JOURNAL OF MANUAL & MANIPULATIVE THERAPY, 2018
VOL. 26, NO. 1, 36–42
https://doi.org/10.1080/10669817.2017.1350328

Taylor & Francis
Taylor & Francis Group



# Prevalence and reliability of treatment-based classification for subgrouping patients with low back pain

Isadora Orlando de Oliveira[a,b], Rodrigo Antunes de Vasconcelos[a,b], Bruna Pilz[a,b], Paulo Eduardo Portes Teixeira[b], Eduarda de Faria Ferreira[b], Wilson Mello[b] and Débora Bevilaqua Grossi

[a]Ribeirão Preto School of Medicine, University of São Paulo, Ribeirão Preto, Brazil; [b]Wilson Mello Institute, Campinas, Brazil

**ABSTRACT**

**Objectives**: To observe the distribution of patients who presented with low back pain (LBP) and to determine the between therapists' interrater reliability of assessments in a private outpatient setting using treatment-based classification (TBC) subgroups.

**Methods**: An observational and methodological study was conducted. Four hundred and twenty-nine patients (231 male; 198 female) presenting LBP symptoms and referred to conservative treatment were assessed by 13 physical therapists who conducted a 60-min examination process utilizing TBC subgroups. Interrater reliability analyses from six raters were assessed using Fleiss' kappa and previously recorded data ($n = 30$).

**Results**: In this study, 65.74% of patients were classified in only one subgroup, the most prevalent being stabilization (21.91%), followed by extension (15.38%), traction (11.89%), flexion (10.96%), manipulation (5.13%), and lateral shift (0.47%). Approximately 20.98% of patients were classified in two subgroups, where the most frequent overlaps were flexion + stabilization (7.46%), extension + stabilization (6.06%), flexion + traction (4.20%), extension + manipulation (1.86%), and 13.29% of patients were not classified in any TBC subgroup. Analysis of interrater reliability showed a kappa value of 0.62 and an overall agreement of 66% between raters.

**Discussion**: LBP is a heterogeneous clinical condition and several classification methods are proposed in the attempt to observe better outcomes for patients. Eighty-five percent of patients assessed were able to be classified when using the TBC assessment and reliability analysis showed a substantial agreement between raters.

**Level of Evidence**: 2c.

**KEYWORDS**
Treatment-based classification; reliability; back pain; lumbar spine; rehabilitation; Musculoskeletal; Evidence-based Practice; Physical Therapy

## Introduction

Low back pain (LBP) is a common dysfunction which affects people of all ages. In Brazil, about 10 million people present disability associated to LBP and data suggest that at least 70% of the population will have a painful episode of it throughout life [1,2]. Worldwide, its lifelong prevalence ranges over 70% on most industrialized countries [3] and due to its disabling symptoms, LBP can be considered a serious injury which leads to high costs not only on the individuals but also the community [4]. The lack of evidence for many clinical interventions associated with this condition [5] and the difficulty to find conclusive treatment options for patients give rise to the development of many treatment options.

Several classification methods have been proposed in the attempt to sort different subgroups of patients and observe larger treatment effects, based on the hypothesis that not all patients would respond to the same treatment methods. Clinicians agree that LBP is a heterogeneous condition, however, there is no consensus on which are the most appropriate and reliable methods to categorize it in order to provide the best outcome for the symptoms presented.

Delitto et al. [6] published in 1995, a classification system with the purpose to provide tailored therapeutic interventions for patients presenting LBP in acute or subacute phase. The clinical decision-making and the definition of the subgroups were based on expert opinions in conjunction with clinical trials at that time [7,8].

Fritz et al. updated this classification system in 2007 by proposing modifications according to most recent evidence [5]. Using this classification approach allows clinicians to identify patients with LBP, who might benefit with specific interventions such as spinal manipulation, stabilization exercises, specific exercises or traction.

Although each interventional subgroup finds itself in different stages of validation and improvement [10–12], case reports and randomized clinical trials in the literature suggest that treatments based on clinical decisions guided by treatment-based classification (TBC) have better clinical outcomes in short and mid-term (up to 1 year) LBP than patients who received unmatched

**CONTACT** Isadora Orlando de Oliveira isadoraooliveira@gmail.com, iooliveira@usp.br

interventions [4,13]. In patients with acute, work-related LBP, this classification method resulted in improved disability and faster return to work status when compared with patients who were submitted to treatment based on evidence-based clinical practice guidelines [14] and also when compared to national Dutch evidence-based clinical guideline, in aspects such as clinical effectiveness and cost-effectiveness was shown equivalent outcomes in patients with chronic LBP [1,9]. As a result, it can be considered as an alternative to help therapists place patients in the most appropriate therapeutic intervention.

Assessment for TBC subgroups may result in patients meeting more than one subgroup treatment criteria, but evidence for the prevalence of these overlaps and for patients classified in one treatment subgroup remain unclear. Also, studies concerning the reliability of the TBC classification among therapists were conducted [15,16] and showed conflicting results.

Stanton et al. [17] previously conducted a prevalence and reliability study with smaller cohorts in Australia and the United States, but additional data are also important to verify regional differences and also contribute to further improvements and modifications on this classification method.

For that reason, the present study aimed to identify the distribution of each subgroup of patients with LBP who were assessed by trained clinicians in a modified version of the TBC system, determining the prevalence of each subgroup in a private practice in Brazil.

## Method

### *Subjects*

A total of 429 patients diagnosed with acute, subacute, or chronic LBP were recruited. The exclusion criteria were patients who had spinal surgery and/or spinal anesthetic blocks in a year period, whose most prominent symptoms were something other than LBP, currently pregnant women, subjects who had spinal fusion or scoliosis rods or screws in the last year or have any known or suspected *red flags*.

An observational and methodological study was conducted and its protocol was approved by the Medical Ethical Committee of the Ribeirão Preto Medical School from the University of São Paulo (14838/2014).

### *Assessment procedure*

Standardized assessments were conducted with all patients by 13 registered physical therapists from a private practice clinic in Brazil with 8–20 years of clinical practice experience and an average of 8 years of applying TBC classification, including over 100 h of training with hands on courses conducted between 2010 and 2012 with one of the TBC developers [18].

Researchers and therapists also remained in constant communication throughout the recruiting period in order to discuss any problems or questions concerning the classification and inclusion in the study.

The examination process contained four phases and takes approximately 60 min to perform (Table 1).

Firstly, four questionnaires – Fear Avoidance Beliefs Questionnaire (FABQ), Oswestry Disability Index (ODI), Numeric Pain Rating Scale (NPRS), Roland Morris Disability Questionnaire (RMDQ) – were administered by the therapist and after that, screening was performed in order to complete medical history, and exclude patients who presented any known or suspected *red flags* [19]. All instruments have been translated and cross-culturally adapted into Brazilian Portuguese and their measurement properties have been tested [20–23].

The assessment also contained questions involving the nature of pain which contributed to detect the source of mechanical pain, such as daily life activities, leisure, sports, work environments, and also specific postures. The physical exam started with gait and posture evaluation to evaluate spinal deformities (scoliosis and lateral shift) and neurological involvement. Spine repetitive active range of motion (ROM) movements in standing

**Table 1.** Treatment-based criteria for placing a patient into a classification.

| Treatment subgroup | Classification criteria |
|---|---|
| Manipulation | Hypomobility of the Lumbar Spine |
| | No symptoms distal to the knee |
| | FABQ-Work < 19 points |
| | Symptom onset < 16 days |
| | Hip internal rotation ROM > 35° (for at least one hip) |
| Stabilization | Age < 40 years old |
| | Greater general flexibility (SLR > 91°) |
| | Positive prone instability test |
| | Hypermobility of the Lumbar Spine |
| | Aberrant movements or 'instability catch' during spinal movements |
| | High frequency of recurrence |
| Flexion | Age > 50 years old |
| | Symptoms centralize with lumbar flexion |
| | Symptoms peripheralize with lumbar extension |
| | Directional preference for flexion in activities of daily life |
| | Imaging evidence of lumbar spine stenosis |
| | Neurogenic claudication signs during gait |
| Extension | Symptoms centralize with lumbar extension |
| | Symptoms peripheralize with lumbar flexion |
| | Directional preference for extension in activities of daily life |
| | Distal symptoms to the buttock |
| Lateral Shift | Visible frontal deviation of the shoulders related to the pelvis |
| | Directional preference or centralization for lateral translational movements of the pelvis |
| Traction | Age > 30 years old |
| | Positive neurological signs |
| | No movements centralize symptoms |
| | Centralization with manual traction |
| | Positive Lasegue's or Cross-Lasegue's sign |
| | FABQ-Work < 21 points |
| Combined | At least two criteria of each subgroup associated with directional preference or centralization phenomenon |

(1–2 movements on each direction according to patients' symptom behavior and ROM restrictions) were performed in flexion, extension, side bending, and pelvic translocation with the intention to see pain behavior and also to set up a standard reference pattern. Pain behavior was defined by the increased or decreased pain in a specific direction and also by the centralization phenomenon, which is an accurate predictor of treatment outcomes, characterized by relocation of distal pain in a proximal direction according to repetitive movements [24].

In the sitting position on an examination table, a neurological exam was performed involving lower quarter screening (dermatomes and myotomes), upper motor neuron reflex (knee and ankle jerk), and the Slump test. After the examination in sitting, the patient laid in supine and following tests were performed: hip and knee passive ROM (flexion, extension, rotations, abduction, and adduction); FABER test, where the patient is positioned in supine and the clinician performs a combined movement of hip flexion, abduction, and external rotation; Lasegue's sign where the patient is positioned in supine and the clinician lifts the patient's leg by the posterior ankle while keeping the knee in a fully extended position; compression, distraction, and thigh thrust for sacroiliac joint provocation tests based on Laslett's classification [25]. Also in the supine position, manual traction was performed to observe pain behavior or centralization phenomenon. For traction response testing, the patient was positioned with both knees flexed 90º and was instructed to relax, while the therapist performed repetitive and gentle manual traction with the forearm placed in the popliteal fossae. Traction classification was composed of criteria which combined clinical prediction rules from two studies in the derivation phase [5,26].

In the prone position, hip internal and external ROM, postero-anterior accessory mobility test, prone instability test, and the sacral thrust test were performed. The last phase of the clinical exam, the patient performed the repetitive movements of partial Mechanical Diagnosis and Therapy (MDT) McKenzie protocol starting to 10–40 repetitive extension movements in the prone position with different loading strategies.

After extension repetitions, all active spine ROM movements were performed in standing position to compare modifications in pain behavior, improvements of the spine-active ROM, directional preference, or centralization phenomenon. The same protocol was used for repetitive flexion movements in supine, in four kneeling position, in sitting and also lateral shifts in standing. After the baseline assessment, therapists evaluated image exams and scans and also informed the patient of the evaluation results, giving instructions using the biopsychosocial model.

Treatment involved a 12-week protocol, with one or two therapy sessions a week, depending on patients' clinical improvements. Re-evaluations were performed every four weeks. As soon as patients presented significant clinical improvement, defined as NPRS scores below 3 points and ODI scores lower than 20%, they were introduced to a functional recovery stage, which consisted in global and local strengthening of the lumbopelvic–hip complex muscles and aerobic exercises.

### Reliability

Six trained raters, who were licensed physical therapists with 10–12 years of clinical practice experience and have 120 h of training in TBC assessment procedures participated in the reliability testing. The reliability testing consisted of analyzing previously recorded data from actual patient cases who participated in the prevalence study ($n = 30$), classifying them according to TBC classification.

Several practice cases (different from those used in this study) were discussed during weekly meetings between researchers and therapists to familiarize the raters with the data and study goals. After the discussion, each of the raters were presented to 30 clinical examination forms that included the patients' history and physical exam data and a form to mark their answers. The vignette technique was used in order to present the cases to assess interrater reliability, since it is a method that can elicit perceptions, opinions, beliefs, and attitudes from responses or comments to stories depicting scenarios and situations [27].

Working independently, each rater was instructed to use the information provided and their TBC experience to assign only one out of six possible classification categories (traction, manipulation, specific exercise, stabilization, combined, or no classification) to each of the 30 cases presented. Combined classification was defined according to the algorithm developed by Stanton et al. [17] in patients who present factors favoring two different subgroups according to the assessment results. Raters were blinded to the judgments of the other raters in this study and to those made in the original assessment.

The raters were licensed physical therapists with 10–12 years of clinical practice experience and had 120 h of training in TBC assessment procedures. They practiced in an outpatient orthopedic setting and had completed a residency program or had at least one specialty certification or post-graduate degree. All raters were informed of the study protocol, gave consent prior to their participation, and were instructed in the use of their knowledge concerning TBC assessment to determine patient classification using previously recorded patient data, which were not collected by the raters who participated on the reliability study and remained anonymous during this procedure.

### Statistical analysis

Prevalence data were calculated according to the individual subgroup criteria, where the percentage of patients

meeting the criteria for each treatment subgroup was calculated as: number of patients meeting subgroup criteria divided by the total of patients ($n = 429$).

Reliability was assessed using Fleiss' kappa with 95% confidence interval (CI). Kappa values of 0.81–1.0 represent almost perfect agreement, values between 0.61 and 0.80 represent substantial agreement, 0.41–0.60 moderate agreement, 0.21–0.40 fair agreement, 0.01–0.20 slight agreement, and those below 0 indicate poor agreement [28].

## Results

Baseline characteristics for the subjects included in the study are provided in Table 2.

Table 3 shows the principal diagnosis received by doctor referrals prior to physical therapy assessment. Concerning the medical diagnosis subjects presented, the most common were LBP (40%), followed by non-specific LBP (17%) and degenerative disc disease (10%); only 4% of subjects who were referred for physical therapy evaluation had an undefined medical diagnosis; patients who did not present any diagnosis were classified as 'undefined'.

The rates for each treatment subgroup for classification using TBC criteria are provided in Table 4. Analyzing the entire sample, 66% of the patients met the criteria for one subgroup ($n = 282$) and 21% of the patients met the criteria for two subgroups ($n = 90$). According to the individual subgroup criteria, the most common treatment subgroup for which the criteria were met was specific exercise, which involve flexion, extension, or lateral shift movements (percentage of 27%), followed by stabilization (percentage of 22%). The most common combination of subgroups for which the criteria were met was flexion associated with stabilization (percentage of 7%) and the most frequent subgroup to be presented in classifications in general was stabilization, which appeared 152 times. The percentage of patients who did not meet the TBC criteria for any of the treatment subgroups was 13%.

Concerning reliability, 6 raters analyzed previous assessment data from 30 subjects and the results of TBC classification presented a kappa value of 0.62 (substantial agreement) and an overall agreement of 66% between raters.

## Discussion

This study demonstrated that 85% of patients assessed in an outpatient private could be classified using the TBC classification criteria. When applied, this assessment resulted in a substantial interrater agreement and more than 60% of the assessed patients meeting the criteria for only one subgroup. On other cases, patients either met the criteria for more than one subgroup (21%) or did not meet the criteria for any subgroups (13%). These results might give therapists an expectation of the cases that they are more likely to see in clinical practice, and also provide essential information for guiding additional research and physical therapists in terms of treatment options and evaluation methods.

**Table 2.** Baseline characteristics of subjects (Mean ± SD).

| | |
|---|---|
| Age, yr | 45.93 ± 13.91 |
| Height, cm | 171 ± 1 |
| Weight, kg | 77.69 ± 16.46 |
| BMI | 26.19 ± 4.11 |
| Chronic LBP (>12 week), % | 75.92 |
| Acute LBP (<12 week)% | 24.08 |
| Currently taking medication for LBP, % | 55.27 |
| Marital Status, % | |
| Married | 81.40 |
| Single | |
| Divorced | 13.06 |
| 1.50 | |
| Widow/Widower | 4.02 |
| Smoker, % | 15.07 |
| College Education, % | 72.36 |
| Currently Employed, % | 94.87 |
| Pain Intensity, NPRS (0–10) | 5.84 ± 2.32 |
| ODI (0–100) | 23.45 ± 9.89 |
| FABQ activity (0–24) | 15.60 ± 6.19 |
| FABQ work (0–42) | 11.41 ± 9.73 |
| RMDQ (0–24) | 10.73 ± 5.82 |

Notes: BMI = Body mass index; LBP = Low back pain; NPRS = Numeric Pain Rating Scale; ODI = Oswestry Disability Index; FABQ = Fear Avoidance Beliefs Questionnaire; RMDQ = Roland Morris Disability Questionnaire.

**Table 3.** Medical diagnosis presented by the subjects referred for PT.

| Medical diagnosis | *N* (%) |
|---|---|
| LBP | 171 (39.67) |
| Non-specific LBP | 74 (17) |
| Disc Herniation | 35 (8.34) |
| -L4/L5 | 11 (2.67) |
| -L5/S1 | 24 (5.67) |
| Sciatica | 11 (2.67) |
| Disc Protusion | 29 (6.67) |
| Spondylolysthesis | 4 (1) |
| Disc Bulging | 5 (1.33) |
| Degenerative Disc Disease | 43 (9.66) |
| Lumbar Spine Degenerative Stenosis | 2 (0.66) |
| Other | 38 (9) |
| Undefined | 17 (4) |

Note: LBP = Low back pain.

**Table 4.** Rates using the treatment-based classification criteria.

| TBC subgroup | Number of patients | % |
|---|---|---|
| *Patients meeting one treatment classification* | | |
| Stabilization | 94 | 21.91 |
| Specific Exercise – Flexion | 47 | 10.96 |
| Specific Exercise – Extension | 66 | 15.38 |
| Specific Exercise – Lateral Shift | 2 | 0.47 |
| Manipulation | 22 | 5.13 |
| Traction | 51 | 11.89 |
| *Patients meeting two treatment classifications* | | |
| Extension + Manipulation | 8 | 1.86 |
| Extension + Stabilization | 26 | 6.06 |
| Extension + Traction | 6 | 1.40 |
| Flexion + Stabilization | 32 | 7.46 |
| Flexion + Traction | 18 | 4.20 |
| *Patients meeting no treatment classification* | | |
| | 57 | 13.29 |
| Total | 429 | 100 |

Apart from the fact that important aspects of the content validity of classification systems are intimately related to their classification categories [29,30], direct comparison of prevalence rates between the current study and previous literature were made with carefulness, since studies diverge in patient populations and outcomes.

When comparing to other work, the prevalence rates were partially similar, except for the manipulation subgroup, where previous studies' presented manipulation rates ranged from 23 to 59% [10,31–33] (present study: 5.13%, 95% CI 2.57–7.69%). Rates from the literature for stabilization prevalence ranged from 24 to 26% [10,31] (present study: 21.91%, 95% CI 19.28–24.55%). For the specific exercise subgroup (flexion, extension, and lateral shift), previous prevalence rates ranged from 21 to 74% [10,31,34,35] (present study: 26.81%, 95% CI 23.05–29.63%), and for traction, prevalence rates in the literature were recently reported to be 9% (present study: 11.89%, 95% CI 9.25–14.52%) [10].

Regarding the manipulation subgroup, there are several reasons that this study presented lower prevalence rates when compared to previous literature is due to the fact that to classify patients in the manipulation subgroup particularly in this study they must present at least four out of five criteria of the clinical prediction rules to be classified as a manipulation subgroup because the presence of four out of five criteria increases the likelihood of success to 95% (positive likelihood ratio = 24.4 (95% CI 4.6–139.4), whereas most studies considered only two criteria in order to classify patients in this subgroup, however, using only two criteria have a small positive likelihood ratio of 1.18 (95% CI 1.09–1.42) [36].

Another reason to use specifically four out of five criteria to perform spinal manipulation is that the majority of the population in this study was chronic patients (75.92%) and one of the main clinical prediction rules for this subgroup targets patients with acute symptoms of less than 16 days duration. Also, since the majority of patients presented with chronic LBP condition, using four of five criteria to perform spinal manipulation was in line with current guidelines that recommend passive interventions avoidance in this specific population [14].

This study population had a mean ODI score of 23.45% ± 9.89. This score shows a lower level of functional disability compared to other studies that enrolled subjects with acute and subacute symptoms and an ODI above 30% [34]. To our knowledge there are only two studies, which investigated TBC in patients with chronic symptoms, with mean ODI scores between 19% and 21%, aligning with the present study [1,37].

Concerning specific exercise subgroup, the therapists involved in this study performed a partial MDT [38,39] assessment. The assessment was considered partial because, repetitive movements in the coronal plane (extension with the pelvis off center, flexion with the pelvis off center, repetitive rotation movements) were excluded from the assessment protocol. Therefore, it is not surprising that the prevalence rates of this study were lower (48%) when compared with prevalence rates of directional preference using exclusively the MDT assessment (74%) [35], but similar to TBC previous literature, where Stanton et al. reported 45.6% and 44% of specific exercise classification performed by physical therapists in Australia and the United States, respectively [17].

This study also showed that approximately 20% of the patients with LBP met the criteria for more than one subgroup when using the TBC criteria, aligning with prior studies [17]. The most common subgroups that we found combined were flexion and stabilization, which represented 35.5% of patients who met two subgroups. Possible reasons for this finding are that in spite of a clear classification of all possible subgroups performed in the assessment, specific exercise was prioritized due to the fact that the criteria for clear classification of specific exercise require centralization with movement and not the more frequent finding of a directional preference.

In this study, approximately 14% of the patients with LBP did not meet the criteria for any of the treatment subgroups when using the individual subgroup criteria; these findings were lower than previous literature, and this could be due to several reasons. For instance, the experience with the TBC criteria prior to the study was long and well explored, resulting in therapists more familiar with the classification system. However, knowing they were participating in a prevalence study concerning TBC classification could have made therapists more eager to classify patients into subgroups. The fact that therapists were trained to think according to the TBC principles, developing their own line of thought during the assessment instead of following a specific algorithm as suggested by previous studies [17,40] could also have contributed to a better performance in classifying patients.

The finding that many patients met the criteria for more than one subgroup or did not have a clear classification suggests that a therapist may begin with the treatment indicated by the TBC assessment and alter the approach as sessions go by according to patients' response. This tailored approach is one of the main strengths of the TBC classification, since it takes into consideration the heterogeneity of LBP, and may assist treatment selection for patients at any moment of the rehabilitation process.

Measuring reliability and agreement between raters for the TBC assessment is crucial in order to compare measurements between therapists, observe classification patterns and consistency, and also explore disagreements among them, contributing to the classification methods' further enhancement. This study presented a substantial agreement ($k = 0.62$) between raters, conflicting with previous studies by Apledoorn et al. (2016) [41], Fritz et al.

(2000) [15] and Stanton et al. (2011) [17] who also examined the interrater reliability of the TBC system for LBP, but found fair to moderate kappa values. Our results aligned to the reliability study by Henry (2012) [16], that found a kappa value of 0.62 among novice raters concerning TBC classification interrater reliability, but limitations must be addressed: raters did not conduct the clinical exam themselves and the patient was not present on the day of the reliability testing. Thus, the present study differs from previous literature, since no algorithm was used to guide raters' classification during reliability testing, which could have improved or worsened the results presented.

Our study has limitations, leaving some questions to be explored by future research. We did not perform an analysis following physical therapy care, as patients with an unclear classification may present inferior treatment outcomes. The same could happen with patients who were classified in a specific subgroup but changed their characteristics during treatment.

Another question that should be considered for future clinical outcomes studies is the association between the manipulation and the specific exercises subgroups due to the fact that some patients who were classified into the manipulation subgroup could achieve better results if associated with specific exercise [40,42].

The results of the present study suggest that performing TBC assessment during the first contact with LBP patients along with the physical examination and medical history could help clinicians develop a more personalized treatment to patients, guiding them to better selection of optimal treatment strategies during rehabilitation process. Alternately, further research may be directed toward analyzing how patients who have experienced this kind of approach responded to treatment.

## Conclusion

Eighty-five percent of patients assessed in a private outpatient setting could be classified using TBC with substantial agreement between rater. Subgrouping patients using this approach is recommended to assist physical therapists in making clinical decisions about therapeutic interventions in patients with LBP. This classification method should be considered by physical therapists, since it contributes to the elaboration of treatment and provides each patient with a tailored type of approach, compatible with their symptoms.

## Disclosure statement

The authors report no declarations of interest.

## Notes on Contributors

***Isadora Orlando de Oliveira*** is a physiotherapist and a research student in Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil and Hospital Israelita Albert Einstein, São Paulo, SP, Brazil. Her research interests include low back pain, treatment-based classification, health economics, and reliability studies.

***Rodrigo Antunes de Vasconcelos*** is a physiotherapist and Head of the Physical Therapy Department, Wilson Mello Institute, Campinas, SP, Brazil and Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil. His research interests include low back pain, treatment-based classification, and ACL injuries.

***Bruna Pilz*** is a physiotherapist in Wilson Mello Institute, Campinas, SP, Brazil and a research student in Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil. Her research interests include low back pain, treatment-based classification, and biopsychosocial factors.

***Paulo Eduardo Portes Teixeira*** is a physiotherapist in Wilson Mello Institute, Campinas, SP, Brazil. His research interests include low back pain, treatment-based classification, and ACL injuries.

***Eduarda de Faria Ferreira*** is a physiotherapist in Wilson Mello Institute, Campinas, SP, Brazil. Her research interest includes low back pain.

***Wilson Mello*** is a medical Doctor and Head of the Wilson Mello Institute, Campinas, SP, Brazil. His research interest includes orthopaedics.

***Débora Bevilaqua Grossi*** is an associate professor in Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil. Her research interests include musculoskeletal rehabilitation and injuries, TMD, headache.

## References

[1] Apeldoorn AT, Bosmans JE, Ostelo RW, et al. Cost-effectiveness of a classification-based system for sub-acute and chronic low back pain. Eur Spine J. 2012;21(7):1290–1300.

[2] Oliveira VC, Furiati T, Sakamoto A, et al. Health locus of control questionnaire for patients with chronic low back pain: psychometric properties of the Brazilian-Portuguese version. Physiother Res Int. 2008;13(1):42–52.

[3] Noriega-Elío M. La polémica sobre las lumbalgias y su relación con el trabajo: estudio retrospectivo en trabajadores con invalidez. Cad Saúde Pública. 2005;21(3):887–897.

[4] Hebert JJ, Koppenhaver SL, Walker BF. Subgrouping patients with low back pain: a treatment-based approach to classification. Sports Health. 2011;3(6):534–542.

[5] Fritz JM, Lindsay W, Matheson JW, et al. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. Spine (Phila Pa 1976). 2007;32(26):E793–E800.

[6] 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(6):470–485.

[7] Delitto A, Cibulka MT, Erhard RE, et al. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. Phys Ther. 1993;73(4):216–222; discussion 223–228.

[8] Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. Phys Ther. 1994;74(12):1093–1100.
[9] Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. Spine. 2003;28(13):1363–1371.
[10] George SZ, Bialosky JE, Donald DA. The centralization phenomenon and fear avoidance beliefs as prognostic factors for acute low back pain: a preliminary investigation involving patients classified for specific exercise. J Orthop Sports Phys Ther. 2005;35:580–588.
[11] Hicks GE, Fritz JM, Delitto A, et al. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. Arch Phys Med Rehabil. 2005;86(9):1753–1762.
[12] Kamper SJ, Maher CG, Hancock MJ, et al. Treatment-based subgroups of low back pain: a guide to appraisal of research studies and a summary of current evidence. Best Pract Res Clin Rheumatol. 2010;24(2):181–191.
[13] Pinto D, Cleland J, Palmer J, et al. Management of low back pain: a case series illustrating the pragmatic combination of treatment- and mechanism-based classification systems. J Man Manip Ther. 2007;15(2):111–122.
[14] Delitto A, George SZ, Van Dillen LR, et al. Orthopaedic section of the American Physical Therapy Association. Low back pain. J Orthop Sports Phys Ther. 2012;42(4): A1–A57.
[15] Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain: interrater reliability and short-term treatment outcomes. Spine. 2000;25(1):106.
[16] Henry SM, Fritz JM, Trombley AR, et al. Reliability of a treatment-based classification system for subgrouping people with low back pain. J Orthop Sports Phys Ther. 2012;42(9):797–805.
[17] Stanton TR, Fritz JM, Hancock MJ, et al. Evaluation of a treatment-based classification algorithm for low back pain: a cross-sectional study. Phys Ther. 2011;91(4):496–509.
[18] Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. Ann Intern Med. 2004;141(12):920–928.
[19] Waddell G. The back pain revolution. New York (NY): Churchill Livingstone; 1999. p 438.
[20] Costa LO, Maher CG, Latimer J, et al. Clinimetric testing of three self-report outcome measures for low back pain patients in Brazil: which one is the best? Spine. 2008;33(22):2459–2463.
[21] Vigatto R, Alexandre NM, Correa Filho HR. Development of a Brazilian Portuguese version of the Oswestry Disability Index: cross-cultural adaptation, reliability, and validity. Spine (Phila Pa 1976). 2007;32(4):481–486.
[22] Roland M, Fairbank J. The Roland-Morris disability questionnaire and the Oswestry disability questionnaire. Spine (Phila Pa 1976). 2000;25(24):3115–3124.
[23] de Souza FS, da Silva Marinho C, Siqueira FB, et al. Psychometric testing confirms that the Brazilian-Portuguese adaptations, the original versions of the Fear-Avoidance Beliefs Questionnaire, and the Tampa Scale of Kinesiophobia have similar measurement properties. Spine (Phila Pa 1976). 2008;33(9):1028–1033.
[24] Donelson R, Silva G, Murphy K. Centralization phenomenon. Its usefulness in evaluating and treating referred pain. Spine (Phila Pa 1976). 1990;15(3):211–213.
[25] Laslett M, Aprill CN, McDonald B. Provocation sacroiliac joint tests have validity in the diagnosis of sacroiliac joint pain. Arch Phys Med Rehabil. 2006;87(6):874.
[26] Cai C, Pua YH, Lim KC. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with mechanical lumbar traction. Eur Spine J. 2009;18(4):554–561.
[27] Jochums Brenda L, Pershey Edward J. Using the vignette method in formative evaluation. Eval Pract. 1993;14(2):155–161.
[28] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.
[29] Fritz JM. Use of a classification approach to the treatment of 3 patients with low back syndrome. Phys Ther. 1998;78(7):766–777.
[30] Murphy DR, Hurwitz EL. A theoretical model for the development of a diagnosis-based clinical decision rule for the management of patients with spinal pain. BMC Musculoskelet Disord. 2007;8:778.
[31] Brennan GP, Fritz JM, Hunter SJ, et al. Identifying subgroups of patients with acute/subacute 'nonspecific' low back pain: results of a randomized clinical trial. Spine (Phila Pa 1976). 2006;31:623–631.
[32] Flynn T, Fritz J, Whitman J. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. Spine (Phila Pa 1976). 2002;27:2835–2843.
[33] Hancock MJ, Maher CG, Latimer J, et al. Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. Eur Spine J. 2008;17(7):936–943.
[34] Browder DA, Childs JD, Cleland JA, et al. Effectiveness of an extension-oriented treatment approach in a subgroup of subjects with low back pain: a randomized clinical trial. Phys Ther. 2007;87:1608–1618.
[35] Long A, Donelson R, Fung T. Does it matter which exercise: a randomized control trial of exercise for low back pain. Spine (Phila Pa 1976). 2004;29:2593–2602.
[36] Costa LO. Four sessions of spinal manipulation, simple exercises and education are not better than usual care for patients with acute low back pain. Evid Based Med. 2016;21(2):69.
[37] Apeldoorn AT, Ostelo RW, van Helvoirt H, et al. A randomized controlled trial on the effectiveness of a classification-based system for subacute and chronic low back pain. Spine (Phila Pa 1976). 2012;37(16):1347–1356.
[38] Busanich BM, Verscheure SD. Does McKenzie therapy improve outcomes for back pain? J Athl Train. 2006;41(1):117–119.
[39] Clare HA, Adams R, Maher CG. A systematic review of efficacy of McKenzie therapy for spinal pain. Aust J Physiother. 2004;50:209–216.
[40] Fritz JM, Brennan GP, Clifford SN, et al. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. Spine (Phila Pa 1976). 2006;31:77–82.
[41] Apeldoorn AT, van Helvoirt H, Ostelo RW, et al. Inter-rater reliability of a modified version of Delitto et al.'s classification-based system for low back pain: a pilot study. J Man Manip Ther. 2016;24(2):98–110.
[42] Cleland JA, Fritz JM, Kulig K, et al. Comparison of the effectiveness of three manual physical therapy techniques in a subgroup of patients with low back pain who satisfy a clinical prediction rule: a randomized clinical trial. Spine(Phila Pa 1976). 2009;34(25):2720–2729.