores indicated that absolute values did not exceed 4.0 (range = -3.3 to 3.6), suggesting the data did not contain extreme outliers.[35] Inspection of agglomeration coefficients from a hierarchical agglomerative cluster analysis of physical impairment based measures indicated the percentage of change was moderate (31.2%) between the 2- and 1-cluster solutions with relatively smaller changes in preceding steps (ie, 21.2% between the 3- and 2-cluster solutions), suggesting a 2-cluster solution is appropriate, which was further supported by visual inspection of plotted agglomeration coefficients.[34,35] Cluster 1 was labeled 'Low Risk Impairment' (n = 119, 81.5%) characterized by individuals who were associated with greater

lumbar mobility (ie, higher values for lumbar flexion, extension, lumbar side bending, and passive SLR) compared to cluster 2. Cluster 2 was labeled 'High Risk Impairment' (n = 25, 17.1%) characterized by individuals who were associated with less lumbar mobility compared to cluster 1 (Figure 1).

[H2] Subgroup Derivation - Discriminant Function Analysis

DFA run with simultaneous entry method with four impairment-based predictors (lumbar flexion: Wilks' $\lambda = .70$, $P < .001$; lumbar extension: Wilks' $\lambda = .89$, $P < .001$; lumbar lateral flexion: Wilks' $\lambda = .79$, $P < .001$; passive SLR: Wilks' $\lambda = .67$, $P < .001$) suggested that each predictor contributed uniquely to impairment-based risk and resulted in one discriminant function which is indicative of two impairment-based risk subgroups. The overall test of the discriminant function was significant (χ^2 (4) = 107.6, Wilks' $\lambda = .46$, $P < .001$) indicating that predictor values were able to discriminate amongst the two impairment-based risk subgroups and accounted for 53.6% (canonical R = .73) of the total relationship between predictors and impairment-based risk subgroups. Correlations between PII measures and generated discriminant functions are provided in Table 2. Passive SLR ROM demonstrated the strongest positive relationship with the discriminant function, whereas lumbar flexion and lateral flexion demonstrated moderate positive relationships and lumbar extension demonstrated the weakest positive relationship. The overall accuracy for classification using the discriminant function was 95.1% using the cross-validated jackknifing technique. The percentages classified correctly were 99.2% for low risk impairment and 76.0% for high-risk impairment.

[H2] Cross-Sectional Subgroup Validation

There were no cluster differences for age, gender, race, current employment, or work-related LBP. As expected, differences in physical impairment based measures between clusters were

detected (Tab.1). Additionally, the 'High Risk Impairment' subgroup had proportionally more individuals who had prior surgery (40.0%) compared to the 'Low Risk Impairment' subgroup (12.7%) (χ^2 (1) = 10.65, *P* = .001). There were no differences between clusters for symptom duration, onset, or location. However, the 'High Risk Impairment' subgroup was associated with greater NPRS (1.5 points, 95% CI = 0.6–2.3) and greater ODI (15.3 points, 95% CI = 8.5–22.1) scores at intake compared to the 'Low Risk Impairment' subgroup (*P* < .001). The 'High Risk Impairment' subgroup also had proportionally more individuals who had positive results for active bilateral SLR, sit-up, and spinal tenderness testing compared to the 'Low Risk Impairment' subgroup (*P* < .001). The distribution of cluster profiles by SBT risk categorization is provided in Figure 2A and 2B. A weak, positive relationship was observed between cluster profiles and SBT risk categorization at intake (r_s = .170, P = .043).

[H2] <u>Longitudinal Subgroup Validation</u>

There was no 3-way interaction effect observed for NPRS scores at 4 weeks ($F_{2,119}$ = 2.91, *P* = .058, partial eta-squared = .047), however a 2-way SBT group-by-time interaction effect was observed ($F_{2,119}$ = 3.61, *P* = .030, partial eta-squared = .057). There was a 3-way cluster profile-by-SBT group-by-time interaction effect observed for ODI scores at 4 weeks ($F_{2,119}$ = 3.47, *P* = .034, partial eta-squared = .055). Specifically, changes in ODI scores after 4 weeks varied across SBT risk categories and this variation differed across impairment clusters (Fig. 3A and 3B).

[H1] DISCUSSION

The purpose of this analysis was to determine if an empirically derived subgrouping approach based on a validated physical impairment measure has potential to add to the prediction of LBP

clinical outcomes when combined with the SBT. Our results suggested that two impairment-based subgroups ('Low Risk Impairment' and 'High Risk Impairment') emerged with different profiles in lumbar region range of motion (Fig. 1). Additionally, baseline differences were observed between High and Low Risk Impairment groups for measures not included in the cluster analysis including the NPRS, ODI, bilateral straight leg raise, active sit up, and spinal tenderness. DFA indicated passive SLR, lumbar flexion and lateral flexion range of motion were the strongest contributors to subgroup membership. Collectively these findings support the discriminant validity of impairment-based subgroups. Additional support for impairment-based subgroups came from low redundancy with the SBT risk categories (ie, divergent validity) and an interaction with SBT risk categories for 4 week disability outcomes. These findings provide proof of concept for incorporating impairment-based subgroups with SBT risk status to improve predictive capabilities for LBP disability outcomes, however further testing in larger samples is needed.

Current subgrouping paradigms have been suggested to improve efficiency and effectiveness of allocating treatment for LBP management.[36–38] In general, current subgrouping approaches can be described as those that aim to: 1) risk stratify treatment based largely on psychosocial prognostic factors[14,39,40] and 2) allocate treatment based on response to a specific intervention approach.[41–45] Our preliminary suggestions for incorporating impairment-based measures into an established risk stratification tool is consistent with previous recommendations for the need to consider modified screening strategies in secondary care settings and recent "state of art" suggestions.[46–49] Specifically, Hodges,[49] recently proposed a "hybrid approach to treatment tailoring for low back pain" consisting of initial risk stratification (prognosis based approach) and the need for additional clinical information that enhances treatment selection.

The current analysis was not powered to detect the clinical impact of impairment based subgroups when combined with SBT risk status specifically by interpreting ANOVA based statistical interaction effects. However, we speculate that identifying impairment-based risk eventually has the potential to inform clinical decision making for the SBT medium risk category. Based on our findings, over half (52.0%) of the high-risk impairment subgroup were categorized as SBT medium risk (ie, greatest SBT risk contributor). Knowing the impairment risk subgroup status for patients who are SBT medium risk has potential to provide more treatment directed information compared to the SBT risk alone. For example, there may be some patients classified as SBT medium risk that could benefit from an intensive impairment focused approach given the concomitant classification as high-risk impairment.

There are also implications of knowing the physical impairment risk status for SBT low and high-risk groups. Based on our findings, 16.0% of the high-risk impairment subgroup were categorized as SBT low risk which may provide important clinical decision making information. For example, overall patients who are SBT low risk are commonly associated with a favorable prognosis,[14,15] however we speculate that those also identified as high risk based on physical impairment may actually have a less favorable trajectory. In addition, 23.5% of the low risk impairment subgroup were categorized as SBT high risk and subsequently could be associated with a more favorable trajectory compared to other patients who are SBT high risk. Approximately 40% of participants had a low/high risk mismatch between physical impairment and SBT classification, which is not an unexpected finding given differences in the measures used to derive the subgroups. Given the current state of evidence, SBT classification should be considered a “primary” risk determinant. This analysis suggests there is an opportunity to refine

SBT classification by considering physical impairment status but additional research is necessary.

Although beyond the scope of this study, we can speculate about how matched treatment could be provided based on our findings (Tab. 3). As previously recommended,[9,50] APTA Clinical Practice Guidelines for LBP[10] provide a resource to guide traditional impairment based treatment decision making, however do not provide clear recommendations for distinguishing treatment approaches across SBT risk groups. Therefore, from a treatment perspective, there is potential for impairment based subgrouping to further inform SBT risk categories, especially for those at medium risk. For example, patients who are SBT medium risk and also categorized as high risk impairment may require increased dosage and intensity of guideline recommended impairment based interventions that are tailored to match a potentially greater quantity and magnitude of physical impairment findings.

There is also potential for impairment based subgrouping to inform SBT low and high risk treatment planning. Although we found most patients who were SBT low risk were also categorized as low risk impairment, a small percentage have potential to match a high risk impairment profile and may require a short trial of individualized guideline recommended impairment focused intervention prior to transitioning to emphasis on self-management. Impairment risk subgroups also could inform the delivery of psychologically informed physical therapist interventions for patients who are SBT high risk. For example, high risk on the SBT and high risk based on physical impairment findings may require intensive impairment based treatment approach as part of their treatment encounter (Tab. 3). Alternatively, the approximately 78% of patients who are SBT high risk who were identified as low risk based on physical impairment findings may benefit from a treatment approach with an exclusive emphasis on

psychological based interventions and may be more amenable to telehealth approaches as there is a lower need for clinic based impairment focused treatment.[51]

The approach described in this manuscript is consistent with the initial stage of a novel statistical clustering approach involving two sequential stages described by Kent and colleagues.[29] In that approach, the first stage involves clustering from within each single domain (eg, physical impairment) and followed by subsequent clustering using patterns from other domains (eg, pain, activity limitation, psychological) to identify subgroups across multiple health domains. Therefore, future studies will further inform this line of research by including similar sequential stage clustering methodology. Additional future research is needed to determine the impact of combining SBT risk and physical impairment information into complex treatment decision-making and if this approach is associated with improved clinical outcomes. This line of research is consistent with previous recommendations for a more flexible, multidimensional, and clinically reasoned approach to profiling patient complexity to inform individualized, patient-centered care.[13] By incorporating two modifiable risk factors a combined impairment and psychosocial subgroup has the potential to help further guide treatment selection and is aligned with an integrated PIP approach. Keefe and colleagues[52] have previously acknowledged the importance of considering physical impairment as a component to an integrated PIP approach and further validation of subgroups comprised of impairment based and psychosocial risk status would be informative to future clinical practice.

[H2] Limitations

There were several limitations to this study. A primary limitation of this study is that we only had 4 week outcome data. Second, response bias may have influenced our findings as several self-report measures (ie, SBT, NPRS, and ODI) were administered in this study. Third,

determining the optimal number of cluster solutions from cluster analysis was based on 'unsupervised', data-driven statistical analyses.[29,53] A potential disadvantage to this approach is that our impairment-based subgroups may not replicate in other samples. Therefore, future studies to validate our prognostic findings and test subgroups against different treatment approaches are recommended. Fourth, we acknowledge this current study was not adequately powered to assess for two or three-way statistical interaction effects, therefore future studies consisting of a larger sample size are required to detect the clinical impact of empirical impairment based clusters when combined with SBT risk groups.

Other limitations to consider when interpreting these study findings include that physical impairment clusters were generated using only range-of-motion measures that do not account for other clinically relevant impairment based tests and measures (eg, passive joint mobility, neurological signs). We observed proportionally more individuals with positive results for active bilateral SLR, sit-up, and spinal tenderness testing (variables not included in cluster analysis) in the high risk impairment subgroup. Not including these and other impairment based variables in the cluster analysis may provide a potential explanation for the large difference in correct classification rates between high (76.0%) and low (99.2%) risk impairment subgroups. Accordingly, the measures we used were only able to account for 53.6% of the total variance for differentiating high and low risk physical impairment group membership. This limited variance could limit the ability to generate reliable classification approaches in external samples.

[H1] CONCLUSIONS

Empirically derived physical impairment subgroups differed on key clinical measures and were not redundant with psychosocial risk category. Furthermore, these subgroups have the potential

to improve prediction of 4-week LBP disability outcomes, but not for pain intensity outcomes. These findings provide direction for how physical impairment and psychosocial risk can be incorporated into LBP assessment.

Author Contributions

Concept / idea / research design: J.M. Beneciuk, S.Z. George

Writing: J.M. Beneciuk, S.Z. George

Data collection: J.M. Beneciuk

Data analysis: J.M. Beneciuk

Project management: J.M. Beneciuk

Fund procurement: J.M. Beneciuk, S.Z. George

Providing participants: J.M. Beneciuk

Providing facilities / equipment: J.M. Beneciuk

Providing institutional liaisons: J.M. Beneciuk

Consultation (including review of manuscript before submitting): S.Z. George

Ethics Approval

Informed consent was obtained in compliance with the University of Florida Institutional Review Board.

Funding

The cohort study was funded by a 2009 award from the Brooks Health System to the University of Florida Foundation (S.Z.G. and J.M.B.). J.M.B. was supported by a National Institutes of Health T32 Neuromuscular Plasticity Research Training Fellowship grant (T32 HD043730).

Disclosures

The authors completed the ICMJE Form for Disclosure of Potential Conflicts of Interest and reported no conflicts of interest.

REFERENCES

1. U.S. Department of Health and Human Services. National Pain Strategy: A Comprehensive Population Health Strategy for Pain. Published 2016. Accessed January 20, 2020. https://iprcc.nih.gov/National_Pain_Strategy/NPS_Main.htm. Accessed September 16, 2020.

2. Interagency Pain Research Coordinating Committee: Federal Pain Research Strategy. 2018;2018(July 11). https://iprcc.nih.gov/Federal-Pain-Research-Strategy/Overview. Accessed September 16, 2020.

3. U.S. Department of Health and Human Services. Pain Management Best Practices Inter-Agency Task Force Report: Updates, Gaps, Inconsistencies, and Recommendations. Published May 2019. Accessed January 4, 2020. https://www.hhs.gov/ash/advisory-committees/pain/reports/index.html. Accessed September 16, 2020.

4. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on research standards for chronic low back pain. *J Pain*. 2014;15:569-585.

5. Chou R, Deyo R, Friedly J, et al. *Noninvasive Treatments for Low Back Pain. Comparative Effectiveness Review*. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 16-EHC004- EF. Agency for Healthcare Research and Quality; 2016. www. effectivehealthcare.ahrq.gov/reports/final.cfm.

6. National Academies of Sciences, Engineering, and Medicine. *The Role of Nonpharmacological Approaches to Pain Management: Proceedings of a Workshop.* National Academies of Sciences, Engineering, and Medicine; 2019. Accessed October 12, 2019. https://doi.org/10.17226/25406

7. 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.

8. Foster NE, Mullis R, Hill JC, et al. Effect of stratified care for low back pain in family practice (IMPaCT Back): a prospective population-based sequential comparison. *Ann Fam Med*. 2014;12:102-111.

9. 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. *Phys Ther*. 2015;95:1120-1134.

10. Delitto A, George SZ, Van Dillen L, et al. Low back pain. *J Orthop Sports Phys Ther*. 2012;42:A1-57.

11. Blanpied PR, Gross AR, Elliott JM, et al. Neck Pain: Revision 2017. *J Orthop Sports Phys Ther*. 2017;47:A1-A83.

12. Yadollahpour N, Zahednejad S, Yazdi M, Esfandiarpour F. Clustering of patients with chronic low back pain in terms of physical and psychological factors: A cross-sectional study based on the STarT Back Screening Tool. *Journal of Back and Musculoskeletal Rehabilitation*. Published online November 29, 2019:1-7. doi:10.3233/BMR-181484

13. Rabey M, Smith A, Kent P, Beales D, Slater H, O'Sullivan P. Chronic low back pain is highly individualised: patterns of classification across three unidimensional subgrouping analyses. *Scand J Pain*. Published online June 29, 2019. doi:10.1515/sjpain-2019-0073

14. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008;59(5):632-641. doi:10.1002/art.23563

15. 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. *Phys Ther*. 2013;93:321-333.

16. Suri P, Delaney K, Rundell SD, Cherkin DC. Predictive Validity of the STarT Back Tool for Risk of Persistent Disabling Back Pain in a U.S Primary Care Setting. *Archives of physical medicine and rehabilitation*. Published online April 3, 2018. doi:10.1016/j.apmr.2018.02.016

17. Medeiros FC, Costa LOP, Added MAN, Salomao EC, Costa L. Longitudinal Monitoring of Patients With Chronic Low Back Pain During Physical Therapy Treatment Using the STarT Back Screening Tool. *The Journal of orthopaedic and sports physical therapy*. 2017;47:314-323.

18. Katzan IL, Thompson NR, George SZ, Passek S, Frost F, Stilphen M. The use of STarT back screening tool to predict functional disability outcomes in patients receiving physical therapy for low back pain. *Spine J*. 2019;19:645-654.

19. Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine (Phila Pa 1976)*. 1992;17:617-628.

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*. 2001;81:776-788.

21. Fritz JM, Piva SR. Physical impairment index: reliability, validity, and responsiveness in patients with acute low back pain. *Spine (Phila Pa 1976)*. 2003;28:1189-1194.

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

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

24. Bolton JE. Accuracy of recall of usual pain intensity in back pain patients. *Pain*. 1999;83:533-539.

25. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine (Phila Pa 1976)*. 2005;30:1331-1334.

26. Hudson-Cook N, Tomes-Nicholson K, Breen A. A revised Oswestry disability questionnaire. In: *In: Roland, M.O., Jenner, J.R. (Eds.), Back Pain: New Approaches to Rehabilitation and Education.* Manchester University Press; 1989.

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

28. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*. 2008;33:90-94.

29. Kent P, Stochkendahl MJ, Christensen HW, Kongsted A. Could the clinical interpretability of subgroups detected using clustering methods be improved by using a novel two-stage approach? *Chiropr Man Therap*. 2015;23:20.

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

31. Zeppieri G, Bialosky J, George SZ. Importance of Outcome Domain for Patients With Musculoskeletal Pain: Characterizing Subgroups and Their Response to Treatment. *Phys Ther*. Published online January 16, 2020. doi:10.1093/ptj/pzaa009

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

33. Ketchen D, Shook C. The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. *Strategic Management Journal*. 1996;17:441-458.

34. Yim O, Ramdeen KT. Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. *The Quantitative Methods for Psychology*. 2015;11:8-21.

35. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 5th ed. Pearson /Allyn and Bacon; 2007.

36. Foster NE, Hill JC, Hay EM. Subgrouping patients with low back pain in primary care: are we getting any better at it? *Man Ther*. 2011;16:3-8.

37. Foster NE, Hill JC, O'Sullivan P, Hancock M. Stratified models of care. *Best practice & research Clinical rheumatology*. 2013;27:649-661.

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

39. Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clin J Pain*. 1998;14:209-215.

40. Traeger AC, Henschke N, Hubscher M, et al. Estimating the Risk of Chronic Pain: Development and Validation of a Prognostic Model (PICKUP) for Patients with Acute Low Back Pain. *PLoS Med*. 2016;13:e1002019.

41. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Phys Ther*. 1995;75:470-485; discussion 485-489.

42. Alrwaily M, Timko M, Schneider M, et al. Treatment-Based Classification System for Low Back Pain: Revision and Update. *Phys Ther*. 2016;96:1057-1066.

43. O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. *Man Ther*. 2005;10:242-255.

44. Van Dillen LR, Sahrmann SA, Norton BJ, et al. Reliability of physical examination items used for classification of patients with low back pain. *Phys Ther*. 1998;78:979-988.

45. Smart KM, Blake C, Staines A, Doody C. The Discriminative validity of "nociceptive," "peripheral neuropathic," and "central sensitization" as mechanisms-based classifications of musculoskeletal pain. *Clin J Pain*. 2011;27:655-663.

46. 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 Med*. 2017;15:13-13.

47. 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. *J Pain*. 2017;18(6):673-686.

48. Beneciuk JM, Robinson ME, George SZ. Subgrouping for patients with low back pain: A multidimensional approach incorporating cluster analysis and the start back screening tool. *Journal of Pain*. 2015;16.

49. Hodges PW. Hybrid Approach to Treatment Tailoring for Low Back Pain: A Proposed Model of Care. *J Orthop Sports Phys Ther*. 2019;49:453-463.

50. Beneciuk JM, George SZ, Greco CM, et al. Targeted interventions to prevent transitioning from acute to chronic low back pain in high-risk patients: development and delivery of a pragmatic training course of psychologically informed physical therapy for the TARGET trial. *Trials*. 2019;20:256.

51. Geraghty AWA, Stanford R, Stuart B, et al. Using an internet intervention to support self-management of low back pain in primary care: findings from a randomised controlled feasibility trial (SupportBack). *BMJ Open*. 2018;8:e016768.

52. Keefe FJ, Main CJ, George SZ. Advancing Psychologically Informed Practice for Patients With Persistent Musculoskeletal Pain: Promise, Pitfalls, and Solutions. *Phys Ther*. 2018;98:398-407.

53. Kent P, Keating JL, Leboeuf-Yde C. Research methods for subgrouping low back pain. *BMC Med Res Methodol*. 2010;10:62.

Table 1. Descriptive Characteristics of Study Sample[ab]

Variable	**Total sample** N = 144	**Low Risk Impairment** n = 119	**High Risk Impairment** n = 25	***P***
Demographic				
Age (y)	41.3 (13.5)	40.6 (13.6)	45.7 (11.9)	.066
Sex (n, % female)	89 (61.8%)	72 (60.5%)	17 (68.0%)	.483
Race				.115
Caucasian/White	108 (75.0%)	90 (75.6%)	18 (72.0%)	
African American/Black	25 (17.4%)	18 (15.1%)	7 (28.0%)	
Other	11 (7.6%)	11 (9.2%)	0 (0.0%)	
Current employment				.871
Employed	94 (66.7%)	78 (66.7%)	16 (66.7%)	
Unemployed	38 (27.0)	31 (26.5%)	7 (29.2%)	
Retired	9 (6.4%)	8 (6.8%)	1 (4.2%)	
Work-related LBP (n, % yes)	19 (13.2%)	15 (12.6%)	4 (16.0%)	.648
Clinical				
Prior surgery (n, % yes)	25 (17.5%)	15 (12.7%)	10 (40.0%)	<.001
Symptom duration				.194
Acute or subacute	73 (51.8%)	63 (54.3%)	10 (40.0%)	
Chronic	68 (48.2%)	53 (45.7%)	15 (60.0%)	
Symptom onset				.126
Gradual	70 (49.3%)	62 (52.5%)	8 (33.3%)	
Sudden	52 (36.6%)	42 (35.6%)	10 (41.7%)	
Traumatic	20 (14.1%)	14 (11.9%)	6 (25.0%)	
Symptom location				.579
LBP only	49 (34.0%)	42 (35.3%)	7 (28.0%)	
LBP and buttock or thigh	70 (48.6%)	58 (48.7%)	12 (48.0%)	
LBP and lower leg	25 (17.4%)	19 (16.0%)	6 (24.0%)	
SBT Risk				.059
Low	53 (36.8%)	49 (41.2%)	4 (16.0%)	
Medium	55 (38.2%)	42 (35.3%)	13 (52.0%)	
High	36 (25.0%)	28 (23.5%)	8 (32.0%)	
Impairment Based Measures				
Lumbar flexion ROM[c]	74.4 (27.4)	81.5 (22.7)	41.5 (18.6)	<.001
Lumbar extension ROM[c]	20.6 (10.8)	22.4 (9.9)	13.1 (5.7)	<.001
Lumbar lateral flexion ROM[c]	20.7 (7.3)	22.2 (6.8)	13.4 (4.2)	<.001
Passive SLR ROM[c]	60.7 (14.1)	64.4 (11.8)	43.1 (10.6)	<.001
Active bilateral SLR (n, % positive)[d]	51 (35.4%)	31 (26.1%)	20 (80.0%)	<.001
Active sit-up (n, % positive)[d]	61 (42.4%)	43 (36.1%)	18 (72.0%)	<.001
Spinal tenderness (n, % positive)[e]	103 (71.5%)	78 (65.5%)	25 (100%)	<.001
Outcome Measures (Intake)				

NPRS (potential range: 0-10)	5.4 (2.0)	5.1 (2.0)	6.6 (1.4)	<.001
ODI (potential range: 0%-100%)	32.5 (16.7)	29.7 (15.8)	45.0 (14.7)	<.001

[a]LBP = low back pain; NPRS = Numerical Pain Rating Scale; ODI = Oswestry Disability Index; ROM = range of motion measured in degrees; SBT = STarT Back Tool; SLR = straight leg raise.
[b]Values are mean (SD) unless otherwise indicated.
[c] indicates measures included in cluster analysis, therefore between-group differences are expected.
[d] Positive test for active bilateral SLR and sit-up defined as inability for patient to maintain position for 5 seconds.
[e] Positive test for spinal tenderness defined as any response other than "No" when asked: "Is that painful?" during application of superficial and firm pressure.
.

Table 2. Coefficients of Impairment Based Measures of the Discriminant Function[a]

Impairment Based Measures	Discriminant Function	
	Standardized Coefficients[b]	Correlation Coefficients[c]
Lumbar flexion ROM	0.493	0.615
Lumbar extension ROM	0.088	0.320
Lumbar lateral flexion ROM	0.459	0.483
Passive SLR ROM	0.688	0.649

[a]ROM = range of motion measured in degrees, SLR = straight leg raise.
[b]Indicates standardized canonical discriminant function coefficients.
[c]Indicates pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions.

Table 3. Potential SBT by Impairment Based Risk Matched Treatments.[a]

Impairment Risk	SBT Risk		
	Low	**Medium**	**High**
Low	Patient education encouraging resumption of activities, facilitate self-management strategies	Follow guideline recommended, impairment-based interventions, lower intensity and/or shorter treatment episode	Increased emphasis on psychological-based interventions, opportunity for remote care delivery via telehealth
High	Brief treatment episode (1-3 sessions) for development of individualized impairment-based activity program, transition to emphasis on self-management	Follow guideline recommended, impairment-based interventions, higher intensity and/or longer treatment episode	Equal emphasis on psychological and impairment-based interventions

[a]SBT = STarT Back Tool.

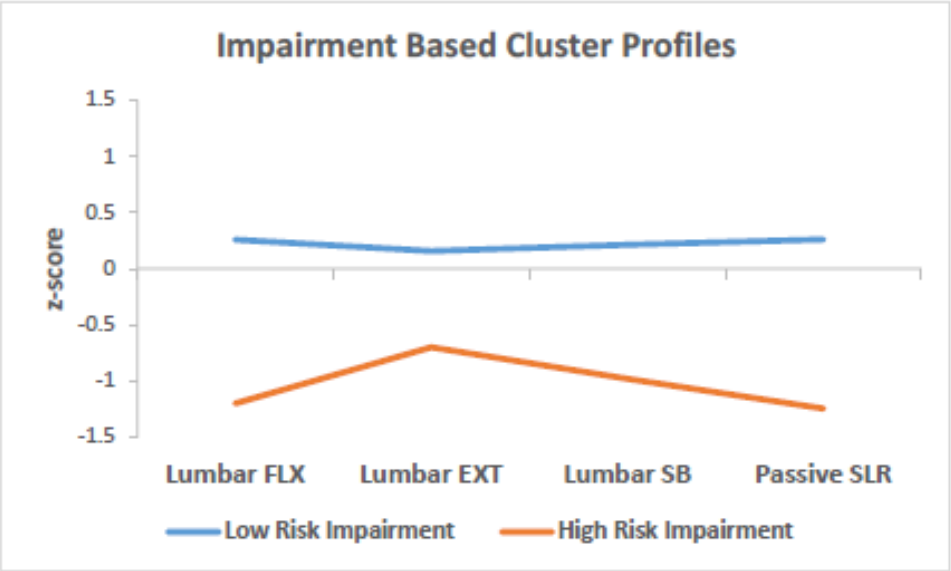

Figure 1. Empirically derived impairment based subgroups. FLX = flexion, EXT = extension, SB = side-bending, SLR = straight leg raise.

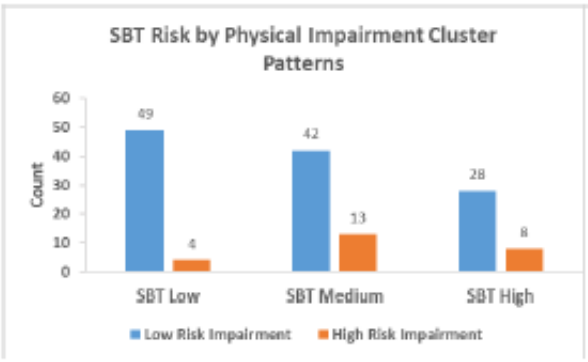

Figure 2A.

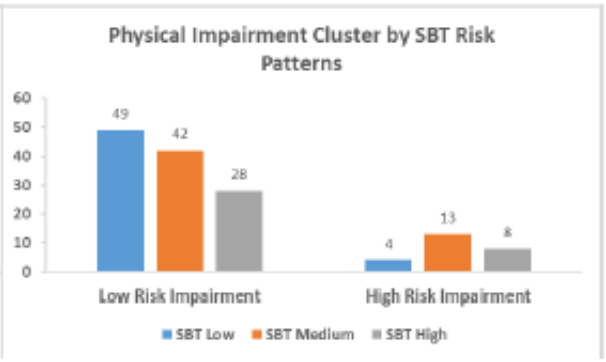

Figure 2B.

Figure 2. Figure 2A (SBT Risk by Physical Impairment Cluster Patterns) and 2B (Physical Impairment Cluster by SBT Risk Patterns). In the 'Low Risk Impairment' subgroup, 41.2% of individuals were categorized as SBT low risk compared to SBT medium and high risk (35.3% and 23.5%, respectively); and in the 'High Risk Impairment' subgroup, 32.0% of individuals were categorized as SBT high risk compared to SBT medium and low risk (52.0% and 16.0%, respectively) (χ^2 (2) = 5.66, P = 0.060). Two-sample proportion z-testing for individual SBT risk group by cluster profile analyses indicated similar proportions for SBT low (z = 0.459, P = 0.645), medium (z = -1.402, P = 0.160) and high (z = -0.948, P = 0.343) risk groups. ODI = Oswestry Disability Index (potential range: 0% to 100%); SBT = STarT Back Tool.

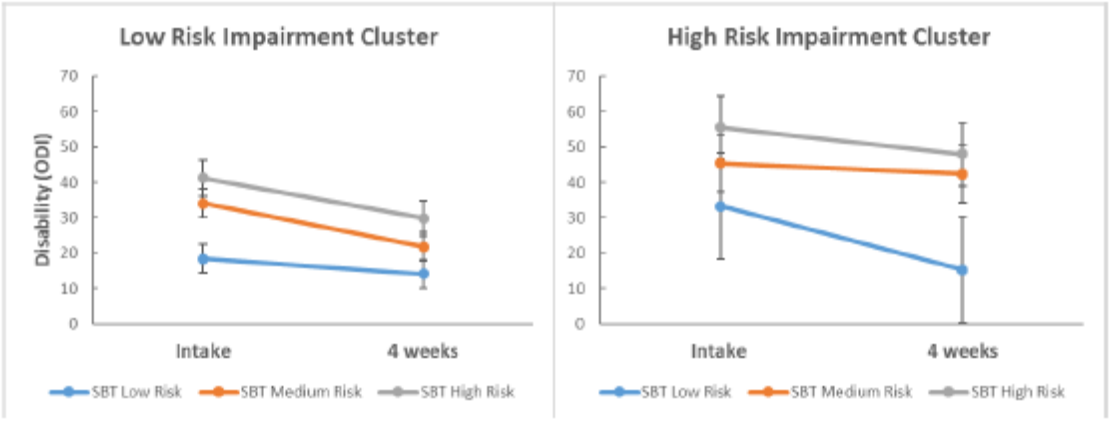

Figure 3A.

Figure 3B.

Figure 3. Figure 3A (Low Risk Impairment Cluster) and 3B (High Risk Impairment Cluster). Results of repeated measures analysis of variance for disability scores using cluster analysis generated physical impairment based subgroups. Tests of simple interaction effects at each impairment cluster suggested the 'Low Risk Impairment' cluster was responsible for the observed three-way interaction ($F_{2,100}$ = 4.34, P = .016, partial eta-squared = .080) [SBT low risk-by-low risk impairment cluster ($F_{1,42}$ = 6.58, P = .014, partial eta-squared = .135), SBT medium risk-by-low risk impairment cluster ($F_{1,34}$ = 29.83, P < .001, partial eta-squared = .467), SBT high risk-by-low risk impairment cluster ($F_{1,24}$ = 14.40, P = .001, partial eta-squared = .375)]. Tests of simple interaction effects for the 'High Risk Impairment' cluster were not statistically significant ($F_{2,19}$ = 2.16, P = .143, partial eta-squared = .185). Error bars indicate 95% confidence intervals. ODI = Oswestry Disability Index (potential range: 0% to 100%); SBT = STarT Back Tool.