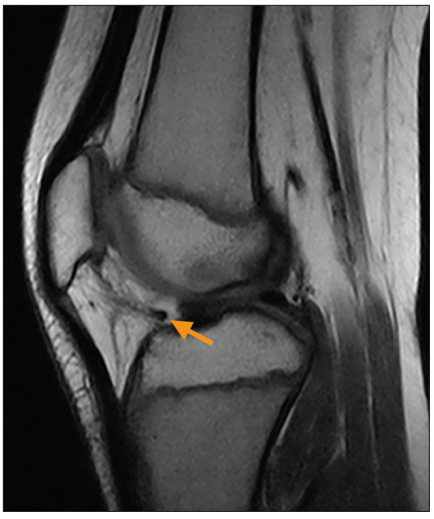


FIGURE 1. Sagittal, T1-weighted magnetic resonance image of the left knee revealing infrapatellar plica synovialis (arrow).

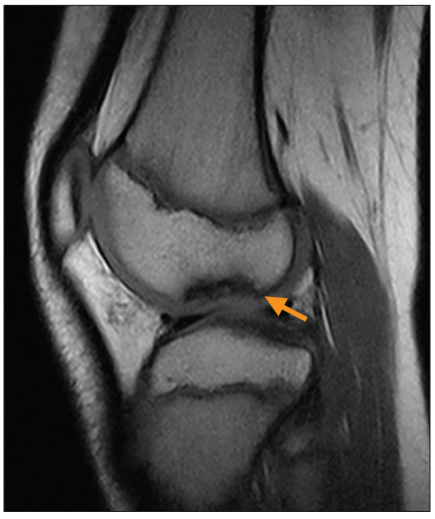


FIGURE 2. Sagittal, T1-weighted magnetic resonance image of the left knee revealing an osteochondral lesion of approximately 18 mm (arrow) on the lateral femoral condyle.

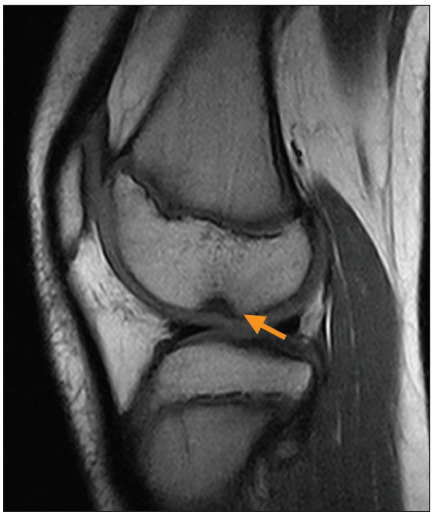


FIGURE 3. Sagittal, T1-weighted magnetic resonance image of the right knee revealing an osteochondral lesion of approximately 10 mm (arrow) on the lateral femoral condyle.

Plica Syndrome and Bilateral Osteochondritis Dissecans

MASSIMO BENEDETTI, PT, MSc, *private practice, Terranuova Bracciolini, Italy.*
MICHELE SPINOSA, PT, MSc, *private practice, Cefalù, Italy.*
FILIPPO MECHELLI, PT, MSc, *private practice, Urbino, Italy.*

A 13-year-old male basketball player presented to a direct-access physical therapy clinic with a chief complaint of left anterolateral knee pain. The patient had no history of trauma or significant medical history. His pain, which began 4 weeks earlier, was exacerbated after playing basketball and decreased with rest.

On physical examination, the patient had a full, pain-free range of active and passive motion and reported no effusion or symptoms of knee locking. Tests for ligamentous laxity and for meniscus tears were negative, whereas the Hughston plica test,[2] mediopatellar plica test,[2] and plica stutter test[2] were all positive. Due to suspicion of plica synovialis, the patient's primary care physician was consulted and recommended a magnetic resonance imaging (MRI) scan.

The MRI scan revealed infrapatellar plica synovialis (**FIGURE 1**) and an osteochondral lesion of the lateral femoral condyle of the left knee (**FIGURE 2**), described as juvenile osteochondritis dissecans (JOCD). Imaging findings and the location of pain suggested that the knee pain originated from the JOCD rather than from plica synovialis.

Considering that patients with JOCD have a 29% incidence of bilateral disease,[1] it was advised to also perform MRI on the right knee, despite it being asymptomatic. The MRI scan of the asymptomatic knee also showed an osteochondral lesion of the lateral femoral condyle (**FIGURE 3**; **FIGURE 4**, scrollable MR images available at www.jospt.org).

After physical therapy treatment (34 visits over a 6-month period) and sport activity intensity reduction (running, but avoiding jumps and sprints), the pain gradually resolved after 6 months, and a follow-up MRI scan revealed no progression of the JOCD defects in both knees.

This clinical case suggests that bilateral involvement of JOCD should be considered in cases where unilateral symptomatic JOCD lesions are present. Early identification of JOCD facilitates successful conservative management. The association of JOCD with plica synovialis requires further investigation. ◉ *J Orthop Sports Phys Ther 2019;49(10):762. doi:10.2519/jospt.2019.8922*

References
1. Cooper T, Boyles A, Samora WP, Klingele KE. Prevalence of bilateral JOCD of the knee and associated risk factors. *J Pediatr Orthop*. 2015;35:507-510. https://doi.org/10.1097/BPO.0000000000000323
2. Magee DJ. *Orthopedic Physical Assessment*. 6th ed. St Louis, MO: Elsevier/Saunders; 2014.

EVAN M. DAY, MS[1] • JASON C. GILLETTE, PhD[2]

Acute Effects of Wedge Orthoses and Sex on Iliotibial Band Strain During Overground Running in Nonfatiguing Conditions

BACKGROUND: Previous research has identified that iliotibial band (ITB) syndrome is more prevalent in females than in males. It has been theorized that high ITB strain rate is a primary etiological factor for developing ITB syndrome. Orthoses are commonly used to influence gait mechanics and may reduce ITB strain rate by influencing alterations in the kinematic chain.

OBJECTIVES: To identify how wedge orthoses and sex affect ITB strain and strain rate.

METHODS: Thirty asymptomatic participants (15 male, 15 female) ran with 7° lateral, 3° lateral, 0° (no wedge), 3° medial, and 7° medial wedges in this within-subject, repeated-measures study. Participants ran overground while data were collected with a motion-capture system and force platform. Iliotibial band strain and strain rate were estimated using a novel 6-degrees-of-freedom musculoskeletal model. A mixed-model multivariate analysis of covariance for between-subject comparison of sex and within-subject comparison of wedge was used.

RESULTS: There were no significant differences in ITB strain or strain rate between wedge conditions. Females had significantly higher ITB strain and strain rate compared to males.

CONCLUSION: Clinicians should be aware that medial wedges may not acutely alter ITB strain or strain rate. Females exhibited greater peak ITB strain and strain rate, potentially due to increased hip internal rotation compared to males. Further research is needed to investigate longitudinal effects of the wedges. *J Orthop Sports Phys Ther 2019;49(10):743-750. Epub 31 Aug 2019. doi:10.2519/jospt.2019.8837*

KEY WORDS: *biomechanics, gait, injury*

Iliotibial band syndrome (ITBS) is the second most prevalent running injury behind patellofemoral pain syndrome and is the leading cause of lateral knee pain in runners.[3,32] Strain rate has been proposed to be a primary factor in the development of ITBS.[13] Adduction and internal rotation at the hip, along with flexion, adduction, and internal rotation of the tibia relative to the femur at the knee, elongate and thus increase strain in the iliotibial band (ITB).[13] Increased ankle eversion is associated with increased tibial internal rotation,[19] which may increase ITB strain and strain rate by altering knee kinematics.

Retrospective studies provide conflicting evidence on the association of biomechanical factors with ITBS. Symptomatic males and females and individuals with previous ITBS exhibit less hip adduction throughout stance than healthy runners, potentially due to a tighter (ie, stiffer) ITB or a potential compensatory mechanism.[11,12,28] However, prospective results show that individuals who developed ITBS exhibit increased hip adduction and knee internal rotation.[27] Individuals with symptomatic ITBS also exhibit greater knee flexion and knee internal rotation velocity,[23] which are thought to elongate the ITB and may result in increased ITB strain rate. Individuals with ITBS also exhibit a more inverted foot at heel strike and greater rearfoot motion throughout stance,[23] but do not exhibit a larger peak eversion angle.[10,22,27] Musculoskeletal models of the ITB have estimated that both strain and strain rate of the ITB are higher in individuals with ITBS.[13,23] Taken together, these studies suggest that increases in ITB strain and strain rate may be due to multiple and potentially competing biomechanical factors. Iliotibial band strain may be altered by proximal factors at the hip or distal factors at the foot, due to its attachment and insertion points.

Sex differences may contribute to the development of ITBS and complicate comparisons of biomechanical factors.[9,29,34] Healthy females tend to run with greater hip internal rotation, hip adduction, and knee abduction, 2 of which are associated with increased ITB strain.[9,34] Prospective studies have identified that females who develop ITBS exhibit greater hip adduction,[27] knee internal rotation,[27] and hip external rotation[29] than their healthy counterparts,

[1]Department of Human Physiology, University of Oregon, Eugene, OR. [2]Department of Kinesiology, Iowa State University, Ames, IA. The study was approved by the Institutional Review Board at Iowa State University. No funding was received for this study. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Evan M. Day, Bowerman Sports Science Clinic, B10 McArthur Court, 1601 University Street, Eugene, OR 97403. E-mail: eday5@uoregon.edu

each of which could increase strain of the ITB. In contrast, males with ITBS exhibit greater ankle adduction than healthy males, which may increase ITB strain due to joint coupling that increases internal rotation of the tibia relative to the femur.[19,29] It remains unknown, however, whether ITB strain and strain rate, potential predictors of ITBS,[13] are higher in healthy females compared to males due to natural kinematic differences.

Orthoses are commonly used in clinical settings for injury prevention and rehabilitation for distance runners.[32] Medial wedges reduce rearfoot eversion (coupled with tibial internal rotation) and reduce knee flexion velocity.[31] Decreases in these movement parameters may reduce ITB strain and strain rate, as the ITB crosses the lateral aspect of the knee and attaches on the tibia.

No study to date has investigated the effects of wedged orthoses and sex on ITB strain and strain rate. Existing literature indicates that medial wedges reduce rearfoot eversion, which may lead to potentially beneficial kinematic adjustments at the knee to reduce ITB strain and strain rate. In addition, previous studies indicate that differences in hip and knee kinematics may increase ITB strain in females. The purpose of this study was to determine the acute effects of wedge orthoses and sex on ITB strain and strain rate. Our first hypothesis was that medial wedges would result in lower ITB strain and strain rate than no wedge. Our second hypothesis was that females would display higher ITB strain and strain rate than males across wedge conditions.

METHODS

Participants

Thirty (15 male, 15 female) healthy recreational and competitive runners were recruited for this study (**TABLE 1**). Exclusion criteria for participants included currently running fewer than 25 km per week, lower extremity injury in the past 3 months, lower extremity surgery in the past year, currently using orthoses, and currently pregnant. Participants provided informed consent before data collection and completed a questionnaire about their running injury history after data collection. This study was approved by the Institutional Review Board at Iowa State University.

Data Collection

Segment coordinate axes were defined by placing retroreflective markers on the right shod foot, leg, and torso.[30] Markers were placed on the right toe, lateral dorsal aspect of the foot, heel, medial and lateral malleoli, anterior and lateral calf, medial and lateral tibial epicondyles, medial and lateral femoral epicondyles, and anterior and lateral thigh. Bilateral markers were placed on the greater trochanters, anterior superior iliac spines, posterior superior iliac spines, and acromia. Single markers were placed on the sacrum and fifth cervical vertebra. Participants performed a static trial by standing with their feet shoulder-width apart. Markers were then removed from the medial malleolus, tibial epicondyle, and femoral epicondyle so that they did not interfere with running movements. The removed markers were recreated during the dynamic trials based on the relative position and orientation of the remaining markers on the segment.

Participants wore spandex or running shorts and their own running shoes. There were 5 wedge conditions: 7° medial, 3° medial, no wedge, 3° lateral, and 7° lateral. Wedge angles were chosen to match other studies that have reported significant intervention effects.[15,18,31] The soft wedges were made from ethylene-vinyl acetate (Shore durometer type A of 75) and donated by Marathon Orthotics, Inc (Eden Prairie, MN) (**FIGURE 1**). The order of conditions was randomized between participants to balance potential effects of fatigue. Participants ran on the treadmill at a self-selected speed for 1 minute before each condition to get used to the new wedge.[25,35]

Dynamic trials were performed on a 30-m runway. Kinematic data were captured at 160 Hz with an 8-camera motion-capture system (Vicon; Oxford Metrics, Yarnton, UK), and kinetic data were captured at 1600 Hz with an in-ground force platform (Advanced Mechanical Technology, Inc, Watertown, MA). Participants were encouraged to rest as long as needed between trials to avoid potential effects of fatigue. Participants were instructed to look straight ahead while running to avoid targeting the force platform. A successful trial was defined as hitting the force platform with the entire right foot without any visual evidence of targeting. Participants were instructed to run at their preferred velocity that was indicative of normal training pace. Trials within ±5% of the preferred running velocity were considered acceptable for analysis. Running velocity was monitored by calculating the average an-

TABLE 1 Participant Characteristics*

	Sex		
	Male	Female	*P* Value
Age, y	21 ± 2	21 ± 2	.877
Mass, kg	72 ± 6	56 ± 5	<.001†
Height, m	1.83 ± 0.05	1.65 ± 0.08	<.001†
Running speed, m/s	4.0 ± 0.3	3.5 ± 0.4	<.001†
Weekly running distance, km	71 ± 26	53 ± 26	.129
Years running	7 ± 3	6 ± 2	.302
Best time in 5000-m run, min	16:27 ± 1:35	20:36 ± 2:54	<.001†

**Values are mean ± SD unless otherwise indicated.*
†Significant difference between sexes.

terior-to-posterior velocity of the sacral marker during the stance phase.[21] Conditions were completed when 5 acceptable trials were recorded.

Data Analysis

A custom MATLAB (Version R2016b; The MathWorks, Inc, Natick, MA) program was used to calculate kinematics, kinetics, and ITB strain. Kinematic parameters were determined throughout the stance phase, defined as the phase when the vertical ground reaction force exceeded 5% body weight.[6] The ankle joint center was calculated as the midpoint between the lateral and medial malleoli, the knee joint center as the midpoint between the medial and lateral femoral epicondyles, and the hip joint center as 25% of the distance between the left and right greater trochanters.[33] Marker data were filtered using a dual-pass, fourth-order, low-pass Butterworth filter with a 20-Hz cutoff frequency.[23] Joint angles were estimated using a Euler/Cardan rotation order of flexion/extension, abduction/adduction, and internal/external rotation.

Joint moments were estimated using an inverse-dynamics approach. A cutoff frequency of 20 Hz was used for force-plate data. Segment masses, centers of mass, and moments of inertia were individually estimated.[7] Joint moments were transformed to the distal segment coordinate system and normalized by body mass. All joint moments are reported as internal moments. The stance phase results were interpolated to 101 points for analysis.

A model of the ITB was developed by adapting the gait2392_simbody model in OpenSim (Version 3.3; https://simtk.org/projects/opensim).[8] The knee joint was modified from a single-degree-of-freedom joint (flexion/extension) to a 3-degrees-of-freedom joint. It was assumed that the ITB followed the same anatomical pathway as the tensor fascia latae.[4,13,20,23] For each of the 3 degrees of freedom of the knee and hip joints, a polynomial equation was fitted by exporting the tensor fascia latae length change as a function of joint angle, as defined in OpenSim (see the **APPENDIX**, available at www.jospt.org).[8] Iliotibial band resting length was individually adjusted to equal the distance that connected markers on the right anterior superior iliac spine, greater trochanter, lateral femoral epicondyle, and lateral tibial epicondyle. Iliotibial band strain in the neutral position was 3%, determined as the common intercept from the 6 ITB length equations when the joint angles were set to zero. Validation of an ITB model is difficult without invasive procedures, and thus several assumptions and limitations are present in the developed model. It was assumed that all individuals' ITBs had the same stiffness and that the distance between the connected markers represented the true length of the ITB. Limitations to the model are that it is passive and thus does not factor in muscle activation. Additionally, there was no set wrapping sphere to model impingement of the ITB against the lateral femoral epicondyle.[23]

Iliotibial band length was modeled as the additive sum of ITB length changes from the 6 joint-angle equations (3 hip, 3 knee) and the ITB resting length. For each of the 6 individual joint angles, a change in ITB length as a result of the joint angle is defined by a polynomial equation (**APPENDIX**).[8] These 6 individual changes in ITB length from each joint angle were then added together to estimate the total length change of the ITB from the 6 degrees of freedom of the hip and knee. This total length change was then added to the ITB resting length to estimate the length of the ITB at each time point. Iliotibial band strain during stance was computed using the following equation: $\text{Strain}_i = \frac{L_i - L_0}{L_0}$, where L_i is the ITB length at data point i and L_0 is the ITB resting length. Iliotibial band strain rate was calculated at each time step using the first central difference method: $\text{Strain rate}_i = \frac{\text{Strain}_{i+1} - \text{Strain}_{i-1}}{\text{Time}_{i+1} - \text{Time}_{i-1}}$.

Maximum values for kinematics, kinetics, ITB strain, and ITB strain rate were averaged across 5 trials per condition. Joint angles assessed were ankle dorsiflexion and inversion; knee flexion, adduction, and internal rotation; and hip flexion, adduction, and internal rotation. Joint moments assessed were ankle plantar flexion and inversion; knee extension and adduction; and hip extension and abduction. The multiple angles and moments were included in the analysis to understand individual contributions to the complex nature of ITB strain. A mixed-model (group by condition) multivariate analysis of covariance with a between-subject comparison of sex and

FIGURE 1. A soft wedge orthosis was inserted into the shoe for testing conditions.

within-subject comparison of wedge condition, and running speed as a covariate, was performed in SPSS (Version 23; IBM Corporation, Armonk, NY). Significant differences were set to $\alpha = .05$. When significant main effects were detected, Bonferroni post hoc comparisons were utilized to test for significant differences between wedge conditions. Partial eta square was used to determine effect sizes, defined as small ($\eta^2 = 0.01$), medium ($\eta^2 = 0.06$), or large ($\eta^2 = 0.14$).[5]

RESULTS

The multivariate analysis of covariance detected a significant wedge-by-sex interaction ($P = .044$, $\eta^2 = 0.188$). There was a significant within-subject main effect of wedge ($P = .017$, $\eta^2 = 0.200$) and a significant between-subject effect of sex ($P = .047$, $\eta^2 = 0.779$).

Iliotibial band strain ($P = .815$, $\eta^2 = 0.014$) and strain rate ($P = .872$, $\eta^2 = 0.011$) were not different between wedge conditions (**TABLE 2**, **FIGURES 2** and **3**). Females had higher strain (5.7% ± 0.6%)

TABLE 2 Kinematic Variables as a Function of Wedge Condition*

	Wedge Condition				
	7° Lateral	3° Lateral	No Wedge	3° Medial	7° Medial
Angle, deg					
Ankle dorsiflexion	23.8 ± 3.9	23.9 ± 3.7	23.8 ± 3.5	23.9 ± 3.5	23.9 ± 3.5
Ankle eversion	9.6 ± 5.4†‡§‖	7.8 ± 4.5‡§‖¶	6.6 ± 5.2‡‖¶	6.7 ± 5.3†‖¶	5.4 ± 5.3†‡§¶
Knee flexion	46.9 ± 6.1	47.6 ± 5.8	46.9 ± 5.9	47.4 ± 5.6	47.1 ± 5.9
Knee adduction	5.6 ± 6.0	5.7 ± 5.6	5.6 ± 5.6	5.5 ± 6.1	5.2 ± 5.4
Knee IR	11.9 ± 5.9	11.9 ± 5.9	12.4 ± 5.9	11.9 ± 6.4	11.8 ± 6.0
Hip flexion	31.3 ± 6.4	31.2 ± 5.9	31.4 ± 5.3	31.5 ± 5.3	30.9 ± 5.8
Hip adduction	13.6 ± 5.5	13.9 ± 4.8	13.8 ± 4.1	13.4 ± 4.7	13.5 ± 4.7
Hip IR	3.2 ± 8.7	2.3 ± 8.5	1.7 ± 7.7	3.5 ± 8.4	3.5 ± 8.6
Iliotibial band					
Strain, %	5.4± 0.9	5.5 ± 0.6	5.5 ± 0.7	5.5 ± 0.8	5.5 ± 0.7
Strain rate, %/s	40.5 ± 12.0	42.1 ± 12.7	39.9 ± 11.6	41.5 ± 13.4	42.5 ± 12.7

Abbreviation: IR, internal rotation.
**Values are mean ± SD.*
†Significantly different (P<.05) from 3° lateral wedge.
‡Significantly different (P<.05) from no wedge.
§Significantly different (P<.05) from 3° medial wedge.
‖Significantly different (P<.05) from 7° medial wedge.
¶Significantly different (P<.05) from 7° lateral wedge.

TABLE 3 Kinetic Variables as a Function of Wedge Condition*

	Wedge Condition				
Moment, Nm/kg	7° Lateral	3° Lateral	No Wedge	3° Medial	7° Medial
Ankle plantar flexion	2.62 ± 0.39†	2.64 ± 0.40	2.68 ± 0.39‡§	2.66 ± 0.40§	2.60 ± 0.39†‖
Ankle inversion	0.44 ± 0.12§	0.43 ± 0.13§	0.42 ± 0.13§	0.41 ± 0.13‡§	0.36 ± 0.14‡‖¶
Knee extension	2.91 ± 0.51	2.91 ± 0.57	2.84 ± 0.52	2.87 ± 0.51	2.91 ± 0.50
Knee adduction	1.00 ± 0.38§	1.02 ± 0.38§	1.01 ± 0.35§	1.04 ± 0.38§	1.10 ± 0.36†‡‖¶
Hip extension	1.91 ± 0.35§	1.94 ± 0.36	1.94 ± 0.35	1.97 ± 0.38	1.99 ± 0.37‡
Hip abduction	1.92 ± 0.30	1.91 ± 0.33	1.90 ± 0.27	1.89 ± 0.27	1.89 ± 0.32

**Values are mean ± SD.*
†Significantly different (P<.05) from no wedge.
‡Significantly different (P<.05) from 7° lateral wedge.
§Significantly different (P<.05) from 7° medial wedge.
‖Significantly different (P<.05) from 3° medial wedge.
¶Significantly different (P<.05) from 3° lateral wedge.

than males (5.3% ± 0.7%) (P = .031, η^2 = 0.161) and higher strain rate (45.1%/s ± 10.0%/s) than males (34.2%/s ± 11.4%/s) (P = .004, η^2 = 0.267) (**FIGURES 2** and **3**).

Maximum ankle eversion angles were significantly different between wedge conditions (P<.001, η^2 = 0.753) (**TABLE 2**). The ankle eversion angles were significantly higher with the 7° lateral and 3° lateral wedges and significantly lower with the 7° medial wedge compared to those with no wedge. Maximum knee adduction, knee internal rotation, and hip internal rotation angles were significantly dependent on sex. Males had greater peak knee adduction (7.6° ± 4.6°) than females (2.3° ± 1.6°) (P = .001, η^2 = 0.344) and greater knee internal rotation (14.8° ± 3.5°) than females (9.0° ± 4.1°) (P = .012, η^2 = 0.213). Females had greater hip internal rotation angles (5.9° ± 4.1°) than males (−0.5° ± 7.0°) (P = .048, η^2 = 0.137).

Ankle plantar flexion, ankle inversion, knee adduction, and hip extension moments were significantly different between wedge conditions (**TABLE 3**). The 7° medial wedge resulted in significantly lower ankle inversion moments than no wedge. In addition, the 7° medial wedge resulted in significantly higher knee adduction moments than no wedge.

DISCUSSION

The purpose of this study was to investigate the acute effects of wedge orthoses and sex on ITB strain and strain rate. Our first hypothesis was that ITB strain and strain rate would decrease with medial wedges. Our results fail to support this hypothesis. Iliotibial band strain and strain rate were not significantly different between wedge conditions (**TABLE 2**). The main reason for this finding is presumably the lack of change in knee and hip kinematics between wedge conditions. Our results support our second hypothesis, that females would have greater ITB strain and strain rate than males.

Greater peak ITB strain and strain rate in females compared to males is a novel finding of this study. Previous studies investigating ITB mechanics have assessed females or males separately,[11,13,27,28] most likely due to sex kinematic differences during running that may affect ITB mechanics.[9,29,34] Previous research has indicated that ITB strain and strain rate are greater in individuals with ITBS.[13,23] Our finding of greater ITB strain and strain rate in females may be a factor that contributes to the higher prevalence of ITBS in females.[13,32] However, due to the acute and nonfatiguing nature of this study, other etiological factors may be at play that contribute to the more prevalent development of ITBS in females.

Our finding of increased peak hip internal rotation in females compared to males is in agreement with previous findings,[9] as are decreased peak knee internal rotation[34] and knee adduction.[9,34] Our finding of no difference in peak hip adduction between sexes is in agreement with one previous report[14] but contrary to other previous studies.[9,34] Males did have a faster running speed than females (**TABLE 1**), but it is unknown whether the difference in running speed is related to a lack of difference in hip adduction angles. While increased peak hip adduction and knee internal rotation have been prospectively identified as indicators of developing ITBS,[27] these kinematic parameters had weak correlations with peak strain and strain rate in a previous study.[13] Based on the ITB length equations used in the current study and identified sex differences, it appears that addressing an increase in peak hip internal rotation angles may be an important factor in reducing ITB strain and strain rate in females.

The differences in ITB strain and strain rate between sexes are of similar magnitude to studies investigating changes in response to step width[20] and

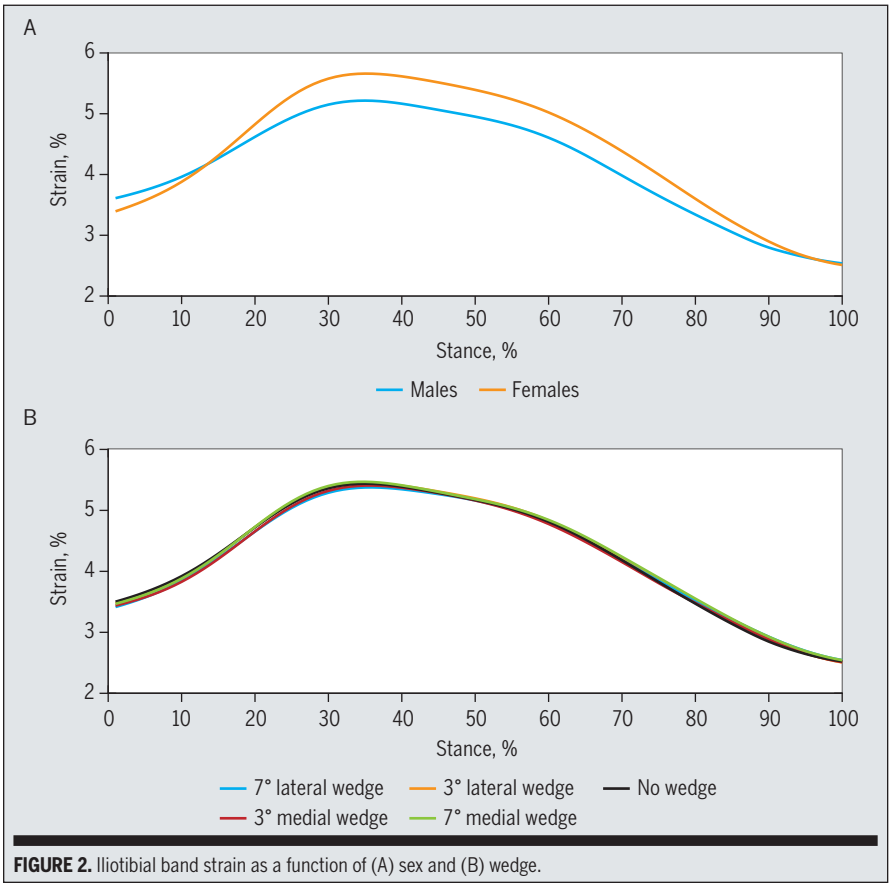


FIGURE 2. Iliotibial band strain as a function of (A) sex and (B) wedge.

stride length.[4] Peak ITB strain was 0.4% lower in males than in females, similar to changes of 0.3% lower with shorter stride length[4] and 0.5% lower with widened step width.[20] The observed 10.9%/s difference in peak strain rate between sexes is larger than reported differences in step width[20] and step length,[4] but slightly less than the reported difference of 12.7%/s between individuals with ITBS and healthy controls.[13] The difference in peak strain rate between males and females being similar to the difference between individuals who developed ITBS and those who did not may suggest that females are more at risk of developing ITBS.

As expected, the type of wedge produced systematic changes in maximum ankle eversion angles, consistent with previous studies.[16,17,26] Eversion angles were reduced by a total of 4.2° between the 7° lateral and 7° medial wedges (**TABLE 2**). According to the ratio of 1° of tibial internal rotation for every 1.2° to 1.8° of calcaneal eversion,[19] there should have been a 2.3° to 3.5° change in knee internal rotation between the 7° lateral and 7° medial wedges. However, our results indicate no changes in knee joint angles, contrary to previous studies that have reported decreases in internal tibial rotation and knee abduction using custom foot orthoses.[17,26,31] Additionally, we also observed no changes in hip kinematics. This is in contrast to Boldt et al,[3] who reported a small but significant decrease in hip adduction excursion of 0.6° using medial wedges. A reason for our differing results may be the type of orthosis used, or the variable subject-specific response to the use of an orthosis.[1] The lack of change in internal tibial rotation resulted in no subsequent kinematic-chain compensations[16] and no change in ITB strain between wedge conditions. Examination of individual participant ITB strain curves in response to each wedge revealed no systematic changes, presumably due to no observed changes in knee or hip kinematics. These results suggest that the use of wedge orthoses may not affect ITB strain or strain rate, even on an individual basis, in healthy runners. While it is unknown whether individuals with symptomatic ITBS would respond differently, our results suggest that wedge orthoses may not be well suited for prevention or rehabilitation of ITBS.

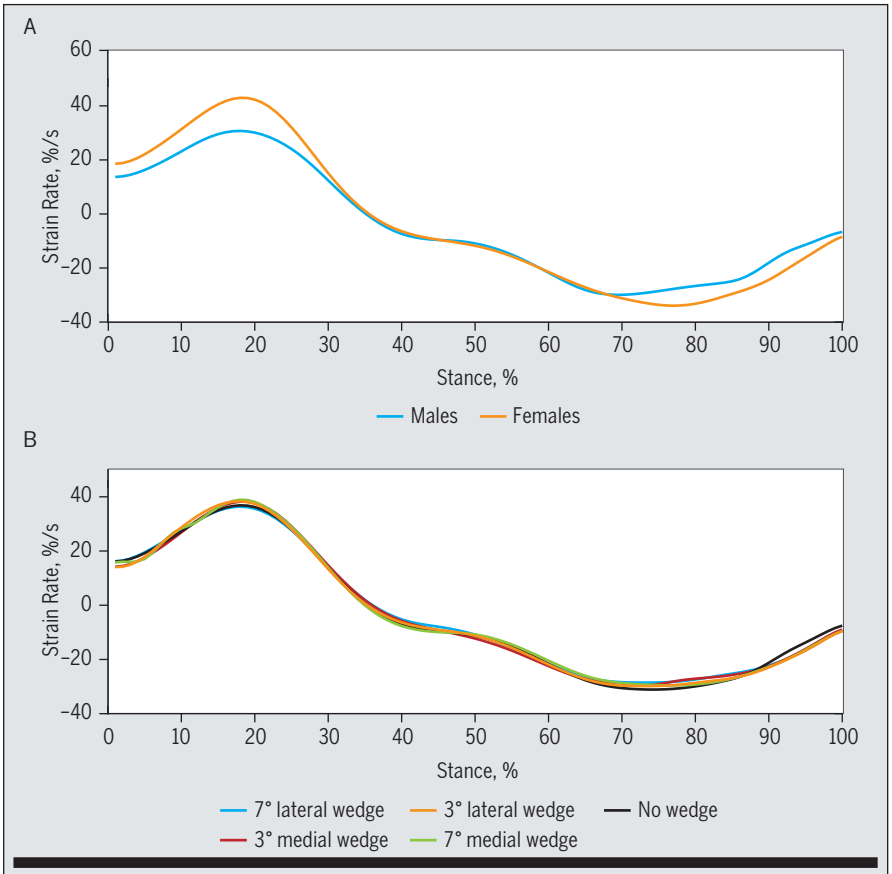


FIGURE 3. Iliotibial band strain rate as a function of (A) sex and (B) wedge.

This study is not without limitations, the first of which is the use of standard nonfitted wedges for all participants. Use of custom foot orthoses may lead to changes not observed in this study. Further, having participants use their own footwear increases external validity, but may have resulted in wedge effects being masked between footwear of varying midsole material. The musculoskeletal model used is limited by intersubject variability and the complexity of the ITB structure. Validation would be difficult without invasive measures, but the strain magnitudes (4%-7%) were within normal in vivo limits of elastic tissue[24] and lower than the failure point identified by cadaver testing.[2] The estimated ITB mechanics also do not reflect whether participants have a tighter or stiffer ITB, or factor in muscle activation, as our model is similar to previous passive tissue ITB models.[4,13,20,23] Last, investigating the effectiveness of wedge orthoses for prevention or rehabilitation of ITBS would be better served with a longitudinal study design that involved symptomatic individuals with ITBS.[11-13,23,28] Additionally, inclusion of a fatigue protocol, which has been associated with ITBS symptom onset, would be of benefit to understanding the effectiveness of the wedge orthosis.[28]

CONCLUSION

A 6-degrees-of-freedom ITB model was developed to analyze the effects of wedge orthoses and sex on ITB strain and strain rate during running. There was no evidence that the

use of medial or lateral wedges affects ITB strain and strain rate. Sex comparisons revealed that healthy females have greater ITB strain and strain rate than males. Clinicians treating patients for prevention or rehabilitation of ITBS should be aware that footwear interventions such as wedge orthoses may not be beneficial. ◉

KEY POINTS

FINDINGS: The acute use of medial or lateral wedge orthoses does not influence iliotibial band (ITB) strain and strain rate. Females tend to exhibit higher ITB strain and strain rate than males.

IMPLICATIONS: Clinicians should be aware that medial and lateral wedges might not cause acute changes in ITB strain or strain rate.

CAUTION: It is speculative which kinematics most strongly influence the higher ITB strain in females. The model of the ITB is difficult to validate, and results should be interpreted while taking into account the assumptions and limitations of the model.

REFERENCES

1. Bates BT. Single-subject methodology: an alternative approach. *Med Sci Sports Exerc*. 1996;28:631-638.
2. Birnbaum K, Siebert CH, Pandorf T, Schopphoff E, Prescher A, Niethard FU. Anatomical and biomechanical investigations of the iliotibial tract. *Surg Radiol Anat*. 2004;26:433-446. https://doi.org/10.1007/s00276-004-0265-8
3. Boldt AR, Willson JD, Barrios JA, Kernozek TW. Effects of medially wedged foot orthoses on knee and hip joint running mechanics in females with and without patellofemoral pain syndrome. *J Appl Biomech*. 2013;29:68-77. https://doi.org/10.1123/jab.29.1.68
4. Boyer ER, Derrick TR. Select injury-related variables are affected by stride length and foot strike style during running. *Am J Sports Med*. 2015;43:2310-2317. https://doi.org/10.1177/0363546515592837
5. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
6. Day EM, Hahn ME. A comparison of metatarsophalangeal joint center locations on estimated joint moments during running. *J Biomech*. 2019;86:64-70. https://doi.org/10.1016/j.jbiomech.2019.01.044
7. de Leva P. Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. *J Biomech*. 1996;29:1223-1230. https://doi.org/10.1016/0021-9290(95)00178-6
8. Delp SL, Loan JP, Hoy MG, Zajac FE, Topp EL, Rosen JM. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Trans Biomed Eng*. 1990;37:757-767. https://doi.org/10.1109/10.102791
9. Ferber R, Davis IM, Williams DS, 3rd. Gender differences in lower extremity mechanics during running. *Clin Biomech (Bristol, Avon)*. 2003;18:350-357. https://doi.org/10.1016/S0268-0033(03)00025-1
10. Ferber R, Noehren B, Hamill J, Davis IS. Competitive female runners with a history of iliotibial band syndrome demonstrate atypical hip and knee kinematics. *J Orthop Sports Phys Ther*. 2010;40:52-58. https://doi.org/10.2519/jospt.2010.3028
11. Foch E, Milner CE. Frontal plane running biomechanics in female runners with previous iliotibial band syndrome. *J Appl Biomech*. 2014;30:58-65. https://doi.org/10.1123/jab.2013-0051
12. Grau S, Krauss I, Maiwald C, Axmann D, Horstmann T, Best R. Kinematic classification of iliotibial band syndrome in runners. *Scand J Med Sci Sports*. 2011;21:184-189. https://doi.org/10.1111/j.1600-0838.2009.01045.x
13. Hamill J, Miller R, Noehren B, Davis I. A prospective study of iliotibial band strain in runners. *Clin Biomech (Bristol, Avon)*. 2008;23:1018-1025. https://doi.org/10.1016/j.clinbiomech.2008.04.017
14. Hannigan JJ, Osternig LR, Chou LS. Sex-specific relationships between hip strength and hip, pelvis, and trunk kinematics in healthy runners. *J Appl Biomech*. 2018;34:76-81. https://doi.org/10.1123/jab.2016-0333
15. Lewinson RT, Fukuchi CA, Worobets JT, Stefanyshyn DJ. The effects of wedged footwear on lower limb frontal plane biomechanics during running. *Clin J Sport Med*. 2013;23:208-215. https://doi.org/10.1097/JSM.0b013e31826b7c83
16. MacLean C, Davis IM, Hamill J. Influence of a custom foot orthotic intervention on lower extremity dynamics in healthy runners. *Clin Biomech (Bristol, Avon)*. 2006;21:623-630. https://doi.org/10.1016/j.clinbiomech.2006.01.005
17. MacLean CL, Davis IS, Hamill J. Influence of running shoe midsole composition and custom foot orthotic intervention on lower extremity dynamics during running. *J Appl Biomech*. 2009;25:54-63. https://doi.org/10.1123/jab.25.1.54
18. MacLean CL, van Emmerik R, Hamill J. Influence of custom foot orthotic intervention on lower extremity intralimb coupling during a 30-minute run. *J Appl Biomech*. 2010;26:390-399. https://doi.org/10.1123/jab.26.4.390
19. McClay I, Manal KT. Coupling parameters in runners with normal and excessive pronation. *J Appl Biomech*. 1997;13:109-124. https://doi.org/10.1123/jab.13.1.109
20. Meardon SA, Campbell S, Derrick TR. Step width alters iliotibial band strain during running. *Sports Biomech*. 2012;11:464-472. https://doi.org/10.1080/14763141.2012.699547
21. Meardon SA, Willson JD, Gries SR, Kernozek TW, Derrick TR. Bone stress in runners with tibial stress fracture. *Clin Biomech (Bristol, Avon)*. 2015;30:895-902. https://doi.org/10.1016/j.clinbiomech.2015.07.012
22. Messier SP, Edwards DG, Martin DF, et al. Etiology of iliotibial band friction syndrome in distance runners. *Med Sci Sports Exerc*. 1995;27:951-960.
23. Miller RH, Lowry JL, Meardon SA, Gillette JC. Lower extremity mechanics of iliotibial band syndrome during an exhaustive run. *Gait Posture*. 2007;26:407-413. https://doi.org/10.1016/j.gaitpost.2006.10.007
24. Monti RJ, Roy RR, Zhong H, Edgerton VR. Mechanical properties of rat soleus aponeurosis and tendon during variable recruitment *in situ*. *J Exp Biol*. 2003;206:3437-3445. https://doi.org/10.1242/jeb.00550
25. Mündermann A, Nigg BM, Humble RN, Stefanyshyn DJ. Consistent immediate effects of foot orthoses on comfort and lower extremity kinematics, kinetics, and muscle activity. *J Appl Biomech*. 2004;20:71-84. https://doi.org/10.1123/jab.20.1.71
26. Mündermann A, Nigg BM, Humble RN, Stefanyshyn DJ. Foot orthotics affect lower extremity kinematics and kinetics during running. *Clin Biomech (Bristol, Avon)*. 2003;18:254-262. https://doi.org/10.1016/S0268-0033(02)00186-9
27. Noehren B, Davis I, Hamill J. ASB Clinical Biomechanics Award winner 2006: prospective study of the biomechanical factors associated with iliotibial band syndrome. *Clin Biomech (Bristol, Avon)*. 2007;22:951-956. https://doi.org/10.1016/j.clinbiomech.2007.07.001
28. Noehren B, Schmitz A, Hempel R, Westlake C, Black W. Assessment of strength, flexibility, and running mechanics in men with iliotibial band syndrome. *J Orthop Sports Phys Ther*. 2014;44:217-222. https://doi.org/10.2519/jospt.2014.4991
29. Phinyomark A, Osis S, Hettinga BA, Leigh R, Ferber R. Gender differences in gait kinematics in runners with iliotibial band syndrome. *Scand J Med Sci Sports*. 2015;25:744-753. https://doi.org/10.1111/sms.12394
30. Robertson DGE, Caldwell GE, Hamill J, Kamen G, Whittlesey SN. *Research Methods in Biomechanics*. Champaign, IL: Human Kinetics; 2004.
31. Stackhouse CL, Davis IM, Hamill J. Orthotic intervention in forefoot and rearfoot strike running patterns. *Clin Biomech (Bristol, Avon)*. 2004;19:64-70. https://doi.org/10.1016/j.clinbiomech.2003.09.002
32. Taunton JE, Ryan MB, Clement DB, McKenzie

DC, Lloyd-Smith DR, Zumbo BD. A retrospective case-control analysis of 2002 running injuries. *Br J Sports Med*. 2002;36:95-101. https://doi.org/10.1136/bjsm.36.2.95

33. Weinhandl JT, O'Connor KM. Assessment of a greater trochanter-based method of locating the hip joint center. *J Biomech*. 2010;43:2633-2636. https://doi.org/10.1016/j.jbiomech.2010.05.023

34. Willson JD, Petrowitz I, Butler RJ, Kernozek TW. Male and female gluteal muscle activity and lower extremity kinematics during running. *Clin Biomech (Bristol, Avon)*. 2012;27:1052-1057. https://doi.org/10.1016/j.clinbiomech.2012.08.008

35. Willwacher S, König M, Potthast W, Brüggemann GP. Does specific footwear facilitate energy storage and return at the metatarsophalangeal joint in running? *J Appl Biomech*. 2013;29:583-592. https://doi.org/10.1123/jab.29.5.583

APPENDIX

The following are the ITB length change (meters) equations used for development of the 6-degrees-of-freedom model to calculate ITB strain. The equations are adapted from the gait2392_simbody model in OpenSim (Version 3.3).[8] The respective joint angles (degrees) are represented by θ in each polynomial equation.

ITB length change (hip flexion) = $-0.0000074401\theta^2 - 0.0006040\theta$
ITB length change (hip adduction) = $-0.0000077796\theta^2 + 0.0008703\theta$
ITB length change (hip internal rotation) = $-0.000000000006371\theta^4 + 0.0000000004483\theta^3 + 0.0000002749\theta^2 - 0.0001012\theta$
ITB length change (knee flexion) = $-0.0000009994\theta^2 + 0.0002354\theta$
ITB length change (knee adduction) = 0.0006308θ
ITB length change (knee internal rotation) = $0.0000003984\theta^2 - 0.0000319 2\theta$

Abbreviation: ITB, iliotibial band.
OpenSim equations used with permission (https://creativecommons.org/licenses/by/3.0/).

HESTER L. DEN BANDT, PT[1-3] • WINIFRED D. PAULIS, PhD[1] • DAVID BECKWÉE, PhD[2,4]
KELLY ICKMANS, PhD[2-4] • JO NIJS, PhD[2-4] • LENNARD VOOGT, PhD[1,3]

Pain Mechanisms in Low Back Pain: A Systematic Review With Meta-analysis of Mechanical Quantitative Sensory Testing Outcomes in People With Nonspecific Low Back Pain

Nonspecific low back pain (LBP) is one of the most common health problems and places an enormous burden on individuals, their families, and society.[1-4] Nonspecific LBP is pain felt at the lower back, between the lower rib and gluteal fold, for which no specific pathophysiological process can be designated.[1]

Current guidelines for nonspecific LBP suggest biopsychosocial approaches and individually tailored interventions, consisting of combinations of education, exercise, and hands-on treatment.[8] In cases where monodisciplinary approaches fall short of success, multidisciplinary biopsychosocial rehabilitation is indicated.[1] Although the success of these interventions is well demonstrated, effect sizes are still generally small and recurrence rates are high.[8,21] There is a clear need for improvements in the management of nonspecific LBP. One suggestion is to better align treatments for LBP with the underlying biological processes.[15,27]

Changes in the neurophysiological processing of nociceptive information may play an important role in nonspecific LBP.[2,15] Amplification of peripheral nociceptive information at the height of the dorsal horn, enhanced processing of nociceptive information within several brain nuclei, and their interrelated connections that together form a "dynamic pain connectome" are taken as important biological processes that should be considered in nonspecific LBP.[28] Enhanced processing of nociceptive information is currently summarized as central sensitization[15,42]—"an amplification of neural signalling within the central nervous system that elicits pain hypersensitivity."[49]

BACKGROUND: Mechanical quantitative sensory testing (QST) assesses sensory functioning and detects functional changes in (central) nociceptive processing. It has been hypothesized that these functional changes might be apparent in people with nonspecific low back pain (LBP), although the results are mixed.

OBJECTIVE: The aim of this systematic review was to examine whether sensory function, measured with QST, was altered in people with nonspecific LBP.

METHODS: This systematic review was conducted according to PRISMA guidelines. Six databases were searched for relevant literature. Studies comparing mechanical QST measures involving people with subacute and chronic LBP and healthy controls were included if (1) pressure pain thresholds (PPTs), (2) temporal summation, or (3) conditioned pain modulation were reported. Risk of bias was assessed using the Newcastle-Ottawa scale. When possible, the results from different studies were pooled.

RESULTS: Twenty-four studies were included. Scores on the Newcastle-Ottawa scale varied between 1 and 6 points. People with nonspecific LBP, compared to healthy controls, had significantly lower PPTs at remote sites and increased temporal summation at the lower back. The PPTs measured at the scapula were significantly lower in patients with nonspecific LBP than in healthy controls (pooled mean difference, 119.2 kPa; 95% confidence interval: 91.8, 146.6 kPa; $P<.001$).

CONCLUSION: The PPT measurements at remote body parts were significantly lower in people with nonspecific LBP compared with healthy controls. Temporal summation and conditioned pain modulation measurements had mixed outcomes.

LEVEL OF EVIDENCE: Therapy, level 3a.
J Orthop Sports Phys Ther 2019;49(10):698-715. Epub 23 Aug 2019. doi:10.2519/jospt.2019.8876

KEY WORDS: *central sensitization, conditioned pain modulation, low back pain, pressure pain threshold, temporal summation*

[1]Department of Physiotherapy, Rotterdam University of Applied Sciences, Rotterdam, the Netherlands. [2]Department of Physiotherapy, Human Physiology and Anatomy, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels, Belgium. [3]Pain in Motion International Research Group, Brussels, Belgium. [4]Department of Physical Medicine and Physiotherapy, University Hospital Brussels, Brussels, Belgium. The review protocol was registered a priori at the International Prospective Register of Systematic Reviews (registration number CRD42017055599). There was no external funding for the preparation of this manuscript. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Hester L. den Bandt, Rochussenstraat 198, 3015 EK Rotterdam, the Netherlands. E-mail: h.l.den.bandt@hr.nl

From a clinical perspective, it is valuable to know whether central sensitization is part of the nonspecific LBP problem. Central sensitization is associated with higher pain intensity, widespread pain, worse prognosis, and lower quality of life.[41,43] Central sensitization is a neurophysiological concept, and the underlying processes cannot directly be measured in clinical practice. Quantitative sensory testing (QST) is used to study altered sensory processing, as a derivative of signs of central sensitization.[2,3]

Central sensitization is suggested to be the dominant pain mechanism in about 25% of the population with nonspecific LBP.[37] A previous narrative review reported on differences between people with chronic LBP and healthy controls in several QST measures. Higher pain thresholds at remote body parts, enhanced temporal summation, and abnormal conditioned pain modulation were interpreted as signs of central sensitization.[4,10,18,23,26,33] A narrative review does not systematically screen the available literature, may not be comprehensive, does not take methodological quality of included studies into account, and does not statistically pool data to generate firm conclusions.

We performed a systematic review to examine whether sensory function, measured with QST, was altered in people with nonspecific LBP compared with healthy controls. We aimed to critically appraise current literature comparing remote pressure pain thresholds (PPTs), local and remote temporal summation, and conditioned pain modulation in people with nonspecific LBP and healthy controls to examine whether sensory functioning, measured with QST, is altered in people with nonspecific LBP.

METHODS

Protocol and Registration

The review protocol was registered a priori at the International Prospective Register of Systematic Reviews (registration number CRD42017055599). This systematic review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (www.prisma-statement.org).

Eligibility Criteria

Studies were included if the following criteria were met: (1) studies involved adults (18 years of age or older) with nonspecific LBP (subacute and chronic) and healthy controls; (2) sensory functioning was determined by using PPT, mechanical temporal summation, and/or conditioned pain modulation measures; and (3) studies had to be written in Dutch, English, or German. Subacute nonspecific LBP is defined as pain that has been present between 6 and 12 weeks.[46] Chronic nonspecific LBP is defined as pain that persists for at least 12 weeks.[1] Various QST procedures are described in the literature. The PPT is defined as the minimum amount of pressure that elicits a painful sensation.[5] Temporal summation is the increased pain response after a series of identical stimuli.[23] Conditioned pain modulation is the increase in PPT after a painful stimulus on a remote body part.[3] To enable meta-analysis, only studies using mechanical procedures were chosen. Central sensitization can be a normal physiological phenomenon during the acute LBP phase, but will resolve in most cases.[16] Studies involving patients with subacute and/or chronic LBP were included in the meta-analysis, as the difference between these 2 groups cannot clearly be delineated from a pain physiological perspective, but rather stems from epidemiological convention. Central sensitization can be apparent in both groups. Studies involving people with sciatica, pelvic problems, pregnancy, whiplash-associated disorders, nonspecific neck pain, fibromyalgia, low back surgery, or any other medical condition besides nonspecific LBP were excluded.

Information Sources and Search Strategy

Literature was searched up to January 7, 2019 in MEDLINE, the Cochrane Library, Google Scholar, Web of Science, CINAHL, and Embase. An information specialist from the medical library of the Erasmus University Medical Center (Rotterdam, the Netherlands) constructed search strategies for the different databases. The main key words were *central sensitization*, *pain threshold*, *hyperalgesia*, *hypoalgesia*, *quantitative sensory testing*, *wind-up*, *conditioned pain modulation*, *low back pain*, *inhibition and facilitation*, and *synonyms*. The search string for MEDLINE is displayed in **APPENDIX A** (available at www.jospt.org).

Study Selection

After removal of duplicates, the titles and abstracts of retrieved articles were screened for relevance by 2 independent investigators (H.d.B. and W.P.). Full-text versions of relevant articles were obtained and assessed for eligibility by the same 2 investigators. If there was uncertainty about whether an article fit the criteria, a third investigator (L.V.) was consulted and made the final decision. Corresponding authors of original studies were contacted in an attempt to obtain extra information when necessary.

Risk of Bias in Individual Studies

Risk of bias was assessed independently by H.d.B. and W.P. The Newcastle-Ottawa quality assessment scale (NOS) for nonrandomized studies, including case-control studies and cohort studies, was used.[48] The NOS has a "star rating system" in which a study is assessed on 3 aspects: selection of the study groups, comparability of the groups, and ascertainment of the exposure or outcome of interest.[48] Each aspect contains several items that can be scored with 1 star (except "comparability," which can have up to 2 stars; see **APPENDIX B**, available at www.jospt.org). This process leads to a score between 0 and 9 stars.[44] Investigators assessed the included studies independently. Interrater agreement was calculated (kappa and 95% confidence interval [CI]) using SPSS Version 24 (IBM Corporation, Armonk, NY). Disagreements were solved through discussion. When necessary, the

third investigator (L.V.) determined the final score.

Data Extraction and Data Items

The following data were extracted from the included articles: authors and year of publication; number of participants; definition of nonspecific LBP; study design; QST measures; location of QST stimuli and temporal summation protocol; PPT, temporal summation, and conditioned pain modulation results; and study conclusions. Data were extracted by both investigators independently. In case of missing data, authors were contacted and requested to provide required information.

Data Management and Meta-analysis

In most articles, results of PPTs, temporal summation, and conditioned pain modulation were reported as mean, 95% CI, standard deviation, and *P* value. All data on PPT outcomes from individual articles were recorded or converted to the unit kilopascals. Studies were grouped based on the applied QST protocol (remote PPT, temporal summation, conditioned pain modulation, or local temporal summation) and further clustered according to the remote body location (scapula, arm, hand, gluteal, lower leg, and lumbar). If a cluster contained at least 2 studies reporting means and standard deviations for patients with nonspecific LBP and healthy controls, a meta-analysis was performed for PPT and temporal summation outcomes using a binary random-effects meta-analysis model. Meta-analyses were performed using Review Manager software (Version 5.3; The Nordic Cochrane Centre, Copenhagen, Denmark). Meta-analyses for temporal summation were pooled based on identical remote body locations, temporal summation protocols, and outcome units. Heterogeneity was assessed using the I^2 statistic. For the interpretation of I^2 values, the following classification was used: 0% to 40%, no heterogeneity; 30% to 60%, moderate; 50% to 90%, substantial; and 75% to 100%, considerable heterogeneity.[13] If heterogeneity was higher than 60% (predetermined) and a subgroup contained at least 3 articles, then studies were pooled according to their NOS score and divided into below average and average or above average scores.[20,40] If the *P* value of "the overall effect" of the meta-analysis was smaller than .05 (predetermined), then the effect was considered significant. Studies not included in the meta-analysis were described separately. Funnel plots were created and inspected for publication bias (asymmetrical figure; **APPENDIX C**, available at www.jospt.org). A meta-analysis was not performed for conditioned pain modulation because of differences in measurement protocols. Some studies used cold or hot water, while other studies used a thermode, as a noxious stimulus. In some studies, the participants had to immerse their foot, leg, or hand in a bucket of ice water.[4,22,25,32] In another study, the participants had to immerse their hand in a bucket of hot water.[12] In one study, the noxious stimulus was applied with a thermode on the dorsal part of the hand.[35] The temporal summation measurements were more uniform across studies. Most of the temporal summation protocols referred to the German Research Network on Neuropathic Pain, and the remaining used temporal summation protocols similar to that of the German Research Network on Neuropathic Pain.[36]

RESULTS

Study Selection

The search yielded 6801 articles. The flow chart of inclusion is shown in **FIGURE 1**. After removing duplicates (n = 4198), the remaining 2603 articles were screened by title and abstract. Full texts of 62 articles were read. Finally, 24 articles were included in this review.[3-7,9,10,12,18,19,22,24-26,30-35,38,39,45,50] The corresponding authors of 2 publications were contacted with the request to provide the required details for meta-analysis. Both authors responded and delivered the required information.

Study Characteristics

Study characteristics are shown in **TABLE 1**. In all studies, different measurements were taken at the same moment. All studies used PPT as an outcome measure, except the study by Meints et al.[24] Seven studies involved temporal summation[3,9,12,26,32,34,45] and 6 studies involved conditioned pain

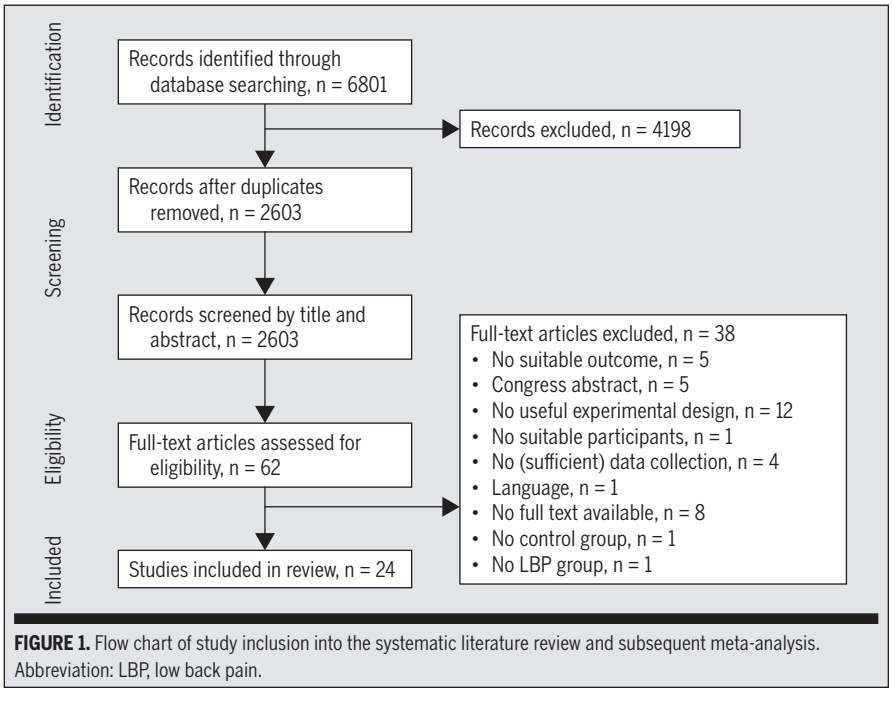


FIGURE 1. Flow chart of study inclusion into the systematic literature review and subsequent meta-analysis. Abbreviation: LBP, low back pain.

modulation.[4,12,22,25,32,35] Eight studies conducted PPT measurements and temporal summation measurements.[3,9,12,22,26,32,34,45] In about half of the studies (n = 13), patients and controls were appropriately matched for age and sex.[4,10,12,22,25,26,30-32,34,35,39,45] In 21 studies, PPTs were taken at both the lower back and a remote body site (eg, forehead, thenar eminence, wrist, hand, infraspinatus, triceps brachii, gluteus maximus, or second toe). In one study, only the lumbar area was tested using conditioned pain modulation.[35] In another study, only the remote hand was tested using temporal summation.[24]

TABLE 1 Characteristics of Included Studies (n = 24)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Blumenstiel et al[3]	n = 23 with chronic back pain (all female); mean ± SD age, 43.4 ± 8.6 y and n = 20 healthy controls (female-male ratio not mentioned); mean ± SD age, 38.3 ± 7.6 y	The presence of back pain for at least 45 d within the last 3 mo	Cross-sectional	PPT and TS	Paraspinal muscles and dorsum of the hand. Chronic back pain: on the most painful area in the back and on the hand (dorsum of the same side of the body as healthy control site) TS protocol: the ratings of single pinprick stimulation were compared with those of a series of 10 repeated pinprick stimuli of the same force (256 mN) over the same area; TS was calculated by dividing the mean ratings of the series by the mean pain ratings of single stimuli	PPT local (back): chronic back pain versus healthy control, 239.3 kPa (95% CI: 200, 287) versus 352 kPa (95% CI: 286, 432); *P*<.01 TS local (back): chronic back pain versus healthy control, 2.36 (95% CI: 1.74, 3.21) versus 3.61 (95% CI: 2.56, 5.11); NS PPT remote (hand): chronic back pain versus healthy control, 345 kPa (95% CI: 301, 394) versus 318 kPa (95% CI: 273, 370); NS TS remote (hand): chronic back pain versus healthy control, 3.57 (95% CI: 2.74, 4.67) versus 2.81 (95% CI: 2.07, 3.82); NS
Corrêa et al[4]	n = 30 with LBP (18 female, 12 male); mean ± SD age, 51 ± 8.7 y and n = 30 healthy controls (18 female, 12 male); mean ± SD age, 47 ± 7.7 y	Classified as having nonspecific chronic LBP using the diagnostic triage, as recommended by the European guidelines, as well as by the APTA guidelines	Case-control	PPT and CPM	Bilateral; 5 cm lateral to the L3 spinous process and 5 cm lateral to the L5 spinous process and TA of the right leg 5 cm from tibial tuberosity	PPT local (lumbar): LBP versus healthy control, 253.0 ± 96.5 kPa versus 342.5 ± 127.7 kPa (95% CI: 40.9, 131.1); *P* = .001 PPT remote (TA): LBP versus healthy control, 262.4 ± 93.1 kPa versus 321.8 ± 84.5 kPa (95% CI: 13.5, 105.4); *P* = .012 CPM: LBP versus healthy control, −47.17 ± 73.3 versus 71.4 ± 83.8 (95% CI: 77.9, 159.2); *P*<.001
Farasyn and Meeusen[7]	n = 87 with nonspecific LBP (39 female, 48 male); mean ± SD age, 43 ± 13 y and n = 64 healthy controls (40 female, 24 male); mean ± SD age, 40 ± 11 y	Subacute LBP as defined by the Dutch guideline for physical therapy	Cross-sectional	PPT	Paravertebral musculature (5 cm lateral) at T6, T10, L1, L3, 4 cm lateral to L5, 3 cm lateral to the iliac crest of the gluteus maximus, gluteus medius, tensor fascia latae, midpoint of the left triceps brachii	PPT local (L3): nonspecific LBP versus healthy control, 5.1 ± 1.3 kg/cm^2 versus 7.7 ± 1.7 kg/cm^2; *P*<.001 PPT local (L5): nonspecific LBP versus healthy control, 7.2 ± 1.6 kg/cm^2 versus 9.5 ± 1.2 kg/cm^2; *P*<.001 PPT remote (triceps brachii): nonspecific LBP versus healthy control, 6.7 ± 1.8 kg/cm^2 versus 7.1 ± 1.7 kg/cm^2; *P* = .119 PPT remote (tensor fascia latae): nonspecific LBP versus healthy control, 6.3 ± 1.5 kg/cm^2 versus 7.1 ± 1.4 kg/cm^2; *P*<.001
Farasyn and Meeusen[6]	n = 48 with LBP (25 female, 23 male); mean ± SD age, 45 ± 13 y and n = 64 healthy controls (38 female, 26 male); mean ± SD age, 40 ± 11 y	Nonspecific LBP is defined as pain for which no disorder in the anatomical structure can be found to sufficiently account for the patient's complaints	Clinical trial	PPT	Midpoint of the left triceps brachii, paravertebral musculature (erector spinae) 5 cm from L1 and L3 and 4 cm from L5, 3 cm below iliac crest from proximal part of the gluteus maximus (back pain–related site)	PPT local (L5): LBP versus healthy control, 7.3 ± 1.7 versus 9.5 ± 1.2 kg/cm^2; *P*<.001 PPT remote (triceps brachii): LBP versus healthy control, 6.9 ± 1.5 versus 7.1 ± 1.7 kg/cm^2; *P* = .457 PPT remote (gluteus maximus): LBP versus healthy control, 6.4 ± 1.6 versus 8.0 ± 1.5 kg/cm^2; *P*<.001

Table continues on page 702.

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Farasyn and Lassat[5]	n = 30 with chronic LBP (female-male ratio unknown); mean ± SD age, 47 ± 13 y and n = 30 healthy controls (female-male ratio unknown); mean ± SD age, 41 ± 11 y	"Simple backache" is defined as LBP that is not attributed to any recognizable pathology like nerve root pain and serious spinal pathologies, such as an infection, tumor, osteoporosis, rheumatoid arthritis, fracture, or inflammation	Cross-sectional	PPT	Erector trunci: T8, T10, L1, L3, 3 cm distal to the iliac crest from the proximal part of the gluteus maximus (superior), trochanter major of the femur (inferior)	PPT local (L1): chronic LBP versus healthy control, 3.71 ± 1.20 kg/cm^2 versus 8.69 ± 1.66 kg/cm^2; $P \leq .001$ PPT local (L3): chronic LBP versus healthy control, 5.29 ± 1.27 kg/cm^2 versus 9.86 ± 1.41 kg/cm^2; $P \leq .001$ PPT remote (T8): chronic LBP versus healthy control, 3.96 ± 1.30 kg/cm^2 versus 7.03 ± 1.50 kg/cm^2; $P \leq .001$ PPT remote (T10): chronic LBP versus healthy control, 3.73 ± 1.10 kg/cm^2 versus 7.77 ± 1.31 kg/cm^2; $P \leq .001$ PPT remote (gluteus maximus, pars superior): chronic LBP versus healthy control, 3.73 ± 1.17 kg/cm^2 versus 9.10 ± 1.83 kg/cm^2; $P \leq .001$ PPT remote (gluteus maximus, pars inferior): chronic LBP versus healthy control, 3.84 ± 0.94 kg/cm^2 versus 8.81 ± 2.01 kg/cm^2; $P \leq .001$
Gerhardt et al[9]	n = 77 with chronic back pain, divided into chronic localized pain: n = 48 (24 female, 24 male); mean ± SD age, 59.7 ± 11.8 y and chronic widespread pain: n = 29 (17 female, 12 male); mean ± SD age, 55.2 ± 8.3 y, and n = 40 healthy controls (17 female, 23 male); mean ± SD age, 61.6 ± 12.0 y	Chronic widespread pain defined according to ACR criteria for chronic back pain plus contralateral limb pain (upper and lower and left and right side of the body) Chronic localized pain is defined as chronic widespread pain criteria not being fulfilled	Cross-sectional	PPT and TS	Paraspinal muscles L1-L5 of the painful low back area, dorsum of pain-free ipsilateral hand, PPT at the hand that was tested at the thenar eminence TS protocol: ratings of single pinprick stimulation were compared with a series of 10 repeated pinprick stimuli of the same force (256 mN) over the same area. The mean ratings of the series were divided by the mean pain ratings of single stimuli and calculated as the TS	PPT local (lumbar): chronic localized pain versus healthy control, 0.72 ± 0.22 kg/cm^2 versus 0.81 ± 0.15 kg/cm^2; $P < .001$ PPT local (lumbar): chronic widespread pain versus healthy control, 0.74 ± 0.19 kg/cm^2 versus 0.81 ± 0.15 kg/cm^2; NS TS local (lumbar): chronic localized pain versus healthy control, 0.38 ± 0.26 versus 0.30 ± 0.22 mN; NS TS local (lumbar): chronic widespread pain versus healthy control, 0.42 ± 0.20 versus 0.30 ± 0.22; $P < .001$ PPT remote (hand dorsum): chronic localized pain versus healthy control, 0.64 ± 0.15 kg/cm^2 versus 0.68 ± 0.12 kg/cm^2; NS PPT remote (hand dorsum): chronic widespread pain versus healthy control, 0.62 ± 0.11 kg/cm^2 versus 0.68 ± 0.12 kg/cm^2; NS TS remote (hand dorsum): chronic localized pain versus healthy control, 0.32 ± 0.31 versus 0.30 ± 0.24; NS TS remote (hand dorsum): chronic widespread pain versus healthy control, 0.34 ± 0.18 versus 0.30 ± 0.24; NS

Table continues on page 703.

Risk of Bias

Results of risk-of-bias assessment are shown in **TABLE 2**. Agreement between the 2 reviewers (κ = 0.69; 95% CI: 0.61, 0.77) was "substantial."[17] Each article could have a maximum score of 9 points on the NOS. None of the 24 articles had a score above 6 points, and the average score was 4. Only 2 articles[6,45] had an adequate case definition. All articles, except those of Blumenstiel et al,[3] Farasyn and Meeusen,[6] and Farasyn and Lassat,[5] used the "same method of ascertainment for cases and controls." None of the articles reported "nonresponse rate." The third

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Giesbrecht and Battié[10]	n = 30 with chronic LBP (all female); mean ± SD age, 41.6 ± 9.7 y and n = 30 healthy controls (all female); mean ± SD age, 42.2 ± 9.5 y	No description	Cross-sectional	PPT	Test sites were measured bilaterally: paraspinal muscles C5, L3, L5, wrist extensor muscle, middle phalanx of the second finger, calf muscle	PPT local (L3 and L5): chronic LBP versus healthy control, 5.9 ± 3.0 lb/cm^2 versus 8.0 ± 2.9 lb/cm^2; P = .008 PPT remote (wrist extensor and second finger): chronic LBP versus healthy control, 5.1 ± 1.6 lb/cm^2 versus 6.1 ± 1.6 lb/cm^2; P = .016 PPT global (L3, L5, wrist extensor, second finger, calf muscle, C5): chronic LBP versus healthy control, 5.6 ± 2.1 lb/cm^2 versus 6.9 ± 2.1 lb/cm^2; P = .018
Goubert et al[12]	n = 16 chronic LBP (8 female, 8 male); mean ± SD age, 46 ± 14 y; median age, 50 y (IQR, 28) and n = 21 healthy controls (12 female, 9 male); mean ± SD age, 38 ± 13 y; median age, 40 y (IQR, 29)	No description	Cross-sectional	PPT, TS, and CPM	Erector spinae at 5 cm lateral to L3 spinous process, quadriceps muscle at the midpoint between SIAS and basis patella, trapezius muscle at the midpoint between acromion and C7 spinous process, and the web between the index finger and thumb (dorsal side of hand) TS protocol: the previously determined mean PPT intensity was applied 10 times at each assessment site and was maintained 1 second before being released. Pressure was increased at 1-second intervals until the previously determined mean PPT intensity was reached, followed by 1 second of rest. After the first, fifth, and10th stimuli, an NRS score of the pressure-induced pain sensation was recorded. The area under the curve of the pain sensation during pulses 1, 5, and 10 when mean PPT was applied 10 times was measured	PPT local (lower back): chronic LBP versus healthy control, 623.70 ± 340.29 kPa versus 715.89 ± 433.45 kPa TS local (lower back): chronic LBP versus healthy control, 12.46 ± 5.57 versus 11.13 ± 6.38 kPa PPT remote (trapezius): chronic LBP versus healthy control, 396.19 ± 167.69 kPa versus 511.91 ± 368.73 kPa PPT remote (hand): chronic LBP versus healthy control, 447.18 ± 223.59 kPa versus 567.81 ± 407.96 kPa PPT remote (quadriceps): chronic LBP versus healthy control, 612.92 ± 248.11 kPa versus 733.54 ± 458.95 kPa TS remote (trapezius): chronic LBP versus healthy control, 12.79 ± 5.58 versus 11.35 ± 5.10 TS remote (hand): chronic LBP versus healthy control, 12.29 ± 6.88 versus 11.98 ± 5.38 TS remote (quadriceps): chronic LBP versus healthy control, 12.5 ± 5.48 versus 11.30 ± 6.17 CPM (VAS) (no CS): chronic LBP versus healthy control, 0.58 ± 0.93 points versus 1.11 ± 1.61 points CPM (VAS) (no CS minus CS): chronic LBP versus healthy control, 0.01 ± 0.35 points versus 0.32 ± 0.72 points

Table continues on page 704.

independent researcher was not required for making final decisions.

Pressure Pain Threshold

The results of the meta-analysis are shown in **FIGURES 2** through **6**. Funnel plots were symmetrical, and no sign of publication bias was noted. The PPT measured at the scapula (**FIGURE 2**) was significantly lower in patients with nonspecific LBP than in healthy controls (pooled mean difference, 119.2 kPa; 95% CI: 91.8, 146.6 kPa; P<.001).[12,19,26,30,32,50] The PPT measured at the arm (**FIGURE 3**) was significantly lower in patients with nonspecific LBP than in healthy controls (mean difference, 36.32 kPa; 95% CI: 2.27, 70.37 kPa; P = .04).[6,7,32,39,50] For PPTs measured at the hand (**FIGURE 4**), heterogeneity was high (I^2 = 97%).[3,9,10,12,22,31,34,45] Subgroup analysis revealed that I^2 values dropped to 6% and 0% when taking into account studies with NOS scores at or above 4 or below 4, respectively. Pooled PPT values of studies with NOS scores of 4 or greater were significantly lower in the group with nonspecific LBP compared to healthy controls (mean difference, 5.20 kPa; 95% CI: 1.32, 9.07 kPa; P = .009). Pooled PPT values of studies

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Laursen et al[18]	n = 10 with chronic LBP (all female); median age, 45 y (range, 28-58) and n = 41 healthy controls (all female); median age, 42 y (range, 25-61)	No description	Cross-sectional	PPT	Midline of abdomen, midline of low back, lateral surface of upper arm, pulpa of forefinger, first joint of dorsal side of forefinger, midpoint of the lower extremity, medial border of the scapula	PPT local (midline low back): chronic LBP versus healthy control: median, 269 kPa versus 520 kPa; *P*<.001 PPT remote (medial scapula): chronic LBP versus healthy control: median, 295 kPa versus 620 kPa; *P*<.001 PPT remote (first joint forefinger): chronic LBP versus healthy control: median, 340 kPa versus 850 kPa; *P*<.001 PPT remote (pulpa forefinger): chronic LBP versus healthy control: median, 408 kPa versus 860 kPa; *P*<.001 PPT remote (below umbilicus): chronic LBP versus healthy control: median, 238 kPa versus 388 kPa; *P*<.001 PPT remote (upper arm): chronic LBP versus healthy control: median, 196 kPa versus 649 kPa; *P*<.001 PPT remote (lower extremity): chronic LBP versus healthy control: median, 392 kPa versus 739 kPa; *P*<.01
Lewis et al[19]	n = 15 with chronic LBP (9 female, 6 male); mean ± SD age, 40.9 ± 11.3 y and n = 15 healthy controls (6 female, 9 male); mean ± SD age, 38.7 ± 12.3 y	LBP was defined as per the IASP	Cross-sectional	PPT	PL4 (lateral to spinous process or immediately adjacent to paraspinal musculature), LPL5 (between PSIS and PIIS), deltoid site	PPT local (PL4): LBP versus healthy control, 462.1 kPa (95% CI: 371.1, 553.1) versus 634.4 kPa (95% CI: 534.5, 734.3); NS PPT remote (LPL5): LBP versus healthy control, 380.9 kPa (95% CI: 299.8, 462) versus 535.9 kPa (95% CI: 441.9, 629.9); NS PPT remote (deltoid site): chronic LBP versus healthy control, 296.2 kPa (95% CI: 227.4, 365) versus 401.9 kPa (95% CI: 283.7, 685.6); NS
Marcuzzi et al[22]	n = 7 with persistent LBP (3 female, 4 male); mean ± SD age, 30.6 ± 11.9 y and n = 43 healthy controls (25 female, 23 male); mean ± SD age, 30.0 ± 9.8 y	NRS ≥2 at 4 mo post inclusion with acute LBP (pain and discomfort localized below the costal margin and above the inferior gluteal folds, with or without leg pain, lasting more than 24 h but less than 3 wk, preceded by a pain-free period of at least 1 mo)	Cohort	PPT, TS, and CPM	Bilaterally at the back, dorsum of the left hand (except for PPT, was tested at the thenar eminence) TS protocol: the perceived magnitude of pain from a single pinprick stimulus (256 mN) on a 101-point NRS was compared with that of a series of 10 pinprick stimuli of the same force to measure TS. The repeated stimuli were delivered at a rate of 1 per second within an area of 1 cm^2; TS was calculated as the mean pain rating from the 5 series of 10 repeated stimuli, divided by the mean pain rating from the 5 single stimuli	PPT local (back): LBP versus healthy control, 374 kPa (SE, 66) versus 457 kPa (SE, 26) TS local (back): LBP versus healthy control, 3.9 (SE, 0.7) versus 2.1 (SE, 0.3); *P* = .671 PPT remote (hand): LBP versus healthy control, 345 kPa (SE, 57) versus 384 kPa (SE, 22) TS remote (hand): LBP versus healthy control, 4.2 (SE, 1.6) versus 1.9 (SE, 0.1); *P* = .072 CPM: LBP versus healthy control, −14.2 (SE, 5.8) versus −13.4 (SE, 2.3); *P* = .348

Table continues on page 705.

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Meints et al[24]	n = 167 with chronic LBP (97 female, 70 male); mean ± SD age, 40.77 ± 12.29 y and n = 33 healthy controls (18 female, 15 male); mean ± SD age, 43.35 ± 10.84 y	No description	Baseline data from longitudinal treatment study	TS	Dorsum of the right middle finger (middle phalanx) TS protocol: mechanical punctate pain was assessed using weighted pinprick stimulators. Participants used an NRS from 0 (no pain) to 100 (worst pain imaginable) to rate the sensation of pain produced by 64-mN, 128-mN, and 256-mN stimulators. The lowest force stimulator that produced a painful sensation was then used to apply a train of 10 stimuli to the skin at a rate of 1 pinprick per second. Participants provided pain ratings for the first, fifth, and 10th stimuli. To calculate TS, the pain intensity rating after the first stimulus was subtracted from the rating after the 10th stimulus	TS remote (hand): LBP versus healthy control, 15.97 ± 17.57 versus 14.64 ± 16.73; *d* = 0.08
Mlekusch et al[25]	n = 34 chronic LBP (17 female, 17 male); mean ± SD age, 50.8 ± 14 y and n = 30 healthy controls (16 female, 14 male); mean ± SD age, 37.4 ± 10.9 y	No description	Case-control	PPT and CPM	Second toe	PPT remote (second toe): chronic LBP versus healthy control, 407.8 ± 178.6 kPa versus 548.8 ± 183.6 kPa; *P*<.001 CPM: chronic LBP versus healthy control, 568.5 ± 238.3 kPa versus 681.0 ± 190.6 kPa; *P* = .025
O'Neill et al[30]	n = 12 with chronic LBP (6 female, 6 male); mean age, 46.4 y and n = 12 healthy controls (age and sex matched); mean age, 47.1 y	No description	Cross-sectional	PPT	TA and infraspinatus	PPT remote (infraspinatus): chronic LBP versus healthy control: median, 4.65 kg (95% CI: 3.50, 6.77) versus 6.40 kg (95% CI: 5.09, 10.00); NS PPT remote (TA): chronic LBP versus healthy control: median, 5.45 kg (95% CI: 4.07, 8.89) versus 8.05 kg (95% CI: 5.55, 10.00); *P*<.05

Table continues on page 706.

with NOS scores less than 4 were significantly higher in the group with nonspecific LBP compared to healthy controls (mean difference, −28.27 kPa; 95% CI: −29.30, −27.24 kPa; *P*<.001).[3,34] The PPT measured at the gluteal site (**FIGURE 5**) was significantly lower in patients with nonspecific LBP than in healthy controls (mean difference, 218.63 kPa; 95% CI: 49.69, 387.57 kPa; *P* = .01).[5-7,19] The PPT measured at the lower leg (**FIGURE 6**) was significantly lower in patients with nonspecific LBP than in healthy controls (mean difference, 68.51 kPa; 95% CI: 19.15, 117.86 kPa; *P* = .007).[4,25,30,31]

Three studies with PPT measurements could not be included in the meta-analysis. Two studies used the "remote site" that did not fit within our subgroups,[33,38] and 1 study presented results by reporting the median.[18] All PPT values (lower back and remote site) of the group with nonspecific LBP in that study[18] were significantly lower than those in healthy controls. Özdolap et al[33] measured PPTs at the lower back, 12 sciatic Valleix points, and the fibromyalgia tender points. All mean PPT values in the group with nonspecific LBP were significantly lower than those in healthy controls. Schenk et al[38] measured PPTs at the lower back and forehead. All PPT values measured at the lower back in people with nonspecific LBP did not differ from those measured in healthy controls, whereas PPT values measured at the forehead were lower (*P* = .049) compared to those in healthy controls.

Temporal Summation

The results of the meta-analysis are shown in **FIGURES 7** and **8**. Funnel plots were symmetrical, and no sign of publication bias was noted. For temporal summation measured at the lower back

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Neziri et al[26]	n = 40 with chronic LBP (19 female, 21 male); mean ± SD age, 50.5 ± 13.2 y and n = 300 healthy controls (148 female, 152 male); mean ± SD age, 47.1 ± 15.6 y	No description	Case-control	PPT	Suprascapular, pulp of second toe, site of the most severe pain at the low back, nonpainful site at the low back, middle of upper border of iliac crest, and corresponding spinous process at low back (controls only)	PPT local (site of the most severe pain at the low back): chronic LBP versus healthy control, 168 ± 113 kPa versus 352 ± 131 kPa; OR = 0.13 (95% CI: 0.07, 0.24); *P*<.001 PPT local (nonpainful site at the low back): chronic LBP versus healthy control, 249 ± 132 kPa versus 352 ± 131 kPa; OR = 0.37 (95% CI: 0.24, 0.57); *P*<.001 PPT remote (suprascapular): chronic LBP versus healthy control, 185 ± 103 kPa versus 302 ± 103 kPa; OR = 0.25 (95% CI: 0.15, 0.40); *P*<.001 After full adjustment for age, sex, body mass index, STAI trait and catastrophizing PPT local (site of the most severe pain at the low back): OR = 0.10 (95% CI: 0.04, 0.18), *P*<.001 PPT local (nonpainful site at the low back): OR = 0.38 (95% CI: 0.22, 0.68), *P*<.001 PPT remote (suprascapular): OR = 0.27 (95% CI: 0.15, 0.51), *P*<.001
O'Sullivan et al[31]	n = 19 with LBP (15 female, 4 male); mean ± SD age, 41.9 ± 13.9 y and n = 19 healthy controls (11 female, 8 male); mean ± SD age, 42.6 ± 14.9 y	Mechanical pain group: LBP associated with reports of specific and consistent mechanical aggravating and easing factors Nonmechanical pain group: LBP was more widespread and ill defined, LBP being more constant, nonremitting, spontaneous, and where minor mechanical loading factors resulted in exaggerated or prolonged pain responses	Cross-sectional	PPT	Dorsal area of the wrist joint line, L5-S1 interspinous space, lateral calcaneus	PPT local (lumbar) mechanical pain: mechanical pain versus healthy control: median, 288.7 kPa (IQR, 289.0) versus 352.7 kPa (IQR, 222.3) PPT remote (wrist) mechanical pain: mechanical pain versus healthy control: median, 302.0 kPa (IQR, 177.3) versus 301.3 kPa (IQR, 141.7) PPT remote (heel) mechanical pain: mechanical pain versus healthy control: median, 315.0 kPa (IQR, 159.0) versus 309.3 kPa (IQR, 151.0) PPT local (lumbar) nonmechanical pain: nonmechanical pain versus healthy control: median, 183.0 kPa (IQR, 115.3) versus 352.7 kPa (IQR, 222.3) PPT remote (wrist) nonmechanical pain: nonmechanical pain versus healthy control: median, 239.7 kPa (IQR, 167.7) versus 301.3 kPa (IQR, 141.7) PPT remote (heel) nonmechanical pain: nonmechanical pain versus healthy control: median, 270.3 kPa (IQR, 109.3) versus 309.3 kPa (IQR, 151.0)

Table continues on page 707.

(**FIGURE 7**), heterogeneity was high (I^2 = 72%). Subgroup analysis revealed that I^2 values dropped to 0% and 3% when considering studies with NOS scores less than 4 and 4 or greater, respectively. Pooled temporal summation values of studies with NOS scores less than 4 were significantly higher in healthy controls compared to patients with nonspecific LBP (mean difference, 1.04; 95% CI: 0.16, 1.93; P = .02). Pooled temporal summation values of studies with NOS scores of 4 or greater were significantly higher in patients with nonspecific LBP compared to healthy controls (mean dif-

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Owens et al[32]	n = 25 with chronic LBP (14 female, 11 male); mean ± SD age, 57.64 ± 10.84 y and n = 25 healthy controls (14 female, 11 male); mean ± SD age, 55.16 ± 7.86 y	No description	Observational	PPT, TS (mechanical), and CPM	TS mechanical pain: back of the nondominant hand and ipsilateral trapezius bilaterally. TS heat pain: the volar surface of the forearm. CPM: with PPT, the dominant dorsal forearm and ipsilateral trapezius TS protocol was assessed using a nylon monofilament. To assess TS, participants were instructed to provide a verbal 0-to-100 rating of pain following a single contact of the monofilament. Then, participants were instructed to provide another 0-to-100 rating of their greatest pain intensity experienced following a series of 10 contacts, which were provided at a rate of 1 contact per second. This procedure was repeated twice at each anatomical location. Pain ratings for single and multiple contacts performed at each anatomical location were averaged across the 2 trials	PPT remote (forearm): chronic LBP versus healthy control, 369.70 ± 217.94 kPa versus 393.16 ± 180.87 kPa PPT remote (trapezius): chronic LBP versus healthy control, 340.80 ± 196.27 kPa versus 412.98 ± 212.67 kPa TS (VAS) mechanical remote (hand): chronic LBP versus healthy control (1 contact), 9.96 ± 16.07 points versus 4.32 ± 5.13 points TS (VAS) mechanical remote (hand): chronic LBP versus healthy control (10 contacts), 25.68 ± 24.63 points versus 10.80 ± 10.92 points TS (VAS) mechanical remote (trapezius): chronic LBP versus healthy control (1 contact), 9.02 ± 13.88 points versus 4.12 ± 3.77 points TS (VAS) mechanical remote (trapezius): chronic LBP versus healthy control (10 contacts), 31.24 ± 29.92 points versus 14.38 ± 15.09 points CPM (forearm): chronic LBP versus healthy control, 402.97 ± 209.65 kPa versus 449.88 ± 213.29 kPa CPM (trapezius): chronic LBP versus healthy control, 398.40 ± 230.01 kPa versus 525.40 ± 246.71 kPa
Özdolap et al[33]	n = 70 with chronic LBP (44 female, 26 male); mean ± SD age, 37.6 ± 10.1 y and n = 62 healthy controls (33 female, 29 male); mean ± SD age, 34.6 ± 9.6 y	No description	Cross-sectional	PPT	18 tender points, as defined by the ACR, for fibromyalgia syndrome, 12 points for the sciatic Valleix (bilateral midpoint of gluteus maximus, midpoint of the gluteal sulcus, midpoint and posterior point of thigh, popliteal fossa, midpoint and posterior point of cruris, midpoint of Achilles tendon), and 4 lumbar paravertebral points (bilateral; 2 cm lateral to the L2 and L4 spinous processes)	PPT local (4 lumbar points): chronic LBP versus healthy control, 18.8 kg/cm^2 versus 28.7 kg/cm^2; P<.001 PPT remote (12 sciatic Valleix points): chronic LBP versus healthy control, 78.5 ± 25.8 kg/cm^2 versus 93.4 ± 26.1 kg/cm^2; P = .001 PPT remote (fibromyalgia tender points): chronic LBP versus healthy control, 87.2 ± 29.5 kg/cm^2 versus 105.0 ± 31.6 kg/cm^2; P = .001

Table continues on page 708.

ference, −0.84; 95% CI: −1.24, −0.44; P<.001).[3,34]

The subgroup with temporal summation measured at the hand (**FIGURE 8**) revealed no significant difference between patients with nonspecific LBP and healthy controls (P = .06).[3,9,22,24,34,45]

Three studies using temporal summation were not included in the meta-analysis because of a different measurement protocol.[12,26,32] Goubert et al[12] reported that the temporal summation value of people with nonspecific LBP was higher (ie, more enhanced) than that in healthy controls. Significance was not described. The temporal summation values reported by Owens et al[32] showed a significantly higher sensitivity in patients with nonspecific LBP compared with healthy controls.

Conditioned Pain Modulation

In 6 studies, a conditioned pain modulation protocol was used. Results were not pooled because of differences between the protocols.[4,12,22,25,32,35] The study by Rabey et al[35] found that more healthy controls showed a significant inhibitory effect than did people with nonspecific LBP. In the study by Corrêa et al,[4] conditioned pain modulation outcomes showed that

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Puta et al[34]	n = 18 with LBP (all female); mean ± SD age, 51.2 ± 4.2 y and n = 16 healthy controls (all female); mean ± SD age, 51.1 ± 5.5 y	No description	Cross-sectional design	PPT and TS	Painful body site: paraspinal T12-L5 and nonpainful body site: hand (palmar) TS protocol was assessed by trains of 10 punctate stimuli. To determine TS, the ratio of the mean pain rating of trains divided by the mean pain rating for a single stimulus was calculated	PPT local (back): chronic LBP versus healthy control, $\log_{10}(152)(2.182 \pm 0.278$ kPa) versus $\log_{10}(197)(2.294 \pm 0.188$ kPa); $P = .19$ TS local (back): chronic LBP versus healthy control, $\log_{10}(2.48)(0.394 \pm 0.205$ kPa) versus $\log_{10}(3.30)(0.519 \pm 0.326$ kPa); $P = .20$ PPT remote (hand): chronic LBP versus healthy control, $\log_{10}(238)(2.376 \pm 0.222$ kPa) versus $\log_{10}(209)(2.321 \pm 0.146$ kPa); $P = .41$ TS remote (hand): chronic LBP versus healthy control, $\log_{10}(2.14)(0.331 \pm 0.245$ kPa) versus $\log_{10}(2.62)(0.419 \pm 0.289$ kPa); $P = .35$
Rabey et al[35]	n = 64 with chronic LBP (35 female, 29 male); mean ± SD age, 34.6 ± 10.6 y and n = 64 healthy controls (35 female, 29 male); mean ± SD age, 33.5 ± 11.0 y	No description	Case-control trial	PPT and CPM (heat noxious stimuli)	Chronic LBP: most painful lumbar region. Healthy controls: over paraspinal muscles adjacent to the L5 spinous process	NRS with concurrent CS: chronic LBP, 7.3 ± 1.4 points (95% CI: 6.9, 7.6); $P \leq .001$ NRS with concurrent healthy control: healthy control, 5.8 ± 1.3 points (95% CI: 5.5, 6.2); $P = .35$
Schenk et al[38]	n = 38 with chronic LBP (all female); mean ± SD age of nurses, 51.9 ± 4.5 y; mean ± SD age of secretaries, 52.7 ± 4.8 y and n = 68 healthy controls (all female); mean ± SD age of nurses, 51.8 ± 4.8 y; mean ± SD age of secretaries, 52.9 ± 5.1 y	No description	Cross-sectional	PPT	Paravertebral muscles, quadratus lumborum, os ilium, iliolumbar ligament, piriformis, greater trochanter, and middle of forehead	PPT local (back): chronic LBP versus healthy control, $P = .68$ PPT remote (forehead): chronic LBP versus healthy control, $P = .049$
Simmonds and Claveau[39]	n = 23 with chronic LBP (12 female, 11 male); mean ± SD age, 43.2 ± 12.9 y and n = 23 healthy controls (12 female, 11 male); mean ± SD age, 43.0 ± 12.4 y	No description	Cross-sectional	PPT (dolorimeter)	L3-L4 interspinous space and on the ulnar border of the forearm	PPT local (back) dolorimeter: chronic LBP versus healthy control, 4.74 ± 2.24 kg/cm^2 versus 5.24 ± 1.76 kg/cm^2; NS PPT remote (arm) dolorimeter: chronic LBP versus healthy control, 5.18 ± 3.38 kg/cm^2 versus 5.52 ± 1.98 kg/cm^2; NS

Table continues on page 709.

PPT values at the lower back and the tibialis anterior in the group with nonspecific LBP were significantly lower compared to those in healthy controls. During conditioned pain modulation, the group with nonspecific LBP demonstrated a statistically significant decrease in the lumbar PPT, while healthy controls demonstrated a significant increase in the lumbar PPT.[4] Goubert et al[12] demonstrated no significant differences between patients with nonspecific LBP and healthy controls. Mlekusch et al[25] and Owens et al[32] showed a normal conditioned pain modulation effect in both groups; PPT values increased after the conditioned pain stimulus in both the group with nonspecific LBP and healthy controls. Marcuzzi et al[22] showed no significant differences between the group with nonspecific LBP and healthy controls.

TABLE 1 Characteristics of Included Studies (n = 24) (continued)

Study	Participants	Definition of Nonspecific LBP	Study Design	Stimulus	Locations and TS Protocol	Results
Tesarz et al[45]	n = 93 with chronic LBP (61 female, 32 male); mean age, 58.2 y (95% CI: 26.3, 60.2) and n = 31 healthy controls (18 female, 13 male); mean age, 60.1 y (95% CI: 55.7, 64.5)	No description	Cross-sectional	PPT and TS	Paraspinal muscles at the height of lumbar segments L1-L5 of the low back area, and on the dorsum of the ipsilateral hand TS protocol: the train of pinprick stimuli was given within a small area of 1 cm^2, and the participant was asked to give a pain rating representing the pain at the end of the train using an NRS. The mean ratings of series divided by the mean pain ratings of single stimuli were calculated as TS	PPT local (back): NSCLBP-W-TE versus healthy control, 0.69 kg/cm^2 (95% CI: 0.65, 0.73) versus 0.77 kg/cm^2 (95% CI: 0.72, 0.83); P = .001 TS local (back): NSCLBP-W-TE versus healthy control, 0.46 kg/cm^2 (95% CI: 0.40, 0.51) versus 0.29 kg/cm^2 (95% CI: 0.20, 0.38); P = .010 PPT remote (hand): NSCLBP-W-TE versus healthy control, 0.61 kg/cm^2 (95% CI: 0.58, 0.64) versus 0.65 kg/cm^2 (95% CI: 0.60, 0.69); P = .006 TS remote (hand): NSCLBP-W-TE versus healthy control, 0.39 kg/cm^2 (95% CI: 0.33, 0.45) versus 0.31 kg/cm^2 (95% CI: 0.22, 0.41); P = .320
Yildiz et al[50]	n = 121 with chronic LBP (81 female, 40 male); mean ± SD age, 36.8 ± 9.9 y and n = 91 healthy controls (65 female, 26 male); mean ± SD age, 34.1 ± 10.2 y	No description	Case-control	PPT	The midpoint of the dorsum of the forearm, the midpoint of the upper trapezius muscle, and the paravertebral muscles at L1, L3, and L5 were examined bilaterally	PPT local (L1): chronic LBP versus healthy control, 6.9 ± 2.3 kg/cm^2 versus 8.1 ± 2.1 kg/cm^2; P<.001 PPT local (L3): chronic LBP versus healthy control, 6.9 ± 2.3 kg/cm^2 versus 8.0 ± 2.2 kg/cm^2; P<.001 PPT local (L5): chronic LBP versus healthy control, 6.9 ± 2.4 kg/cm^2 versus 8.0 ± 2.1 kg/cm^2; P<.001 PPT remote (forearm): chronic LBP versus healthy control, 7.2 ± 2.1 kg/cm^2 versus 7.7 ± 2.1 kg/cm^2; P = .089 PPT remote (trapezius): chronic LBP versus healthy control, 5.6 ± 2.2 kg/cm^2 versus 7.0 ± 2.4 kg/cm^2; P<.001

Abbreviations: ACR, American College of Rheumatology; APTA, American Physical Therapy Association; CI, confidence interval; CPM, conditioned pain modulation; CS, central sensitization; IASP, International Association for the Study of Pain; IQR, interquartile range; LBP, low back pain; NRS, numeric rating scale; NS, nonsignificant; NSCLBP-W-TE, nonspecific chronic low back pain without trauma exposure; OR, odds ratio; PIIS, posterior inferior iliac spine; PPT, pressure pain threshold; PSIS, posterior superior iliac spine; SE, standard error; SIAS, spina iliaca anterior superior; STAI, State-Trait Anxiety Inventory; TA, tibialis anterior; TS, temporal summation; VAS, visual analog scale.

DISCUSSION

The present systematic review and meta-analysis critically appraised the current literature on mechanical QST measurements in patients with nonspecific LBP in order to examine signs of altered sensory functioning in this population. The meta-analysis found that overall PPT measurements at remote body parts were significantly lower in patients with nonspecific LBP compared with healthy controls. This finding is indicative of central sensitization in people with nonspecific LBP.[11] In the studies with superior methodological quality, temporal summation was enhanced in the lumbar region, but not at remote sites, in people with nonspecific LBP compared to healthy controls. Regarding conditioned pain modulation in patients with nonspecific LBP, the findings were mixed. Although we did not find a clear picture of signs of central sensitization in people with nonspecific LBP, the available literature regarding mechanical somatosensory functioning provides some evidence of central sensitization in people with nonspecific LBP.

Central sensitization is a phenomenon characterized by enhanced nociceptive processing combined with disturbed top-down modulation. Quantitative sensory testing measures objectify these neurophysiological processes and are used to draw conclusions about the way the sensory systems process different stimuli. In this study, only a small number of studies used temporal summation and/or conditioned pain modulation, which hampered conclusions about changes in this type of QST measurement and may explain the inconsistent results, underscoring the importance of

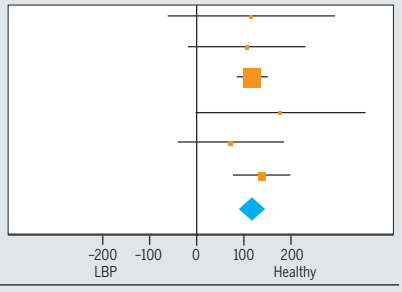
TABLE 2 Results of Risk-of-Bias Assessment of the Selected Studies (n = 24)*

	Selection				Comparability		Exposure			
Study	**1**	**2**	**3**	**4**	**1a**	**1b**	**1**	**2**	**3**	**Total Score**
Blumenstiel et al[3]			X		X					2
Corrêa et al[4]				X	X	X		X		4
Farasyn and Meeusen[7]				X				X		2
Farasyn and Meeusen[6]	X			X	X					3
Farasyn and Lassat[5]			X	X						2
Gerhardt et al[9]		X		X			X	X		4
Giesbrecht and Battié[10]			X	X	X	X	X	X		6
Goubert et al[12]			X	X	X	X		X		5
Laursen et al[18]			X	X				X		3
Lewis et al[19]				X				X		2
Marcuzzi et al[22]		X	X			X	X	X		5
Meints et al[24]			X					X		2
Mlekusch et al[25]		X	X	X	X	X		X		6
O'Neill et al[30]		X		X	X	X		X		5
Neziri et al[26]		X	X	X	X	X		X		6
O'Sullivan et al[31]		X		X	X	X		X		5
Owens et al[32]			X	X	X	X		X		5
Özdolap et al[33]				X			X	X		3
Puta et al[34]					X	X		X		3
Rabey et al[35]			X		X	X		X		4
Schenk et al[38]			X	X			X	X		4
Simmonds and Claveau[39]					X	X		X		3
Tesarz et al[45]	X			X	X	X	X	X		6
Yildiz et al[50]								X		1

**Selection: (1) adequacy of the case definition, (2) representativeness of the cases, (3) selection of controls, (4) definition of controls; Comparability: (1a) study controls for age and/or sex, (1b) questionnaire; Exposure: (1) ascertainment of exposure, (2) same method of ascertainment for cases and controls, (3) nonresponse rate.*

Pressure Pain Threshold: Scapula

	Healthy		LBP			
Study	**Mean ± SD kPa**	**Total, n**	**Mean ± SD kPa**	**Total, n**	**Weight**	**MD IV, Random (95% Confidence Interval)**
Goubert et al[12]	511.91 ± 368.73	21	396.19 ± 167.69	16	2.4%	115.72 (−62.11, 293.55)
Lewis et al[19]	401.9 ± 213.44	15	296.2 ± 124.24	15	4.8%	105.70 (−19.28, 230.68)
Neziri et al[26]	302 ± 103	300	185 ± 103	40	65.0%	117.00 (83.02, 150.98)
O'Neill et al[30]	686.5 ± 242.2	12	509.1 ± 206.7	12	2.3%	177.40 (−2.75, 357.55)
Owens et al[32]	412.98 ± 212.67	25	340.8 ± 196.27	25	5.8%	72.18 (−41.26, 185.62)
Yildiz et al[50]	686.47 ± 235.36	91	549.17 ± 215.75	121	19.7%	137.30 (75.52, 199.08)
Total*		464		229	100.0%	119.20 (91.80, 146.60)

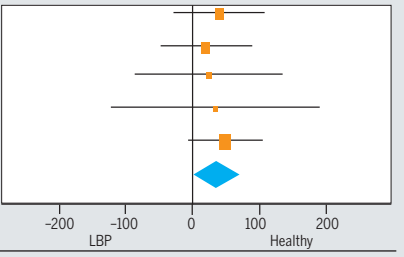


Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference.
**Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 1.45$, $df = 5$ ($P = .92$), $I^2 = 0\%$. Test for overall effect: $z = 8.53$ ($P<.001$).*

FIGURE 2. Pooled results of pressure pain thresholds for the scapula cluster.

Pressure Pain Threshold: Arm

	Healthy		LBP			
Study	**Mean ± SD kPa**	**Total, n**	**Mean ± SD kPa**	**Total, n**	**Weight**	**MD IV, Random (95% Confidence Interval)**
Farasyn and Meeusen[7]	696.27 ± 166.71	32	657.05 ± 176.52	87	24.6%	39.22 (−29.43, 107.87)
Farasyn and Meeusen[6]	696.27 ± 166.71	32	676.66 ± 147.1	58	24.3%	19.6 (−49.45, 88.67)
Owens et al[32]	393.16 ± 180.87	25	369.7 ± 217.94	25	9.4%	23.46 (−87.56, 134.48)
Simmonds and Claveau[39]	541.33 ± 194.17	23	507.98 ± 331.46	23	4.7%	33.35 (−123.64, 190.34)
Yildiz et al[50]	755.11 ± 205.94	91	706.08 ± 205.94	121	37.0%	49.03 (−6.98, 105.04)
Total*		203		314	100.0%	36.32 (2.27, 70.37)

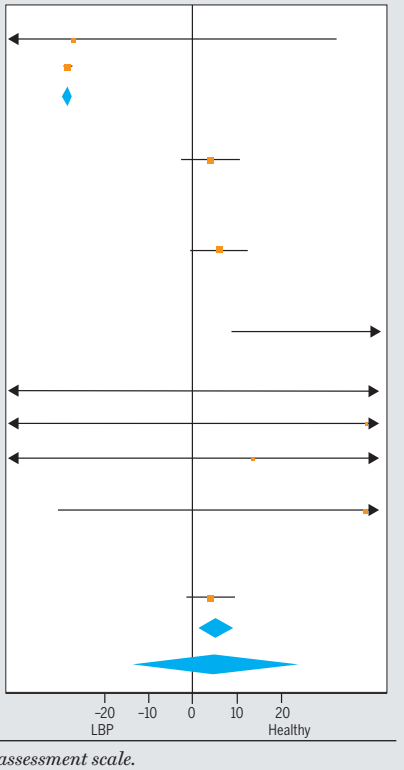


Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference.
**Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 0.48$, $df = 4$ ($P = .98$), $I^2 = 0\%$. Test for overall effect: $z = 2.09$ ($P = .04$).*

FIGURE 3. Pooled results of pressure pain thresholds for the arm cluster.

Pressure Pain Threshold: Hand

	Healthy		LBP			
Subgroup/Study	**Mean ± SD kPa**	**Total, n**	**Mean ± SD kPa**	**Total, n**	**Weight**	**MD IV, Random (95% Confidence Interval)**
NOS score <4						
Blumenstiel et al[3]	318 ± 96.15	20	345 ± 101.75	23	6.5%	−27.00 (−86.20, 32.20)
Puta et al[34]	209.41 ± 1.4	16	237.68 ± 1.67	18	17.6%	−28.27 (−29.30, −27.24)
Subtotal*		36		41	24.1%	−28.27 (−29.30, −27.24)
NOS score ≥4						
Gerhardt et al[9] (chronic localized pain)	66.69 ± 11.78	20	62.76 ± 14.71	48	17.3%	3.93 (−2.70, 10.56)
Gerhardt et al[9] (chronic widespread pain)	66.69 ± 11.78	20	60.8 ± 10.8	29	17.3%	5.89 (−0.60, 12.38)
Giesbrecht and Battié[10]	271.34 ± 71.17	30	226.86 ± 71.17	30	10.8%	44.48 (8.46, 80.50)
Goubert et al[12]	567.81 ± 407.96	21	447.18 ± 223.59	16	0.8%	120.63 (−85.40, 326.66)
Marcuzzi et al[22]	384 ± 144.26	43	345 ± 150.81	7	2.2%	39.00 (−80.75, 158.75)
O'Sullivan et al[31] (mechanical pain)	319.9 ± 85.8	10	306.3 ± 107.5	17	4.8%	13.60 (−60.15, 87.35)
O'Sullivan et al[31] (nonmechanical pain)	319.9 ± 85.8	9	280.9 ± 91.4	19	5.3%	39.00 (−30.51, 108.51)
Tesarz et al[45]	63.74 ± 13.36	31	59.82 ± 14.28	93	17.4%	3.92 (−1.61, 9.45)
Subtotal†		184		259	75.9%	5.20 (1.32, 9.07)
Total‡		220		300	100.0%	5.00 (−14.05, 24.05)

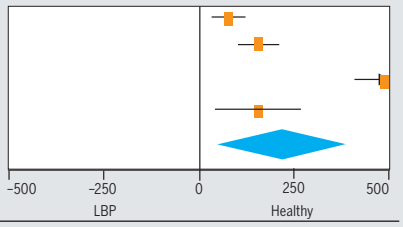


Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference; NOS, Newcastle-Ottawa quality assessment scale.
**Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 0.00$, $df = 1$ ($P = .97$), $I^2 = 0\%$. Test for overall effect: $z = 53.68$ ($P<.001$).*
†Heterogeneity: $\tau^2 = 2.06$, $\chi^2 = 7.42$, $df = 7$ ($P = .39$), $I^2 = 6\%$. Test for overall effect: $z = 2.63$ ($P = .009$).
‡Heterogeneity: $\tau^2 = 535.74$, $\chi^2 = 323.99$, $df = 9$ ($P<.001$), $I^2 = 97\%$. Test for overall effect: $z = 0.51$ ($P = .61$). Test for subgroup differences: $\chi^2 = 267.34$, $df = 1$ ($P<.001$), $I^2 = 99.6\%$.

FIGURE 4. Pooled results of pressure pain thresholds for the hand cluster.

Pressure Pain Threshold: Gluteal

Study	Healthy Mean ± SD kPa	Healthy Total, n	LBP Mean ± SD kPa	LBP Total, n	Weight	MD IV, Random (95% Confidence Interval)
Farasyn and Meeusen[7]	696.27 ± 137.3	64	617.82 ± 147.1	87	25.8%	78.45 (32.77, 124.13)
Farasyn and Meeusen[6]	784.53 ± 147.1	64	627.63 ± 156.91	58	25.7%	156.90 (102.78, 211.02)
Farasyn and Lassat[5]	863.97 ± 197.11	30	376.58 ± 92.18	30	25.0%	487.39 (409.52, 565.26)
Lewis et al[19]	535.9 ± 169.74	15	380.9 ± 146.45	15	23.5%	155.00 (41.55, 268.45)
Total*		173		190	100.0%	218.63 (49.69, 387.57)

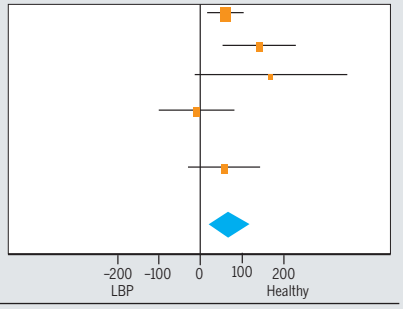


Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference.
**Heterogeneity: $\tau^2 = 28200.30$, $\chi^2 = 79.52$, $df = 3$ ($P<.001$), $I^2 = 96\%$. Test for overall effect: $z = 2.54$ ($P = .01$).*

FIGURE 5. Pooled results of pressure pain thresholds for the gluteal cluster.

Pressure Pain Threshold: Lower Leg

Study	Healthy Mean ± SD kPa	Healthy Total, n	LBP Mean ± SD kPa	LBP Total, n	Weight	MD IV, Random (95% Confidence Interval)
Corrêa et al[4]	321.8 ± 84.5	30	262.4 ± 93.1	30	36.3%	59.40 (14.41, 104.39)
Mlekusch et al[25]	548.8 ± 183.6	30	407.8 ± 178.6	34	19.3%	141.00 (52.00, 230.00)
O'Neill et al[30]	771.5 ± 216.4	12	602.3 ± 244.4	12	6.3%	169.20 (-15.50, 353.90)
O'Sullivan et al[31] (mechanical pain)	336.5 ± 119.3	10	346.5 ± 114.6	17	18.6%	-10.00 (-101.84, 81.84)
O'Sullivan et al[31] (nonmechanical pain)	336.5 ± 119.3	9	280.5 ± 90.8	19	19.6%	56.00 (-31.99, 143.99)
Total*		91		112	100.0%	68.51 (19.15, 117.86)



Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference.
**Heterogeneity: $\tau^2 = 1221.39$, $\chi^2 = 6.69$, $df = 4$ ($P = .15$), $I^2 = 40\%$. Test for overall effect: $z = 2.72$ ($P = .007$).*

FIGURE 6. Pooled results of pressure pain thresholds for the lower-leg cluster.

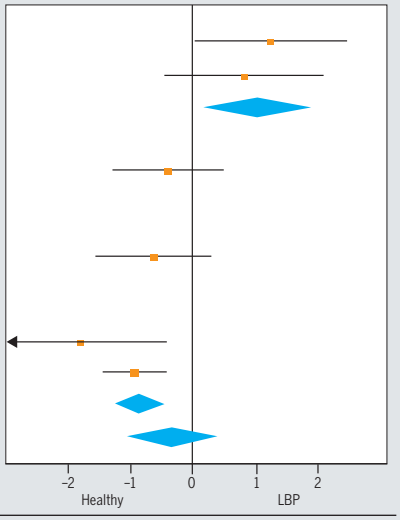
conducting a meta-analysis. Inconsistent findings regarding QST measurements may also be due to the presence of subgroups within the population with nonspecific LBP. Only 2 of the included studies separately reported on localized and widespread pain. Therefore, subgroup analyses were not possible. The present review was not designed to reveal or refute subgroups within people with nonspecific LBP. There is a need for more studies using more extended QST measurements in order to determine the existence of different QST profiles in patients with nonspecific LBP.

As mechanical QST measurements are most often used in studies of patients with nonspecific LBP, this review is limited to studies using mechanical QST measurements only. How the somatosensory system responds to thermal and electrical stimuli in people with nonspecific LBP and central sensitization remains to be examined. Finally, it is currently unknown whether the different results in these static (PPT) and dynamic (temporal summation and conditioned pain modulation) measurements can be explained by methodological issues (eg, smaller sample sizes and different protocols) or by underlying physiological differences. Notably, a clear definition of nonspecific LBP was not reported in most studies.

The strength of this review is that it is the first meta-analysis to study and summarize QST measurements in people with nonspecific LBP. It should be taken into consideration that many of the included studies were rated as having low to moderate methodological quality. Based on their narrative analysis of the literature, Roussel et al[37] concluded that signs of central sensitization may be present in patients with LBP. The results of our meta-analysis confirm that PPTs at remote body parts are significantly lower and temporal summation at the lower back is enhanced in patients with nonspecific LBP compared to healthy controls. This conclusion could be strengthened by studies with higher methodological quality. Because the reported standard error of measurement of QST measures may vary between measured populations, measured body parts, and different protocols, it is difficult to compare scores and evaluate the

Temporal Summation: Lumbar

	Healthy		LBP			
Subgroup/Study	**Mean ± SD NRS**	**Total, n**	**Mean ± SD NRS**	**Total, n**	**Weight**	**MD IV, Random (95% Confidence Interval)**
NOS score <4						
Blumenstiel et al[3]	3.61 ± 2.4	20	2.36 ± 1.52	23	14.8%	1.25 (0.03, 2.47)
Puta et al[34]	3.3 ± 2.12	16	2.48 ± 1.6	18	14.3%	0.82 (−0.45, 2.09)
Subtotal*		36		41	29.1%	1.04 (0.16, 1.93)
NOS score ≥4						
Gerhardt et al[9] (chronic localized pain)	2.0 ± 1.66	20	2.4 ± 1.82	48	18.1%	−0.40 (−1.29, 0.49)
Gerhardt et al[9] (chronic widespread pain)	2.0 ± 1.66	20	2.63 ± 1.59	29	17.7%	−0.63 (−1.56, 0.30)
Marcuzzi et al[22]	2.1 ± 0.55	43	3.9 ± 1.85	7	13.3%	−1.80 (−3.18, −0.42)
Tesarz et al[45]	1.95 ± 1.04	31	2.88 ± 1.81	93	21.8%	−0.93 (−1.45, −0.41)
Subtotal†		114		177	70.9%	−0.84 (−1.24, −0.44)
Total‡		150		218	100.0%	−0.32 (−1.07, 0.42)

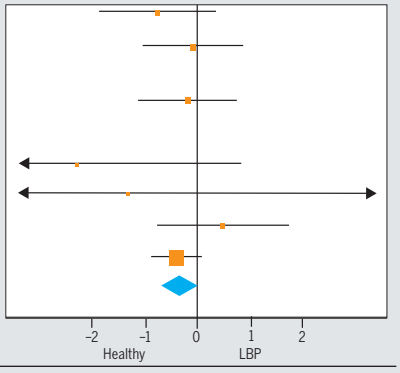


Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference; NOS, Newcastle-Ottawa quality assessment scale; NRS, numeric rating scale.
**Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 0.23$, $df = 1$ ($P = .63$), $I^2 = 0\%$. Test for overall effect: $z = 2.32$ ($P = .02$).*
†Heterogeneity: $\tau^2 = 0.01$, $\chi^2 = 3.11$, $df = 3$ ($P = .38$), $I^2 = 3\%$. Test for overall effect: $z = 4.14$ ($P<.001$).
‡Heterogeneity: $\tau^2 = 0.60$, $\chi^2 = 18.12$, $df = 5$ ($P = .003$), $I^2 = 72\%$. Test for overall effect: $z = 0.85$ ($P = .40$). Test for subgroup differences: $\chi^2 = 14.61$, $df = 1$ ($P<.001$), $I^2 = 93.2\%$.

FIGURE 7. Pooled results of temporal summation for the lumbar cluster.

Temporal Summation: Hand

	Healthy		LBP			
Study	**Mean ± SD NRS**	**Total, n**	**Mean ± SD NRS**	**Total, n**	**Weight**	**MD IV, Random (95% Confidence Interval)**
Blumenstiel et al[3]	2.81 ± 1.69	20	3.57 ± 2.03	23	9.9%	−0.76 (−1.87, 0.35)
Gerhardt et al[9] (chonic localized pain)	1.99 ± 1.74	20	2.09 ± 2.04	48	13.4%	−0.09 (−1.05, 0.86)
Gerhardt et al[9] (chronic widespread pain)	1.99 ± 1.74	20	2.19 ± 1.51	29	13.8%	−0.19 (−1.13, 0.75)
Marcuzzi et al[22]	1.9 ± 0.66	43	4.2 ± 4.23	7	1.2%	−2.30 (−5.44, 0.84)
Meints et al[24]	14.64 ± 16.73	33	15.97 ± 17.57	167	0.3%	−1.33 (−7.63, 4.97)
Puta et al[34]	2.62 ± 1.95	16	2.14 ± 1.76	18	7.8%	0.48 (−0.77, 1.73)
Tesarz et al[45]	2.04 ± 1.04	31	2.45 ± 1.51	93	53.6%	−0.41 (−0.89, 0.07)
Total*		183		385	100.0%	−0.33 (−0.68, 0.02)



Abbreviations: IV, inverse variance; LBP, low back pain; MD, mean difference; NRS, numeric rating scale.
**Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 4.21$, $df = 6$ ($P = .65$), $I^2 = 0\%$. Test for overall effect: $z = 1.85$ ($P = .06$).*

FIGURE 8. Pooled results of temporal summation for the hand cluster.

magnitude of pooled differences properly. However, the pooled difference for PPTs measured at the scapula (mean difference, 119.2 kPa; 95% CI: 91.8, 146.6 kPa) exceeds the range of previously reported standard error of measurement of 18.2 to 52 kPa.[47]

The results of this study should be interpreted with caution, as we only included several types of observational study design that compared groups of patients with nonspecific LBP to healthy controls. Additionally, we currently lack clear cut-off scores for QST measurement that would enable health care professionals

to make sound judgments in individual cases. However, health care professionals should be aware that altered sensory processing may be present in patients with nonspecific LBP, and that this might require a different treatment approach.[29]

CONCLUSION

The PPTs at remote body parts and temporal summation at the lower back differed between people with nonspecific LBP and healthy controls. Results of studies using conditioned pain modulation measurement were mixed. ◉

KEY POINTS

FINDINGS: In people with nonspecific low back pain, altered sensory functioning was demonstrated. This was present in the pressure pain threshold measurements at remote body parts.

IMPLICATIONS: Health care professionals should be aware that sensory processing may be enhanced in patients with nonspecific low back pain, which may require a different treatment approach.

CAUTION: The results of this systematic review are based on cross-sectional studies that compared groups of people with low back pain to healthy controls; therefore, no conclusion on an individual level can be made.

ACKNOWLEDGMENTS: *The authors thank W. Bramer of the Erasmus University Medical Center (Rotterdam, the Netherlands) for constructing the search strategies for the different databases.*

REFERENCES

1. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*. 2006;15 suppl 2:S192-S300. https://doi.org/10.1007/s00586-006-1072-1
2. Arendt-Nielsen L, Morlion B, Perrot S, et al. Assessment and manifestation of central sensitisation across different chronic pain conditions. *Eur J Pain*. 2018;22:216-241. https://doi.org/10.1002/ejp.1140
3. Blumenstiel K, Gerhardt A, Rolke R, et al. Quantitative sensory testing profiles in chronic back pain are distinct from those in fibromyalgia. *Clin J Pain*. 2011;27:682-690. https://doi.org/10.1097/AJP.0b013e3182177654
4. Corrêa JB, Costa LO, de Oliveira NT, Sluka KA, Liebano RE. Central sensitization and changes in conditioned pain modulation in people with chronic nonspecific low back pain: a case–control study. *Exp Brain Res*. 2015;233:2391-2399. https://doi.org/10.1007/s00221-015-4309-6
5. Farasyn A, Lassat B. Cross friction algometry (CFA): comparison of pressure pain thresholds between patients with chronic non-specific low back pain and healthy subjects. *J Bodyw Mov Ther*. 2016;20:224-234. https://doi.org/10.1016/j.jbmt.2015.09.005
6. Farasyn A, Meeusen R. Effect of roptrotherapy on pressure-pain thresholds in patients with subacute nonspecific low back pain. *J Musculoskelet Pain*. 2007;15:41-53. https://doi.org/10.1300/J094v15n01_06
7. Farasyn A, Meeusen R. The influence of non-specific low back pain on pressure pain thresholds and disability. *Eur J Pain*. 2005;9:375-381. https://doi.org/10.1016/j.ejpain.2004.09.005
8. Foster NE, Anema JR, Cherkin D, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet*. 2018;391:2368-2383. https://doi.org/10.1016/S0140-6736(18)30489-6
9. Gerhardt A, Eich W, Janke S, Leisner S, Treede RD, Tesarz J. Chronic widespread back pain is distinct from chronic local back pain: evidence from quantitative sensory testing, pain drawings, and psychometrics. *Clin J Pain*. 2016;32:568-579. https://doi.org/10.1097/AJP.0000000000000300
10. Giesbrecht RJ, Battié MC. A comparison of pressure pain detection thresholds in people with chronic low back pain and volunteers without pain. *Phys Ther*. 2005;85:1085-1092. https://doi.org/10.1093/ptj/85.10.1085
11. Giesecke T, Gracely RH, Grant MA, et al. Evidence of augmented central pain processing in idiopathic chronic low back pain. *Arthritis Rheum*. 2004;50:613-623. https://doi.org/10.1002/art.20063
12. Goubert D, Danneels L, Graven-Nielsen T, Descheemaeker F, Meeus M. Differences in pain processing between patients with chronic low back pain, recurrent low back pain, and fibromyalgia. *Pain Physician*. 2017;20:307-318.
13. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Oxford, UK: The Cochrane Collaboration; 2009.
14. Hoy D, Bain C, Williams G, et al. A systematic review of the global prevalence of low back pain. *Arthritis Rheum*. 2012;64:2028-2037. https://doi.org/10.1002/art.34347
15. Hush JM, Stanton TR, Siddall P, Marcuzzi A, Attal N. Untangling nociceptive, neuropathic and neuroplastic mechanisms underlying the biological domain of back pain. *Pain Manag*. 2013;3:223-236. https://doi.org/10.2217/pmt.13.11
16. Klyne DM, Moseley GL, Sterling M, Barbe MF, Hodges PW. Are signs of central sensitization in acute low back pain a precursor to poor outcome? *J Pain*. 2019;20:994-1009. https://doi.org/10.1016/j.jpain.2019.03.001
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174. https://doi.org/10.2307/2529310
18. Laursen BS, Bajaj P, Olesen AS, Delmar C, Arendt-Nielsen L. Health related quality of life and quantitative pain measurement in females with chronic non-malignant pain. *Eur J Pain*. 2005;9:267-275. https://doi.org/10.1016/j.ejpain.2004.07.003
19. Lewis C, Souvlis T, Sterling M. Sensory characteristics of tender points in the lower back. *Man Ther*. 2010;15:451-456. https://doi.org/10.1016/j.math.2010.03.006
20. Lievense AM, Bierma-Zeinstra SM, Verhagen AP, van Baar ME, Verhaar JA, Koes BW. Influence of obesity on the development of osteoarthritis of the hip: a systematic review. *Rheumatology (Oxford)*. 2002;41:1155-1162. https://doi.org/10.1093/rheumatology/41.10.1155
21. Machado LA, Kamper SJ, Herbert RD, Maher CG, McAuley JH. Analgesic effects of treatments for non-specific low back pain: a meta-analysis of placebo-controlled randomized trials. *Rheumatology (Oxford)*. 2009;48:520-527. https://doi.org/10.1093/rheumatology/ken470
22. Marcuzzi A, Wrigley PJ, Dean CM, Graham PL, Hush JM. From acute to persistent low back pain: a longitudinal investigation of somatosensory changes using quantitative sensory testing—an exploratory study. *Pain Rep*. 2018;3:e641. https://doi.org/10.1097/PR9.0000000000000641
23. Meeus M, Nijs J. Central sensitization: a biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. *Clin Rheumatol*. 2007;26:465-473. https://doi.org/10.1007/s10067-006-0433-9
24. Meints SM, Mawla I, Napadow V, et al. The relationship between catastrophizing and altered pain sensitivity in patients with chronic low-back pain. *Pain*. 2019;160:833-843. https://doi.org/10.1097/j.pain.0000000000001461
25. Mlekusch S, Neziri AY, Limacher A, Jüni P, Arendt-Nielsen L, Curatolo M. Conditioned pain modulation in patients with acute and chronic low back pain. *Clin J Pain*. 2016;32:116-121. https://doi.org/10.1097/AJP.0000000000000238
26. Neziri AY, Curatolo M, Limacher A, et al. Ranking of parameters of pain hypersensitivity according to their discriminative ability in chronic low back pain. *Pain*. 2012;153:2083-2091. https://doi.org/10.1016/j.pain.2012.06.025
27. Nijs J, Apeldoorn A, Hallegraeff H, et al. Low back pain: guidelines for the clinical classification

of predominant neuropathic, nociceptive, or central sensitization pain. *Pain Physician*. 2015;18:E333-E346.
28. Nijs J, Loggia ML, Polli A, et al. Sleep disturbances and severe stress as glial activators: key targets for treating central sensitization in chronic pain patients? *Expert Opin Ther Targets*. 2017;21:817-826. https://doi.org/10.1080/14728222.2017.1353603
29. Nijs J, Meeus M, Cagnie B, et al. A modern neuroscience approach to chronic spinal pain: combining pain neuroscience education with cognition-targeted motor control training. *Phys Ther*. 2014;94:730-738. https://doi.org/10.2522/ptj.20130258
30. O'Neill S, Manniche C, Graven-Nielsen T, Arendt-Nielsen L. Generalized deep-tissue hyperalgesia in patients with chronic low-back pain. *Eur J Pain*. 2007;11:415-420. https://doi.org/10.1016/j.ejpain.2006.05.009
31. O'Sullivan P, Waller R, Wright A, et al. Sensory characteristics of chronic non-specific low back pain: a subgroup investigation. *Man Ther*. 2014;19:311-318. https://doi.org/10.1016/j.math.2014.03.006
32. Owens MA, Bulls HW, Trost Z, et al. An examination of pain catastrophizing and endogenous pain modulatory processes in adults with chronic low back pain. *Pain Med*. 2016;17:1452-1464. https://doi.org/10.1093/pm/pnv074
33. Özdolap Ş, Sarikaya S, Köktürk F. Evaluation of pain pressure threshold and widespread pain in chronic low back pain. *Turk J Phys Med Rehabil*. 2014;60:32-36. https://doi.org/10.5152/tftrd.2014.71602
34. Puta C, Schulz B, Schoeler S, et al. Somatosensory abnormalities for painful and innocuous stimuli at the back and at a site distinct from the region of pain in chronic back pain patients. *PLoS One*. 2013;8:e58885. https://doi.org/10.1371/journal.pone.0058885
35. Rabey M, Poon C, Wray J, Thamajaree C, East R, Slater H. Pro-nociceptive and anti-nociceptive effects of a conditioned pain modulation protocol in participants with chronic low back pain and healthy control subjects. *Man Ther*. 2015;20:763-768. https://doi.org/10.1016/j.math.2015.02.011
36. Rolke R, Magerl W, Campbell KA, et al. Quantitative sensory testing: a comprehensive protocol for clinical trials. *Eur J Pain*. 2006;10:77-88. https://doi.org/10.1016/j.ejpain.2005.02.003
37. Roussel NA, Nijs J, Meeus M, Mylius V, Fayt C, Oostendorp R. Central sensitization and altered central pain processing in chronic low back pain: fact or myth? *Clin J Pain*. 2013;29:625-638. https://doi.org/10.1097/AJP.0b013e31826f9a71
38. Schenk P, Laeubli T, Klipstein A. Validity of pressure pain thresholds in female workers with and without recurrent low back pain. *Eur Spine J*. 2007;16:267-275. https://doi.org/10.1007/s00586-006-0124-x
39. Simmonds MJ, Claveau Y. Measures of pain and physical function in patients with low back pain. *Physiother Theory Pract*. 1997;13:53-65. https://doi.org/10.3109/09593989709036448
40. Slavin RE. Best evidence synthesis: an intelligent alternative to meta-analysis. *J Clin Epidemiol*. 1995;48:9-18. https://doi.org/10.1016/0895-4356(94)00097-A
41. 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. https://doi.org/10.1097/AJP.0b013e318215f16a
42. Smart KM, Blake C, Staines A, Doody C. Self-reported pain severity, quality of life, disability, anxiety and depression in patients classified with 'nociceptive', 'peripheral neuropathic' and 'central sensitisation' pain. The discriminant validity of mechanisms-based classifications of low back (±leg) pain. *Man Ther*. 2012;17:119-125. https://doi.org/10.1016/j.math.2011.10.002
43. Smart KM, Blake C, Staines A, Thacker M, Doody C. Mechanisms-based classifications of musculoskeletal pain: part 1 of 3: symptoms and signs of central sensitisation in patients with low back (±leg) pain. *Man Ther*. 2012;17:336-344. https://doi.org/10.1016/j.math.2012.03.013
44. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25:603-605. https://doi.org/10.1007/s10654-010-9491-z
45. Tesarz J, Gerhardt A, Leisner S, Janke S, Treede RD, Eich W. Distinct quantitative sensory testing profiles in nonspecific chronic back pain subjects with and without psychological trauma. *Pain*. 2015;156:577-586. https://doi.org/10.1097/01.j.pain.0000460350.30707.8d
46. van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine (Phila Pa 1976)*. 1997;22:2128-2156. https://doi.org/10.1097/00007632-199709150-00012
47. Walton DM, MacDermid JC, Nielson W, Teasell RW, Chiasson M, Brown L. Reliability, standard error, and minimum detectable change of clinical pressure pain threshold testing in people with and without acute neck pain. *J Orthop Sports Phys Ther*. 2011;41:644-650. https://doi.org/10.2519/jospt.2011.3666
48. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed August 19, 2019.
49. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. *Pain*. 2011;152:S2-S15. https://doi.org/10.1016/j.pain.2010.09.030
50. Yildiz SH, Ulaşli AM, Özdemir Erdoğan M, et al. Assessment of pain sensitivity in patients with chronic low back pain and association with HTR2A gene polymorphism. *Arch Rheumatol*. 2017;32:3-9. https://doi.org/10.5606/ArchRheumatol.2017.5846

VIEW Videos on *JOSPT*'s Website

Videos posted with select articles on the *Journal*'s website (**www.jospt.org**) show how conditions are diagnosed and interventions performed. To view the associated videos for an article, click on **Supplementary Material** and scroll down to stream the videos online or download them to your computer or device.

APPENDIX A

SEARCH STRING FOR MEDLINE

("Central Nervous System Sensitization"/ OR hyperalgesia/ OR "Neural Inhibition"/ OR "pain threshold"/ OR hypersensitivity/ OR (sensitization* OR sensitisation* OR desensitization* OR desensitisation* OR hyperalgesi* OR hypoalgesi* OR (central* ADJ3 sensitivit*) OR hyperexcitab* OR (pain ADJ6 (modulat*)) OR ((inhibit* OR facilitat*) ADJ3 mechanism*) OR ((nerve OR neural*) ADJ3 inhibit*) OR (pain ADJ3 (threshold*)) OR algometr* OR hypersensitiv* OR (summat*) OR (quantitativ* ADJ3 sensor* ADJ3 test*) OR qst OR habituat* OR (cognit* ADJ6 modulat*)).ab,ti.) AND ("low back pain"/ OR "back pain"/ OR (((backpain OR backache)) OR (back ADJ3 pain*) OR lowback OR (low* ADJ back) OR ((lumbo* OR lumba*) ADJ6 pain*))) NOT (exp animals/ NOT humans/)

APPENDIX B

THE NEWCASTLE-OTTAWA QUALITY ASSESSMENT SCALE[44]: CASE-CONTROL STUDY

Selection

1. Is the case definition adequate?
2. Representativeness of the cases
3. Selection controls
4. Definition of controls

Comparability

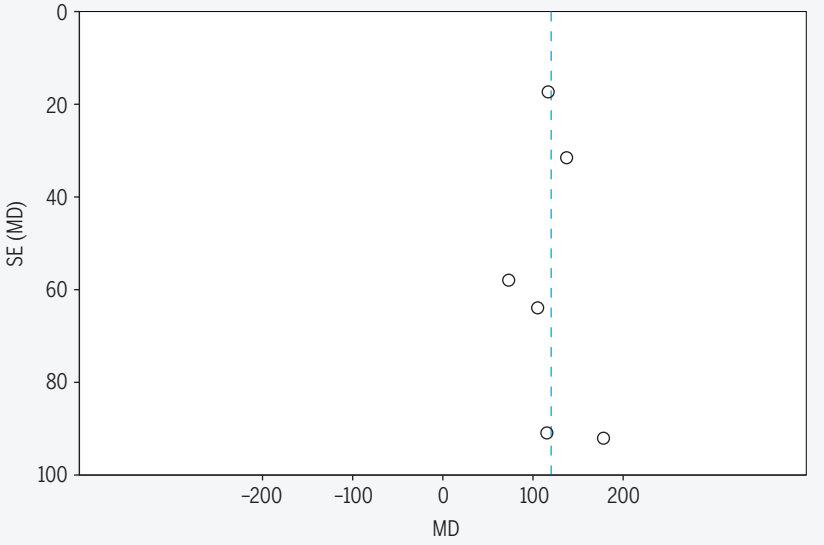
1. Study controls for (select the most important factor: we chose to match by age and sex)
2. Study controls for any additional factor (this criterion can be modified to indicate specific controls for a second important factor)

Exposure

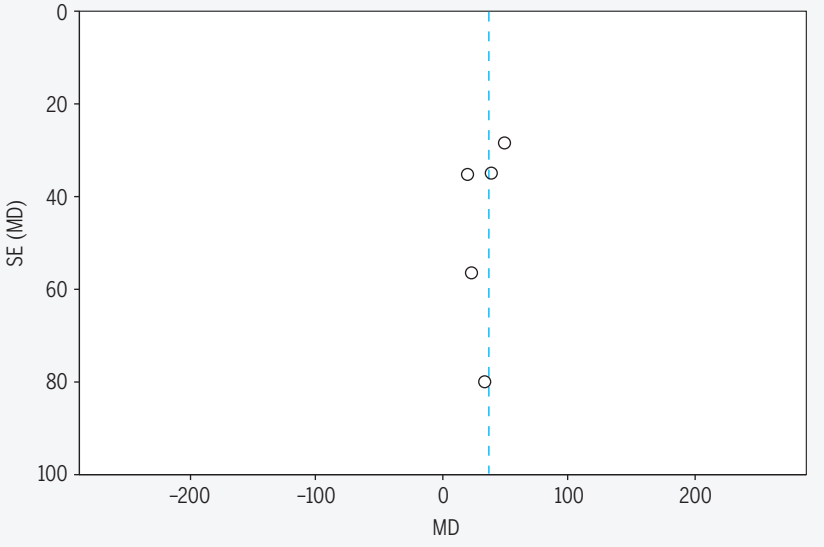
1. Ascertainment of exposure
2. Same method of ascertainment for cases and controls
3. Nonresponse rate

APPENDIX C

FUNNEL PLOTS

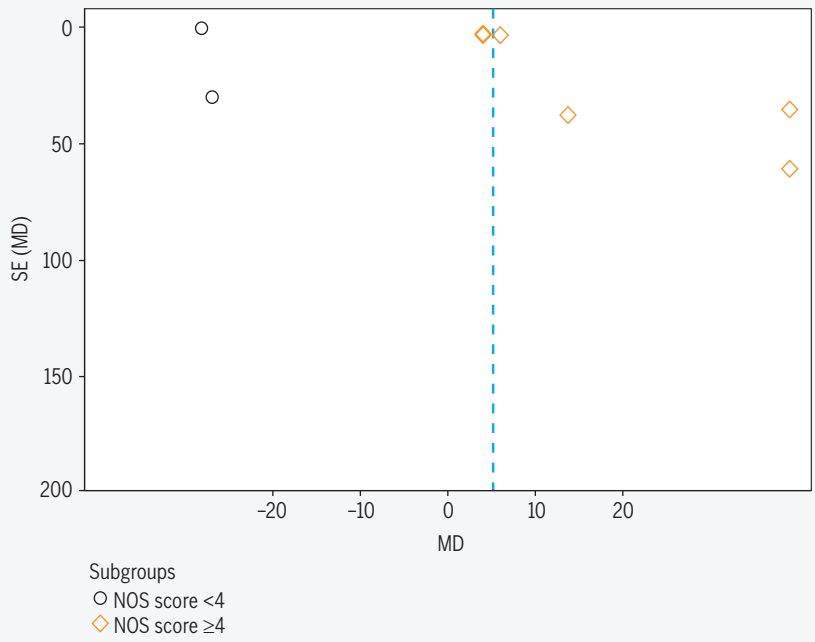


Funnel plot of pressure pain thresholds for the scapula cluster. Abbreviations: MD, mean difference; SE, standard error.

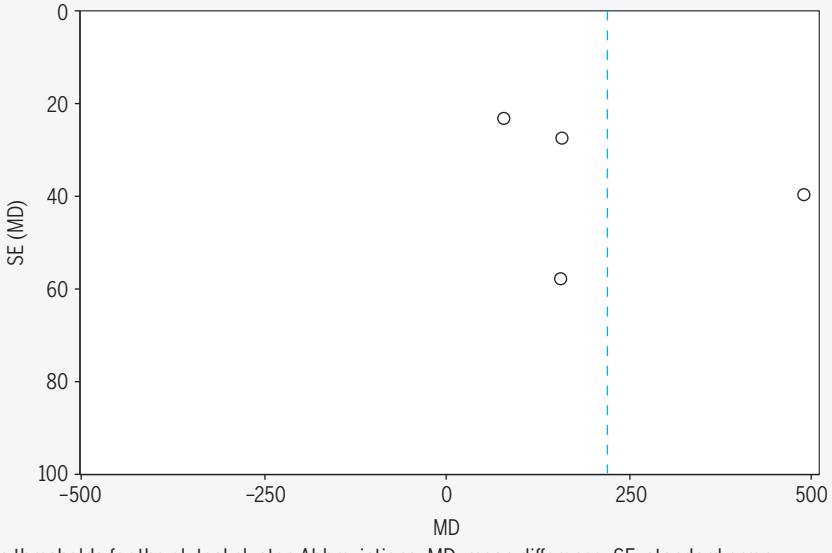


Funnel plot of pressure pain thresholds for the arm cluster. Abbreviations: MD, mean difference; SE, standard error.

APPENDIX C

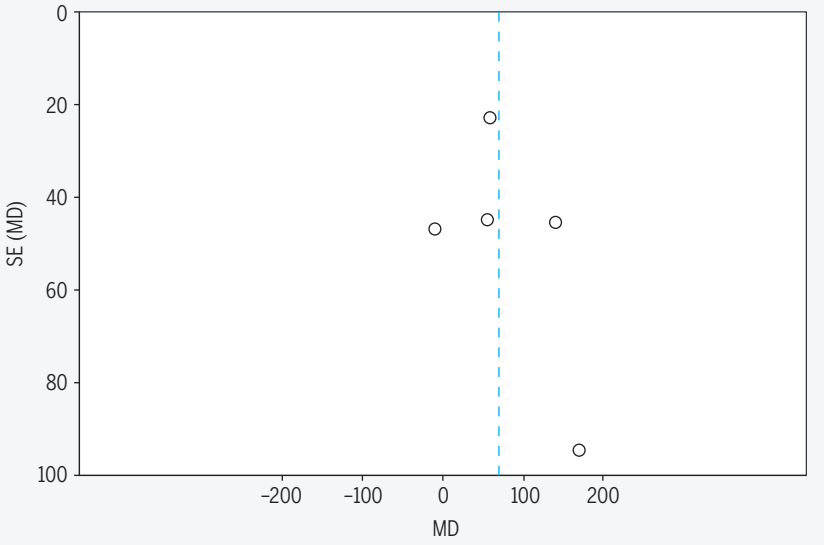


Funnel plot of pressure pain thresholds for the hand cluster. Abbreviations: MD, mean difference; NOS, Newcastle-Ottawa quality assessment scale; SE, standard error.

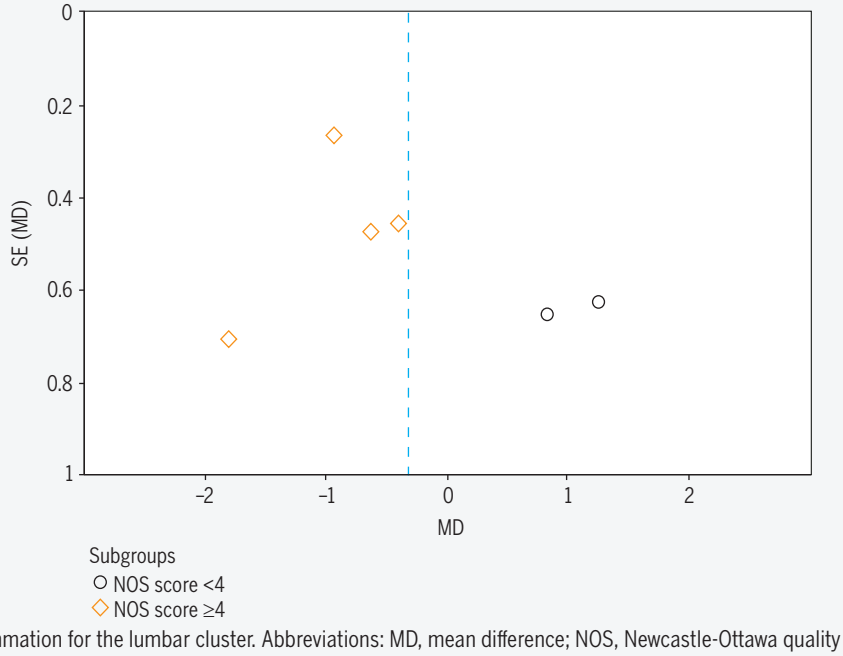


Funnel plot of pressure pain thresholds for the gluteal cluster. Abbreviations: MD, mean difference; SE, standard error.

APPENDIX C

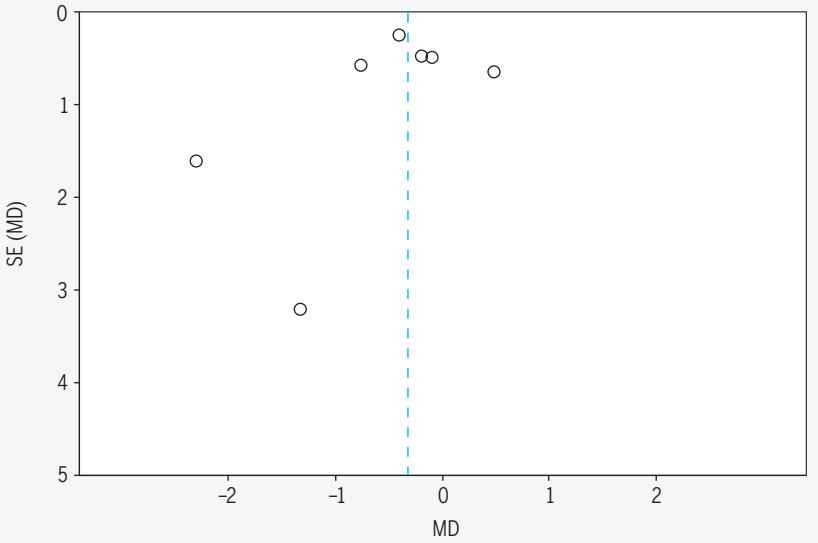


Funnel plot of pressure pain thresholds for the lower-leg cluster. Abbreviations: MD, mean difference; SE, standard error.



Funnel plot of temporal summation for the lumbar cluster. Abbreviations: MD, mean difference; NOS, Newcastle-Ottawa quality assessment scale; SE, standard error.

APPENDIX C



Funnel plot of temporal summation for the hand cluster. Abbreviations: MD, mean difference; SE, standard error.

TRYNTSJE FOKKEMA, MSc[1] • **ROBERT-JAN DE VOS,** MD, PhD[2]
SITA M.A. BIERMA-ZEINSTRA, PhD[1,2] • **MARIENKE VAN MIDDELKOOP,** PhD[1]

Opinions, Barriers, and Facilitators of Injury Prevention in Recreational Runners

BACKGROUND: Effective injury prevention measures for running-related injuries (RRIs) have not yet been identified. More insight into the opinions of runners about injury prevention might help to develop effective injury prevention programs that are supported by the target population.

OBJECTIVES: To describe the opinions of recreational runners on different components of injury prevention, and to identify the barriers to and facilitators of injury prevention in adult recreational runners.

METHODS: In this comparative cross-sectional study, a single questionnaire was sent to 2378 recreational runners. The questionnaire contained questions about their interests, actions undertaken, and perceived barriers to and facilitators of injury prevention. Descriptive analyses were used to examine differences with regard to sex, age, and previous RRIs.

RESULTS: One thousand thirty-four adult recreational runners (43.5%) responded to the questionnaire. Runners with previous RRIs were more likely to rate injury prevention as very useful than runners who had never sustained an RRI (76.8% versus 63.6%, $P<.001$). In total, 81.8% of the participants indicated that they already performed preventive measures, including changes to training schedules (65.4%) and warming up and cooling down (57.8%). Most frequently reported barriers to injury prevention were "not knowing what to do" (45.2%) and "no history of RRI" (34.6%). The most important facilitator was an injury (60.1%). Women more often preferred information via a trainer or running store than did men, while men more frequently preferred websites or e-mail.

CONCLUSION: The majority of runners rated injury prevention as important. To increase effectiveness, future prevention programs should be developed with an awareness of the barriers and facilitators experienced by adult runners. *J Orthop Sports Phys Ther 2019;49(10):736-742. Epub 23 Aug 2019. doi:10.2519/jospt.2019.9029*

KEY WORDS: *injury prevention program, opinions, RRI, running-related injuries*

Running is a sport that is practiced frequently and is still growing in popularity.[20] This is probably because running is an easily accessible and inexpensive sport that can yield fast improvements in physical fitness.[21,23] However, a major drawback of running is the high number of running-related injuries (RRIs). A systematic review[8] from 2015 showed that injury proportions range from 3.2% to 84.9% in adult runners in studies with a follow-up time or recall period between 1 day and lifetime. These percentages indicate a necessity for effective RRI prevention measures.[8] In the last few decades, several randomized trials on RRI prevention have been performed.[2,3,7,10,11,19,22] However, in most trials, there was no significant reduction in the number of RRIs.

According to the Translating Research into Injury Prevention Practice (TRIPP) framework of Finch,[5] identifying etiologic factors that are readily modifiable and consistent with a biological mechanism is important to preventing RRIs. As suggested by Bertelsen et al,[1] insights into how factors influence the dose-response relationship between running participation and injuries will likely increase the understanding of the etiology of RRIs. However, insight into the behavioral context in which injury prevention measures will be implemented is necessary for running injury prevention.[5] Taking the attitudes about, barriers to, and facilitators of injury prevention in athletes into account when designing and implementing injury prevention measures may increase the odds of successful injury prevention. Saragiotto et al[18] explored the beliefs of recreational runners about the most important risk factors for RRIs. Runners think that RRIs are mainly related to (1) training, (2) running shoes, and (3) exceeding the limits of the body. These factors should be considered when developing new injury prevention strategies. To increase our understanding of the attitudes about, barriers to, and facilitators of injury prevention, this exploratory study aimed to (1) describe the opinions of adult recreational runners on different components of injury prevention and compare the opinions of different subgroups of runners, and to (2) identify the barriers to and facilitators of injury prevention in runners.

METHODS

This study is part of the INtervention Study on Prevention of Injuries in Runners at Erasmus [Medical Center] (INSPIRE) trial, a randomized

[1]Department of General Practice, Erasmus University Medical Center, Rotterdam, the Netherlands. [2]Department of Orthopaedics and Sports Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands. The study was approved by the Medical Ethics Committee of the Erasmus University Medical Center (MEC-2016-292). This study was funded by the Netherlands Organization for Health Research and Development (536001001). The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Tryntsje Fokkema, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, the Netherlands. E-mail: t.fokkema@erasmusmc.nl

controlled trial (RCT) on the effectiveness of a multifactorial online RRI prevention program.[6] Recreational runners 18 years of age or older who registered in 2017 for 1 of 3 selected running events (distances ranging from 5 to 42.2 km) were invited to participate in the trial. Participants in the intervention group were given access to the online injury prevention program, which consisted of information on evidence-based risk factors and advice on how to reduce injury risk. Participants in the control group followed their regular preparation for the running event. With 3 follow-up questionnaires, the effectiveness of the prevention program on the number of RRIs was evaluated. In the INSPIRE trial, an RRI was defined as an injury of the muscles, joints, tendons, and/or bones in the lower back or lower extremities (hip, groin, thigh, knee, leg, ankle, foot, and toes) that was caused by running. At least 1 of the following criteria had to be met: (1) the injury was severe enough to cause a reduction in running distance, speed, duration, or frequency for at least 1 week; (2) the injury led to a visit to a doctor and/or physical therapist; or (3) medication was necessary to reduce symptoms as a result of the injury. More details on the INSPIRE trial are published elsewhere.[6] The INSPIRE trial was funded by the Netherlands Organization for Health Research and Development (536001001) and was performed in collaboration with Golazo Sports, a company that organizes large running events in the Netherlands. This study was approved by the Medical Ethics Committee of the Erasmus University Medical Center (MEC-2016-292). The participants signed an informed-consent form before participating, and their rights were protected.

Approximately 7 months after the running event, all participants in the intervention group and control group received an implementation questionnaire containing questions about their interests, preventive actions undertaken, and barriers to and facilitators of injury prevention. For the present study, only data from these implementation questionnaires were used.

The implementation questionnaire consisted of 4 sections. First, information about the runners was collected: sex, date of birth, years of running experience, average running frequency and training volume per week, and previous RRIs. The second section contained questions about RRI prevention. The runners were asked about the factors they thought were important in RRI prevention: healthy lifestyle, running clothes, running shoes, progression of the training program, running technique, running surface, and/or other. Attitudes toward the usefulness of RRI prevention was also captured in this section (very useful, a little useful, or not useful). Participants were asked whether they ever searched for RRI prevention measures (yes or no). Next, they were asked whether they actively performed RRI prevention measures themselves (yes or no). If so, more information on the type of measures was obtained: healthier lifestyle, changes to the training schedule, warming up/cooling down, stretching, changes to clothes, changes to shoes, insoles/orthotics, bandages/braces/taping, compression socks, running surface, changes in running technique, and/or other. In the last section, information on barriers to and facilitators of RRI prevention was obtained. The runners who did not perform preventive measures were asked about the most important barriers to injury prevention (never had an injury, no time, not useful, not amusing, not motivated, does not fit into my training schedule, do not know what to do, and/or other) and facilitators of injury prevention (an RRI, attractive offer of information on prevention, better access to information on RRI prevention, integration into daily training, more knowledge of effectiveness, improving running performance, financial compensation, free supplies for RRI prevention, and/or other). Finally, participants were asked for their preferred ways to receive information on RRI prevention (mobile application, website, e-mail, trainer, running store, magazine, health professional, and/or other).

Differences in characteristics between the participants in the INSPIRE trial who did and did not respond to the implementation questionnaire were determined using independent-samples *t* tests and chi-square tests. For all data collected, means and standard deviations (continuous data) or frequencies and percentages (categorical data) were calculated. To test the impact of the injury prevention program of the INSPIRE trial on the answers to the implementation questionnaire, the responses of participants in the intervention group were compared with those of the control group. Furthermore, subgroup analyses were performed for sex, age (younger than 35 years of age, 35 to 50 years of age, and older than 50 years of age), and previous injuries (yes or no). Subgroup differences were tested using chi-square tests. Analyses were performed in SPSS Statistics Version 24 (IBM Corporation, Armonk, NY), and a P value less than .05 was regarded as statistically significant.

RESULTS

In total, 2378 adult recreational runners participated in the INSPIRE trial, of whom 43.5% (1034 runners) completed the implementation questionnaire. The runners who completed the questionnaire were on average older (44.1 ± 12.5 years versus 39.8 ± 11.2 years, P<.001), had more running experience (7.5 ± 8.8 years versus 5.8 ± 6.9 years, P<.001), and were more often male (55.8% versus 50.4%, P = .014) than the runners who did not respond to this questionnaire. The characteristics of the participants in this study are shown in **TABLE 1**.

Almost three quarters of the participants (74.1%; 95% confidence interval [CI]: 71.3%, 76.7%) rated injury prevention as very useful (**TABLE 2**). Progression of the training program (94.4%; 95% CI: 92.8%, 95.7%), running shoes (76.4%; 95% CI: 73.7%, 78.9%), and running technique (55.8%; 95% CI: 52.7%, 58.9%)

were reported to be the most important aspects of injury prevention. The majority of the participants (68.4%; 95% CI: 65.4%, 71.2%) actively searched for information about injury prevention, and 81.8% (95% CI: 79.3%, 84.1%) performed preventive measures themselves. Preventive measures most often included changes to training schedules (65.4%; 95% CI: 62.0%, 68.6%), warming up and cooling down (57.8%; 95% CI: 54.4%, 61.1%), and stretching (49.8%; 95% CI: 46.3%, 53.2%). The most important barriers reported by runners who did not perform injury prevention were "not knowing what to do" (45.2%; 95% CI: 38.0%, 52.6%) and no history of RRI (34.6%; 95% CI: 27.9%, 41.9%) (**TABLE 3**). The most important reported reason to start injury prevention was an RRI (60.1%; 95% CI: 52.7%, 67.1%). The most important ways to receive information about injury prevention were through mobile applications (49.3%; 95% CI: 46.2%, 52.4%) and websites (45.4%; 95% CI: 42.3%, 48.5%).

Of all responses, only 2 showed a significant difference between participants in the intervention group and those in the control group of the INSPIRE trial: runners in the intervention group performed injury prevention measures more often than participants in the control group (84.4% versus 79.5%, P = .041) and more often preferred to receive information through an app (52.7% versus 46.2%, P = .036).

The results of the subgroup analyses are presented in **TABLES 2** and **3**. Men more often preferred to receive information on injury prevention through websites (49.2% versus 40.5%, P = .005) or e-mail (36.4% versus 29.3%, P = .017) than women, while women more frequently preferred to receive the information personally via a trainer (43.5% versus 31.0%, P<.001) or at a running store (19.0% versus 11.8%, P = .001). More runners younger than 35 years of age would start taking injury prevention measures if they would receive financial compensation (15.2% versus 0.0% and 1.8%, P<.001) or free supplies (34.8% versus 9.2% and 12.3%, P<.001) for injury prevention. Runners with a history of RRI more often experienced a lack of motivation (25.2% versus 12.3%, P = .032) and "not knowing what to do" (59.1% versus 23.3%, P<.001) as barriers to injury prevention than did runners who had not suffered an RRI in the past.

DISCUSSION

The aims of this study were to describe the opinions of adult recreational runners on different components of injury prevention and compare the opinions of different subgroups of runners, and to identify the barriers to and facilitators of injury prevention in these runners. The large majority of participants regarded injury prevention as very useful. The most important barriers for injury prevention were "not knowing what to do" and "no history of RRI," while sustaining an RRI was the most important facilitator of injury prevention. Mobile applications and websites were the most preferred ways to receive information on injury prevention.

Injury prevention is important to recreational runners. In the present study, almost 70% of the runners reported actively searching for information on injury prevention, and over 80% reported performing injury prevention measures themselves. However, the number of RRIs among recreational runners is high, indicating that the injury prevention measures undertaken may not have the intended effect.[8] In this study, recreational runners' opinions on the most important aspects of injury prevention were comparable to those reported by Saragiotto et al[18] regarding risk factors. In both studies, training, running technique, and running shoes were regarded as important aspects for injury prevention. Some of these aspects correspond to known risk factors for RRIs; for example, different aspects of training and running technique.[12,16,24] However, the fact that

TABLE 1 Characteristics of the Study Participants (n = 1034)*

		Sex		Age, y			History of RRI	
	All	Male	Female	<35	35-50	>50	Yes	No
Total, n (%)	1034 (100)	577 (55.8)	457 (44.2)	303 (29.3)	381 (36.8)	350 (33.8)	820 (79.3)	214 (20.7)
Sex (male), n (%)	577 (55.8)	577 (100)	0 (0)	108 (35.3)	214 (56.2)	255 (72.9)	471 (57.4)	106 (49.5)
Age, y	44.1 ± 12.5	47.8 ± 11.9	39.4 ± 23.1	28.7 ± 4.0	43.5 ± 4.4	58.0 ± 5.5	44.6 ± 12.5	42.2 ± 12.4
BMI, kg/m^2	23.7 ± 2.9	24.1 ± 2.7	23.1 ± 3.1	23.1 ± 3.0	23.7 ± 3.1	24.1 ± 2.6	23.8 ± 3.0	23.4 ± 2.9
Running experience, y	7.5 ± 8.8	9.1 ± 10.1	5.4 ± 6.2	3.7 ± 3.5	6.1 ± 6.3	12.3 ± 11.8	7.8 ± 9.2	6.2 ± 6.9
Running frequency, times/wk	2.4 ± 1.2	2.5 ± 1.2	2.3 ± 1.2	2.1 ± 1.1	2.5 ± 1.3	2.6 ± 1.0	2.5 ± 1.2	2.1 ± 1.1
Running distance, km/wk	22.7 ± 15.8	25.6 ± 16.8	19.1 ± 13.5	18.3 ± 14.6	23.5 ± 16.7	25.8 ± 14.7	23.6 ± 15.7	19.5 ± 15.4
Previous RRI (yes), n (%)	820 (79.3)	471 (81.6)	349 (76.4)	236 (77.9)	292 (76.6)	292 (83.4)	820 (100)	0 (0)

Abbreviations: BMI, body mass index; RRI, running-related injury.
**Values are mean ± SD unless otherwise indicated.*

running shoes were also regarded as an important aspect for injury prevention is probably because shoe manufacturers and running stores generally aim to convince runners that wearing a certain type of shoe can prevent injuries. There is an ongoing debate regarding the relationship between running shoes and RRIs; nevertheless, it has never been demonstrated that RRIs can be prevented by wearing a certain type of shoe or by matching shoe type to foot morphology.[9,15] Therefore, future injury prevention programs should be designed with awareness of the perceptions of runners about the most important aspects of injury prevention.[5] Runners should also be informed that there is evidence against the effectiveness of injury prevention via the "prescription" of specific shoes based on the runner's foot type.[15] However, more research is needed to increase our understanding of how and why RRIs occur and to optimize both the content and context of injury prevention measures.[1]

Compared with runners who had suffered an RRI, runners without a history of RRI seemed less interested in injury prevention than runners who had an RRI in the past (ie, they rated the usefulness of injury prevention lower and performed fewer preventive actions themselves). Furthermore, an RRI was rated as the

TABLE 2 Opinions of the Participants About the Importance of Running Injury Prevention and Performing Preventive Measures*

		Sex			Age, y				History of RRI		
	All	Male	Female	*P* Value	<35	35-50	>50	*P* Value	Yes	No	*P* Value
Total, n	1034	577	457		303	381	350		820	214	
How useful is injury prevention?											
Very	74.1	75.0	72.9	.418	69.0	74.8	77.7	.147	76.8	63.6	<.001
A little	25.0	23.7	26.5		29.7	24.4	21.4		22.4	34.6	
Not	1.0	1.2	0.7		1.3	0.8	0.9		0.7	1.9	
What is important for injury prevention?											
Healthy lifestyle	44.8	45.2	44.2	.740	44.6	42.8	47.1	.494	43.0	51.4	.029
Clothes	7.6	6.6	9.0	.152	10.9	7.9	4.6	.010	7.7	7.5	.919
Shoes	76.4	72.1	81.8	<.001	82.2	75.1	72.9	.015	75.7	79.0	.320
Progression of the training schedule	94.4	93.9	95.0	.473	96.7	93.7	93.1	.110	94.6	93.5	.505
Technique	55.8	53.2	59.1	.059	62.7	56.7	48.9	.002	56.3	53.7	.495
Surface	36.5	35.5	37.6	.484	43.9	32.0	34.9	.004	37.3	33.2	.263
Other	7.4	9.0	5.5	.509	4.6	6.8	10.6	.322	7.1	8.9	.199
Do you actively search for injury prevention measures yourself?											
Yes	68.4	67.6	69.4	.542	66.3	68.5	70.0	.603	73.9	47.2	<.001
Do you take injury prevention measures yourself?											
Yes	81.8	80.6	83.4	.250	78.2	82.9	83.7	.149	86.0	65.9	<.001
Which injury prevention measures do you take?†											
Healthier lifestyle	37.0	38.3	35.4	.394	32.9	34.5	43.0	.029	36.6	39.0	.588
Changes to the training schedule	65.4	63.0	68.2	.112	65.0	61.7	69.6	.121	66.8	58.2	.049
Warming up and cooling down	57.8	55.9	60.1	.219	53.6	52.5	66.9	<.001	57.4	59.6	.641
Stretching	49.8	48.0	52.0	.246	48.5	48.4	52.2	.582	49.2	52.5	.479
Changes to clothes	9.2	8.8	9.7	.655	8.0	8.2	11.3	.326	8.9	10.6	.524
Changes to shoes	41.1	32.5	51.7	<.001	49.4	38.6	37.2	.009	40.6	44.0	.453
Insoles or orthotics	26.2	28.8	23.1	.060	18.6	23.7	35.2	<.001	28.4	15.6	.002
Bandages, braces, or tape	7.3	4.9	10.2	.003	8.9	6.3	7.2	.523	8.1	3.5	.059
Compression socks	21.4	21.9	20.7	.672	20.3	22.2	21.5	.864	22.3	17.0	.165
Changes in running surface	24.9	26.2	23.4	.336	29.1	21.2	25.6	.099	25.1	24.1	.804
Changes in running technique	24.2	27.1	20.7	.032	21.9	24.4	25.9	.564	25.2	19.1	.123
Other	12.9	12.9	12.9	.496	11.4	17.4	9.2	.517	12.9	12.8	.352

Abbreviation: RRI, running-related injury.
**Values are percent unless otherwise indicated.*
†Only participants who reported taking injury prevention measures (n = 846) were asked this question.

most important facilitator for injury prevention. Therefore, runners with a history of RRI may have a higher intrinsic motivation for injury prevention. However, runners with a history of RRI may also benefit most from injury prevention measures, because a previous RRI is the most important risk factor for a new RRI.[17,20,21] Therefore, future research on injury prevention could possibly target runners with a previous RRI.

A high percentage of runners (81.8%) performed injury prevention measures. This may be partly related to the fact that the runners participated in an RCT on injury prevention. Runners who are not interested in injury prevention may not have participated in this RCT, and the injury prevention program may have motivated runners in the intervention group to perform injury prevention measures. However, the high percentage of runners in the control group (79.5%) who performed injury prevention measures indicates that many recreational runners perform injury prevention measures. This is important to realize when designing a new RCT on injury prevention. It might make it more difficult to test the effectiveness of injury prevention measures, as it is unlikely that a control group would include only runners who have never performed any injury prevention measure.

TABLE 3 Opinions of the Participants on Barriers to and Facilitators of Injury Prevention*

		Sex			Age, y				History of RRI		
	All	Male	Female	*P* Value	<35	35-50	>50	*P* Value	Yes	No	*P* Value
Total, n	1034	577	457		303	381	350		820	214	
What are the barriers to injury prevention?†											
No history of RRI	34.6	28.6	43.4	.036	33.3	36.9	33.3	.886	1.7	86.3	<.001
No time	11.7	8.0	17.1	.058	18.2	9.2	7.0	.118	13.9	8.2	.237
Not effective	5.9	6.3	5.3	.777	6.1	6.2	5.3	.974	7.8	2.7	.148
Not amusing	11.7	12.5	10.5	.679	13.6	7.7	14.0	.461	14.8	6.8	.099
Not motivated	20.2	21.4	18.4	.614	22.7	12.3	26.3	.129	25.2	12.3	.032
Does not fit in training schedule	7.4	6.3	9.2	.448	7.6	7.7	7.0	.989	8.7	5.5	.413
Not knowing what to do	45.2	46.4	43.4	.684	48.5	43.1	43.9	.800	59.1	23.3	<.001
Other	9.0	8.1	9.2	.447	6.1	9.2	12.3	.422	10.4	6.8	.555
What are the facilitators of injury prevention?†											
An injury	60.1	58.0	63.2	.482	68.2	61.5	49.1	.095	46.1	82.2	<.001
Attractive information	17.6	20.5	13.2	.192	19.7	16.9	15.8	.839	23.5	8.2	.007
Better access to information	28.7	31.3	25.0	.353	31.8	24.6	29.8	.645	35.7	17.8	.008
Integration in daily training	28.2	27.7	28.9	.850	24.2	30.8	29.8	.671	30.4	24.7	.391
More knowledge on effectiveness	31.9	33.9	28.9	.472	27.3	27.7	42.1	.141	37.4	23.3	.043
Improvement in performance	28.2	30.4	25.0	.423	27.3	24.6	33.3	.554	26.1	31.5	.421
Financial compensation	5.9	7.1	3.9	.360	15.2	0.0	1.8	<.001	3.5	9.6	.082
Free supplies for injury prevention	19.1	17.9	21.1	.585	34.8	9.2	12.3	<.001	23.5	12.3	.058
Other	1.6	2.7	0.0	.558	1.5	1.5	1.8	.419	0.0	4.1	.187
What are your preferred ways to receive information on injury prevention?											
Mobile application	49.3	48.0	51.0	.342	50.2	54.9	42.6	.004	49.4	49.1	.933
Website	45.4	49.2	40.5	.005	44.6	44.1	47.4	.628	44.9	47.2	.544
E-mail	33.3	36.4	29.3	.017	30.0	27.6	42.3	<.001	34.1	29.9	.241
Trainer	36.6	31.0	43.5	<.001	39.9	33.3	37.1	.197	37.8	31.8	.103
Running store	15.0	11.8	19.0	.001	22.4	14.4	9.1	<.001	15.5	13.1	.380
Running magazine	13.7	13.2	14.4	.556	13.2	13.9	14.0	.950	13.9	13.1	.757
Health professional	12.9	12.0	14.0	.329	15.8	10.0	13.4	.069	14.0	8.4	.029
Other	4.0	4.3	3.5	.454	3.3	5.0	3.4	.538	4.1	3.3	.530

Abbreviation: RRI, running-related injury.
**Values are percent unless otherwise indicated.*
†Only participants who reported not taking injury prevention measures (n = 188) were asked this question.

Because the most frequently mentioned barrier was "not knowing what to do," future prevention measures should include clear and practical information on injury prevention. An important facilitator was "more knowledge on the effectiveness of the prevention program." Unfortunately, it is impossible to provide such information on a new injury prevention measure that has yet to be tested. However, runners could be informed that the injury prevention measures are, for example, related to risk factors for RRIs and are therefore designed to decrease the number of RRIs. Also, the preferred ways to receive information on injury prevention should be taken into account. Running is an individual sport, and most runners preferred to receive information on injury prevention in an individual way. Mobile applications and websites were the preferred ways to receive information on injury prevention, and, therefore, future injury prevention measures could be delivered via these mediums. Personal ways to deliver information (eg, via a trainer or at a running store) might also be used when targeting women.

Strengths and Limitations

A strength of this study is that, to our knowledge, it is the first to investigate the barriers to and facilitators of injury prevention in adult recreational runners. Another strength is the large sample size. Nevertheless, some limitations need to be addressed. First, only runners who registered for a running event were included in this study. Even though runners from all levels participated in the selected running event, participants of running events may be more fanatic runners than runners who do not participate in running events, which may have caused some bias in the results. Second, all runners in this study participated in an RCT on injury prevention, which may have biased the results. Because runners who are not interested in injury prevention would probably not participate in an RCT on injury prevention, the percentages of runners who rated injury prevention as useful and who performed injury prevention measures might be higher than in the general running population. Furthermore, runners in the intervention group of the INSPIRE trial received information about injury prevention, which may have biased their opinion on important aspects of injury prevention. Another limitation is that the questionnaire used multiple-choice answers. These answer options might have biased the participants' answers to the questions on opinions, barriers, and facilitators by restricting them, as opposed to open-ended questions. However, open-ended questions are known to have a higher rate of missing data.[13] Additionally, we provided an "other" option at the end of each question regarding opinions, barriers, and facilitators, which was open ended and allowed the runners to reflect on their personal beliefs. A fourth limitation is that knowledge of some potential contributors to injury prevention, like nutrition and sleep, was not assessed.[4,14] Another limitation is the relatively low response rate to the implementation questionnaire. More than 50% of the participants in the INSPIRE trial did not respond, which may have biased the results of the current study. There were significant differences between the runners who did and did not respond to the implementation questionnaire. Responders were more often male and relatively older runners. However, it should be mentioned that these differences were very small (less than 4 years in age and slightly more than 5% more men) and may therefore not be of relevance when designing a prevention program. Finally, we did not correct for multiple testing. However, all significant differences between subgroups were large (5.6%-84.6%) and therefore relevant.

CONCLUSION

The majority of adult recreational runners reported that injury prevention is important and performed injury prevention measures themselves. According to the TRIPP framework,[5] it is important to take into account the ideas of runners about injury prevention, as well as the experienced barriers to and facilitators of the implementation of injury prevention measures. We suggest presenting future injury prevention programs on a mobile application and/or website. For women, it might be beneficial to also offer the opportunity to receive information on injury prevention personally (eg, via a trainer or at a running store). Because "not knowing what to do" was the most important reported barrier to injury prevention, future injury prevention programs should contain clear and practical information that runners can easily apply to their training. Finally, future injury prevention programs may primarily target runners with a history of RRI, because these runners seem more motivated to perform preventive measures than runners with no history of RRI. ◉

KEY POINTS

FINDINGS: The majority of adult recreational runners rated injury prevention as very important and performed injury prevention measures themselves. The most frequently reported barriers to injury prevention were "not knowing what to do" and "no history of running-related injury," while the most important facilitator was sustaining an injury.

IMPLICATIONS: To increase effectiveness, future prevention programs should be developed with awareness of the opinions, and experienced barriers and facilitators, of runners.

CAUTION: The runners in this study participated in a randomized controlled trial on injury prevention; runners who were not interested in injury prevention may not have participated. Therefore, the percentages of runners who rated injury prevention as useful and who performed injury prevention measures might be overestimated.

ACKNOWLEDGMENTS: *The authors would like to thank all runners who participated in the INSPIRE trial, and Golazo Sports for its collaboration.*

REFERENCES

1. Bertelsen ML, Hulme A, Petersen J, et al. A framework for the etiology of running-related injuries. *Scand J Med Sci Sports*. 2017;27:1170-1180. https://doi.org/10.1111/sms.12883
2. Bredeweg SW, Zijlstra S, Bessem B, Buist I. The effectiveness of a preconditioning programme on preventing running-related injuries in novice runners: a randomised controlled trial. *Br J Sports Med*. 2012;46:865-870. https://doi.org/10.1136/bjsports-2012-091397
3. Buist I, Bredeweg SW, van Mechelen W, Lemmink KA, Pepping GJ, Diercks RL. No effect of a graded training program on the number of running-related injuries in novice runners: a randomized controlled trial. *Am J Sports Med*. 2008;36:33-39. https://doi.org/10.1177/0363546507307505
4. De Souza MJ, Williams NI, Nattiv A, et al. Misunderstanding the Female Athlete Triad: refuting the IOC consensus statement on Relative Energy Deficiency in Sport (RED-S). *Br J Sports Med*. 2014;48:1461-1465. https://doi.org/10.1136/bjsports-2014-093958
5. Finch C. A new framework for research leading to sports injury prevention. *J Sci Med Sport*. 2006;9:3-9; discussion 10. https://doi.org/10.1016/j.jsams.2006.02.009
6. Fokkema T, de Vos RJ, van Ochten JM, et al. Preventing running-related injuries using evidence-based online advice: the design of a randomised-controlled trial. *BMJ Open Sport Exerc Med*. 2017;3:e000265. https://doi.org/10.1136/bmjsem-2017-000265
7. Hespanhol LC, Jr., van Mechelen W, Verhagen E. Effectiveness of online tailored advice to prevent running-related injuries and promote preventive behaviour in Dutch trail runners: a pragmatic randomised controlled trial. *Br J Sports Med*. 2018;52:851-858. https://doi.org/10.1136/bjsports-2016-097025
8. Kluitenberg B, van Middelkoop M, Diercks R, van der Worp H. What are the differences in injury proportions between different populations of runners? A systematic review and meta-analysis. *Sports Med*. 2015;45:1143-1161. https://doi.org/10.1007/s40279-015-0331-x
9. Knapik JJ, Trone DW, Tchandja J, Jones BH. Injury-reduction effectiveness of prescribing running shoes on the basis of foot arch height: summary of military investigations. *J Orthop Sports Phys Ther*. 2014;44:805-812. https://doi.org/10.2519/jospt.2014.5342
10. Malisoux L, Chambon N, Delattre N, Gueguen N, Urhausen A, Theisen D. Injury risk in runners using standard or motion control shoes: a randomised controlled trial with participant and assessor blinding. *Br J Sports Med*. 2016;50:481-487. https://doi.org/10.1136/bjsports-2015-095031
11. Malisoux L, Chambon N, Urhausen A, Theisen D. Influence of the heel-to-toe drop of standard cushioned running shoes on injury risk in leisure-time runners: a randomized controlled trial with 6-month follow-up. *Am J Sports Med*. 2016;44:2933-2940. https://doi.org/10.1177/0363546516654690
12. Malisoux L, Nielsen RO, Urhausen A, Theisen D. A step towards understanding the mechanisms of running-related injuries. *J Sci Med Sport*. 2015;18:523-528. https://doi.org/10.1016/j.jsams.2014.07.014
13. Millar M, Dillman D. Do mail and Internet surveys produce different item nonresponse rates? An experiment using random mode assignment. *Survey Pract*. 2012;5:011. https://doi.org/10.29115/SP-2012-0011
14. Mountjoy M, Sundgot-Borgen JK, Burke LM, et al. IOC consensus statement on Relative Energy Deficiency in Sport (RED-S): 2018 update. *Br J Sports Med*. 2018;52:687-697. https://doi.org/10.1136/bjsports-2018-099193
15. Napier C, Willy RW. Logical fallacies in the running shoe debate: let the evidence guide prescription. *Br J Sports Med*. 2018;52:1552-1553. https://doi.org/10.1136/bjsports-2018-100117
16. Nielsen RO, Buist I, Sørensen H, Lind M, Rasmussen S. Training errors and running related injuries: a systematic review. *Int J Sports Phys Ther*. 2012;7:58-75.
17. Saragiotto BT, Yamato TP, Hespanhol Junior LC, Rainbow MJ, Davis IS, Lopes AD. What are the main risk factors for running-related injuries? *Sports Med*. 2014;44:1153-1163. https://doi.org/10.1007/s40279-014-0194-6
18. Saragiotto BT, Yamato TP, Lopes AD. What do recreational runners think about risk factors for running injuries? A descriptive study of their beliefs and opinions. *J Orthop Sports Phys Ther*. 2014;44:733-738. https://doi.org/10.2519/jospt.2014.5710
19. Theisen D, Malisoux L, Genin J, Delattre N, Seil R, Urhausen A. Influence of midsole hardness of standard cushioned shoes on running-related injury risk. *Br J Sports Med*. 2014;48:371-376. https://doi.org/10.1136/bjsports-2013-092613
20. van der Worp MP, ten Haaf DS, van Cingel R, de Wijer A, Nijhuis-van der Sanden MW, Staal JB. Injuries in runners; a systematic review on risk factors and sex differences. *PLoS One*. 2015;10:e0114937. https://doi.org/10.1371/journal.pone.0114937
21. van Gent RN, Siem D, van Middelkoop M, van Os AG, Bierma-Zeinstra SM, Koes BW. Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review. *Br J Sports Med*. 2007;41:469-480; discussion 480. https://doi.org/10.1136/bjsm.2006.033548
22. van Mechelen W, Hlobil H, Kemper HC, Voorn WJ, de Jongh HR. Prevention of running injuries by warm-up, cool-down, and stretching exercises. *Am J Sports Med*. 1993;21:711-719. https://doi.org/10.1177/036354659302100513
23. Verhagen E. Prevention of running-related injuries in novice runners: are we running on empty? *Br J Sports Med*. 2012;46:836-837. https://doi.org/10.1136/bjsports-2012-091505
24. Zadpoor AA, Nikooyan AA. The relationship between lower-extremity stress fractures and the ground reaction force: a systematic review. *Clin Biomech (Bristol, Avon)*. 2011;26:23-28. https://doi.org/10.1016/j.clinbiomech.2010.08.005

STEVEN J. KAMPER, PhD[1]

Confidence Intervals: Linking Evidence to Practice

J Orthop Sports Phys Ther 2019;49(10):763-764. doi:10.2519/jospt.2019.0706

Two previous Evidence in Practice articles described the shift in clinical research toward using between-group differences as the measure of treatment effectiveness.[2,3] One key advantage to reporting the between-group difference (the effect estimate), as opposed to only providing a *P* value from a hypothesis test, is that it tells the reader about the size of the effect.

The effect in a study is called an "estimate" because the data are collected from a sample of people, not from everyone in the population. But what readers need is a measure of the treatment effect in the whole population—the concept of *generalizability* will be covered in the next Evidence in Practice article. The estimate of treatment effect provided by a study is associated with some error, so it is important to know something about how precise the estimate is. An effect estimate and its confidence interval give the reader important information about the size, spread, and direction of the population treatment effect.

Confidence intervals can be applied to estimates from many different types of research questions, for example, about treatment effectiveness, prevalence, risk or prognostic factors, or diagnostic test accuracy. The same principles apply to confidence intervals in any study, but for simplicity, this article focuses on treatment effectiveness studies.

What Are Confidence Intervals?

Confidence intervals span a range of values above and below an effect estimate. The mathematics behind calculating a confidence interval are strongly influenced by the size of the sample and the variability in the treatment effect. A small study will generally have a wider confidence interval compared to a large study. If the treatment is very effective for some people and the control is more effective for others, then the confidence interval will be wide.

A 95% confidence interval tells the reader the following: if the same treatments were compared in 100 randomized controlled trials in the same population, and the researchers generated 100 confidence intervals, the true between-group difference would fall within 95 of those confidence intervals. Five of the 100 confidence intervals would not contain the true between-group difference.

Because researchers typically only conduct a study on the same population once, the reader cannot be completely sure that the confidence interval contains the true effect—hence, 95% confidence. Researchers usually report the 95% confidence interval, but other intervals can be calculated. A 90% confidence interval would be narrower (more precise), but the reader would have less confidence that it contained the true effect. A 99% confidence interval would be wider (less precise), but the reader would be more confident that the interval contained the true population estimate.

Confidence intervals are commonly misinterpreted. A confidence interval is not the range of effects that 95% of patients in the population will fall into. It is also not strictly true to say that there is a 95% chance that the confidence interval includes the true population effect.

The confidence interval is the range of effects that will most likely contain the true mean effect of treatment, compared to the control. The reader can be confident about the size of the mean effect of treatment when the confidence interval is narrow, but unsure when the confidence interval is wide. The true population mean treatment effect is more likely to be close to the effect estimate (in the middle of the confidence interval) rather than near either end of the range.

Using Confidence Intervals in Practice

The confidence interval gives the range of plausible effects clinicians and patients can expect from a treatment. This information should form an important part of the conversation when coming to a shared treatment decision.

The concept of clinically meaningful effects, described in 2 previous Evidence in Practice articles,[2,3] is also relevant when interpreting confidence intervals. If the lower end of the confidence interval is lower than the clinically meaningful effect, it is possible that the treatment does not have a worthwhile benefit, regardless of the size of the effect estimate.

[1]School of Public Health, University of Sydney, Camperdown, Australia; Centre for Pain, Health and Lifestyle, Australia.

When a confidence interval crosses the line of no effect (ie, contains the value of zero when assessing between-group differences), it is possible that the treatment is not more effective than the control (**FIGURE**).

A study with a wide confidence interval does not provide useful information about the effectiveness of a treatment (**FIGURE**). This is partly why small studies are not very helpful when making decisions about treatment, and why well-conducted meta-analyses are valuable. Appropriately combining results from several studies in a single analysis (ie, meta-analysis) narrows the confidence interval and provides a more precise estimate of treatment effect. Researchers in the physical therapy field are reporting effect estimates with confidence intervals more regularly,[1] helping readers to accurately interpret the evidence.

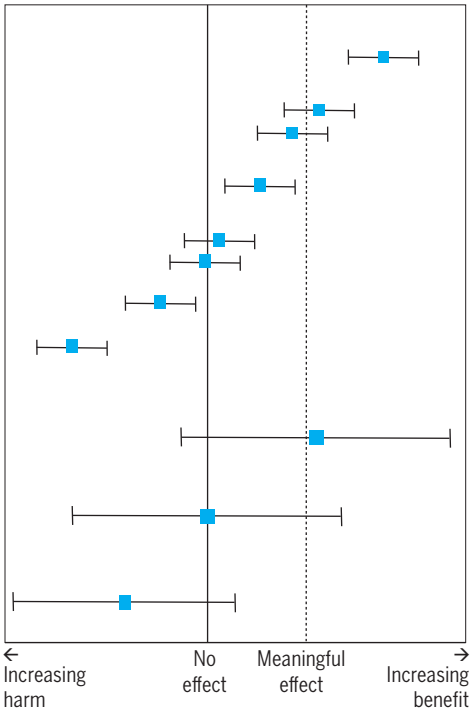


FIGURE. Interpreting confidence intervals. Adapted with permission from Kamper.[4]

Conclusion

Reporting confidence intervals is part of a shift from judging treatment effectiveness solely by *P* values to estimating the size of an effect. Confidence intervals give the reader critical information about the precision of an effect estimate reported in a trial. Integrating information about the likely effect and its precision, along with understanding the concept of clinical meaningfulness, helps the clinician engage patients in an informed, shared decision-making process.

REFERENCES

1. Freire APCF, Elkins MR, Ramos EMC, Moseley AM. Use of 95% confidence intervals in the reporting of between-group differences in randomized controlled trials: analysis of a representative sample of 200 physical therapy trials. *Braz J Phys Ther*. 2019;23:302-310. https://doi.org/10.1016/j.bjpt.2018.10.004
2. Kamper SJ. Interpreting outcomes 2—statistical significance and clinical meaningfulness: linking evidence to practice. *J Orthop Sports Phys Ther*. 2019;49:559-560. https://doi.org/10.2519/jospt.2019.0704
3. Kamper SJ. Interpreting outcomes 3—clinical meaningfulness: linking evidence to practice. *J Orthop Sports Phys Ther*. 2019;49:677-678. https://doi.org/10.2519/jospt.2019.0705
4. Kamper SJ. Showing confidence (intervals). *Braz J Phys Ther*. 2019;23:277-278. https://doi.org/10.1016/j.bjpt.2019.01.003

HENDRIKUS ANTONIUS "RIK" KRANENBURG, PT[1,2] • ROB TYER, PT[3] • MAARTEN SCHMITT, PT, PhD[4] • GERT JAN LUIJCKX, MD, PhD[5]
CEES VAN DER SCHANS, PT, PhD[1,2,6] • NATHAN HUTTING, PT, PhD[7,8] • ROGER KERRY, PT, PhD[9]

Effects of Head and Neck Positions on Blood Flow in the Vertebral, Internal Carotid, and Intracranial Arteries: A Systematic Review

BACKGROUND: Manual therapy interventions targeting the neck include various positions and movements of the craniocervical region. The hemodynamic changes in various spinal positions potentially have clinical relevance.

OBJECTIVES: To investigate the effects of craniocervical positions and movements on hemodynamic parameters (blood flow velocity and/or volume) of cervical and craniocervical arteries.

METHODS: A search of 4 databases (PubMed, Embase, CINAHL, and Index to Chiropractic Literature) and, subsequently, a hand search of reference lists were conducted. Full-text experimental and quasi-experimental studies on the influence of cervical positions on blood flow of the vertebral, internal carotid, and basilar arteries were eligible for this review. Two independent reviewers selected and extracted the data using the double-screening method.

RESULTS: Of the 1453 identified studies, 31 were included and comprised 2254 participants. Most studies mentioned no significant hemodynamic changes during maximal rotation (n = 16). A significant decrease in hemodynamics was identified for the vertebral artery, with a hemodynamic decrease in the position of maximum rotation (n = 8) and combined movement of maximum extension and maximum rotation (n = 4). A similar pattern of decreased hemodynamics was also identified for the internal carotid and intracranial arteries. Three studies focused on high-velocity thrust positioning and movement. None of the studies reported hemodynamic changes. The synthesized data suggest that in the majority of people, most positions and movements of the craniocervical region do not affect blood flow.

CONCLUSION: The findings of this systematic review suggest that craniocervical positioning may not alter blood flow as much as previously expected.

LEVEL OF EVIDENCE: Therapy, level 2a.
J Orthop Sports Phys Ther 2019;49(10):688-697. Epub 5 Jul 2019. doi:10.2519/jospt.2019.8578

KEY WORDS: *blood velocity, cervical spine, hemodynamic*

Manual therapy interventions for the management of people with head and neck pain utilize various positions and movements of the craniocervical region. These interventions have rarely been associated with adverse events.[26,34] Exact incidence rates of adverse events are unknown, and causality between intervention and adverse event is debated.[5,34] Variables such as specific techniques, screening tests, and patient characteristics have been studied in an attempt to enhance the safety of treatment. Unfortunately, studies have been unable to identify specific variables related to the risk of adverse events.[14,21,22,26] However, a suspicion that high-velocity thrust (HVT) techniques may be associated with adverse events remains.[23,54]

Understanding the clinical relevance of arterial pathologies is essential for health care professionals working with the cervical spine.[42] The broad range of pathologies relevant to clinical reasoning and selecting appropriate interventions are defined by the umbrella term *cervical arterial dysfunction* (CAD).[25] This includes arterial events ranging from atherosclerotic disease to mechanical trauma of vessels. One of the most frequently described adverse events following cervical treatment techniques is arterial dissection.[26] Although many other pathological processes are of concern, dissection serves as a useful model to understand the relationship between cervical movement and arterial pathology. The pathophysiology of a dissection is not completely clear. A

[1]Research Group on Healthy Aging, Allied Health Care and Nursing, Hanze University of Applied Sciences, Groningen, the Netherlands. [2]Department of Rehabilitation, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [3]Intermediate Musculoskeletal Assessment and Treatment Team, Connect Health/Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [4]Research Center Innovations in Care, Rotterdam University of Applied Sciences, Rotterdam, the Netherlands. [5]Department of Neurology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [6]Department of Health Psychology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [7]Center for Physiotherapy and Manual Therapy, Rijen, the Netherlands. [8]Department of Occupation and Health, HAN University of Applied Sciences, Nijmegen, the Netherlands. [9]Division of Physiotherapy and Rehabilitation Sciences, University of Nottingham, Nottingham, UK. No funding was received for this study. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Rik Kranenburg, Hanze University of Applied Sciences, PO Box 3109, 9701DC Groningen, the Netherlands. E-mail: h.a.kranenburg@pl.hanze.nl

dissection is characterized by separation of the inner layer (tunica intima) from the middle and outer layers of the arterial wall due to mechanical stress. This separation can lead to a partial or full occlusion of an artery and obstruct blood flow to the brain. Occlusion of 1 artery may not result in direct brain perfusion problems because of the bilateral supply to the brain. In both dissection and nondissection events, a semi-solid, coagulated mass of red and white blood cells can be formed (embolus), eventually, as a consequence, leading to a critical arterial blockage, resulting in a stroke.[3,8]

Several movements of the cervical spine have been postulated to alter the amount of blood flow volume or velocity (hemodynamics) in the cervical vessels.[31] For example, cervical end-range rotation has been reported to be associated with increased stress at the walls of the vertebral artery and internal carotid artery.[31] The hemodynamic parameters of blood flow volume and velocity are considered as robust proxy measures of mechanical stress on vessels and are commonly used to investigate mechanical stress on arteries.[37] Movement-induced stress could potentially initiate acute pathologies such as dissection, or embolus formation in atherosclerotic pathologies. Due to the unique anatomy of the upper cervical spine, roughly half of cervical rotation occurs at the atlanto-occipital joint. The potential mechanical stress on cervical arteries occurring during rotation of the upper cervical spine could potentially compromise the arterial wall of a CAD event in progress.[51] It seems unlikely that a healthy artery would be traumatized by a therapeutic intervention alone.[51] However, an increase of force (such as a cervical manipulation, mobilization, or repeated active movement) during naturally occurring arterial stresses might act as either a causative or exacerbating factor leading to a central neurovascular event (eg, stroke).[10,11]

A commonly described symptom of CAD pathologies is neck or head pain, for which patients may seek assistance from a manipulative physical therapist for evaluation and treatment for relief of pain and improvement of function. Therefore, it is plausible that a CAD is not an adverse event of the treatment itself, but exists in situ prior to treatment.[3] Understanding the mechanical stress that each cervical position or movement puts on the cervical arterial arteries could potentially enhance diagnostic reasoning and the safety of cervical therapeutic interventions.[3]

Mechanical stress on cervical arteries during cervical mobilization or cervical manipulative techniques has been hypothesized as a cause of CAD, especially in patients with pre-existent vascular pathologies.[9,15] Insight into mechanical factors, such as cervical artery blood flow during positions and movements of the cervical spine, may help to decrease the risk for occurrence of CAD after cervical spinal mobilization or manipulation. The aim of this systematic review was to analyze the effects of cervicocranial positions and movements on hemodynamic parameters (blood flow velocity and/or volume) of the cervical and craniocervical arteries.

METHODS

Literature Search

A systematic search was performed in PubMed, Embase, CINAHL, and Index to Chiropractic Literature (ICL) in August 2018. No date range was set. The search strategies developed by 2 authors (H.A.K. and N.H.) were reviewed and adjusted for each database by a senior librarian. All individual search strategies are provided in **APPENDIX A** (available at www.jospt.org). Subsequently, additional literature was identified by related articles (PubMed function; https://www.ncbi.nlm.nih.gov/pubmed/) and by hand searching reference lists of articles included in the review. Additionally, 3 experts who published multiple studies on this topic were asked whether they felt we missed relevant studies. A gray literature search was not performed.

Study Selection

The following inclusion criteria were set a priori: (1) experimental and quasi-experimental research on the influence of cervical positions on blood flow of the vertebral, basilar, and internal carotid arteries; (2) values of blood flow velocity or blood flow volume were described in neutral and altered cervical positions; (3) assessed adult participants; and (4) published in the English language.

Identification

To identify eligible studies, the "double-screening" method was used.[46] All retrieved records were uploaded to Refworks (https://www.refworks.com), and duplicates were removed. The first and second authors (H.A.K. and R.T.) individually determined the eligibility of the articles and, to facilitate interrater reliability, discussed the results after each of the first 5 potentially eligible articles. Articles were scored as "included," "provisionally included," "excluded," or "incomplete" where article titles or abstracts were missing. Differences were discussed, and when the reviewers disagreed the study was included for full-text analysis. A similar procedure was repeated for the full-text articles, and disagreements over inclusion were resolved by a third author. Where an article did not provide adequate information to determine its eligibility, the authors were contacted via e-mail.

Risk-of-Bias Assessment

Because no tool exists to appraise the quality or bias of observational studies or studies for which a reference test does not exist, a modified tool was developed. The foundations of the tool were based on the *Cochrane Handbook for Systematic Reviews of Interventions*, Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2), COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN), and A MeaSurement Tool to Assess systematic Reviews (AMSTAR).[20,24,45,56] With this tool, we critically appraised the selection

bias, attrition bias, reporting bias, and other bias.[20] The tool consisted of 7 parts: (1) specific objectives or hypotheses (other bias), (2) eligibility criteria for participants (selection bias), (3) sample size (other bias), (4) detailed description of interventions for each group (other bias), (5) test conditions similar for all measurements (other bias), (6) prespecified primary and secondary outcome measures (attrition bias), and (7) all of the predefined outcomes were specified in the Results section (reporting bias). The COSMIN was used to weight the sample size (item 3). Two authors with clinical and content-specific expertise (H.A.K. and M.S.) appraised all articles individually.[13] Disagreements were discussed first, then, if no consensus could be reached, a third author was asked to determine the final methodological score.

Data Synthesis and Subgroup Analyses

A data-extraction sheet was composed based on participant characteristics (eg, age and pathologies), the intervention itself (eg, test position, cervical position, cervical artery, and device), and the effect on blood flow (blood velocity or blood volume before, during, and after intervention). Collected data were analyzed using descriptive techniques.

Subgroup analyses were set a priori and made between (1) healthy patients versus patients with vascular pathologies and other pathologies, (2) different positions of the cervical spine, and (3) neutral position and treatment positions.

RESULTS

The results of the search are presented in the **FIGURE**. Of the 1453 identified studies, 67 were considered potentially relevant and reviewed in full text, and all disagreements were resolved by consensus. Of the remaining articles, most were excluded due to language restrictions. Finally, 31 articles met the inclusion criteria and were analyzed by H.A.K. and R.T. Results were compared and discussed without the necessity of a third reviewer.

Study Characteristics

Characteristics of the included studies are summarized in **TABLE 1**.

Participants

The 31 studies included data on 2254 patients, of whom 1162 were male. However, in 4 studies, with a total of 91 individuals, no sex was specified.[28,30,38,57] Overall, the mean age of participants was reported in 25 studies and was 55 years, ranging from 17 to 98 years.

Measurements

The majority (n = 26)[1,2,4,6,7,12,18,19,27-30,32,33,35,36,38,40,41,43,44,47-50,57] of the 31 included studies used a color duplex sonography device to measure flow velocities and flow volumes. The remaining 5 studies used magnetic resonance angiography (n = 3)[17,52,53] and magnetic resonance imaging (n = 2).[39,55]

Participants were mostly tested in a supine position (n = 17).[1,2,4,6,12,17,28-30,39-41,44,50,52,53,55] Other test positions included sitting (n = 7)[7,18,19,32,36,49,57] and prone (n = 1),[33] or were not mentioned (n = 6).[27,35,38,43,47,48]

For the vertebral artery, maximum rotation (n = 18)[2,17,18,29,30,32,33,36,39-41,43,49,50,52,53,55,57] and the combination of maximum rotation and maximum extension (n = 6)[2,6,29,40,49,50] were the cervical positions tested most frequently. Vascular test maneuvers as described by Wallenberg or de Kleyn, which are all combinations of maximum rotation and maximum extension, were included in the latter position.[6] Other cervical positions in which the vertebral artery was tested were maximum rotation and distraction; maximum rotation at C1-C2[53]; rotation at 5° to 15°[50]; rotation at 30°[36]; rotation at 45°[18,29,39,57]; rotation at 60°[35]; maximum extension[2,27,41,50,55]; maximum extension and 45° of rotation[1]; maximum extension, maximum rotation, and distraction[29]; premanipulative positions at C1-C2[2,4,12]; maximum flexion and maximum rotation[49]; distraction[7]; and a posttest in neutral.[57]

For the carotid artery, maximum rotation (n = 4)[40,52,53,55] was also most frequently tested, followed by maximum extension and maximum rotation (n = 2).[28,40] Other described cervical positions

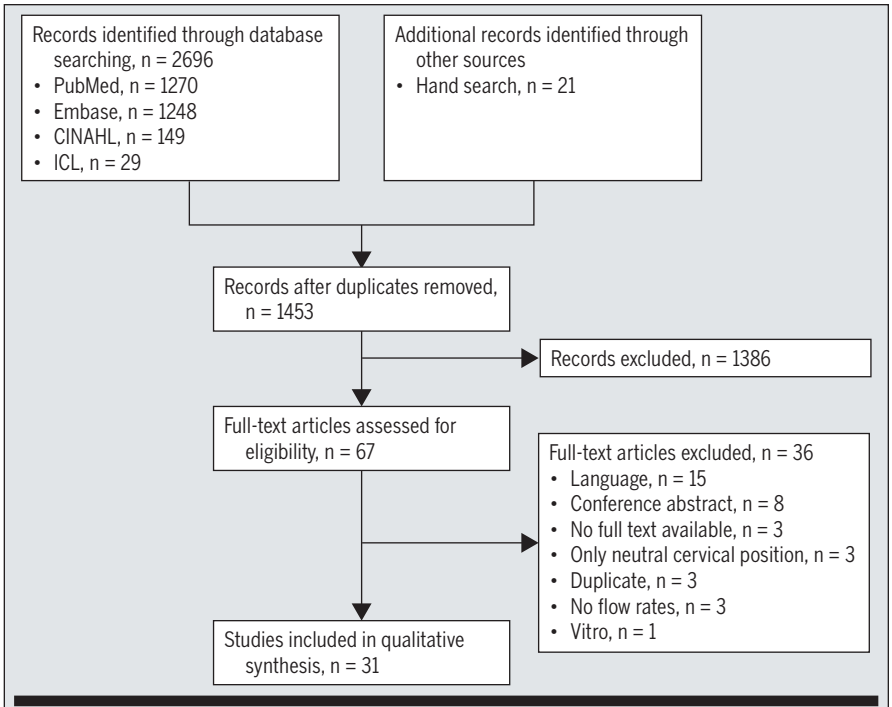


FIGURE. PRISMA flow chart: selection process of relevant studies Abbreviation: ICL, Index to Chiropractic Literature..

TABLE 1 Study Characteristics

Study	Artery	Section	Cervical Positions	Population and Sex	Hemodynamic Effect	Age, y*	Test Position	Device
Hedera et al[19]	BA, ACA, MCA, PCA	P1, transtemporal, suboccipital	Neutral, maximum rotation, maximum extension plus maximum rotation	41 healthy participants (23 male, 18 female) with asymmetry of VA <75%	NSC	47.9 ± 14.1	Sitting	CDS
				11 healthy participants (7 male, 4 female) with asymmetry of VA >75%	SD	47.3 ± 13.8		
Sturzenegger et al[48]	BA	P1	Neutral, maximum rotation, maximum extension, maximum flexion	14 patients with suspected VBI (6 male, 8 female)	NSC, SD	57 (range, 34-76)	Unknown	CDS
Thiel et al[50]	VA	C3-C5	Neutral, maximum rotation, rotation at 5°-15°, maximum extension, maximum extension plus maximum rotation	30 healthy participants (17 male, 13 female)	NSC	28.3 ± 5.3 (range, 19-40)	Supine	CDS
				12 chiropractic patients with a positive Wallenberg test (3 male, 9 female)	NSC, SD	47.4 ± 14.4 (range, 25-68)		
Weintraub and Khoury[55]	VA, CA, BA	NA	Neutral, maximum rotation, maximum extension	64 patients with suspected ischemic cerebrovascular disease (20 male, 44 female)	NSC, SD	70.9 (range, 21-97)	Supine	MRI
				30 healthy patients (10 male, 20 female)	NSC, SD	66.3 (range, 22-80)		
Côté et al[6]	VA	C3-C5	Neutral, maximum extension plus maximum rotation	30 healthy participants (17 male, 13 female)	NSC	28.3 ± 5.3	Supine	CDS
				12 patients with a positive Wallenberg test and dizziness (3 male, 9 female)	NSC, SD	47.4 ± 14.4		
Petersen et al[38]	BA	C0-C1, suboccipital window	Neutral, maximum rotation	46 patients with VBI (28 male, 18 female)	SD	62 ± 1.5 (range, 41-83)	Unknown	CDS
				25 healthy young participants (sex unknown)	NSC	26 ± 0.48 (range, 22-30)		
				15 healthy elderly participants (sex unknown)	SD	59 ± 2.06 (range, 50-75)		
Licht et al[30]	VA	Origin and foramen, C6	Neutral, maximum rotation, rotation at 45°	20 healthy participants (sex unknown)	NSC	Unknown	Supine	CDS
Rivett et al[40]	VA, CA	C3-C5	Neutral, maximum rotation, rotation at 45°, maximum extension plus maximum rotation	10 patients with a positive premanipulative test (2 male, 8 female)	NSC, SD	37.9 ± 13.0 (range, 24-65)	Supine	CDS
				10 healthy participants (2 male, 8 female)	NSC, SD	32.7 ± 10.3 (range, 20-47)		
Li et al[27]	VA	C0-C1	Neutral, maximum rotation, maximum extension, maximum extension plus maximum rotation	27 healthy elderly participants (21 male, 6 female)	NSC	62 (range, 60-72)	Unknown	CDS
				23 healthy participants (all male)	NSC	21 (range, 19-22)		
Licht et al[29]	VA	Origin and foramen, C6	Neutral, maximum rotation, rotation at 45°, maximum extension plus maximum rotation, maximum extension plus maximum rotation plus distraction	20 chiropractic patients with positive vascular premanipulative tests (5 male, 15 female)	NSC	Median, 44 (range, 27-74)	Supine	CDS

Table continues on page 692.

TABLE 1 Study Characteristics (continued)

Study	Artery	Section	Cervical Positions	Population and Sex	Hemodynamic Effect	Age, y*	Test Position	Device
Haynes and Milne[18]	VA	C2	Neutral, maximum rotation, rotation at 45°	20 patients with neck-related symptoms (9 male, 11 female)	NSC	39 ± 4.2 (range, 20-52)	Sitting	CDS
Haynes et al[17]	VA	NA	Neutral, maximum rotation	8 healthy participants (6 male, 2 female)	NSC	44.4 ± 14.1 (range, 25-61)	Supine	MRA
Licht et al[28]	CA	Unknown	Neutral, maximum extension plus maximum rotation	11 patients with a positive vascular premanipulative test (sex unknown)	NSC	Unknown	Supine	CDS
Mitchell[33]	VA	C0-C1	Neutral, maximum rotation	120 healthy participants (60 male, 60 female)	SD	Range, 20-30	Prone	CDS
Rivett et al[41]	VA	C1-C2, C2-C3	Neutral, maximum rotation, maximum extension	20 healthy participants (8 male, 12 female)	NSC	35.5 ± 9.3 (range, 24-55)	Supine	CDS
Sakaguchi et al[43]	VA	C4-C6	Neutral, maximum rotation	1108 patients referred for neurovascular examination (710 male, 398 female)	SD	61.4 ± 12.9	Unknown	CDS
Zaina et al[57]	VA	C1-C2, C5-C6	Neutral, maximum rotation, rotation at 45°	20 healthy participants (sex unknown)	NSC, SD	32.7 ± 8.82	Sitting	CDS
Arnold et al[2]	VA	C3-C5	Neutral, maximum rotation, maximum extension, maximum extension plus maximum rotation, premanipulative position	22 healthy participants (8 male, 14 female)	NSC, SD	35 ± 10.5	Supine	CDS
Mitchell et al[32]	VA	C0-C1	Neutral, maximum rotation	30 healthy participants (all female)	NSC, SD	21	Sitting	CDS
Ozdemir et al[36]	VA	C2-C6	Neutral, maximum rotation, rotation at 30°	28 patients with cervical degenerative changes (11 male, 17 female)	NSC, SD	51 (range, 44-76)	Sitting	CDS
				24 patients with clinically proven VBI (14 male, 10 female)	NSC, SD	47 (range, 36-58)		
				20 healthy participants (8 male, 12 female)	NSC, SD	36 (range, 19-40)		
Sultan et al[49]	VA, MCA, PCA	Above C6, P1, P2	Maximum rotation, maximum extension plus maximum rotation, maximum flexion plus maximum rotation	46 patients with suspected positional VBI (16 male, 30 female)	NSC	69 (range, 32-98)	Sitting	CDS
Bowler et al[4]	VA	C2-C3	Neutral, premanipulative position	14 healthy participants (3 male, 11 female)	NSC	31 ± 10.76 (range, 19-49)	Supine	CDS
	CA	2 cm proximal to bifurcation			SD			
Thomas et al[53]	VA, CA, TCI	NA	Neutral, maximum rotation, maximum rotation plus distraction, maximum rotation at C1-C2, distraction	20 healthy participants (10 male, 10 female)	NSC	33.1 ± 11.9 (range, 21-59)	Supine	MRA
Quesnele et al[39]	VA	C1-C2	Neutral, maximum rotation, rotation at 45°, manipulation at C1-C2	10 healthy participants (all male)	NSC	26.8 ± 1.6 (range, 24-30)	Supine	MRI
Erhardt et al[12]	VA	V3	Neutral, premanipulative position, manipulation at C1-C2	23 healthy participants (9 male, 14 female)	NSC	40 (range, 27-69)	Supine	CDS
Thomas et al[52]	VA, CA	NA	Neutral, maximum rotation	20 healthy participants (10 male, 10 female)	NSC, SD	33.1 ± 11.9 (range, 21-59)	Supine	MRA
Siwach et al[47]	ACA, MCA, PCA	Unknown	Neutral, maximum extension, maximum flexion	50 spondylosis patients (23 male, 27 female)	NSC	45.4 ± 11.9 (range, 20-70)	Unknown	CDS
Saracoglu et al[44]	CA	2 cm proximal to bifurcation	Neutral, semi-Fowler extension plus 10° of collateral rotation	28 patients during thyroid surgery (6 male, 22 female)	SI, NSC, SD	39.1 ± 9.8 (range, 18-50)	Supine	CDS

Table continues on page 693.

TABLE 1 Study Characteristics (continued)

Study	Artery	Section	Cervical Positions	Population and Sex	Hemodynamic Effect	Age, y*	Test Position	Device
Araz Server et al[1]	VA	V1, V2, V3, V4	Neutral, maximum rotation, maximum extension plus rotation at 45°	21 patients with vestibular symptoms (3 male, 18 female)	NSC, SD	45.5 ± 11.1	Supine	CDS
				21 healthy participants (5 male, 16 female)	NSC	41.3 ± 9.2		
Creighton et al[7]	VA	C6 transverse foramen	Neutral, traction	30 individuals (healthy or patients unclear; sex unknown)	NSC	36.6 (range, 21-57)	Sitting	CDS
Niewiadomski et al[35]	VA	Unknown	Neutral, rotation at 60°	50 patients with vertigo and/or hearing loss and vessel anomalies (20 male, 30 female)	NSC	49.9 (range, 17-79)	Unknown	CDS
				50 healthy participants (26 male, 24 female)	NSC	44.4 (range, 20-71)		

Abbreviations: ACA, anterior cerebral artery; BA, basilar artery; CA, carotid artery; CDS, color duplex sonography; MCA, middle cerebral artery; MRA, magnetic resonance angiography; MRI, magnetic resonance imaging; NA, not applicable; NSC, no significant change; PCA, posterior cerebral artery; SD, significant decrease; SI, significant increase; TCI, total cerebral input; VA, vertebral artery; VBI, vertebrobasilar insufficiency.
**Values are mean or mean ± SD unless otherwise indicated.*

for the carotid artery were maximum rotation and distraction, maximum rotation at C1-C2, rotation at 45°, maximum extension, premanipulative positions, a semi-Fowler position, extension and 10° of collateral rotation, and a posttest in neutral.

The intracranial arteries were most frequently tested in maximum rotation (n = 5)[19,38,48,49,53] and maximum extension (n = 3).[47,48,55] The other cervical positions for this artery included maximum rotation and distraction,[53] maximum rotation at C1-C2,[53] extension and maximum rotation,[19,49] maximum flexion,[47,48] maximum flexion and maximum rotation,[49] distraction,[53] and a posttest in neutral.[38]

Hemodynamic Changes

Fourteen studies[7,12,17,18,27-30,35,39,41,47,49,53] reported no significant hemodynamic changes in any of the included cervical positions, whereas 2 studies[33,43] reported a significant hemodynamic decrease in all of the included cervical positions. The majority of studies reported no significant hemodynamic changes during maximum rotation (n = 16).[2,17,18,29,30,32,36,39-41,43,49,50,52,53,57] The significant changes most commonly identified for the vertebral artery were a hemodynamic decrease in maximum rotation (n = 8)[2,32,33,36,40,43,52,55] and in the combined movement of maximum extension and maximum rotation (n = 4).[2,6,40,50] A similar pattern was also identified for maximum rotation and the combined movement of maximum extension and maximum rotation in relation to the hemodynamics of the internal carotid and intracranial arteries. One study mentioned an increase in peak flow velocity and time-averaged mean flow velocity in the carotid artery.[44] However, this was post induction in a presurgery situation.

A specification of all cervical positions combined with hemodynamic changes by artery can be found in **APPENDIX B** (available at www.jospt.org).

Subgroup Analyses

Twenty-two studies used groups with healthy participants.[1,2,4,6,12,17,19,27,30,32,33,35,36,38-41,50,52,53,55,57] Twelve studies used groups with people with vascular pathology.[6,28,29,35,36,38,40,43,48-50,55] Five studies mentioned nonvascular participant groups.[1,18,36,44,47] One study did not report whether the participants were healthy or had a pathology.[7] A comparison of the groups of people with vascular pathology and groups of other patients shows that there were proportionally no differences.

Manipulations were mentioned for the vertebral artery only.[12,39] Both studies scored well in our risk-of-bias assessment, except for a risk of bias in Quesnel et al[39] due to moderate sample size. Quesnele et al[39] included 10 healthy participants and Erhardt et al[12] 23 participants (**TABLE 2**). The relationship between premanipulative position and the vertebral artery was reported in 3 studies[2,4,12] and the carotid artery in 1 study.[4] Only Arnold et al[2] reported that a premanipulative position significantly decreased the velocity and resistance index. However, this relationship was not found for both arteries in left and right rotation. Bowler et al[4] mentioned a significant decrease in the resistance index, but not in peak systolic velocity, end diastolic velocity, and mean velocity. The other study mentioned no significant difference in flow velocities or resistance index.[12]

Risk-of-Bias Assessment

The results are presented in **TABLE 2**. No studies were scored as having a high risk of bias. Five articles[2,7,33,35,43] were scored as having no risk of bias, and no article scored positive on more

TABLE 2 Risk of Bias

	Item*							
Study	1†	2†	3‡	4†	5†	6†	7†	Total Score
Araz Server et al[1]	0	0	2	0	0	0	0	2
Arnold et al[2]	0	0	0	0	0	0	0	0
Bowler et al[4]	0	0	3	0	0	0	0	3
Côté et al[6]	0	0	2	1	0	0	0	3
Creighton et al[7]	0	0	0	0	0	0	0	0
Erhardt et al[12]	0	0	3	0	0	0	0	3
Haynes and Milne[18]	0	0	3	0	0	0	0	3
Haynes et al[17]	0	0	3	0	0	0	0	3
Hedera et al[19]	0	0	1	0	0	0	0	1
Li et al[27]	1	0	1	0	0	0	0	2
Licht et al[28]	0	0	3	0	0	0	0	3
Licht et al[29]	0	0	3	0	0	0	0	3
Licht et al[30]	1	0	3	0	0	0	0	4
Mitchell et al[32]	0	0	3	0	0	0	0	3
Mitchell[33]	0	0	0	0	0	0	0	0
Niewiadomski et al[35]	0	0	0	0	0	0	0	0
Ozdemir et al[36]	0	0	1	0	0	0	0	1
Petersen et al[38]	0	0	1	0	0	0	0	1
Quesnele et al[39]	0	0	3	0	0	0	0	3
Rivett et al[40]	0	0	3	0	0	0	0	3
Rivett et al[41]	0	0	3	0	0	0	0	3
Sakaguchi et al[43]	0	0	0	0	0	0	0	0
Saracoglu et al[44]	0	0	2	0	0	0	0	2
Siwach et al[47]	0	0	1	0	0	0	0	1
Sturzenegger et al[48]	1	0	3	0	0	0	0	4
Sultan et al[49]	1	0	2	0	0	0	0	3
Thiel et al[50]	0	0	2	0	0	0	0	2
Thomas et al[52]	0	0	3	0	0	0	0	3
Thomas et al[53]	0	0	3	0	0	0	0	3
Weintraub and Khoury[55]	1	0	0	0	0	0	0	1
Zaina et al[57]	1	0	3	0	0	0	0	4

**Items: (1) Specific objectives or hypotheses; (2) Eligibility criteria for participants; (3) Sample size; (4) Detailed description of interventions for each group; (5) Test conditions similar for all measurements; (6) Prespecified primary and secondary outcome measures; (7) All of the predefined outcomes were specified in the Results section.*

†Zero points are awarded for answers of "yes" or "not applicable," and 1 point is awarded for answers of "no" or "can't answer."

‡Scoring is as follows: "adequate sample size" (100 or more participants), 0 points; "good sample size" (50-99 participants), 1 point; "moderate sample size" (30-49 participants), 2 points; "small sample size" (fewer than 30 participants), 3 points.

than 2 of the 7 parts of the assessment tool. Risk of bias due to a moderate or small sample size was found in 20 studies.[1,4,6,12,17,18,28-30,32,39-41,44,48-50,52,53,57] Risk of bias due to inadequate sample size (item 3) was found in 15 studies.[4,12,17,18,28-30,32,39-41,48,52,53,57] Risk of bias as a result of inadequately described objective or hypothesis was found in 6 studies.[27,30,48,49,55,57] One study did not provide a detailed description of the interventions for each group.[6]

DISCUSSION

The data synthesized from 31 experimental and quasi-experimental studies suggest that in most people, craniocervical positions and movements had no effect on blood flow. In a small proportion of the groups "healthy subjects," "vascular patients," and "other patients," blood flow does decrease during some movements, specifically maximum rotation and/or maximum extension. The positions and movements utilized in HVT techniques do not seem to alter blood flow. A clinical implication from this review is that the relationship between craniocervical movement and alterations in blood flow does not seem to be as obvious as previous data suggested. Considering blood flow as a robust measure of vessel stress, based on these data, it is unlikely that head and neck movement alone, even if forceful, could mechanistically explain the etiology of adverse events that has conventionally been purported to be related to therapeutic interventions.

Hemodynamic parameters act as a proxy measure for mechanical stress on cervical arteries. The rationale for vessel stress in healthy persons and patients with vascular pathology is similar. When stress is applied to a vessel, the diameter changes and can alter the blood flow velocity or volume. Therefore, when a cervical positional change puts stress on a vessel, it should theoretically also change the hemodynamics. Most studies reported no change in hemodynamic parameters during all tested movements and positions, in both healthy and vascular/other groups. Some studies reported hemodynamic changes during maximum rotation and maximum extension when performed in either isolation or when combined. More positions were found to influence hemodynamic parameters in studies that included people with vascular pathology and other patients. Overall, the pattern of hemodynamic responses to cervical position and movement seems to be a naturally occurring phenomenon related to the anatomy of the cervicocranial

region. This conclusion is supported by both the high proportion of studies that demonstrated no changes at all in any group and the proportion that showed changes in healthy participants. The differences in hemodynamic parameters between healthy and vascular/other participants are only in terms of the number of positions where changes were identified. Conventional thought within the domain of manual therapy has been that rapid, forceful interventions such as HVT techniques are considered to constitute a higher risk for neurovascular events resulting from cervical arterial compromise. However, we found that studies that focused specifically on HVT reported no hemodynamic changes. Furthermore, studies that reported positioning and movement were ambiguous in reporting hemodynamic changes.

Various studies investigated hemodynamics in single or multiple cervical positions, in a single artery, or in relation to treatment technique.[25,31] However, these data had not been previously synthesized. Our findings are similar to the conclusions of previous reviews on this topic. Mitchell[31] conducted a meta-analysis of data from 9 studies (n = 204 participants) and reported that contralateral rotation was the movement most commonly associated with a reduction of flow parameters. This occurred more in patients than it did in healthy participants. Mitchell[31] also reported that studies in which patients experienced symptom reproduction (specifically for vertebral artery insufficiency) during the compromising movement did not establish an association between flow change and symptoms. This observation could have implications for the validity of testing procedures that rely on this underlying mechanism, for example, functional positional tests. In our review, the recording of symptom reproduction in the included studies was insufficient to allow us to draw any conclusions in line with those of Mitchell.[31] This might be explained by the broader inclusion criteria and the studies published after 2009. We included 25 studies for the vertebral artery, versus 9 in Mitchell's study.[31] Hutting et al[22] reviewed 4 blood flow studies (n = 1271) to examine the concept of diagnostic accuracy of functional positional testing. They, too, were unable to establish a relationship between flow changes and symptom reproduction. The aim of these vascular integrity test procedures is to unilaterally compress an artery to test the contralateral blood supply. However, our data suggest that testing based on this mechanism may not be a valid construct. Therefore, the rationale and value of the tests should be questioned. Hemodynamic patterns in Mitchell's study[31] were in agreement with those found in the current review.

The present data have potential clinical implications for the use of therapeutic interventions for the management of people with head and neck pain. There appear to be no consistently reported positions that induce greater hemodynamic responses than others. The 2 studies that focused on HVT did not find a hemodynamic effect, either.[12,39] However, it cannot be ruled out that rapid, forceful movements may trigger vascular-wall trauma, which is not identifiable by the measurement parameters assessed in the current review. We therefore cannot conclude that all interventions are equally safe, especially because the 2 studies had a moderate and inadequate sample size.[12,39] This point is in agreement with the key developments highlighted in the latest International Federation of Orthopaedic Manipulative Physical Therapists practice framework, which promotes a more holistic consideration of risk management, including factors other than just the effect of a specific intervention (eg, underlying pathology, cardiovascular risk factors, etc).[42] The present data support this reasoning, which suggests that adverse events related to cervical spine interventions might be the result of something other than the therapeutic positioning or movement of the head and neck. Clinicians should be mindful, however, that there may be small subgroups of the population with underlying arterial pathology, in whom small hemodynamic changes may be sufficient to induce or exacerbate serious neurovascular compromise. Therefore, it might be wise to choose treatment techniques first in positions with less than 45° of cervical rotation, as the data from the included studies are most consistent in these positions.

Limitations

We considered a number of possibilities to provide a meaningful quality assessment, but due to the wide variation of study type, no available reference standard for what constitutes high quality in the constituent variables of these particular methods, and a lack of focus toward a specific intervention or diagnosis, a suitable validated tool was not available. Given the importance of assessing the risk of bias, the authors developed a new tool, as suggested in the *Cochrane Handbook for Systematic Reviews of Interventions*,[20,24] based on the Delphi principle. The primary concept of the tool was based on literature and reviewed in 2 more rounds.[16] Five studies[2,7,33,35,43] were scored as having no risk of bias, and none of the others were scored as having a risk of bias on more than 2 of the 7 points. In general, no study was scored as having a high risk of bias. The most reported bias was a small sample size. Although this quality tool was developed thoughtfully, it did not detect ambiguities in the study of Niewiadomski et al.[35] The authors did not present all data to substantiate their conclusions and did not respond to an e-mail requesting further explanation. A second limitation is the lack of quantifiable change in terms of unit measurement. The heterogeneity and variety of flow and velocity parameters precluded a standardized unit for comparisons or judgments of effect size. Due to this methodological diversity, we decided to conduct a high-quality synthesis instead of a meta-analysis.[13] Further, there is no a priori reference standard for what constitutes significant change when using

blood flow parameters as a proxy measure for vessel stress.

In future research, we advise authors to report all data available, such as standard deviations, confidence intervals, and all hemodynamic outcomes. The availability of these parameters would enhance the ability to perform a meta-analysis.

CONCLUSION

Our results suggest that in most people—healthy people as well as patients with vascular pathologies—craniocervical positions do not alter cervical blood flow. This includes vascular test positions, premanipulative positions, and manipulations. ◉

KEY POINTS

FINDINGS: A key clinical implication from this review is that the relationship between craniocervical movement and blood flow does not seem to be as strong as previously suggested.

IMPLICATIONS: Although the majority of the included studies found no significant decrease in end-range positions, the data were most consistent in positions with less than 45° of rotation. Therefore, clinicians should initially consider treatment techniques within this range.

CAUTION: The absence of a reference standard and the heterogeneity of data made it impossible to calculate effect sizes and perform a meta-analysis.

REFERENCES

1. Araz Server E, Tuna Edizer D, Yigit Ö, Görkem Yasak A, Erdim Ç. Relationship between vertebral artery blood flow in different head positions and vertigo. *Acta Otolaryngol*. 2018;138:1-5. https://doi.org/10.1080/00016489.2017.1373849
2. Arnold C, Bourassa R, Langer T, Stoneham G. Doppler studies evaluating the effect of a physical therapy screening protocol on vertebral artery blood flow. *Man Ther*. 2004;9:13-21. https://doi.org/10.1016/S1356-689X(03)00087-0
3. Biller J, Sacco RL, Albuquerque FC, et al. Cervical arterial dissections and association with cervical manipulative therapy: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2014;45:3155-3174. https://doi.org/10.1161/STR.0000000000000016
4. Bowler N, Shamley D, Davies R. The effect of a simulated manipulation position on internal carotid and vertebral artery blood flow in healthy individuals. *Man Ther*. 2011;16:87-93. https://doi.org/10.1016/j.math.2010.07.007
5. Church EW, Sieg EP, Zalatimo O, Hussain NS, Glantz M, Harbaugh RE. Systematic review and meta-analysis of chiropractic care and cervical artery dissection: no evidence for causation. *Cureus*. 2016;8:e498. https://doi.org/10.7759/cureus.498
6. Côté P, Kreitz BG, Cassidy JD, Thiel H. The validity of the extension-rotation test as a clinical screening procedure before neck manipulation: a secondary analysis. *J Manipulative Physiol Ther*. 1996;19:159-164.
7. Creighton DS, Marsh D, Gruca M, Walter M. The application of a pre-positioned upper cervical traction mobilization to patients with painful active cervical rotation impairment: a case series. *J Back Musculoskelet Rehabil*. 2017;30:1053-1059. https://doi.org/10.3233/BMR-169644
8. Debette S. Pathophysiology and risk factors of cervical artery dissection: what have we learnt from large hospital-based cohorts? *Curr Opin Neurol*. 2014;27:20-28. https://doi.org/10.1097/WCO.0000000000000056
9. Debette S, Compter A, Labeyrie MA, et al. Epidemiology, pathophysiology, diagnosis, and management of intracranial artery dissection. *Lancet Neurol*. 2015;14:640-654. https://doi.org/10.1016/S1474-4422(15)00009-5
10. Debette S, Leys D. Cervical-artery dissections: predisposing factors, diagnosis, and outcome. *Lancet Neurol*. 2009;8:668-678. https://doi.org/10.1016/S1474-4422(09)70084-5
11. Dittrich R, Rohsbach D, Heidbreder A, et al. Mild mechanical traumas are possible risk factors for cervical artery dissection. *Cerebrovasc Dis*. 2007;23:275-281. https://doi.org/10.1159/000098327
12. Erhardt JW, Windsor BA, Kerry R, et al. The immediate effect of atlanto-axial high velocity thrust techniques on blood flow in the vertebral artery: a randomized controlled trial. *Man Ther*. 2015;20:614-622. https://doi.org/10.1016/j.math.2015.02.008
13. Grindem H, Mansournia MA, Øiestad BE, Ardern CL. Was it a good idea to combine the studies? Why clinicians should care about heterogeneity when making decisions based on systematic reviews. *Br J Sports Med*. 2019;53:399-401. https://doi.org/10.1136/bjsports-2018-099516
14. Haldeman S, Kohlbeck FJ, McGregor M. Risk factors and precipitating neck movements causing vertebrobasilar artery dissection after cervical trauma and spinal manipulation. *Spine (Phila Pa 1976)*. 1999;24:785-794. https://doi.org/10.1097/00007632-199904150-00010
15. Hartkamp NS, Petersen ET, Chappell MA, et al. Relationship between haemodynamic impairment and collateral blood flow in carotid artery disease. *J Cereb Blood Flow Metab*. 2018;38:2021-2032. https://doi.org/10.1177/0271678X17724027
16. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32:1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x
17. Haynes MJ, Cala LA, Melsom A, Mastaglia FL, Milne N, McGeachie JK. Vertebral arteries and cervical rotation: modeling and magnetic resonance angiography studies. *J Manipulative Physiol Ther*. 2002;25:370-383. https://doi.org/10.1067/mmt.2002.126130
18. Haynes MJ, Milne N. Color duplex sonographic findings in human vertebral arteries during cervical rotation. *J Clin Ultrasound*. 2001;29:14-24. https://doi.org/10.1002/1097-0096(200101)29:1<14::AID-JCU3>3.0.CO;2-H
19. Hedera P, Bujdáková J, Traubner P. Blood flow velocities in basilar artery during rotation of the head. *Acta Neurol Scand*. 1993;88:229-233. https://doi.org/10.1111/j.1600-0404.1993.tb04224.x
20. Higgins JP, Altman DG, Sterne JA. Assessing risk of bias in included studies. In: Higgins JP, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. Oxford, UK: The Cochrane Collaboration; 2011:ch 8.
21. Hutting N, Scholten-Peeters GG, Vijverman V, Keesenberg MD, Verhagen AP. Diagnostic accuracy of upper cervical spine instability tests: a systematic review. *Phys Ther*. 2013;93:1686-1695. https://doi.org/10.2522/ptj.20130186
22. Hutting N, Verhagen AP, Vijverman V, Keesenberg MD, Dixon G, Scholten-Peeters GG. Diagnostic accuracy of premanipulative vertebrobasilar insufficiency tests: a systematic review. *Man Ther*. 2013;18:177-182. https://doi.org/10.1016/j.math.2012.09.009
23. IFOMPT Standards Committee. Glossary of terms. In: *Educational Standards in Orthopaedic Manipulative Therapy*. Auckland, New Zealand: International Federation of Orthopaedic Manipulative Physical Therapists; 2016:27-36.
24. Katikireddi SV, Egan M, Petticrew M. How do systematic reviews incorporate risk of bias assessments into the synthesis of evidence? A methodological study. *J Epidemiol Community Health*. 2015;69:189-195. https://doi.org/10.1136/jech-2014-204711
25. Kerry R, Taylor AJ, Mitchell J, McCarthy C. Cervical arterial dysfunction and manual therapy: a critical literature review to inform professional practice. *Man Ther*. 2008;13:278-288. https://doi.org/10.1016/j.math.2007.10.006
26. Kranenburg HA, Schmitt MA, Puentedura EJ, Luijckx GJ, van der Schans CP. Adverse events associated with the use of cervical spine manipulation or mobilization and patient characteristics: a systematic review. *Musculoskelet Sci Pract*. 2017;28:32-38. https://doi.org/10.1016/j.msksp.2017.01.008

27. Li YK, Zhang YK, Lu CM, Zhong SZ. Changes and implications of blood flow velocity of the vertebral artery during rotation and extension of the head. *J Manipulative Physiol Ther*. 1999;22:91-95. https://doi.org/10.1016/S0161-4754(99)70113-2
28. Licht PB, Christensen HW, Høilund-Carlsen PF. Carotid artery blood flow during premanipulative testing. *J Manipulative Physiol Ther*. 2002;25:568-572. https://doi.org/10.1067/mmt.2002.128367
29. Licht PB, Christensen HW, Høilund-Carlsen PF. Is there a role for premanipulative testing before cervical manipulation? *J Manipulative Physiol Ther*. 2000;23:175-179. https://doi.org/10.1016/S0161-4754(00)90247-1
30. Licht PB, Christensen HW, Høilund-Carlsen PF. Vertebral artery volume flow in human beings. *J Manipulative Physiol Ther*. 1999;22:363-367. https://doi.org/10.1016/S0161-4754(99)70080-1
31. Mitchell J. Vertebral artery blood flow velocity changes associated with cervical spine rotation: a meta-analysis of the evidence with implications for professional practice. *J Man Manip Ther*. 2009;17:46-57. https://doi.org/10.1179/106698109790818160
32. Mitchell J, Keene D, Dyson C, Harvey L, Pruvey C, Phillips R. Is cervical spine rotation, as used in the standard vertebrobasilar insufficiency test, associated with a measureable change in intracranial vertebral artery blood flow? *Man Ther*. 2004;9:220-227. https://doi.org/10.1016/j.math.2004.03.005
33. Mitchell JA. Changes in vertebral artery blood flow following normal rotation of the cervical spine. *J Manipulative Physiol Ther*. 2003;26:347-351. https://doi.org/10.1016/S0161-4754(03)00074-5
34. Nielsen SM, Tarp S, Christensen R, Bliddal H, Klokker L, Henriksen M. The risk associated with spinal manipulation: an overview of reviews. *Syst Rev*. 2017;6:64. https://doi.org/10.1186/s13643-017-0458-y
35. Niewiadomski P, Bielińska M, Pietkiewicz P, Olszewski J. Diagnostic evaluation of neck torsion test in objective examination in patients with vertigo and/or hearing impairment. *Otolaryngol Pol*. 2017;71:18-28. https://doi.org/10.5604/01.3001.0010.5313
36. Ozdemir H, Cihangiroglu M, Berilgen S, Bulut S. Effects of cervical rotation on hemodynamics in vertebral arteries. *J Diagn Med Sonogr*. 2005;21:384-391. https://doi.org/10.1177/8756479305278982
37. Peng SL, Shih CT, Huang CW, Chiu SC, Shen WC. Optimized analysis of blood flow and wall shear stress in the common carotid artery of rat model by phase-contrast MRI. *Sci Rep*. 2017;7:5253. https://doi.org/10.1038/s41598-017-05606-4
38. Petersen B, von Maravic M, Zeller JA, Walker ML, Kömpf D, Kessler C. Basilar artery blood flow during head rotation in vertebrobasilar ischemia. *Acta Neurol Scand*. 1996;94:294-301. https://doi.org/10.1111/j.1600-0404.1996.tb07068.x
39. Quesnele JJ, Triano JJ, Noseworthy MD, Wells GD. Changes in vertebral artery blood flow following various head positions and cervical spine manipulation. *J Manipulative Physiol Ther*. 2014;37:22-31. https://doi.org/10.1016/j.jmpt.2013.07.008
40. Rivett DA, Sharples KJ, Milburn PD. Effect of premanipulative tests on vertebral artery and internal carotid artery blood flow: a pilot study. *J Manipulative Physiol Ther*. 1999;22:368-375. https://doi.org/10.1016/S0161-4754(99)70081-3
41. Rivett DA, Sharples KJ, Milburn PD. Reliability of ultrasonographic measurement of vertebral artery blood flow. *NZ J Physiother*. 2003;31:119-128.
42. Rushton A, Rivett D, Carlesso L, Flynn T, Hing W, Kerry R. International framework for examination of the cervical region for potential of Cervical Arterial Dysfunction prior to orthopaedic manual therapy intervention. *Man Ther*. 2014;19:222-228. https://doi.org/10.1016/j.math.2013.11.005
43. Sakaguchi M, Kitagawa K, Hougaku H, et al. Mechanical compression of the extracranial vertebral artery during neck rotation. *Neurology*. 2003;61:845-847. https://doi.org/10.1212/01.wnl.0000078081.12097.ae
44. Saracoglu A, Altun D, Yavru A, Aksakal N, Sormaz IC, Camci E. Effects of head position on cerebral oxygenation and blood flow velocity during thyroidectomy. *Turk J Anaesthesiol Reanim*. 2016;44:241-246. https://doi.org/10.5152/TJAR.2016.77598
45. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62:1013-1020. https://doi.org/10.1016/j.jclinepi.2008.10.009
46. Shemilt I, Khan N, Park S, Thomas J. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. *Syst Rev*. 2016;5:140. https://doi.org/10.1186/s13643-016-0315-4
47. Siwach RC, Kumar S, Kumar M. Assessment of cranial blood flow in cervical spondylosis patients. *Indian J Clin Anat Physiol*. 2016;3:460-463. https://doi.org/10.5958/2394-2126.2016.00105.5
48. Sturzenegger M, Newell DW, Douville C, Byrd S, Schoonover K. Dynamic transcranial Doppler assessment of positional vertebrobasilar ischemia. *Stroke*. 1994;25:1776-1783. https://doi.org/10.1161/01.str.25.9.1776
49. Sultan MJ, Hartshorne T, Naylor AR. Extracranial and transcranial ultrasound assessment in patients with suspected positional 'vertebrobasilar ischaemia'. *Eur J Vasc Endovasc Surg*. 2009;38:10-13. https://doi.org/10.1016/j.ejvs.2008.12.006
50. Thiel H, Wallace K, Donat J, Yong-Hing K. Effect of various head and neck positions on vertebral artery blood flow. *Clin Biomech (Bristol, Avon)*. 1994;9:105-110. https://doi.org/10.1016/0268-0033(94)90032-9
51. Thomas LC. Cervical arterial dissection: an overview and implications for manipulative therapy practice. *Man Ther*. 2016;21:2-9. https://doi.org/10.1016/j.math.2015.07.008
52. Thomas LC, McLeod LR, Osmotherly PG, Rivett DA. The effect of end-range cervical rotation on vertebral and internal carotid arterial blood flow and cerebral inflow: a sub analysis of an MRI study. *Man Ther*. 2015;20:475-480. https://doi.org/10.1016/j.math.2014.11.012
53. Thomas LC, Rivett DA, Bateman G, Stanwell P, Levi CR. Effect of selected manual therapy interventions for mechanical neck pain on vertebral and internal carotid arterial blood flow and cerebral inflow. *Phys Ther*. 2013;93:1563-1574. https://doi.org/10.2522/ptj.20120477
54. Wand BM, Heine PJ, O'Connell NE. Should we abandon cervical spine manipulation for mechanical neck pain? Yes. *BMJ*. 2012;344:e3679. https://doi.org/10.1136/bmj.e3679
55. Weintraub MI, Khoury A. Critical neck position as an independent risk factor for posterior circulation stroke. A magnetic resonance angiographic analysis. *J Neuroimaging*. 1995;5:16-22. https://doi.org/10.1111/jon19955116
56. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529-536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009
57. Zaina C, Grant R, Johnson C, Dansie B, Taylor J, Spyropolous P. The effect of cervical rotation on blood flow in the contralateral vertebral artery. *Man Ther*. 2003;8:103-109. https://doi.org/10.1016/S1356-689X(02)00155-8

APPENDIX A

SEARCH STRATEGIES

PubMed

("Neck"[Mesh] OR "Rotation"[Mesh] OR "Musculoskeletal Manipulations"[Mesh] OR "Cervical spine"[Title/Abstract] OR Neck[Title/Abstract] OR Head[Title/Abstract] OR Mobilization[Title/Abstract] OR Mobilisation[Title/Abstract] OR Extension*[Title/Abstract] OR Flexion*[Title/Abstract] OR Rotation*[Title/Abstract] OR Distraction*[Title/Abstract] OR Manipulation*[Title/Abstract] OR Midrange*[Title/Abstract] OR Mid-range*[Title/Abstract] OR Premanipulat*[Title/Abstract] OR Pre-manipulat*[Title/Abstract]) AND ("Vertebral Artery"[Mesh] OR "Carotid Arteries"[Mesh] OR "Carotid Artery, internal"[Mesh] OR "Vertebrobasilar Insufficiency"[Mesh] OR Vertebral Arter*[Title/Abstract] OR Carotid Arter*[Title/Abstract] OR vertebrobasilar insufficienc*[Title/Abstract]) AND ("Hemodynamics"[Mesh] OR "Blood Circulation"[Mesh] OR "Blood Flow Velocity"[Mesh] OR "Regional Blood Flow"[Mesh] OR "blood supply" [Subheading] OR hemodynamic*[Title/Abstract] OR Blood flow*[Title/Abstract] OR Bloodflow[Title/Abstract] OR Blood circulat*[Title/Abstract] OR Blood suppl*[Title/Abstract] OR "Flow in"[Title/Abstract] OR Inflow*[Title/Abstract] OR Flow velocit*[Title/Abstract] OR Arterial pressur*[Title/Abstract] OR Test*[Title/Abstract]) AND ("Ultrasonography"[Mesh] OR "Ultrasonography, Doppler"[Mesh] OR "Ultrasonography, Doppler, Color"[Mesh] OR "Magnetic Resonance Angiography"[Mesh] OR "Magnetic Resonance Imaging"[Mesh] OR "Coronary Angiography"[Mesh] OR "Ultrasonography"[Title/Abstract] OR "Ultrasound"[Title/Abstract] OR "Ultrasound imaging"[Title/Abstract] OR "Medical sonography"[Title/Abstract] OR "Ultrasonic imaging"[Title/Abstract] OR "Echography"[Title/Abstract] OR "Doppler"[Title/Abstract] OR "MRI"[Title/Abstract] OR "Magnetic Resonance Imaging"[Title/Abstract] OR "MRA"[Title/Abstract] OR "Magnetic Resonance Angiography"[Title/Abstract])

CINAHL

((MH ("Neck" OR "Rotation") OR TI ("Cervical spine" OR Neck OR Head OR Mobilization OR Mobilisation OR Extension* OR Flexion* OR Rotation* OR Distraction* OR Manipulation* OR Midrange* OR Mid-range* OR Premanipulat* OR Pre-manipulat*) OR AB ("Cervical spine" OR Neck OR Head OR Mobilization OR Mobilisation OR Extension* OR Flexion* OR Rotation* OR Distraction* OR Manipulation* OR Midrange* OR Mid-range* OR Premanipulat* OR Pre-manipulat*)) AND ((MH ("Vertebral Artery" OR "Carotid Arteries") OR TI (Vertebral Arter* OR Carotid Arter* OR vertebrobasilar insufficienc*) OR AB (Vertebral Arter* OR Carotid Arter* OR vertebrobasilar insufficienc*)) AND ((MH ("Hemodynamics" OR "Blood Circulation" OR "Blood Flow Velocity") OR TI (hemodynamic* OR Blood flow* OR Bloodflow OR Blood circulat* OR Blood suppl* OR "Flow in" OR Inflow* OR Flow velocit* OR Arterial pressur* OR Test*) OR AB (hemodynamic* OR Blood flow* OR Bloodflow OR Blood circulat* OR Blood suppl* OR "Flow in" OR Inflow* OR Flow velocit* OR Arterial pressur* OR Test*)) AND ((MH ("Ultrasonography" OR "Ultrasonography, Doppler, Color" OR "Ultrasonography, Doppler" OR "Magnetic Resonance Angiography" OR "Magnetic Resonance Imaging" OR "Coronary Angiography") OR TI ("Ultrasonography" OR "Ultrasound" OR "Ultrasound imaging" OR "Medical sonography" OR "Ultrasonic imaging" OR "Echography" OR "Doppler" OR "MRI" OR "Magnetic Resonance Imaging" OR "MRA" OR "Magnetic Resonance Angiography") OR AB ("Ultrasonography" OR "Ultrasound" OR "Ultrasound imaging" OR "Medical sonography" OR "Ultrasonic imaging" OR "Echography" OR "Doppler" OR "MRI" OR "Magnetic Resonance Imaging" OR "MRA" OR "Magnetic Resonance Angiography"))

Embase

('neck'/mj OR 'rotation'/exp OR 'Cervical spine':ab,ti OR Neck:ab,ti OR Head:ab,ti OR Mobilization:ab,ti OR Mobilisation:ab,ti OR Extension:ab,ti OR Extensions:ab,ti OR Flexion:ab,ti OR Flexions:ab,ti OR Rotation:ab,ti OR Rotations:ab,ti OR Distraction:ab,ti OR Distractions:ab,ti OR Manipulation:ab,ti OR Manipulations:ab,ti OR Midrange:ab,ti OR Mid-range:ab,ti OR Premanipulation:ab,ti OR Pre-manipulation:ab,ti) AND ('vertebral artery'/exp OR 'carotid artery'/exp OR 'basilar artery'/exp OR 'Vertebral Artery':ab,ti OR 'Vertebral Arteries':ab,ti OR 'Carotid Artery':ab,ti OR 'Carotid Arteries':ab,ti OR 'Basilar Artery':ab,ti OR 'Basilar Arteries':ab,ti OR 'vertebrobasilar insufficiency':ab,ti OR 'vertebrobasilar insufficiencies':ab,ti) AND ('hemodynamics'/mj OR 'arterial circulation'/exp OR 'blood flow velocity'/exp OR Hemodynamic:ab,ti OR Hemodynamics:ab,ti OR 'arterial circulation':ab,ti OR 'Blood flow':ab,ti OR 'Blood flowing':ab,ti OR Bloodflow:ab,ti OR Bloodflowing:ab,ti OR 'Blood circulation':ab,ti OR 'Blood supply':ab,ti OR 'Blood supplies':ab,ti OR 'Blood supplying':ab,ti OR 'Flow in':ab,ti OR Inflow:ab,ti OR 'Flow velocity':ab,ti OR 'Flow velocities':ab,ti OR 'Arterial pressure':ab,ti OR 'Arterial pressures':ab,ti) AND ('echography'/exp OR 'Doppler echocardiography'/exp OR 'Doppler ultrasonography'/exp OR 'magnetic resonance angiography'/exp OR 'nuclear magnetic resonance imaging'/exp OR 'coronary angiography'/exp OR 'Ultrasonography':ab,ti OR 'Ultrasound':ab,ti OR 'Ultrasound imaging':ab,ti OR 'Medical sonography':ab,ti OR 'Ultrasonic imaging':ab,ti OR 'Echography':ab,ti OR 'Doppler':ab,ti OR 'MRI':ab,ti OR 'Magnetic Resonance Imaging':ab,ti OR 'MRA':ab,ti OR 'Magnetic Resonance Angiography':ab,ti OR 'coronary angiography':ab,ti OR 'coronary angiographies':ab,ti)

Index to Chiropractic Literature

All Fields:Neck OR All Fields:\\\"Musculoskeletal Manipulations\\\" OR All Fields:Manipulations OR All Fields:Cervical spine OR All Fields:Head OR All Fields:Mobilization OR All Fields:Mobilisation OR All Fields:Extension OR All Fields:Flexion OR All Fields:Rotation OR All Fields:Distraction OR All Fields:Manipulation OR All Fields:Midrange OR All Fields:Mid-range OR All Fields:Premanipulat OR All Fields:Pre-manipulat AND All Fields:Vertebral Artery OR All Fields:Carotid Arteries OR All Fields:Carotid Artery, internal OR All Fields:Vertebrobasilar Insufficiency OR All Fields:Vertebral Arter* OR All Fields:Carotid Arter* OR All Fields:vertebrobasilar insufficienc* OR All Fields:VBI AND All Fields:Hemodynamics OR All Fields:Blood Circulation OR All Fields:Blood Flow Velocity OR All Fields:Regional Blood Flow OR All Fields:blood supply OR All Fields:hemodynamic* OR All Fields:Blood flow* OR All Fields:Bloodflow OR All Fields:Blood circulat* OR All Fields:Blood suppl* OR All Fields:Flow in OR All Fields:Inflow* OR All Fields:Flow velocit* OR All Fields:Arterial pressur* OR All Fields:Test* AND All Fields:Ultrasonography OR All Fields:Ultrasonography, Doppler OR All Fields:Ultrasonography, Doppler, Color OR All Fields:Coronary Angiography OR All Fields:Ultrasound OR All Fields:Ultrasound imaging OR All Fields:Medical sonography OR All Fields:Ultrasonic imaging OR All Fields:Echography OR All Fields:Doppler OR All Fields:MRI OR All Fields:Magnetic Resonance Imaging OR All Fields:MRA OR All Fields:Magnetic Resonance Angiography

APPENDIX B

How to Read These Tables

In the first main row of the table below, the vertebral hemodynamic effects of a maximum cervical rotation to the left are summarized and specified. The second column indicates that some studies (references 2, 9, 13, 20, 27) state that there is no significant change in peak systolic velocity for the left vertebral artery. However, 1 study (reference 18) states that there is a significant decrease in peak systolic velocity for this movement and artery. As you continue, you can see that references 2, 8, 9, 13, 20, 22, 27, and 32 state that there is no significant change, and that references 18 and 21 state that there is a significant decrease, in the right vertebral artery during the same movement.

Cervical Positional Influences on Vertebral Arterial Velocity or Volume

	PSV	EDV	Mean FV*	Mean Peak FV	AD	PSV/EDV Ratio	RI	BFV	BFV Rate	BF Velocity Ratio
Rotation										
Maximum to left										
LVA	NSC[2,9,13,20,27] SD[18]	NSC[2,9,13,27]	NSC[13,20] SD[16,31]	NSC[9] SD[15]	SD[23]		NSC[2,9]	NSC[14,29,30] SD[18,31]	NSC[32]	
RVA	NSC[2,8,9,13,20,22,27,32] SD[18,21]	NSC[8,9,13,22,27] SD[2,21]	NSC[8,13,20,22] SD[16,31]	NSC[9,15]	NSC[22] SD[23]	NSC[22,28]	NSC[9,22] SD[2,21]	NSC[14,18,22,30] SD[29,31]	NSC[32]	
Maximum to right										
LVA	NSC[2,8,9,13,20,22,27,32] SD[18,21]	NSC[8,9,13,22,27] SD[2,21]	NSC[8,13,20,22] SD[16]	NSC[9,15]	NSC[22] SD[23]	NSC[22,28]	NSC[9,21,22] SD[2]	NSC[14,22,29,30] SD[18]		
RVA	NSC[2,9,13,20,27] SD[18]	NSC[2,9,13,27]	NSC[13,20] SD[16]	NSC[9,15]	SD[23]		NSC[2,9]	NSC[14,18,29,30]		
Maximum to left plus distraction										
LVA								NSC[30]		
RVA								NSC[30]		
Maximum to right plus distraction										
LVA								NSC[30]		
RVA								NSC[30]		
C1-C2 maximum to left										
LVA								NSC[30]		
RVA								NSC[30]		
C1-C2 maximum to right										
LVA								NSC[30]		
RVA								NSC[30]		
5°-15° to left										
RVA						NSC[28]				
5°-15° to right										
RVA						NSC[28]				
30° to left										
LVA	SD[18]							SD[18]		
RVA	SD[18]							NSC[18]		
30° to right										
LVA	SD[18]							SD[18]		
RVA	SD[18]							NSC[18]		
45° to left										
LVA	NSC[9,13,20]	NSC[9,13]	NSC[13,20]	NSC[9]			NSC[9]	NSC[14]		
RVA	NSC[9,13,20,21,32]	NSC[9,13,21]	NSC[13,20]	NSC[9]			NSC[9,21]	NSC[14]	NSC[32]	

Table continues on page A3.

APPENDIX B

	PSV	EDV	Mean FV*	Mean Peak FV	AD	PSV/EDV Ratio	RI	BFV	BFV Rate	BF Velocity Ratio
45° to right										
LVA	NSC[9,13,20,21,32]	NSC[9,13,21]	NSC[13,20]	NSC[9]			NSC[9,21]	NSC[14]	NSC[32]	
RVA	NSC[9,13,20]	NSC[9,13]	NSC[13,20]	NSC[9]			NSC[9]	NSC[14]		
60° to left										
LVA	NSC[17]	NSC[17]								
RVA	NSC[17]	NSC[17]								
60° to right										
LVA	NSC[17]	NSC[17]								
RVA	NSC[17]	NSC[17]								
Extension										
Maximum										
LVA	NSC[2,11,22]	NSC[2,11,22]	NSC[11,22] SD[31]		NSC[22]	NSC[28]	NSC[2,11,22]	NSC[22] SD[31]	NSC[22]	
RVA	NSC[2,11,22]	NSC[2,11,22]	NSC[11,22] SD[31]		NSC[22]	NSC[28]	NSC[2,11,22]	NSC[22] SD[31]	NSC[22]	
Maximum plus rotation at 45° to left										
LVA					NSC[1]			SD[1] NSC[1]		
RVA					NSC[1]			NSC[1]		
Maximum plus rotation at 45° to right										
LVA					NSC[1]			NSC[1]		
RVA					NSC[1]			NSC[1] SD[1]		
Maximum extension plus maximum rotation†										
Left										
LVA	NSC[2,11,13,27]	NSC[2,11,21,27]	NSC[11,13]				NSC[2,11] SD[2]			NSC[4]
RVA	NSC[2,11,13,27] SD[21]	NSC[2,11,21,27] SD[21]	NSC[11,13]			NSC[21,28] SD[21]	NSC[2,11]			SD[4]
Right										
LVA	NSC[2,13,27] SD[21]	NSC[2,21,27] SD[21]	NSC[13]			SD[28] NSC[21]	NSC[2]			NSC[4]
RVA	NSC[2,13,27]	NSC[2,27]	NSC[13]				NSC[2] SD[2]			NSC[4]
Maximum extension plus maximum rotation plus distraction										
Left										
LVA	NSC[13]	NSC[13]					NSC[13]			
RVA	NSC[13]	NSC[13]					NSC[13]			
Right										
LVA	NSC[13]	NSC[13]					NSC[13]			
RVA	NSC[13]	NSC[13]					NSC[13]			

Table continues on page A4.

APPENDIX B

	PSV	EDV	Mean FV*	Mean Peak FV	AD	PSV/EDV Ratio	RI	BFV	BFV Rate	BF Velocity Ratio
Premanipulative position										
Left										
LVA	NSC[2,3]	NSC[3] SD[2]	NSC[3]				NSC[2] SD[3]			
RVA	NSC[3] SD[2]	NSC[3] SD[2]	NSC[3]				SD[2,3]			
Right										
LVA	NSC[3] SD[2]	NSC[3] SD[2]	NSC[3]				NSC[3] SD[2]			
RVA	NSC[2,3,7]	NSC[3,7] SD[2]	NSC[3,7]	NSC[3,7]			NSC[2,7] SD[3]			
C1-C2 manipulation										
Left										
LVA	NSC[20]		NSC[20]							
RVA	NSC[20]		NSC[20]							
Right										
LVA	NSC[20]		NSC[20]							
RVA	NSC[720]	NSC[7]	NSC[720]	NSC[7]			NSC[7]			
Maximum flexion plus maximum rotation										
Left										
LVA	NSC[27]	NSC[27]								
RVA	NSC[27]	NSC[27]								
Right										
LVA	NSC[27]	NSC[27]								
RVA	NSC[27]	NSC[27]								
Distraction										
LVA	NSC[5]	NSC[5]						NSC[30]		
RVA	NSC[5]	NSC[5]						NSC[30]		
Posttest (neutral)										
LVA	SD[32]							NSC[30]	NSC[32]	
RVA	NSC[32]							NSC[30]	NSC[32]	

Abbreviations: AD, arterial diameter; BF, blood flow; BFV, blood flow volume; EDV, end diastolic velocity; FV, flow velocity; LVA, left vertebral artery; NSC, no significant change; PSV, peak systolic velocity; RI, resistance index; RVA, right vertebral artery; SD, significant decrease.

**Time averaged.*

†Including the maneuvers as described by Wallenberg or de Kleyn.[6]

Cervical Positional Influences on Carotid Arterial Velocity or Volume

	PSV	EDV	Mean FV*	AD	RI	BFV	CCV
Rotation							
Maximum to left							
LVA			NSC[31]			NSC[29-31]	
RVA	NSC[21] SD[21]	NSC[21] SD[21]	NSC[31]		NSC[21]	NSC[30,31] SD[29,30]	
Maximum to right							
LVA	NSC[21]	NSC[21]	NSC[31]		NSC[21]	NSC[29,31]	
RVA			NSC[31]			NSC[29-31]	

Table continues on page A5.

APPENDIX B

	PSV	EDV	Mean FV*	AD	RI	BFV	CCV
Maximum to left plus distraction							
LVA						NSC[30]	
RVA						NSC[30]	
Maximum to right plus distraction							
LVA						NSC[30]	
RVA						NSC[30]	
C1-C2 maximum to left							
LVA						NSC[30]	
RVA						NSC[30]	
C1-C2 maximum to right							
LVA						NSC[30]	
RVA						NSC[30]	
45° to left							
RVA	NSC[21]	NSC[21]			NSC[21]		
45° to right							
LVA	NSC[21]	NSC[21]			NSC[21]		
Extension							
Maximum							
LVA			NSC[31]			NSC[31]	
RVA			NSC[31]			NSC[31]	
Maximum extension plus maximum rotation†							
Left							
LVA	NSC[12]		NSC[12]				
RVA	NSC[12] SD[21]	SD[21]	NSC[12]		NSC[21]		
Right							
LVA	NSC[12,21]	NSC[21]	NSC[12]		NSC[21]		
RVA	NSC[12]		NSC[12]				
Distraction							
LVA						NSC[30]	
RVA						NSC[30]	
Semi-Fowler							
Extension plus 10° of collateral rotation post induction							
LVA	SI[24]		SI[24]	NSC[24]			SD[24]
Extension plus 10° of collateral rotation post surgery							
LVA	SD[24]		SD[24]	NSC[24]			SD[24]
Posttest (neutral)							
LVA						NSC[30]	
RVA						NSC[30]	

Abbreviations: AD, arterial diameter; BFV, blood flow volume; CCV, common carotid volume; EDV, end diastolic velocity; FV, flow velocity; LVA, left vertebral artery; NSC, no significant change; PSV, peak systolic velocity; RI, resistance index; RVA, right vertebral artery; SD, significant decrease; SI, significant increase.

**Time averaged.*

†Including the maneuvers as described by Wallenberg or de Kleyn.[6]

APPENDIX B

Cervical Positional Influences on Intracranial Arterial Velocity or Volume

	PSV	EDV	Mean FV*	RI	Mean Peak FV	BFV
Rotation						
Maximum to left						
LVA	NSC[27]	NSC[27]	NSC[19,26,31] SD[10,19]	NSC[10]		NSC[30,31]
RVA	NSC[27]	NSC[27]	NSC[19,26,31] SD[10,19]	SD[10]		NSC[30,31]
Maximum to right						
LVA	NSC[27]	NSC[27]	NSC[19,26,31] SD[10,19]	SD[10]		NSC[30,31]
RVA	NSC[27]	NSC[27]	NSC[19,26,31] SD[10,19]	NSC[10]		NSC[30,31]
Maximum to left plus distraction						
LVA						NSC[30]
RVA						NSC[30]
Maximum to right plus distraction						
LVA						NSC[30]
RVA						NSC[30]
C1-C2 maximum to left						
LVA						NSC[30]
RVA						NSC[30]
C1-C2 maximum to right						
LVA						NSC[30]
RVA						NSC[30]
Extension						
Maximum						
LVA			NSC[31] SD[26]		NSC[25]	NSC[31]
RVA			NSC[31] SD[26]		NSC[25]	NSC[31]
Maximum extension plus maximum rotation[†]						
Left						
LVA	NSC[27]	NSC[27]		SD[10]		
RVA	NSC[27]	NSC[27]		NSC[10]		
Right						
LVA				NSC[10]		
RVA				SD[10]		
Flexion						
Maximum						
LVA					NSC[25,26]	
RVA					NSC[25,26]	
Maximum plus maximum rotation to left						
LVA	NSC[27]	NSC[27]				
RVA	NSC[27]	NSC[27]				
Maximum plus maximum rotation to right						
LVA	NSC[27]	NSC[27]				
RVA	NSC[27]	NSC[27]				

Table continues on page A7.

APPENDIX B

	PSV	EDV	Mean FV*	RI	Mean Peak FV	BFV
Distraction						
LVA						NSC[30]
RVA						NSC[30]
Posttest (neutral)						
LVA			NSC[19]			
RVA			NSC[19]			

Abbreviations: AD, arterial diameter; BFV, blood flow volume; EDV, end diastolic velocity; FV, flow velocity; LVA, left vertebral artery; NSC, no significant change; PSV, peak systolic velocity; RI, resistance index; RVA, right vertebral artery; SD, significant decrease.
**Time averaged.*
†Including the maneuvers as described by Wallenberg or de Kleyn.[6]

References

1. Araz Server E, Tuna Edizer D, Yigit Ö, Görkem Yasak A, Erdim Ç. Relationship between vertebral artery blood flow in different head positions and vertigo. *Acta Otolaryngol*. 2018;138:1-5. https://doi.org/10.1080/00016489.2017.1373849
2. Arnold C, Bourassa R, Langer T, Stoneham G. Doppler studies evaluating the effect of a physical therapy screening protocol on vertebral artery blood flow. *Man Ther*. 2004;9:13-21. https://doi.org/10.1016/S1356-689X(03)00087-0
3. Bowler N, Shamley D, Davies R. The effect of a simulated manipulation position on internal carotid and vertebral artery blood flow in healthy individuals. *Man Ther*. 2011;16:87-93. https://doi.org/10.1016/j.math.2010.07.007
4. Côté P, Kreitz BG, Cassidy JD, Thiel H. The validity of the extension-rotation test as a clinical screening procedure before neck manipulation: a secondary analysis. *J Manipulative Physiol Ther*. 1996;19:159-164.
5. Creighton DS, Marsh D, Gruca M, Walter M. The application of a pre-positioned upper cervical traction mobilization to patients with painful active cervical rotation impairment: a case series. *J Back Musculoskelet Rehabil*. 2017;30:1053-1059. https://doi.org/10.3233/BMR-169644
6. de Kleyn A, Nieuwenhuyse P. Schwindelanfalle und Nystagmus bei einer bestimmten Stellung des Kopfes. *Acta Oto-Laryngol*. 1927;11:155-157. https://doi.org/10.3109/00016482709120074
7. Erhardt JW, Windsor BA, Kerry R, et al. The immediate effect of atlanto-axial high velocity thrust techniques on blood flow in the vertebral artery: a randomized controlled trial. *Man Ther*. 2015;20:614-622. https://doi.org/10.1016/j.math.2015.02.008
8. Haynes MJ, Cala LA, Melsom A, Mastaglia FL, Milne N, McGeachie JK. Vertebral arteries and cervical rotation: modeling and magnetic resonance angiography studies. *J Manipulative Physiol Ther*. 2002;25:370-383. https://doi.org/10.1067/mmt.2002.126130
9. Haynes MJ, Milne N. Color duplex sonographic findings in human vertebral arteries during cervical rotation. *J Clin Ultrasound*. 2001;29:14-24. https://doi.org/10.1002/1097-0096(200101)29:1<14::AID-JCU3>3.0.CO;2-H
10. Hedera P, Bujdáková J, Traubner P. Blood flow velocities in basilar artery during rotation of the head. *Acta Neurol Scand*. 1993;88:229-233. https://doi.org/10.1111/j.1600-0404.1993.tb04224.x
11. Li YK, Zhang YK, Lu CM, Zhong SZ. Changes and implications of blood flow velocity of the vertebral artery during rotation and extension of the head. *J Manipulative Physiol Ther*. 1999;22:91-95. https://doi.org/10.1016/S0161-4754(99)70113-2
12. Licht PB, Christensen HW, Høilund-Carlsen PF. Carotid artery blood flow during premanipulative testing. *J Manipulative Physiol Ther*. 2002;25:568-572. https://doi.org/10.1067/mmt.2002.128367
13. Licht PB, Christensen HW, Høilund-Carlsen PF. Is there a role for premanipulative testing before cervical manipulation? *J Manipulative Physiol Ther*. 2000;23:175-179. https://doi.org/10.1016/S0161-4754(00)90247-1
14. Licht PB, Christensen HW, Høilund-Carlsen PF. Vertebral artery volume flow in human beings. *J Manipulative Physiol Ther*. 1999;22:363-367. https://doi.org/10.1016/S0161-4754(99)70080-1
15. Mitchell J, Keene D, Dyson C, Harvey L, Pruvey C, Phillips R. Is cervical spine rotation, as used in the standard vertebrobasilar insufficiency test, associated with a measureable change in intracranial vertebral artery blood flow? *Man Ther*. 2004;9:220-227. https://doi.org/10.1016/j.math.2004.03.005
16. Mitchell JA. Changes in vertebral artery blood flow following normal rotation of the cervical spine. *J Manipulative Physiol Ther*. 2003;26:347-351. https://doi.org/10.1016/S0161-4754(03)00074-5
17. Niewiadomski P, Bielińska M, Pietkiewicz P, Olszewski J. Diagnostic evaluation of neck torsion test in objective examination in patients with vertigo and/or hearing impairment. *Otolaryngol Pol*. 2017;71:18-28. https://doi.org/10.5604/01.3001.0010.5313
18. Ozdemir H, Cihangiroglu M, Berilgen S, Bulut S. Effects of cervical rotation on hemodynamics in vertebral arteries. *J Diagn Med Sonogr*. 2005;21:384-391. https://doi.org/10.1177/8756479305278982
19. Petersen B, von Maravic M, Zeller JA, Walker ML, Kömpf D, Kessler C. Basilar artery blood flow during head rotation in vertebrobasilar ischemia. *Acta Neurol Scand*. 1996;94:294-301. https://doi.org/10.1111/j.1600-0404.1996.tb07068.x
20. Quesnele JJ, Triano JJ, Noseworthy MD, Wells GD. Changes in vertebral artery blood flow following various head positions and cervical spine manipulation. *J Manipulative Physiol Ther*. 2014;37:22-31. https://doi.org/10.1016/j.jmpt.2013.07.008
21. Rivett DA, Sharples KJ, Milburn PD. Effect of premanipulative tests on vertebral artery and internal carotid artery blood flow: a pilot study. *J Manipulative Physiol Ther*. 1999;22:368-375. https://doi.org/10.1016/S0161-4754(99)70081-3
22. Rivett DA, Sharples KJ, Milburn PD. Reliability of ultrasonographic measurement of vertebral artery blood flow. *NZ J Physiother*. 2003;31:119-128.
23. Sakaguchi M, Kitagawa K, Hougaku H, et al. Mechanical compression of the extracranial vertebral artery during neck rotation. *Neurology*. 2003;61:845-847. https://doi.org/10.1212/01.wnl.0000078081.12097.ae

APPENDIX B

24. Saracoglu A, Altun D, Yavru A, Aksakal N, Sormaz IC, Camci E. Effects of head position on cerebral oxygenation and blood flow velocity during thyroidectomy. *Turk J Anaesthesiol Reanim*. 2016;44:241-246. https://doi.org/10.5152/TJAR.2016.77598

25. Siwach RC, Kumar S, Kumar M. Assessment of cranial blood flow in cervical spondylosis patients. *Indian J Clin Anat Physiol*. 2016;3:460-463. https://doi.org/10.5958/2394-2126.2016.00105.5

26. Sturzenegger M, Newell DW, Douville C, Byrd S, Schoonover K. Dynamic transcranial Doppler assessment of positional vertebrobasilar ischemia. *Stroke*. 1994;25:1776-1783. https://doi.org/10.1161/01.str.25.9.1776

27. Sultan MJ, Hartshorne T, Naylor AR. Extracranial and transcranial ultrasound assessment in patients with suspected positional 'vertebrobasilar ischaemia'. *Eur J Vasc Endovasc Surg*. 2009;38:10-13. https://doi.org/10.1016/j.ejvs.2008.12.006

28. Thiel H, Wallace K, Donat J, Yong-Hing K. Effect of various head and neck positions on vertebral artery blood flow. *Clin Biomech (Bristol, Avon)*. 1994;9:105-110. https://doi.org/10.1016/0268-0033(94)90032-9

29. Thomas LC, McLeod LR, Osmotherly PG, Rivett DA. The effect of end-range cervical rotation on vertebral and internal carotid arterial blood flow and cerebral inflow: a sub analysis of an MRI study. *Man Ther*. 2015;20:475-480. https://doi.org/10.1016/j.math.2014.11.012

30. Thomas LC, Rivett DA, Bateman G, Stanwell P, Levi CR. Effect of selected manual therapy interventions for mechanical neck pain on vertebral and internal carotid arterial blood flow and cerebral inflow. *Phys Ther*. 2013;93:1563-1574. https://doi.org/10.2522/ptj.20120477

31. Weintraub MI, Khoury A. Critical neck position as an independent risk factor for posterior circulation stroke. A magnetic resonance angiographic analysis. *J Neuroimaging*. 1995;5:16-22. https://doi.org/10.1111/jon19955116

32. Zaina C, Grant R, Johnson C, Dansie B, Taylor J, Spyropolous P. The effect of cervical rotation on blood flow in the contralateral vertebral artery. *Man Ther*. 2003;8:103-109. https://doi.org/10.1016/S1356-689X(02)00155-8

FLÁVIA CORDEIRO MEDEIROS, PT, PhD[1] • **LEONARDO OLIVEIRA PENA COSTA,** PT, PhD[1]
INDIARA SOARES OLIVEIRA, PT, PhD[1] • **LUCÍOLA DA CUNHA MENEZES COSTA,** PT, PhD[1]

A Responsiveness Analysis of the Subgroups for Targeted Treatment (STarT) Back Screening Tool in Patients With Nonspecific Low Back Pain

BACKGROUND: The Subgroups for Targeted Treatment (STarT) Back Screening Tool (SBST) screens patients with low back pain and directs them to different levels of physical therapy treatment. The SBST is also used to monitor changes in a range of modifiable prognostic factors. However, the current evidence on the responsiveness of the SBST is limited.

OBJECTIVES: To test the responsiveness of the SBST at 6 weeks and 6 months.

METHODS: This measurement property study is a secondary analysis of data from 2 previous studies that included 348 participants with nonspecific low back pain. All participants were assessed at baseline, 6 weeks, and 6 months. To detect clinical changes, the SBST was compared to 3 one-dimensional constructs: global perceived effect, disability, and pain intensity. To assess responsiveness, we tested 15 specific predefined hypotheses based on correlation, effect size, and receiver operating characteristic curve analyses. If 75% or more of the hypotheses were accepted, then responsiveness was considered to be high.

RESULTS: Most of the hypotheses were accepted. Testing the SBST as a continuous score, 85.7% and 87.5% of the hypotheses were accepted at 6 weeks and 6 months, respectively. For medium- and high-risk subgroups, 85.7% and 87.5% of the hypotheses were accepted at 6 weeks and 6 months. The low-risk subgroup had 42.9% of the hypotheses accepted at 6 weeks and 100% of the hypotheses accepted at 6 months.

CONCLUSION: The SBST had high responsiveness at 6 weeks in subgroups of patients with a medium and high risk, and poor responsiveness in those with a low risk, of persistent disability. The SBST has high responsiveness in all SBST subgroups at 6 months. Clinicians can confidently use the SBST to measure changes over time in terms of subgroups.

LEVEL OF EVIDENCE: Longitudinal clinical measurement, level 1b. *J Orthop Sports Phys Ther 2019;49(10):725-735. Epub 23 Aug 2019. doi:10.2519/jospt.2019.8776*

KEY WORDS: *clinical change, measurement properties, responsiveness, STarT Back Tool*

The Subgroups for Targeted Treatment (STarT) Back Screening Tool (SBST) is a simple model used to provide stratified care in patients with low back pain.[18,34] Using the SBST involves 2 stages[18,34]: stratifying patients into risk subgroups of persistent disability, and matching effective treatments to each of these subgroups.[18,19] The matched treatment for patients in the low-risk subgroup consists of 1 personalized consultation, which involves advice and education.[18,19] The matched treatment for patients in the medium-risk subgroup includes advice, education, exercises (general and/or specific), and manual therapy, if necessary.[18,19] Patients in the high-risk subgroup are matched to the same treatment regimen prescribed to those in the medium-risk subgroup, along with additional psychological components.[18,19]

The SBST consists of 9 items subdivided into physical prognostic factors (eg, referred pain, comorbidities, and disability) and psychosocial prognostic factors (eg, bothersomeness, catastrophizing, fear, anxiety, and depression).[18] All of these prognostic factors are modifiable and associated with poor prognosis.[18] Two prospective cohort studies found that, when the SBST was applied in patients with nonspecific low back pain at more than 1 time point, more than half of the patients had a change in subgroup during the assessments.[21,22] These studies indicate that the SBST could be used in patients with acute or chronic low back pain to monitor physical therapy treatment and to highlight changes in prognostic factors.[21,22]

In both clinical practice and research, it is important to have tools capable of

[1]Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de São Paulo, Tatuapé, Brazil. This study was approved by the Research Ethics Committee of the Universidade Cidade de São Paulo (CAAE 14386513.4.0000.0064/25315713.7.0000.0064). Dr Medeiros's PhD scholarship was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (number 1501795) and by the São Paulo Research Foundation (grant 2016/18575-8). Dr Oliveira's PhD scholarship was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. This study did not receive commercial sponsorship but was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (finance code 001). The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Flávia Cordeiro Medeiros, Universidade Cidade de São Paulo, Rua Cesário Galero 448, Tatuapé 03071-100 SP, Brazil. E-mail: flaviamedeiros.fisio@gmail.com

identifying whether the patient's clinical outcomes have changed over time and of measuring this change. Therefore, instruments with good responsiveness are needed.[13] According to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN), responsiveness is the ability of the instrument to detect clinical change over time.[13,24] The SBST has been translated and adapted to more than 25 cultures and languages,[3,35] but no study has tested all of its measurement properties.[35] Although the SBST was developed to screen and categorize patients for targeted treatments, we hypothesize that the instrument may be able to measure clinical changes over time. Only a study conducted in Iran, using the Persian version of the SBST, has measured the responsiveness of the SBST in patients with lumbar central canal stenosis.[4] Because the SBST has been used in patients with low back pain of any duration, we tested the SBST's responsiveness in patients with nonspecific low back pain, regardless of the duration of symptoms (both acute and chronic).[5,18,19,27,38]

It is necessary to test the responsiveness of the SBST with a current methodology based on a priori hypothesis testing, following the recommendations of the COSMIN.[26] Furthermore, responsiveness of the SBST still needs to be tested in a developing country that presents different cultural and social aspects.[15,33] Therefore, we aimed to test the responsiveness of the SBST at 6 weeks and 6 months in a sample of patients with low back pain residing in Brazil. Our second objective was to identify the specific responsiveness of the SBST for each subgroup.

METHODS

Design

This study of measurement properties is a secondary analysis that used pooled data from 2 previous data sets.[21,22] One data set included patients with acute low back pain from emergency departments,[22] and another included patients with chronic low back pain from physical therapy departments.[21] The data collection of these studies was performed sequentially from June 2013 to September 2015, with no time interval between these studies. Both studies were approved by the Research Ethics Committee of the Universidade Cidade de São Paulo (CAAE 14386513.4.0000.0064/25315713.7.0000.0064). Both previous studies had the same objectives: (1) to test the risk stratification in acute and chronic populations, (2) to investigate the prediction of clinical outcomes, and (3) to monitor changes in patients in SBST subgroups over 6 months.[21,22]

Participants and Interventions

Participants between 18 and 80 years of age, with nonspecific low back pain and who were seeking treatment in public hospitals of a metropolitan area of a middle-income country (São Paulo, Brazil), were included. Nonspecific low back pain was defined as pain or discomfort between the costal margins and the inferior gluteal folds, with or without referred pain to the lower limbs.[37] Patients with serious spine pathologies, radicular and infectious conditions, pregnancy, or kidney diseases were excluded from the study.

In this sample, patients were not targeted for treatment directed at each subgroup of the SBST. The received treatments differed in each data set, depending on whether the patients were seen in an emergency or physical therapy department.[21,22] One data set included only patients with acute low back pain from emergency departments, who received treatment consisting of medication and advice for a period of 5 to 20 minutes.[22] Additionally, these patients sometimes had exams and/or referral to other health services.[22] The second data set included patients with chronic low back pain from physical therapy departments.[21] These patients received 10 sessions of physical therapy treatment ranging from general and specific spinal exercises to manual therapy techniques.[1,21]

Measures

The SBST The SBST is a 9-item questionnaire that stratifies patients with low back pain into 3 risk groups (low, medium, and high) that represent their prognosis with regard to disability.[18,34] After stratification, patients are referred to different levels of treatment based on each subgroup.[19] Patients classified as low risk receive a pain education session, medium-risk patients are referred to physical therapy based on clinical practice guidelines, and high-risk patients receive psychosocial components that are added to the physical therapy treatment.[19]

The SBST is composed of 9 items, represented by physical and psychosocial prognostic factors. The first 8 items ask about the patient's symptoms during the last 2 weeks, with response categories of agree and disagree, the affirmative answer counting for 1 point. The ninth item is answered on a 5-point Likert scale (not at all, slightly, moderately, very much, and extremely), with only the latter 2 response categories adding to the final score.[18,31,32] Respondents who answer "agree" on a maximum of 3 items are classified as low risk. Those who agree with more than 3 items, but no more than 3 psychosocial items (items 5-9), are considered to be medium risk. Finally, those who agree with more than 3 psychosocial items (items 5-9) are considered to be high risk.[6]

Global Perceived Effect Scale The Global Perceived Effect scale (GPE) evaluates the patient's overall perception of recovery based on the following question: "Compared to the beginning of this episode of low back pain, how would you describe your back today?" Patients answer using an 11-point Likert scale ranging from −5 (extremely worse) to +5 (fully recovered). Higher scores indicate better patient recovery.[10] The Brazilian version of the GPE showed a very high level of reliability (intraclass correlation coefficient $[ICC]_{2,1}$ = 0.90; 95% confidence interval [CI]: 0.84, 0.93). The construct validity of the GPE was confirmed by significant

correlations with the Functional Rating Index (r = −0.37), the Patient-Specific Functional Scale (r = −0.33), and the Roland-Morris Disability Questionnaire (RMDQ; r = −0.42). No ceiling or floor effects were detected.[10]

The RMDQ The RMDQ is a 24-item questionnaire that measures the level of disability of individuals with low back pain. Each item corresponds to an activity that may or may not be associated with low back pain. Patients are asked to think about how they are feeling on the day of evaluation and answer all items with "yes" or "no." The total score ranges from 0 to 24 points and is determined by the sum of the items marked "yes." The higher the score, the greater the disability.[7,10,11] The Brazilian version of the RMDQ showed a very high level of internal consistency (Cronbach α = .92) and reliability ($ICC_{2,1}$ = 0.95; 95% CI: 0.93, 0.97).[11] The construct validity of the RMDQ was highly correlated with the Functional Rating Index (r = 0.80) and was moderately correlated with pain intensity (r = 0.55). No ceiling or floor effects were detected.[11]

Numeric Pain-Rating Scale The numeric pain-rating scale (NPRS) measures the intensity of pain perceived by the patient based on the last 7 days. The assessment is performed using an 11-point Likert scale ranging from 0 to 10, with 0 being "no pain" and 10 being "the worst possible pain."[8,10] The Brazilian version of the NPRS has a very high level of reliability ($ICC_{2,1}$ = 0.94; 95% CI: 0.90, 0.96). The construct validity of the NPRS was confirmed by significant correlations with the Functional Rating Index (r = 0.63), the Patient-Specific Functional Scale (r = −0.45), and the RMDQ (r = 0.55). No ceiling or floor effects were detected.[8,10]

Procedures

Consecutive patients seeking treatment in emergency departments for acute low back pain and in physical therapy departments for chronic low back pain of 3 public hospitals in São Paulo were approached by 2 research assistants in the waiting room of both departments from Monday to Friday. All eligible patients received information about the study procedures and signed a consent form if they agreed to participate. During the evaluation, data on the demographic and anthropometric characteristics of the patients, clinical outcomes of disability and pain, and the classification of risk of unfavorable prognosis were collected face to face.

All patients were reassessed after 6 weeks and after 6 months via phone call, during which data were collected on clinical outcomes regarding global perceived effect, disability, pain, and risk classification by the SBST.

Responsiveness Measure

Responsiveness was measured to determine whether the SBST, when used in more than 1 assessment, was able to detect clinical changes at 6 weeks and 6 months. Our primary objective of this study was to test responsiveness using the total score of the SBST. This decision was made because statistical analysis using continuous outcomes is more precise compared to categorical data (ie, using the categories of the SBST). Moreover, the total score of the SBST is used to categorize patients into the 3 subgroups. The second objective was to test the responsiveness using each of the subgroups (low, medium, and high risk). This decision was made to identify whether, in isolation, all subgroups remain responsive.

Instruments that have good responsiveness are able to distinguish patients who have changed their health status from those who have not.[24] It is widely known that patients with acute low back pain are more likely to improve compared with patients with chronic low back pain.[9] We decided to combine these 2 data sets, including both patients with acute and patients with chronic low back pain, in order to have a wide spectrum of change (patients who have changed and patients who have not) to test responsiveness.

To detect clinical changes, the SBST was compared to 3 one-dimensional constructs: global perceived effect (measured by the GPE), disability (measured by the RMDQ), and pain intensity (measured by the NPRS). These outcomes were chosen for comparison because they measure the clinical outcomes that are most relevant for patients with low back pain and are considered responsive.[7,8,30]

We assessed responsiveness at 6 weeks and at 6 months after initial assessment. The interval of 6 weeks was chosen because the prognosis of patients with acute or chronic low back pain is very favorable during the first 6 weeks,[9] and most patients tend to recover in this period, regardless of whether they receive treatment.[9] Furthermore, previous studies show that more than 50% of patients with nonspecific low back pain will change classifications, according to the SBST subgroup definitions, between the initial assessment and 6 weeks later.[21-23] In contrast, the 6-month assessment was included in the study to investigate whether the SBST can be considered responsive after a longer period, given that longitudinal studies usually assess responsiveness at 6 weeks and 6 months.[13]

According to the COSMIN, criterion and construct approaches should be used to measure responsiveness.[13] The criterion approach is appropriate when the measured construct is compared to a gold standard,[13] and the construct approach when the tool of interest is compared to tools that are not a gold standard. For lack of a gold standard, this study used the construct approach. To evaluate responsiveness of the SBST using the construct approach, it was necessary to test specific a priori hypotheses.[13]

To assess responsiveness of the SBST, the hypotheses were initially tested using the total score for all patients, then tested to assess the responsiveness of each subgroup of the SBST, classified as low, medium, or high risk according to SBST scores (**TABLE 1**). According to the criteria of de Boer et al,[12] responsiveness is high when at least 75% of the hypotheses are accepted, moderate when 50% to 74% of

the hypotheses are accepted, and poor when less than 50% of the hypotheses are accepted.[12] The hypotheses tested in this study were described in terms of correlation, effect size, and receiver operating characteristic curve tests and are specified in **TABLE 1**.

Statistical Analysis

The normality of the data was inspected using histograms, and descriptive analyses were presented for the sample. All analyses were performed on the total sample of the study and the low-, medium-, and high-risk subgroups. We also

TABLE 1 A Priori Hypothesis Testing to Evaluate the Responsiveness of the SBST Compared to the GPE, RMDQ, and NPRS in Patients With Low Back Pain

Hypothesis	All Patients	Low Risk	Medium Risk	High Risk
6-wk responsiveness (change from baseline to 6 wk)				
1. The correlation between the changes in SBST and GPE scores for the 6-wk follow-up will be positive (r>0.40). This hypothesis is based on a previous study that investigated the correlation between another global measure of clinical improvement and the SBST and showed a moderate correlation[38]	Yes	No	Yes	Yes
2. The correlation between the changes in SBST and RMDQ scores will be positive (r>0.60). This hypothesis is based on previous studies showing that the SBST,[18,35] despite being a multidimensional tool, was developed to stratify patients with low back pain based on the prognosis of disability,[18,35] just as the RMDQ assesses the outcome of disability	Yes	No	Yes	Yes
3. The correlation between the changes in SBST and NPRS scores will be positive (r>0.50). This hypothesis is based on previous studies that show that the SBST is able to predict pain intensity; however, when compared to other outcomes, the SBST does not have as much influence on pain[5,21]	Yes	No	Yes	Yes
4. The effect size for the change in SBST score will be low (0.50 or less). This hypothesis is based on a previous study conducted in the English culture, which showed that the SBST does not present a high effect size[38]	No	No	No	No
5. The area under the curve for the change in SBST score compared to the GPE score at 6-wk follow-up will be 0.70 or greater. This hypothesis is based on a previous study[38] conducted in the English culture that presented high values and used similar instruments to those in this analysis. Patients were classified as "improved" or "did not improve" for the calculation of the receiver operating characteristic curve	Yes	Yes	Yes	Yes
6. The area under the curve for the change in SBST score compared to the change in RMDQ score will be 0.70 or greater. This hypothesis has the same justification as hypothesis 10	Yes	Yes	Yes	Yes
7. The area under the curve for the change in SBST score compared to the change in NPRS score will be 0.70 or greater. This hypothesis presents the same justification as hypothesis 10	Yes	Yes	Yes	Yes
Number of accepted hypotheses	6/7	3/7	6/7	6/7
6-mo responsiveness (change from baseline to 6 mo)				
8. The correlation between the change in SBST score and the GPE score for the 6-mo follow-up will be similar to hypothesis 1 (ie, a positive correlation of r>0.40). This hypothesis is based on previous prognostic studies on low back pain.[9] These studies show that the prognosis of patients with low back pain is favorable in the first 6 weeks, but after that the improvement progresses more slowly, and patients remain relatively stable for up to 6 mo[9]	Yes	Yes	Yes	Yes
9. The correlation between the changes in SBST and RMDQ scores will be similar to hypothesis 2 (ie, a positive correlation of r>0.60). This hypothesis presents the same explanation as hypotheses 2 and 8	Yes	Yes	Yes	Yes
10. The correlation between the changes in SBST and NPRS scores will be similar to hypothesis 3 (ie, a positive correlation of r>0.50). This hypothesis presents the same justification as hypotheses 3 and 8	Yes	Yes	Yes	Yes
11. The correlation between the SBST and RMDQ change scores will be greater than the correlation between the SBST and NPRS change scores and between the SBST and GPE at 6 wk and 6 mo. This hypothesis is based on previous studies that show that disability is the outcome with greater association with the SBST[19]	Yes	Yes	Yes	Yes
12. The effect size for the change in SBST score will be low (ie, 0.50 or less). This hypothesis has the same justification as hypothesis 4[38]	No	Yes	No	No
13. The area under the curve for the change in SBST score compared to the GPE score at the 6-mo follow-up will be 0.70 or greater. This hypothesis has the same justification as hypothesis 5	Yes	Yes	Yes	Yes
14. The area under the curve for the change in SBST score compared to the change in RMDQ score will be 0.70 or greater. This hypothesis presents the same justification as hypothesis 5	Yes	Yes	Yes	Yes
15. The area under the curve for the change in SBST score compared to the change in NPRS score will be 0.70 or greater. This hypothesis presents the same justification as hypothesis 5	Yes	Yes	Yes	Yes
Number of accepted hypotheses	7/8	8/8	7/8	7/8

Abbreviations: GPE, Global Perceived Effect scale; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.

conducted a sensitivity analysis by separating the acute and chronic data sets to check whether duration of symptoms influenced responsiveness (**APPENDIX** available at www.jospt.org).

Correlation We performed a correlation analysis between the SBST change score (the difference between the baseline and final scores) and the GPE scores at 6 weeks and 6 months, and RMDQ, and NPRS change scores from baseline to 6 weeks (time point 1) and from baseline to 6 months (time point 2). Pearson or Spearman correlation coefficients were calculated, depending on the normality of the data. Correlation strength was interpreted according to the following criteria: $r<0.30$ indicates low correlation, $r\geq0.30$ to $r<0.60$ indicates moderate correlation, and $r\geq0.60$ indicates strong correlation.[6]

Effect Size Effect size was calculated to assess SBST responsiveness and was calculated with a CI of 84% of the sample, because nonoverlapping 84% CIs are equivalent to a Z score at a .05 level.[36] The effect size for the SBST was calculated using the mean SBST change score, between initial assessment and follow-ups, divided by the standard deviation of the initial SBST assessment.[20] Using the same formula described above, the effect sizes for the outcomes of disability and pain intensity were also calculated for possible comparison between the tools. For interpretation of the results, effect-size values of 0.2 or less were considered low, effect-size values between 0.50 and 0.80 were considered moderate, and effect-size values of 0.80 or above were considered high. The higher the effect size, the more responsive the tool is.

The receiver operating characteristic curve was calculated to show the probability of correct discrimination between patients who improved and those who did not improve. The change score of the SBST was classified as a dependent variable for all analyses and was compared to the outcomes of the GPE, RMDQ, and NPRS. For the discrimination between patients who improved and those who did not improve, cutoff scores were created based on the literature to transform continuous scores into dichotomous outcomes. This allowed for the identification of recovery indices for each outcome. Based on previous studies, a cutoff score of 3 was used for the outcome of global perceived effect, with patients who scored 3 or more being classified as improved and those with scores lower than 3 as not improved.[10] The disability outcome was dichotomized between improved (based on a score less than 7 in the follow-up assessments) and did not improve (based on a score greater than or equal to 7 in the follow-up assessments).[2,36] For pain intensity (NPRS), an improvement above 30% between initial assessment and follow-up assessments was categorized as improved. For changes in pain less than or equal to 30%, the patient was categorized as not improved.[14,29] The analysis was based on the area under the curve (AUC), which theoretically varies between 0.5 (no discriminant accuracy) and 1.0 (perfect discriminant accuracy). All analyses were performed using SPSS (IBM Corporation, Armonk, NY).

RESULTS

From June 2013 to September 2015, a total of 464 patients with low back pain were approached at emergency and physical therapy departments. After screening for inclusion, 116 patients were excluded and 348 patients with acute and chronic low back pain were included. At 6 weeks, 96% of the 348 were reassessed, and at 6 months 92% were reassessed. The main reasons for loss at follow-ups were (1) patients did not answer the telephone calls, even if these calls were made at different times of the day, and (2) patients changed their telephone number. In addition, 1 patient died from problems unrelated to low back pain.

TABLE 1 presents the hypotheses formulated a priori and the number of those accepted when analyzed as a single sample and when stratified by subgroups of the SBST. Fifteen hypotheses were tested in this study. Seven of these hypotheses tested responsiveness at 6 weeks and 8 tested the responsiveness of the SBST at 6 months. Most of the hypotheses were accepted. When the analyses were performed using all patients, a total of 85.7% and 87.5% of the hypotheses of responsiveness were accepted at 6 weeks and 6 months, respectively. These results represent high responsiveness. When used in the subgroups of patients classified as medium and high risk, the SBST also had high responsiveness (85.7% and 87.5% of the hypotheses were accepted at 6 weeks and 6 months, respectively). However, only 42.9% of the hypotheses tested in patients classified as low risk were accepted at 6 weeks, indicating poor responsiveness. The responsiveness at 6 months in low-risk patients was high, with 100% of hypotheses accepted. Our sensitivity analysis results in the groups with either acute or chronic patients were not different compared to the main analysis. Although the responsiveness estimates in the acute sample were slightly higher than those in the chronic sample, the proportion of accepted hypotheses remained above the 75% threshold. For this reason, the results of this sensitivity analysis were omitted from this manuscript (**APPENDIX**).

TABLE 2 presents the characteristics of the sample at baseline. More than half of the patients were female (56.6%), with a total mean age of 42 years. A total of 57.5% of the patients had acute low back pain and 42.5% had chronic low back pain. Almost half of the patients (46.6%) had been taking medication for low back pain; however, less than 25% of the sample exercised regularly. These patients presented moderate levels of pain intensity and disability, with mean scores of 7.6 and 14.6 points, respectively. Almost half of the patients (44.5%) were classified as being at high risk according to the SBST. The characteristics of the sample are further described in **TABLE 2**.

TABLE 3 presents the a priori hypotheses of correlation and the Pearson (r) or Spearman (rho) correlations between the SBST and the GPE, RMDQ, and NPRS.

Both the continuous SBST score and each subgroup score of the tool showed moderate to high correlation with the outcomes of global perceived effect, disability, and pain intensity. The magnitude of the correlations varied from rho = 0.38 (for the correlation between the low-risk subgroup of the SBST and the GPE at 6 weeks) to $r = 0.76$ (for the correlation between the high-risk subgroup of the SBST and the RMDQ at 6 months). Although classified as a moderate correlation, only the low-risk subgroup of the SBST showed correlations that were lower than expected at 6 weeks and refuted the hypotheses for all outcomes analyzed.

TABLE 4 presents the within-group mean difference for prognostic risk (SBST), disability (RMDQ), and pain intensity (NPRS). Data for all patients included in the study and the stratified data for each SBST subgroup were also described. We can see that the risk subgroups of the SBST reflect the level of clinical symptoms of the patients at baseline, after 6 weeks, and after 6 months. The lower the risk of the subgroups, the lower the mean score for all outcomes. The within-group improvement was similar when compared at 6 weeks and 6 months. The effect sizes calculated to analyze responsiveness were strong (above 0.8) at both 6 weeks and 6 months for all outcomes and in all subgroups, except for the low-risk subgroup, in which the effect size was 0.75 at 6 weeks and 0.40 at 6 months for the SBST. The hypotheses formulated for the effect size were accepted only for the low-risk subgroup of the SBST at 6 months.

In **TABLE 5**, AUC values are specified to test the capacity of the tool to distinguish patients who recovered from those who did not recover. We can see that the estimated value of the AUC was above 0.70 for the comparison between the SBST and all outcomes, both at 6 weeks and 6 months. All hypotheses related to the AUC were accepted.

TABLE 2 Demographic Data and Clinical Characteristics of the Participants*

Variable	All Patients (n = 348)	Acute Patients (n = 200)	Chronic Patients (n = 148)
Sex, n (%)			
Male	151 (43.4)	109 (54.5)	42 (28.4)
Female	197 (56.6)	91 (45.5)	106 (71.6)
Age, y	41.8 ± 12.8	39.3 ± 13.2	45.1 ± 11.6
Duration of symptoms, days†	17 (1076)	6 (14)	48 (101)‡
Weight, kg	72.6 ± 14.5	73.4 ± 15.2	71.4 ± 13.3
Height, m	1.66 ± 0.1	1.68 ± 0.1	1.64 ± 0.1
Marital status, n (%)			
Single	163 (46.8)	101 (50.5)	62 (41.9)
Married	129 (37.1)	69 (34.5)	60 (40.5)
Divorced	28 (8.0)	12 (6.0)	16 (10.8)
Widow/widower	17 (4.9)	10 (5.0)	7 (4.7)
Other	11 (3.2)	8 (4.0)	3 (2.0)
Education, n (%)§			
Primary	120 (35.2)	61 (30.8)	59 (41.3)
Secondary	155 (45.5)	98 (49.5)	57 (39.9)
Undergraduate degree	55 (16.1)	32 (16.2)	23 (16.1)
Postgraduate studies	11 (3.2)	7 (3.5)	4 (2.8)
Use of medication, n (%)	162 (46.6)	92 (46.0)	70 (47.3)
Previous episode of LBP, n (%)	175 (50.3)	138 (69.0)	37 (25.0)
Exercises regularly, n (%)	82 (24.4)	43 (21.5)	39 (26.4)
Smoker, n (%)	55 (15.8)	35 (17.5)	20 (13.5)
Currently on sick leave, n (%)	25 (7.2)	15 (7.5)	10 (6.8)
Receiving financial compensation while on sick leave, n (%)	8 (2.3)	4 (2.0)	4 (2.7)
Pain intensity (0-10)	7.6 ± 1.9	7.7 ± 2.2	7.5 ± 1.7
Disability (0-24)	14.6 ± 6.0	15.4 ± 6.1	13.7 ± 5.8
Total SBST score (0-9)	5.4 ± 2.3	5.3 ± 2.4	5.6 ± 2.3
SBST subgroup, n (%)			
Low risk	77 (22.1)	46 (23.0)	31 (20.9)
Medium risk	116 (33.3)	64 (32.0)	52 (35.1)
High risk	155 (44.5)	90 (45.0)	65 (43.9)

Abbreviations: LBP, low back pain; SBST, Subgroups for Targeted Treatment Back Screening Tool.
**Values are mean ± SD unless otherwise indicated.*
†Values are median (interquartile range).
‡Values are months.
§Data are from 341 patients due to missing data from 7 patients.

DISCUSSION

The aim of this study was to investigate the responsiveness of the SBST as a clinical outcome measure in patients with nonspecific low back pain. Our results show that the SBST presents high responsiveness, both at 6 weeks and 6 months, for the outcomes of global perceived effect, disability, and pain intensity using the total score. The scores of the high- and medium-risk SBST subgroups also indicate high responsiveness at 6 weeks and 6 months, while the score of the low-risk SBST subgroup indicates poor responsiveness at 6 weeks, but high responsiveness at 6 months. These results suggest that clinicians can confidently use the SBST in clinical practice to monitor changes over time.

According to the COSMIN, the most accepted way to evaluate responsiveness

is to use hypothesis testing that is specific and formulated a priori.[26] This was the first study to follow these guidelines when testing responsiveness of the SBST. A total of 15 hypotheses were initially formulated, described, and then tested according to the direction and size of the expected correlation, effect size, and receiver operating characteristic curve tests. The absence of a gold standard for comparison with the SBST may be a limitation of the tool and the study. However, we used the most accepted tools in the literature[7] that measured the main outcomes for low back pain.[8] Another possible limitation of this study is that the data set came from 2 previous studies that were not initially developed to measure the responsiveness of the SBST.[21,22] The sample included in our study was not homogeneous; that is, there were both acute and chronic patients with low back pain. Also, the treatments received by these patients were different.[21,22] Therefore, some caution is needed while interpreting the generalizability of our findings, specifically to populations other than individuals with back pain residing in Brazil. Finally, we observed that, although the treatments were not stratified according to the SBST, patients from all risk groups improved over time, especially the high-risk subgroup. This finding can be explained by a larger regression to the mean, as high-risk patients experience more pain intensity and disability compared with patients from medium- and low-risk groups.[16] Future studies comparing the effects and costs of matched versus unmatched treatments using the SBST are needed.

The responsiveness of the SBST, assessed by correlation tests, was positive. The outcome of disability showed a stronger correlation with the SBST compared to global perceived effect and pain intensity. This was expected, considering that the SBST was developed to measure the unfavorable prognosis of disability.[18] However, only the low-risk subgroup had a smaller correlation than expected over 6 weeks. Based on a previous study,[38] it was expected that the effect size of the SBST would be low to moderate (0.50 or less). However, the SBST only had a low effect size (0.40) in the low-risk subgroup at 6 months. The other subgroups had a high effect size (greater than 0.80), which was higher than expected at both 6 weeks and 6 months. Finally, the responsiveness of the SBST, as assessed by the AUC, was higher than expected (0.70 or greater) in all subgroups at 6 weeks and 6 months. Overall, the SBST had high responsiveness, with greater than 75% of the hypotheses accepted. However, in patients classified as low risk at 6 weeks, the SBST had poor responsiveness, with less than 50% of the hypotheses accepted. This could be explained by the very favorable prognosis that patients classified as low risk had during the first 6 weeks, combined with little change in score (ranging from 0 to 3 points), making it difficult to detect large clinical changes. Another possible explanation is that the analyses performed with the low-risk patients had the lowest number of patients included (n = 74) compared to the medium- and high-risk subgroups. According to COSMIN recommendations, the optimal number of patients for assessing responsiveness is 100.[13]

A recent systematic review[3] of cross-cultural adaptation and measurement properties summarized 11 different SBST versions, translated into 10 languages. The methodological quality of these studies was considered low.[3] Also, none of the studies evaluated all SBST measurement properties.[3] The Brazilian version of the SBST had good measurement properties compared to the original version of the instrument.[31,32] The reliability of the Brazilian version yielded an ICC of 0.79 (95% CI: 0.63, 0.95). The internal consistency measured by Cronbach's alpha was .74 for the total score and .72 for the psychosocial subscale, and the standard error of measurement was 1.9%, demonstrating excellent reliability, agreement, and internal consistency.[31,32] The construct validity of the Brazilian version of the SBST showed a good correlation with the disability tools

TABLE 3 Pearson (*r*) or Spearman (rho) Correlations Between the SBST and the 1-Dimensional Tools for Global Perceived Effect, Disability, and Pain Intensity

		6 wk		6 mo	
Comparison Instrument	**A Priori Hypothesis**	**n**	**Correlation**	**n**	**Correlation**
Change in SBST versus GPE	*r*>0.40				
All patients		333	rho = 0.48	320	rho = 0.48
Low risk		74	rho = 0.38	71	rho = 0.49
Medium risk		112	rho = 0.53	106	rho = 0.73
High risk		147	rho = 0.66	143	rho = 0.71
Change in SBST versus change in RMDQ	*r*>0.60				
All patients		333	*r* = 0.66	320	*r* = 0.75
Low risk		74	*r* = 0.51	71	*r* = 0.71
Medium risk		112	*r* = 0.63	106	*r* = 0.74
High risk		147	*r* = 0.71	143	*r* = 0.76
Change in SBST versus change in NPRS	*r*>0.50				
All patients		333	*r* = 0.55	320	*r* = 0.57
Low risk		74	*r* = 0.40	71	*r* = 0.59
Medium risk		112	*r* = 0.56	106	*r* = 0.58
High risk		147	*r* = 0.65	143	*r* = 0.64

Abbreviations: GPE, Global Perceived Effect scale; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.

(r = 0.61 with the Oswestry Disability Index and r = 0.70 with the RMDQ), but weak correlations with the Fear-Avoidance Beliefs Questionnaire work and physical activity subscales (r = 0.18 and r = 0.28, respectively). Finally, the results of discriminant validity suggest that the Brazilian version of the SBST is able to discriminate patients with low back pain and disability (AUC = 0.88) from those with low back pain and fear-avoidance beliefs (AUC = 0.66).[31,32] The results of the present study included, the measurement properties of the Brazilian version of the SBST meet the recommendations of the COSMIN.[13]

We found that only 1 study, conducted for the Persian version, evaluated the responsiveness of the SBST in patients with low back pain (due to spinal stenosis).[4] Although the study concluded that the SBST is capable of detecting clinical change in this population, responsiveness was only assessed by the statistically significant differences between the SBST score at baseline and post intervention.[4] This type of measurement is not considered adequate, because the statistical test can be influenced by the homogeneity and size of the sample. In addition, the

TABLE 4 Baseline, 6-Week, and 6-Month Values for the Clinical Outcomes of All Patients and of Those in Each SBST Subgroup*

Follow-up/Instrument	n	Baseline	Follow-up	Mean Difference†	Effect Size‡
6 wk					
SBST (0-9)					
All patients	333	5.4 ± 2.32	3.4 ± 3.00	2.1 ± 2.63	0.89 (0.80, 0.97)
Low risk	74	2.1 ± 0.88	1.4 ± 1.77	0.7 ± 1.71	0.75 (0.43, 1.08)
Medium risk	112	5.1 ± 1.10	2.8 ± 2.54	2.3 ± 2.42	2.09 (1.80, 2.38)
High risk	147	7.4 ± 1.22	4.8 ± 3.13	2.6 ± 2.91	2.11 (1.91, 2.39)
RMDQ (0-24)					
All patients	333	14.6 ± 5.99	9.0 ± 7.69	5.6 ± 6.99	0.99 (0.90, 1.08)
Low risk	74	7.9 ± 4.89	3.8 ± 4.96	4.0 ± 6.02	0.82 (0.62, 1.03)
Medium risk	112	14.5 ± 4.63	7.9 ± 6.83	6.6 ± 6.94	1.43 (1.23, 1.63)
High risk	147	18.1 ± 4.23	11.7 ± 8.06	6.4 ± 7.35	1.51 (1.30, 1.71)
NPRS (0-10)					
All patients	333	7.6 ± 1.95	4.6 ± 3.29	3.0 ± 3.20	1.53 (1.40, 1.65)
Low risk	74	6.2 ± 1.92	3.2 ± 2.88	3.0 ± 3.14	1.57 (1.30, 1.84)
Medium risk	112	7.4 ± 1.93	4.2 ± 3.05	3.1 ± 3.18	1.63 (1.41, 1.85)
High risk	147	8.5 ± 1.48	5.7 ± 3.34	2.8 ± 3.26	1.91 (1.66, 2.17)
6 mo					
SBST (0-9)					
All patients	320	5.4 ± 2.34	3.4 ± 3.23	2.0 ± 2.98	0.87 (0.77, 0.97)
Low risk	71	2.1 ± 0.88	1.7 ± 2.45	0.4 ± 2.24	0.40 (−0.03, 0.82)
Medium risk	106	5.1 ± 1.12	3.1 ± 2.89	2.0 ± 2.78	1.79 (1.46, 2.13)
High risk	143	7.3 ± 1.24	4.4 ± 3.43	2.9 ± 3.10	2.35 (2.05, 2.65)
RMDQ (0-24)					
All patients	320	14.5 ± 6.01	7.6 ± 8.06	6.9 ± 8.01	1.15 (1.04, 1.26)
Low risk	71	7.6 ± 4.61	3.3 ± 6.08	4.2 ± 6.65	0.92 (0.68, 1.16)
Medium risk	106	14.3 ± 4.53	7.4 ± 7.55	6.9 ± 7.56	1.52 (1.29, 1.75)
High risk	143	18.2 ± 4.27	9.9 ± 8.44	8.3 ± 8.70	1.93 (1.69, 2.17)
NPRS (0-10)					
All patients	320	7.6 ± 1.96	4.6 ± 3.46	3.0 ± 3.54	1.53 (1.38, 1.67)
Low risk	71	6.3 ± 1.94	3.0 ± 3.21	3.2 ± 3.67	1.65 (1.34, 1.97)
Medium risk	106	7.4 ± 1.23	4.7 ± 3.27	2.6 ± 3.47	2.14 (1.76, 2.53)
High risk	143	8.5 ± 1.45	5.4 ± 3.47	3.1 ± 3.53	2.17 (1.88, 2.46)

Abbreviations: NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.
**Values are mean ± SD unless otherwise indicated.*
†Change score from baseline to follow-up.
‡Values in parentheses are 84% confidence interval.

hypothesis was not formulated a priori.[25,26] A study that was not included in the above systematic review, but that also tested the responsiveness of the SBST,[38] compared change in SBST score to the clinical change of 1-dimensional tools.[38] Those results corroborate the results of the present study: the SBST is responsive in a different culture in a high-income population and is able to detect clinical change in pain intensity, global change, disability, pain catastrophizing, and fear of movement. However, the SBST was not responsive to detect clinical change in patients with depression symptoms.[38] Although there are only 2 studies that tested SBST responsiveness, the results presented to date have been positive, independent of population, and corroborate our study.[4,38]

TABLE 5 Estimated Values Comparing SBST Variation With the Clinical Outcomes of Global Perceived Effect, Disability, and Pain Severity at Follow-up

Comparison Instrument	n	AUC*	*P* Value
Baseline to 6 wk			
Change in SBST by GPE cutoff score†			
All patients	333	0.77 (0.71, 0.82)	<.001
Low risk	74	0.71 (0.56, 0.86)	.009
Medium risk	112	0.79 (0.68, 0.89)	<.001
High risk	147	0.84 (0.77, 0.90)	<.001
Change in SBST by RMDQ cutoff score‡			
All patients	333	0.75 (0.70, 0.80)	<.001
Low risk	74	0.82 (0.70, 0.95)	<.001
Medium risk	112	0.85 (0.78, 0.92)	<.001
High risk	147	0.87 (0.81, 0.93)	<.001
Change in SBST by NPRS cutoff score§			
All patients	333	0.79 (0.74, 0.84)	.003
Low risk	74	0.71 (0.58, 0.83)	<.001
Medium risk	112	0.77 (0.68, 0.86)	<.001
High risk	147	0.87 (0.81, 0.93)	<.001
Baseline to 6 mo			
Change in SBST by GPE cutoff score†			
All patients	320	0.83 (0.78, 0.87)	<.001
Low risk	71	0.75 (0.62, 0.89)	.001
Medium risk	106	0.90 (0.84, 0.96)	<.001
High risk	143	0.89 (0.84, 0.95)	<.001
Change in SBST by RMDQ cutoff score‡			
All patients	320	0.83 (0.79, 0.88)	<.001
Low risk	71	0.98 (0.95, 1.00)	<.001
Medium risk	106	0.91 (0.85, 0.97)	<.001
High risk	143	0.93 (0.89, 0.98)	<.001
Change in SBST by NPRS cutoff score§			
All patients	320	0.79 (0.74, 0.84)	<.001
Low risk	71	0.82 (0.80, 0.92)	<.001
Medium risk	106	0.79 (0.69, 0.88)	<.001
High risk	143	0.86 (0.80, 0.92)	<.001

Abbreviations: AUC, area under the curve; GPE, Global Perceived Effect scale; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.
**Values in parentheses are 95% confidence interval.*
†Patients who scored 3 or more points (−5 to +5) were classified as "improved."
‡Patients who scored less than 7 points (0-24) were classified as "improved."
§Patients with improvement of more than 30% between baseline and follow-up were classified as "improved."

Our results are important for clinical practice for a number of reasons. First, the SBST is largely used in clinical practice and is recommended by clinical practice guidelines, such as the National Institute for Health and Care Excellence guidelines.[28] Second, the SBST is shorter than most of the available outcome measures for low back pain,[34] making it easier to monitor patients with low back pain over time. Third, other risk-stratification tools for patients with low back pain, such as the Örebro Musculoskeletal Pain Screening Questionnaire, were not found to be as responsive as the SBST.[17] Therefore, the SBST could be applied during the reassessment of patients to measure changes in the most important clinical outcomes for patients with low back pain.[7,8] One of the areas in clinical practice that would benefit the most from this reassessment approach is physical therapy, in which patients often return for treatment sessions. A recent study showed that the SBST, when applied during reassessment as part of physical therapy, was able to monitor the unfavorable prognostic factors that were present for each patient and to direct the treatment to modify these factors.[21] Future research should determine the benefits of using the SBST and test the implementation of the SBST in different health sectors.

CONCLUSION

The SBST had high responsiveness measured in a single sample at 6 weeks and 6 months follow-up. Specific responsiveness to subgroups of patients with medium risk and high risk of persistent disability were high, and poor responsiveness to a low-risk subgroup at 6 weeks. All specific SBST

subgroups had high responsiveness at 6 months. Clinicians can confidently use the SBST to measure changes over time in terms of non-specific low back pain.

KEY POINTS

FINDINGS: The Subgroups for Targeted Treatment (STarT) Back Screening Tool (SBST) showed high responsiveness and was able to measure changes in the short and medium term compared with the most important clinical outcomes for patients with low back pain.

IMPLICATIONS: The SBST can measure clinical change in patients with low back pain and can be used in research as well as in clinical practice.

CAUTION: Patients with acute and with chronic low back pain residing in Brazil composed the sample included in the study. Also, the treatment received by these patients was not stratified according to the SBST; that is, 57% of patients received emergency medical care and 43% of patients received 10 sessions of physical therapy care. Therefore, caution is needed while interpreting the generalizability of our findings.

ACKNOWLEDGMENTS: *We wish to thank Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP grant 2013/02075-8 and 2016/18575-8) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for funding this study. The authors thank the patients of the physical therapy and emergency departments of the following hospitals: Irmandade Santa Casa de Misericórdia de São Paulo, Centro Hospitalar do Município de Santo André, and Hospital Municipal do Tatuapé (Dr Cármino Caricchio).*

REFERENCES

1. Added MA, Costa LO, de Freitas DG, et al. Kinesio Taping does not provide additional benefits in patients with chronic low back pain who receive exercise and manual therapy: a randomized controlled trial. *J Orthop Sports Phys Ther*. 2016;46:506-513. https://doi.org/10.2519/jospt.2016.6590
2. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ*. 2006;332:1080. https://doi.org/10.1136/bmj.332.7549.1080
3. Al Zoubi FM, Eilayyan O, Mayo NE, Bussières AE. Evaluation of cross-cultural adaptation and measurement properties of STarT Back Screening Tool: a systematic review. *J Manipulative Physiol Ther*. 2017;40:558-572. https://doi.org/10.1016/j.jmpt.2017.07.005
4. Azimi P, Shahzadi S, Azhari S, Montazeri A. A validation study of the Iranian version of STarT Back Screening Tool (SBST) in lumbar central canal stenosis patients. *J Orthop Sci*. 2014;19:213-217. https://doi.org/10.1007/s00776-013-0506-y
5. Beneciuk JM, Fritz JM, George SZ. The STarT Back Screening Tool for prediction of 6-month clinical outcomes: relevance of change patterns in outpatient physical therapy settings. *J Orthop Sports Phys Ther*. 2014;44:656-664. https://doi.org/10.2519/jospt.2014.5178
6. Campbell MJ, Swinscow TDV. Correlation and regression. In: *Statistics at Square One*. 11th ed. New York, NY: Wiley-Blackwell; 2009:ch 11.
7. Chiarotto A, Boers M, Deyo RA, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain*. 2018;159:481-495.
8. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J*. 2015;24:1127-1142. https://doi.org/10.1007/s00586-015-3892-3
9. Costa LC, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LO. The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ*. 2012;184:E613-E624. https://doi.org/10.1503/cmaj.111271
10. 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 (Phila Pa 1976)*. 2008;33:2459-2463. https://doi.org/10.1097/BRS.0b013e3181849dbe
11. Costa LO, Maher CG, Latimer J, Ferreira PH, Pozzi GC, Ribeiro RN. Psychometric characteristics of the Brazilian-Portuguese versions of the Functional Rating Index and the Roland Morris Disability Questionnaire. *Spine (Phila Pa 1976)*. 2007;32:1902-1907. https://doi.org/10.1097/BRS.0b013e31811eab33
12. de Boer MR, Moll AC, de Vet HC, Terwee CB, Völker-Dieben HJ, van Rens GH. Psychometric properties of vision-related quality of life questionnaires: a systematic review. *Ophthalmic Physiol Opt*. 2004;24:257-273. https://doi.org/10.1111/j.1475-1313.2004.00187.x
13. de Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge, UK: Cambridge University Press; 2011.
14. Farrar JT, Young JP, Jr., LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94:149-158. https://doi.org/10.1016/s0304-3959(01)00349-9
15. Ferreira G, Costa LM, Stein A, Hartvigsen J, Buchbinder R, Maher CG. Tackling low back pain in Brazil: a wake-up call. *Braz J Phys Ther*. 2019;23:189-195. https://doi.org/10.1016/j.bjpt.2018.10.001
16. Herbert R, Jamtvedt G, Hagen KB, Mead J. *Practical Evidence-Based Physiotherapy*. 2nd ed. Edinburgh, UK: Elsevier/Churchill Livingstone; 2011.
17. Hilfiker R, Knutti IA, Raval-Roland B, Rivier G, Crombez G, Opsommer E. Validity and responsiveness of the French version of the Örebro Musculoskeletal Pain Screening Questionnaire in chronic low back pain. *Eur Spine J*. 2016;25:2741-2749. https://doi.org/10.1007/s00586-016-4635-9
18. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008;59:632-641. https://doi.org/10.1002/art.23563
19. 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*. 2011;378:1560-1571. https://doi.org/10.1016/S0140-6736(11)60937-9
20. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53:459-468. https://doi.org/10.1016/S0895-4356(99)00206-1
21. Medeiros FC, Costa LOP, Added MAN, Salomão EC, Costa LCM. Longitudinal monitoring of patients with chronic low back pain during physical therapy treatment using the STarT Back Screening Tool. *J Orthop Sports Phys Ther*. 2017;47:314-323. https://doi.org/10.2519/jospt.2017.7199
22. Medeiros FC, Costa LOP, Oliveira IS, Oshima RK, Costa LCM. The use of STarT Back Screening Tool in emergency departments for patients with acute low back pain: a prospective inception cohort study. *Eur Spine J*. 2018;27:2823-2830. https://doi.org/10.1007/s00586-018-5586-0
23. Melloh M, Elfering A, Egli Presland C, et al. Identification of prognostic factors for chronicity in patients with low back pain: a review of screening instruments. *Int Orthop*. 2009;33:301-313. https://doi.org/10.1007/s00264-008-0707-8
24. Mokkink LB, Prinsen CA, Bouter LM, de Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016;20:105-113. https://doi.org/10.1590/bjpt-rbf.2014.0143
25. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737-745. https://doi.org/10.1016/j.jclinepi.2010.02.006

26. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN Checklist Manual. Amsterdam, the Netherlands: Vrije Universiteit Medical Center; 2012.

27. Morsø L, Kent P, Manniche C, Albert HB. The predictive ability of the STarT Back Screening Tool in a Danish secondary care setting. *Eur Spine J*. 2014;23:120-128. https://doi.org/10.1007/s00586-013-2861-y

28. National Institute for Health and Care Excellence. Low Back Pain and Sciatica in Over 16s: Assessment and Management. London, UK: National Institute for Health and Care Excellence; 2016.

29. Ostelo RW, 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. https://doi.org/10.1097/BRS.0b013e31815e3a10

30. Pengel LH, Refshauge KM, Maher CG. Responsiveness of pain, disability, and physical impairment outcomes in patients with low back pain. *Spine (Phila Pa 1976)*. 2004;29:879-883. https://doi.org/10.1097/00007632-200404150-00011

31. Pilz B, Vasconcelos RA, Marcondes FB, Lodovichi SS, Mello W, Grossi DB. The Brazilian version of STarT Back Screening Tool – translation, cross-cultural adaptation and reliability. *Braz J Phys Ther*. 2014;18:453-461. https://doi.org/10.1590/bjpt-rbf.2014.0028

32. Pilz B, Vasconcelos RA, Teixeira PP, et al. Construct and discriminant validity of STarT Back Screening Tool – Brazilian version. *Braz J Phys Ther*. 2017;21:69-73. https://doi.org/10.1016/j.bjpt.2016.12.006

33. Sharma S, Abbott JH, Jensen MP. Why clinicians should consider the role of culture in chronic pain. *Braz J Phys Ther*. 2018;22:345-346. https://doi.org/10.1016/j.bjpt.2018.07.002

34. Sowden G, Hill JC, Morso L, Louw Q, Foster NE. Advancing practice for back pain through stratified care (STarT Back). *Braz J Phys Ther*. 2018;22:255-264. https://doi.org/10.1016/j.bjpt.2018.06.003

35. Traeger A, McAuley JH. STarT Back Screening Tool. *J Physiother*. 2013;59:131. https://doi.org/10.1016/S1836-9553(13)70170-X

36. Tryon WW. Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: an integrated alternative method of conducting null hypothesis statistical tests. *Psychol Methods*. 2001;6:371-386. https://doi.org/10.1037//1082-989X.6.4.371

37. Waddell G. *The Back Pain Revolution*. 2nd ed. New York, NY: Elsevier/Churchill Livingstone; 2004.

38. Wideman TH, Hill JC, Main CJ, Lewis M, Sullivan MJ, Hay EM. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. *Pain*. 2012;153:2182-2191. https://doi.org/10.1016/j.pain.2012.06.010

APPENDIX

A PRIORI HYPOTHESIS TESTING TO EVALUATE THE RESPONSIVENESS OF THE SBST COMPARED TO THE GPE, RMDQ, AND NPRS IN PATIENTS WITH LOW BACK PAIN

	Acute Patients				Chronic Patients			
Hypothesis	**All Patients**	**Low Risk**	**Medium Risk**	**High Risk**	**All Patients**	**Low Risk**	**Medium Risk**	**High Risk**
6-wk responsiveness (change from baseline to 6 wk)								
1. The correlation between the changes in SBST and GPE scores for the 6-wk follow-up will be positive (r>0.40). This hypothesis is based on a previous study that investigated the correlation between another global measure of clinical improvement and the SBST and showed a moderate[38]	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
2. The correlation between the changes in SBST and RMDQ scores will be positive (r>0.60). This hypothesis is based on previous studies showing that the SBST,[18,35] despite being a multidimensional tool, was developed to stratify patients with low back pain based on the prognosis of disability,[18,35] just as the RMDQ assesses the outcome of disability	Yes	No	Yes	Yes	Yes	Yes	No	Yes
3. The correlation between the changes in SBST and NPRS scores will be positive (r>0.50). This hypothesis is based on previous studies that show that the SBST is able to predict pain intensity; however, when compared to other outcomes, the SBST does not have as much influence on pain[5,21]	Yes	No	Yes	Yes	Yes	No	Yes	Yes
4. The effect size for the change in SBST score will be low (0.50 or less). This hypothesis is based on a previous study conducted in the English culture, which showed that the SBST does not present a high effect size[38]	No	No	No	No	No	No	No	No
5. The area under the curve for the change in SBST score compared to the GPE score at 6-wk follow-up will be 0.70 or greater. This hypothesis is based on a previous study[38] conducted in the English culture that presented high values and used similar instruments to those in this analysis. Patients were classified as "improved" or "did not improve" for the calculation of the receiver operating characteristic curve	Yes	No	Yes	Yes	Yes	No	Yes	Yes
6. The area under the curve for the change in SBST score compared to the change in RMDQ score will be 0.70 or greater. This hypothesis has the same justification as hypothesis 10	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. The area under the curve for the change in SBST score compared to the change in NPRS score will be 0.70 or greater. This hypothesis presents the same justification as hypothesis 10	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Number of accepted hypotheses	6/7	3/7	6/7	6/7	6/7	2/7	5/7	6/7
6-mo responsiveness (change from baseline to 6 mo)								
8. The correlation between the change in SBST score and the GPE score for the 6-mo follow-up will be similar to hypothesis 1 (ie, a positive correlation of r>0.40). This hypothesis is based on previous prognostic studies on low back pain.[9] These studies show that the prognosis of patients with low back pain is favorable in the first 6 wk, but after that the improvement progresses more slowly, and patients remain relatively stable for up to 6 mo[9]	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
9. The correlation between the changes in SBST and RMDQ scores will be similar to hypothesis 2 (ie, a positive correlation of r>0.60). This hypothesis presents the same explanation as hypotheses 2 and 8	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10. The correlation between the changes in SBST and NPRS scores will be similar to hypothesis 3 (ie, a positive correlation of r>0.50). This hypothesis presents the same justification as hypotheses 3 and 8	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
11. The correlation between the SBST and RMDQ change scores will be greater than the correlation between the SBST and NPRS change scores and between the SBST and GPE at 6 wk and 6 mo. This hypothesis is based on previous studies that show that disability is the outcome with greater association with the SBST[19]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. The effect size for the change in SBST score will be low (ie, 0.50 or less). This hypothesis has the same justification as hypothesis 4[38]	No	No	No	No	No	Yes	No	No

Table continues on page C2.

APPENDIX

Hypothesis	Acute Patients				Chronic Patients			
	All Patients	Low Risk	Medium Risk	High Risk	All Patients	Low Risk	Medium Risk	High Risk
13. The area under the curve for the change in SBST score compared to the GPE score at the 6-mo follow-up will be 0.70 or greater. This hypothesis has the same justification as hypothesis 5	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
14. The area under the curve for the change in SBST score compared to the change in RMDQ score will be 0.70 or greater. This hypothesis presents the same justification as hypothesis 5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
15. The area under the curve for the change in SBST score compared to the change in NPRS score will be 0.70 or greater. This hypothesis presents the same justification as hypothesis 5	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Number of accepted hypotheses	7/8	7/8	7/8	7/8	7/8	4/8	7/8	7/8

Abbreviations: GPE, Global Perceived Effect scale; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.

PEARSON (*r*) OR SPEARMAN (rho) CORRELATIONS BETWEEN THE SBST AND THE 1-DIMENSIONAL TOOLS FOR GLOBAL PERCEIVED EFFECT, DISABILITY, AND PAIN INTENSITY

Comparison Instrument	A Priori Hypothesis	6 wk		6 mo	
		n	Correlation	n	Correlation
Acute patients					
Change in SBST versus GPE	*r*>0.40				
All patients		185	rho = 0.54	175	rho = 0.62
Low risk		43	rho = 0.43	41	rho = 0.43
Medium risk		60	rho = 0.61	55	rho = 0.74
High risk		82	rho = 0.66	79	rho = 0.77
Change in SBST versus change in RMDQ	*r*>0.60				
All patients		185	*r* = 0.69	175	*r* = 0.75
Low risk		43	*r* =0.42	41	*r* = 0.67
Medium risk		60	*r* = 0.73	55	*r* = 0.75
High risk		82	*r* = 0.76	79	*r* = 0.77
Change in SBST versus change in NPRS	*r*>0.50				
All patients		185	*r* = 0.58	175	*r* = 0.54
Low risk		43	*r* = 0.38	41	*r* = 0.63
Medium risk		60	*r* = 0.64	55	*r* = 0.60
High risk		82	*r* = 0.67	79	*r* = 0.60
Chronic patients					
Change in SBST versus GPE	*r*>0.40				
All patients		148	rho = 0.42	146	rho = 0.52
Low risk		31	rho = 0.27	30	rho = 0.35
Medium risk		52	rho = 0.43	51	rho = 0.66
High risk		65	rho = 0.64	65	rho = 0.64
Change in SBST versus change in RMDQ	*r*>0.60				
All patients		148	*r* = 0.69	146	*r* = 0.76
Low risk		31	*r* = 0.71	30	*r* = 0.73
Medium risk		52	*r* = 0.56	51	*r* = 0.73
High risk		65	*r* = 0.73	65	*r* = 0.76
Change in SBST versus change in NPRS	*r*>0.50				
All patients		148	*r* = 0.56	146	*r* = 0.58
Low risk		31	*r* = 0.46	30	*r* = 0.45
Medium risk		52	*r* = 0.52	51	*r* = 0.53
High risk		65	*r* = 0.65	65	*r* = 0.70

Abbreviations: GPE, Global Perceived Effect scale; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.

APPENDIX

BASELINE, 6-WEEK, AND 6-MONTH VALUES FOR THE CLINICAL OUTCOMES OF ALL PATIENTS AND OF THOSE IN EACH SBST SUBGROUP*

	Acute Patients					Chronic Patients				
Follow-up/Instrument	**n**	**Baseline**	**Follow-up**	**MD†**	**Effect Size‡**	**n**	**Baseline**	**Follow-up**	**MD†**	**Effect Size‡**
6 wk										
SBST (0-9)										
All patients	185	5.4 ± 2.38	3.6 ± 3.26	1.7 ± 2.61	0.72 (0.61, 0.84)	148	5.6 ± 2.26	3.1 ± 2.63	2.5 ± 2.59	1.10 (0.97, 1.23)
Low risk	43	1.9 ± 0.83	1.2 ± 1.86	0.7 ± 1.82	0.84 (0.36, 1.32)	31	2.3 ± 0.91	1.7 ± 1.62	0.6 ± 1.59	0.67 (0.22, 1.13)
Medium risk	60	5.0 ± 1.03	2.8 ± 2.78	2.2 ± 2.55	2.16 (1.71, 2.62)	52	5.3 ± 1.19	2.9 ± 2.25	2.4 ± 2.28	2.03 (1.65, 2.41)
High risk	82	7.4 ± 1.10	5.5 ± 3.09	1.9 ± 2.88	1.71 (1.30, 2.12)	65	7.3 ± 1.36	3.9 ± 2.97	3.4 ± 2.73	2.53 (2.18, 2.89)
RMDQ (0-24)										
All patients	185	15.5 ± 6.03	8.6 ± 7.91	6.9 ± 7.26	1.15 (1.03, 1.28)	148	13.5 ± 5.76	8.9 ± 7.44	4.7 ± 6.44	0.81 (0.68, 0.94)
Low risk	43	8.6 ± 5.52	3.0 ± 4.13	5.6 ± 5.99	1.02 (0.78, 1.25)	31	6.9 ± 4.89	5.0 ± 5.79	1.8 ± 5.53	0.50 (0.12, 0.88)
Medium risk	60	15.8 ± 4.33	7.3 ± 7.99	8.4 ± 7.53	1.95 (1.63, 2.27)	52	13.1 ± 4.58	8.6 ± 6.55	4.5 ± 5.56	0.99 (0.75, 1.23)
High risk	82	19.0 ± 3.95	12.4 ± 8.09	6.6 ± 7.55	1.67 (1.37, 1.97)	65	17.1 ± 4.35	10.9 ± 8.03	6.1 ± 7.13	1.40 (1.11, 1.69)
NPRS (0-10)										
All patients	185	7.7 ± 2.12	4.5 ± 3.59	3.2 ± 3.43	1.51 (1.34, 1.68)	148	7.5 ± 1.72	4.8 ± 2.88	2.7 ± 2.87	1.57 (1.37, 1.76)
Low risk	43	6.3 ± 1.97	2.4 ± 2.98	3.8 ± 3.12	1.94 (1.59, 2.29)	31	6.1 ± 1.89	4.2 ± 2.57	1.9 ± 2.86	1.01 (0.62, 1.40)
Medium risk	60	7.5 ± 2.22	4.0 ± 3.46	3.5 ± 3.70	1.57 (1.57, 1.87)	52	7.2 ± 1.53	4.5 ± 2.52	2.8 ± 4.42	1.80 (1.49, 2.12)
High risk	82	8.6 ± 1.63	5.9 ± 3.42	2.7 ± 3.34	1.64 (1.32, 1.96)	65	8.3 ± 1.25	5.3 ± 3.22	3.0 ± 3.16	2.42 (1.97, 2.86)
6 mo										
SBST (0-9)										
All patients	175	5.3 ± 2.40	2.9 ± 3.11	2.5 ± 2.84	0.86 (0.78, 1.00)	146	5.6 ± 2.72	4.0 ± 3.27	1.5 ± 3.07	0.67 (0.52, 0.83)
Low risk	41	1.9 ± 0.83	1.0 ± 2.04	0.9 ± 1.95	1.06 (0.53, 1.59)	30	2.3 ± 0.92	2.7 ± 2.66	−0.4 ± 2.24	0.40 (−1.10, 1.30)
Medium risk	55	5.0 ± 1.04	2.4 ± 2.78	2.6 ± 2.66	3.11 (2.02, 3.02)	51	5.2 ± 1.19	3.8 ± 2.86	1.4 ± 2.80	1.17 (0.70, 1.64)
High risk	79	7.3 ± 1.13	4.1 ± 3.24	3.2 ± 3.05	2.84 (2.41, 3.27)	65	7.4 ± 1.36	4.8 ± 3.63	2.5 ± 3.14	1.86 (1.45, 2.27)
RMDQ (0-24)										
All patients	175	15.4 ± 6.10	6.6 ± 8.13	8.8 ± 8.30	1.45 (1.30, 1.60)	146	13.5 ± 5.76	8.8 ± 7.98	4.7 ± 7.06	0.82 (0.68, 0.96)
Low risk	41	8.2 ± 5.19	2.1 ± 5.54	6.0 ± 6.51	1.16 (0.87, 1.45)	30	6.9 ± 3.70	4.9 ± 6.77	2.0 ± 1.12	0.54 (0.11, 0.98)
Medium risk	55	15.4 ± 4.22	5.7 ± 7.36	9.7 ± 7.36	2.29 (1.92, 2.66)	51	13.1 ± 4.58	9.1 ± 7.42	4.0 ± 5.78	0.87 (0.62, 1.13)
High risk	79	19.1 ± 3.99	9.4 ± 8.76	9.7 ± 9.06	2.43 (2.06, 2.80)	65	17.1 ± 4.35	10.5 ± 8.09	6.6 ± 7.89	1.51 (1.19, 1.83)
NPRS (0-10)										
All patients	175	7.7 ± 1.13	3.8 ± 5.56	4.0 ± 3.65	1.85 (1.67, 2.04)	146	7.5 ± 1.73	5.7 ± 3.03	1.8 ± 3.03	1.06 (0.85, 1.26)
Low risk	41	6.3 ± 1.98	1.9 ± 2.75	4.4 ± 3.38	2.24 (1.86, 2.62)	30	6.2 ± 1.91	4.6 ± 3.17	1.5 ± 3.42	0.80 (0.33, 1.27)
Medium risk	55	7.5 ± 2.31	3.8 ± 3.65	3.7 ± 3.91	1.60 (1.27, 1.92)	51	7.2 ± 1.54	5.7 ± 2.46	1.5 ± 2.49	0.97 (0.65, 1.29)
High risk	79	8.7 ± 1.59	4.8 ± 3.50	3.9 ± 3.61	2.45 (2.09, 2.81)	65	8.3 ± 1.26	6.1 ± 3.30	2.2 ± 3.23	1.76 (1.30, 2.21)

Abbreviations: MD, mean difference; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.

**Values are mean ± SD unless otherwise indicated.*

†Change score from baseline to follow-up.

‡Values in parentheses are 84% confidence interval.

APPENDIX

ESTIMATED VALUES COMPARING SBST VARIATION WITH THE CLINICAL OUTCOMES OF GLOBAL PERCEIVED EFFECT, DISABILITY, AND PAIN SEVERITY AT FOLLOW-UP

	Acute Patients			Chronic Patients		
Comparison Instrument	**n**	**AUC***	***P* Value**	**n**	**AUC***	***P* Value**
Baseline to 6 wk						
Change in SBST by GPE cutoff score†						
All patients	185	0.71 (0.64, 0.78)	<.001	148	0.70 (0.61, 0.78)	<.001
Low risk	43	0.68 (0.49, 0.86)	.58	31	0.62 (0.42, 0.82)	.26
Medium risk	60	0.77 (0.66, 0.90)	<.001	52	0.72 (0.58, 0.85)	.01
High risk	82	0.75 (0.64, 0.87)	<.001	65	0.80 (0.64, 0.90)	<.001
Change in SBST by RMDQ cutoff score‡						
All patients	185	0.77 (0.71, 0.84)	<.001	148	0.73 (0.64, 0.81)	<.001
Low risk	43	0.79 (0.57, 1.00)	.01	31	0.86 (0.72, 0.99)	<.001
Medium risk	60	0.88 (0.79, 0.96)	<.001	52	0.82 (0.71, 0.93)	<.001
High risk	82	0.88 (0.79, 0.97)	<.001	65	0.84 (0.74, 0.93)	<.001
Change in SBST by NPRS cutoff score§						
All patients	185	0.81 (0.75, 0.87)	<.001	148	0.79 (0.72, 0.87)	<.001
Low risk	43	0.73 (0.58, 0.88)	.01	31	0.68 (0.43, 0.93)	.18
Medium risk	60	0.85 (0.75, 0.94)	<.001	52	0.68 (0.53, 0.84)	.04
High risk	82	0.86 (0.78, 0.95)	<.001	65	0.86 (0.78, 0.95)	<.001
Baseline to 6 mo						
Change in SBST by GPE cutoff score†						
All patients	175	0.83 (0.77, 0.89)	<.001	146	0.77 (0.69, 0.86)	<.001
Low risk	41	0.77 (0.59, 0.95)	<.001	30	0.60 (0.40, 0.81)	.35
Medium risk	55	0.93 (0.85, 1.00)	<.001	51	0.88 (0.76, 0.99)	<.001
High risk	79	0.90 (0.83, 0.96)	<.001	65	0.86 (0.75, 0.96)	<.001
Change in SBST by RMDQ cutoff score‡						
All patients	175	0.82 (0.76, 0.89)	<.001	146	0.84 (0.77, 0.90)	<.001
Low risk	41	1.00 (1.00, 1.00)	<.001	30	0.95 (0.88, 1.00)	<.001
Medium risk	55	0.86 (0.74, 0.99)	<.001	51	0.94 (0.88, 1.00)	<.001
High risk	79	0.93 (0.88, 0.98)	<.001	65	0.94 (0.87, 1.00)	<.001
Change in SBST by NPRS cutoff score§						
All patients	175	0.77 (0.69, 0.84)	<.001	146	0.80 (0.72, 0.89)	<.001
Low risk	41	0.89 (0.76, 1.00)	<.001	30	0.68 (0.47, 0.88)	.16
Medium risk	55	0.80 (0.68, 0.93)	<.001	51	0.71 (0.54, 0.88)	.04
High risk	79	0.81 (0.72, 0.91)	<.001	65	0.93 (0.87, 0.99)	<.001

Abbreviations: AUC, area under the curve; GPE, Global Perceived Effect scale; NPRS, numeric pain-rating scale; RMDQ, Roland-Morris Disability Questionnaire; SBST, Subgroups for Targeted Treatment Back Screening Tool.
**Values in parentheses are 95% confidence interval.*
†Patients who scored 3 or more points (−5 to +5) were classified as "improved."
‡Patients who scored less than 7 points (0-24) were classified as "improved."
§Patients with improvement of more than 30% between baseline and follow-up were classified as "improved."

EVAN O. NELSON, DPT[1] • **MICHAEL RYAN,** PhD[2]
ERIN AUFDERHEIDE, MD[3] • **BRYAN HEIDERSCHEIT,** PT, PhD, FAPTA[4]

Development of the University of Wisconsin Running Injury and Recovery Index

BACKGROUND: Runners experience a high proportion of overuse injuries, with extended recovery periods involving a gradual, progressive return to preinjury status. A running-specific patient-reported outcome (PRO) measure does not exist, and a questionnaire assessing critical elements of runners' recovery processes may have excellent psychometric properties.

OBJECTIVES: To develop a valid, reliable, and responsive evaluative PRO measure to assess longitudinal change in running ability after running-related injury (RRI) for clinical practice and research applications.

METHODS: Self-identified runners and selected experts participated in an iterative, 6-step development process of the University of Wisconsin Running Injury and Recovery Index (UWRI) in this longitudinal clinical measurement study. Content-related validity was assessed using open comments. Reproducibility was assessed using Cronbach's alpha, the intraclass correlation coefficient (ICC), and standard error of measurement (SEM). An anchor-based construct validity assessment measured the association between the change in UWRI score and global rating of change (GROC). Responsiveness assessments included floor and ceiling effects.

RESULTS: The 9-item UWRI assesses running ability following an RRI, with the maximum score of 36 indicating a return to preinjury running ability. The UWRI demonstrated acceptable internal consistency (α = .82), test-retest reliability (ICC = 0.93), and SEM (1.47 points). Change in UWRI score was moderately correlated with the GROC (r = 0.61; 95% confidence interval: 0.4, 0.76). Floor and ceiling effects were absent. Completion required 3 minutes 15 seconds.

CONCLUSION: The UWRI is a reliable PRO measure and is responsive to changes in running function following an RRI, with minimal administrative burden.

LEVEL OF EVIDENCE: Therapy, level 2c.
J Orthop Sports Phys Ther 2019;49(10):751-760. Epub 3 Aug 2019. doi:10.2519/jospt.2019.8868

KEY WORDS: *patient-reported outcome measure, psychometric assessment, running injury*

Running-related injuries (RRIs) include a heterogeneous collection of musculoskeletal injuries whose pain and physical limitations cause distance runners to miss competitive events, reduce training time, and receive medical care.[22,30,48,52] To explain the RRI heterogeneity and variable precipitating factors, a conceptual framework has been proposed that incorporates tissue-specific load capacity and running-imposed load.[4] Additionally, this conceptual framework acknowledges that stress, fatigue, or psychological factors may contribute to an individual's injury susceptibility. Athletic injuries, including RRIs, produce an emotional response, influenced by internal factors (eg, fear, confidence, motivation) or external factors (eg, coaches, teammates, parents, the competition schedule), that continues through the recovery process.[1,15,45,57] Providers balance numerous considerations during rehabilitation, and the recovery from such injuries may not demonstrate equivalent rates of improvement for the different types of running (eg, daily runs, long runs, speed training, threshold intervals, races).[2,42] Clinical assessments of running ability must account for an injury's disparate impact on different types of running and the dynamic nature of the multiple factors influencing the recovery process.

Patient-reported outcome (PRO) measures are the principal tool used by physical therapists to quantify limitations in functional ability and determine change in patient status due to clinical care.[19,24] It is essential that PRO measures demonstrate efficiency and strong measurement properties, because these assessments influence subsequent clinical care decisions.[41] A running-specific PRO measure does not currently exist. In the absence of a running-specific measure,

[1]Department of Family Medicine and Community Health, University of Wisconsin-Madison, Madison, WI. [2]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, Canada. [3]Section of Emergency Medicine, Baylor College of Medicine, Houston, TX. [4]Department of Orthopedics and Rehabilitation; Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI. This study was approved by the Health Sciences and the Education and Social/Behavioral Sciences Institutional Review Boards at the University of Wisconsin-Madison. This study was funded via the University of Wisconsin Institute for Clinical and Translational Research (Clinical and Translational Science Award program) through the US National Institutes of Health National Center for Advancing Translational Sciences (grant UL1TR000427). The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Bryan Heiderscheit, University of Wisconsin, School of Medicine and Public Health, Department of Orthopedics and Rehabilitation, 1685 Highland Avenue, MFCB 6136, Madison, WI 53705. E-mail: heiderscheit@ortho.wisc.edu

PRO measures specific to a body region (eg, knee, hip) or condition (eg, Achilles tendinopathy, patellofemoral pain) are used to assess functional limitations imposed by RRIs.[5,21,37,40,51,58] The psychometric properties of these PRO measures have not been assessed in a running population, and these measures may have limited evaluative capability, because the majority of items assess low-demand activities or fail to assess the multiple internal and external factors that influence recovery following an RRI.[16,35,36] A running-specific PRO measure could potentially create a valid, reliable, and responsive assessment of running ability to measure RRI severity in prospective clinical studies and standardize clinical effectiveness evaluations in practice and research.[11,29,47]

The purpose of this study was to develop a valid, reliable, and responsive evaluative PRO measure to assess longitudinal change in running ability after an RRI for clinical practice and research applications.

METHODS

The development of this new instrument consisted of 6 steps: (1) item generation, (2) item reduction, (3) item clarification and content validation, (4) preliminary evaluation and revision, (5) reliability assessment, and (6) psychometric assessment. For each step, an RRI was defined as a musculoskeletal problem resulting from running that required the individual to prematurely stop running while training or during a competitive event or to miss a training session or competitive event.[31,56] Rather than employ strict criteria, research participants were allowed to self-identify as a runner in an attempt to represent the diversity within the target population. Runners who suffered at least 1 RRI were recruited from running clubs, groups, and retail stores in Wisconsin and the surrounding states (steps 1-3 and 5). We recruited physicians, physical therapists, and running research experts with at least 5 years of experience conducting patient care in a specialized running clinic or publishing articles about RRIs in peer-reviewed journals through individual solicitation (step 3). Runners experiencing a current RRI were recruited from the UW Health Sports Medicine Runners Clinic (steps 4 and 6). The UW Health Sports Medicine Runners Clinic comprises physical therapists who are running experts and who use a shared decision-making process when performing running evaluations, including video gait assessments, and implementing multimodal intervention plans using a combination of therapeutic exercise, running gait modification, training recommendations, and footwear recommendations.[43] Patients frequently receive concurrent physical therapy care from a separate provider in the UW Health Sports Rehabilitation Clinic.

Questionnaire Development

Semi-structured telephone interviews conducted by a single interviewer (E.A.) explored the spectrum of the RRI experience and investigated how etiological factors influence the recovery process. Questions surveyed the injury description and associated pain, activities while injured, training and racing following injury, emotional impact, and recovery and expectations, and included an open-ended request for additional information. All responses were recorded, transcribed, and coded before generating partially overlapping potential items. Using an ordinal scale, participants assessed the relevance (0 is no, 1 is yes) and importance (1 is not important, 5 is very important) of each potential item. Items were ranked by the importance product, calculated as the sum of the importance scores divided by the sum of the relevance scores across participants.[26] Importance product ranking and expert review reduced the item pool to develop a draft questionnaire. Written comments related to the comprehension, clarity, and ease of responding to the draft questionnaire were sought during item clarification. To assess content validity, participants were asked whether the draft questionnaire comprehensively assessed the RRI recovery process.[55]

Psychometric Assessment

The response distribution for each item was analyzed. Pearson correlation coefficients were calculated to determine how each item correlated with the total score, as well as the impact of removing any single item. An exploratory factor analysis using maximum likelihood with promax rotation was performed.[55] To include systematic sources of error, the intraclass correlation coefficient (ICC) for agreement evaluated test-retest reliability between paired assessments.[32] Though response memory was possible between assessments, physiologic or neuromuscular change was unlikely, and longer response periods would have increased the likelihood that a change in running ability might occur. The standard error of measurement (SEM) for agreement was calculated to evaluate score stability between administrations, including systematic error, as $SD \times \sqrt{1 - ICC}$.[55] The coefficient (Cronbach's) alpha correlation was calculated to evaluate internal consistency for the collective tool and each identified factor.[6,55] Anchor-based validation assessments are a common approach to evaluate patient-reported change in measures that cannot be compared to a gold standard.[55] In the absence of a universal assessment of running ability, multiple comparison measures were used throughout this project. The Lower Extremity Functional Scale (LEFS) is a 20-item PRO measure, with the 80-point maximum composite score indicating higher functional ability.[5] The numeric pain-rating scale (NPRS) is an 11-point (0-10) ordinal response scale with end-point descriptors of "no pain at all" and "worst possible pain."[9] The global rating of change (GROC) is an 11-point (−5 to +5) ordinal response scale with end-point descriptors of "very much worse" and "completely recovered."[23] The strength of the associations between outcome measures was used to evaluate construct validity using the Pearson correlation coefficient.

Correlation coefficients were considered good ($r>0.75$), moderate ($0.50<r\leq0.75$), fair ($0.25<r\leq0.50$), and no association ($r\leq0.25$).[44] Responsiveness was assessed as the average change in PRO score reported by individuals reporting improvement (GROC score of +4 or +5) and slight improvement (GROC score of +2 or +3), and was compared using a between-subject and within-subject approach, respectively.[9,10] Statistical significance was set at $P<.05$. All analyses were performed using the psych, irr, psychometric, and xlsx packages in R software (R Foundation for Statistical Computing, Vienna, Austria).[13,14,17,46,49]

The testing protocols were approved by the Health Sciences and the Education and Social/Behavioral Sciences Institutional Review Boards at the University of Wisconsin-Madison. Informed consent was obtained, and the rights of participants were protected in accordance with institutional policies.

RESULTS

Step 1: Item Generation

Sixteen runners who had suffered at least 1 RRI agreed to be interviewed. Five participants were elite or professional runners, and most had experience running on a team with an official coach at the high school (14/16), college (13/16), or postcollegiate (5/16) level. The authors conducted axial coding using interview question categories to identify how runners assess recovery. Data saturation occurred when interviews of recreational runners did not produce new themes. Forty-two potential items were generated.

Step 2: Item Reduction

A unique sample of 79 runners who had suffered at least 1 RRI completed paper surveys containing 42 potential items and importance product questions (ie, relevance and importance). Items were ranked by importance product, and items above the median were consistent with the primary factors identified in step 1. Items requiring skilled assessment were removed from the item pool. The RRI's impact in daily life was a critical recovery component identified in step 1; however, all items assessing specific nonrunning activities (eg, squatting, walking, etc) were below the median and consolidated into a single item. Items assessing cross-training or physical fitness were removed because they do not directly assess running ability. The conceptual framework incorporating physical symptoms, running performance, and psychological responses was maintained when content consolidation reduced the pool to 12 items. Individual interviews with 6 experts affirmed the relevance of 9 items, but suggested the removal of 3 items assessing interval training, event participation, and individual running goals, as these were not generally applicable or unlikely to assess the spectrum of recovery. Expert panelists provided consensus agreement that the remaining items comprehensively assessed clinically relevant components of postinjury running ability. The beta version of the 9-item University of Wisconsin Running Injury and Recovery Index (UWRI beta) included a 7-point numerical response, with unique written end-point descriptions for each item's response scale.[33] The instructions specified a 7-day recall period, and a hard-stop question restricted the completion of items 4 through 9 unless the individual ran during the preceding 7 days. A question to enforce the recall period was included because an individual may experience improvement when the medical management team advises abstinence from running.

Step 3: Item Clarification and Content Validation

A unique sample of 31 runners who had suffered at least 1 RRI provided comments regarding UWRI beta item clarity and content validity using an electronic form. Three research team members (E.N., M.R., G.T.) who were runners performed triangulation (by analyst) using open coding of all comments until data saturation. The UWRI beta was clear, easy to understand, and comprehensively assessed the recovery of running ability post injury. Participants suggested a 5-point response scale because it would be difficult to differentiate perceived changes in running status on the 7-point scale. A 5-point ordinal response system using checkboxes and written anchors at the midpoint and end points was created for each question.[20] Editorial changes that did not alter the item concept were made to 5 items, because greater than 10% of the sample commented on item syntax. Three participants recommended assessing concepts identified in step 1 that were below the median in step 2. Each item was scored from 0 to 4, with a maximum score of 36 indicating no deficit in running ability. If an individual had not run during the 7 days preceding completion of the UWRI beta, items 4 through 9 were scored as 0.

Step 4: Preliminary Evaluation and Revision

Individuals seeking physical therapy care at the UW Health Sports Medicine Runners Clinic for an RRI were invited to participate in a prospective, preliminary assessment of the UWRI beta. Inclusion criteria in this phase included being at least 14 years of age, able to read and write English, and of generally good health. Exclusion criteria included surgery for the injury in the past 12 months, rheumatologic disease, systemic connective tissue disorders, or clinical diagnosis of depression. A priori sample-size calculation indicated that 23 individuals would be needed to detect a correlation greater than zero, assuming $\alpha = .05$, $\beta = .20$, and $r = 0.55$.

Participants completed the UWRI beta, LEFS, and NPRS at the initial visit, prior to receiving physical therapy care, as is common practice in this population.[5,8,9] To assess reliability, all participants were asked to complete the UWRI beta electronically 48 to 72 hours following the initial assessment. The UWRI beta, LEFS, NPRS, and GROC were electronically

completed 8 weeks following the initial assessment, because measurable functional change is likely to occur in 8 weeks.[7,12,39]

Thirty individuals (20 female) agreed to participate and entered the study following the informed-consent process (**FIGURE**). One individual withdrew for unknown reasons before completing any assessments, and 13 individuals were lost to follow-up, as they did not respond to requests to complete the short-term or 8-week assessment. Baseline characteristics were not different in those lost to follow-up. The age range of the participants was large and a variety of chronic RRIs reduced running volume (**TABLE 1**). The median UWRI beta score was 14 (range, 0-27) at the initial assessment and 24 (range, 6-34) at the final assessment, creating a statistically significant change in the Wilcoxon signed-rank test of paired data: 10 (range, −17 to 27) points (P = .008). The median LEFS score was 69 (range, 42-79) at the initial assessment and 77 (range, 46-80) at the final assessment, equating to a median change of 5.5 (range, −20 to 24) points that was not statistically significant (P = .08). The median NPRS score was 3 (range, 0-7) at the initial evaluation, 2 (range, 0-8) at the final assessment, and had a median change of −1 (range, −4 to 3) points that was not significant (P = .37). The median GROC score was 3 (range, −5 to 5). A good correlation was observed between the change in UWRI beta score over 8 weeks and the GROC (r = 0.75; 95% confidence interval [CI]: 0.39, 0.91; P = .001). The UWRI beta and LEFS demonstrated moderate correlation for change over the course of the study (r = 0.55; 95% CI: 0.05, 0.82; P = .03), as well as at the 8-week assessment (r = 0.73; 95% CI: 0.35, 0.9; P = .002). A correlation was not observed between the UWRI beta and LEFS at the initial assessment (r = 0.04; 95% CI: −0.33, 0.4; P = .82). A correlation between the UWRI beta and NPRS was not observed at any assessment. Individuals reporting significant clinical improvement on the GROC (+4 or +5) demonstrated a median UWRI change of 11.5 (range, 9-20) points and a median LEFS change of 5.5 (range, −2 to 15) points.[27] Three minutes 15 seconds were required to complete the UWRI beta.

Eighteen individuals completed the UWRI beta at the initial evaluation and at the day 2 assessment. The difference in their mean ± SD scores was 2.6 ± 5.7 and not significant (P = .07). The UWRI beta scores were moderately correlated between the 2 assessments (ICC = 0.56; 95% CI: 0.16, 0.81; P = .004). Cronbach's alpha for reproducibility showed acceptable internal consistency (α = .75; 95% CI: .28, .93). The preliminary assessment revealed that the test-retest re-

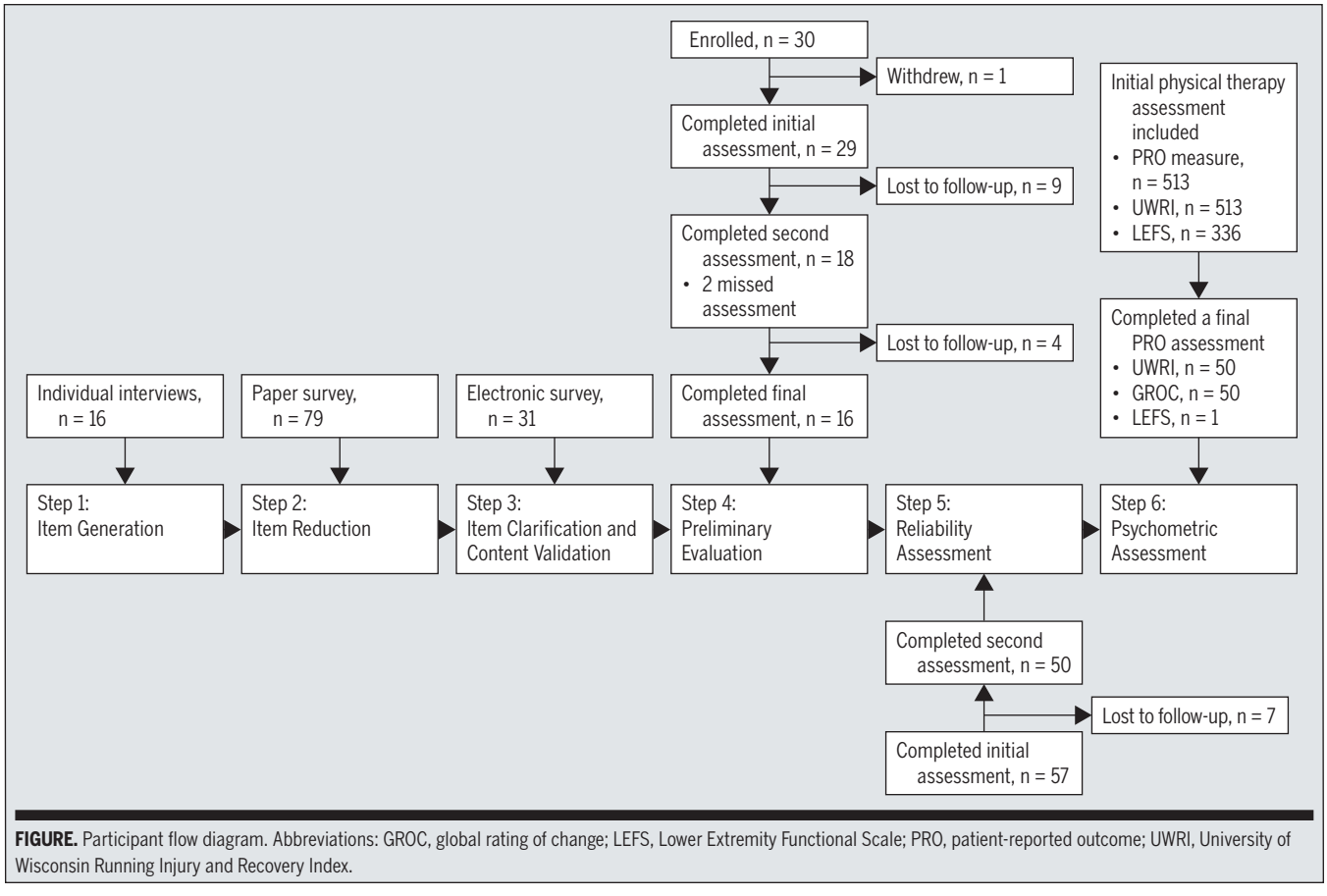


FIGURE. Participant flow diagram. Abbreviations: GROC, global rating of change; LEFS, Lower Extremity Functional Scale; PRO, patient-reported outcome; UWRI, University of Wisconsin Running Injury and Recovery Index.

liability of the UWRI beta was adversely affected, because items 4 through 9 were only presented if the person had run in the past 7 days. The UWRI beta was revised to create the UWRI, presenting all items on each administration (**TABLE 2**; **APPENDIX,** available at www.jospt.org). Written anchors accompanied each response option, and participants could select "unable to run" if the condition severity or medical recommendation prevented running.

Step 5: Reliability Assessment

Fifty-seven English-literate, self-identified adult runners in generally good health and with a single, current RRI participated in an evaluation of the UWRI test-retest reliability using a web-based survey. The sole exclusion criterion was having surgery at the injury site in the past 12 months. The web-based survey dissociated this assessment from clinical care because common RRI interventions, including gait retraining or educational interventions, may have immediate effects that influence perceived running ability.[3,34] Fifty participants reported demographic information, injury duration and location, and current and preinjury running volume, and completed the UWRI electronically on 2 occasions separated by 24 to 72 hours (**TABLE 1**). The UWRI test-retest reliability was excellent, with an ICC of 0.93 (95% CI: 0.89, 0.96; $P<.001$). Scores on the UWRI were not different between the 2 administrations ($P = .12$). The SEM was 1.47 points.

Step 6: Psychometric Assessment

Participants Retrospective chart review of 513 patients from November 2012 to October 2017 was used to conduct a psychometric assessment. The UWRI and LEFS were administered according to standard clinical procedure during the initial encounter, and data on age, sex, running experience, and current running volume were collected (**TABLE 1**). The UWRI and GROC were completed electronically at least 8 weeks after the initial encounter. Follow-up data are limited to 50 patients because a systematic process was not implemented until May 2017.

Item Analysis Responses to each UWRI item were distributed across the range of potential responses, and the composite UWRI score was centered at the middle value (**TABLE 3**). All items were correlated with the composite UWRI score at the initial ($r = 0.51$-0.72) and final ($r = 0.55$-0.86) assessments. Each item was responsive to change, and the item-specific

TABLE 1 Participant Demographics*

	Step 4: Preliminary Evaluation	Step 5: Reliability Assessment	Step 6: Psychometric Assessment
n	30	50	513
Age, y	36.9 ± 11.0 (15-57)	33.4 ± 8.4 (20-54)	35.4 ± 12.9 (12-74)
Sex, n (%)			
Female	20 (66.7)	32 (64)	296 (57.7)
Male	10 (33.3)	18 (36)	217 (42.3)
Running experience, y	2.4 ± 1.2	9.9 ± 7.2	11.5 ± 10.7
Symptom duration, mo	8.8 ± 9.9	2.8 ± 1.7	NA
Current running volume, km/wk	21.9 ± 26.1	31.7 ± 26.3	27.6 ± 21.1
Preinjury running volume, km/wk	37.2 ± 24.6	57.7 ± 35.4	NA
Injury location, n (%)	40 (100)	50 (100)	NA
Lumbopelvic	4 (10)	4 (8)	
Hip or thigh	6 (15)	10 (20)	
Knee	14 (35)	8 (16)	
Leg or calf	6 (15)	4 (8)	
Ankle	2 (5)	7 (14)	
Foot	8 (20)	13 (26)	
Other	0 (0)	4 (8)	

Abbreviation: NA, not available.
**Values are mean ± SD or mean ± SD (range) unless otherwise indicated.*

TABLE 2 Items Included in the University of Wisconsin Running Injury and Recovery Index

Item	Included in Beta Version	Included in Final Version
How does your running injury impact your ability to perform daily activities?	Yes	Yes
How frustrated are you by your running injury?	Yes	Yes
How much recovery have you made from your running injury?	Yes	Yes
Have you run in the past 7 days?	Yes	No
How much pain do you experience while running?	Yes	Yes
How much pain do you experience during the 24 hours following a run?	Yes	Yes
How has your weekly mileage or weekly running time changed as a result of your injury?	Yes	Yes
How has the distance of your longest weekly run changed as a result of your injury?	Yes	Yes
How has your running pace or speed changed as a result of your injury?	Yes	Yes
How does your injury affect your confidence to increase the duration or intensity of your running?	Yes	Yes

change was correlated with the change in composite UWRI score (r = 0.63-0.85).

Internal Consistency The exploratory factor analysis found that 51% of the variance was explained by 2 moderately correlated factors (r = 0.46). Factor 1 (items 6-8) explained 26% of the variance and factor 2 explained 25% of the variance (**TABLE 4**). Cronbach's alpha for the UWRI composite score was .82 (**TABLE 5**).

Construct Validity In addition to the initial UWRI, 50 people completed the GROC and a second evaluation of the UWRI. A statistically significant, moderate correlation was observed between the change in UWRI score and the patient-reported change on the GROC (r = 0.61; 95% CI: 0.4, 0.76; P<.001). A paired-samples t test showed a statistically significant difference between the initial and final UWRI scores (P<.001). The correlation of the changes in UWRI and LEFS scores was not calculated due to the limited quantity of paired data.

Responsiveness Individuals reporting significant clinical improvement demonstrated greater change in UWRI score than those reporting slight clinical improvement (P = .01) (**TABLE 5**).[27] At the initial assessment, the maximum score was achieved in 1/451 (0.21%) and 44/354 (12.43%) participants completing the UWRI and LEFS, respectively, whereas the minimum UWRI or LEFS score was not achieved. At the initial evaluation, the average LEFS score was 70.75 out of 80 possible points.

TABLE 3 Item Analysis of the University of Wisconsin Running Injury and Recovery Index From Step 6: Psychometric Assessment*

Item	Initial	Final	Change
1	3.12 ± 0.89	3.80 ± 0.63	0.16 ± 1.78
2	1.51 ± 1.10	2.50 ± 1.18	0.76 ± 1.23
3	1.85 ± 1.04	2.70 ± 0.82	1.04 ± 1.14
4	2.39 ± 1.00	3.33 ± 0.50	0.60 ± 0.91
5	2.53 ± 0.95	3.00 ± 0.67	0.52 ± 1.05
6	1.44 ± 1.13	2.30 ± 1.16	0.92 ± 1.12
7	1.50 ± 1.21	2.20 ± 1.23	1.16 ± 1.21
8	1.97 ± 1.26	3.10 ± 0.74	0.60 ± 1.22
9	1.23 ± 1.07	2.60 ± 1.26	1.44 ± 1.63
Score	17.71 ± 6.14	27.07 ± 6.63	7.37 ± 7.84

**Values are mean ± SD.*

TABLE 4 Factor Loading of University of Wisconsin Running Injury and Recovery Index Items From Step 6: Psychometric Assessment Initial Evaluations

Item	Factor 1	Factor 2
1	−0.04	0.53
2	0.25	0.51
3	−0.21	0.71
4	0.03	0.66
5	−0.08	0.69
6	0.95	−0.05
7	0.95	−0.06
8	0.68	0.10
9	0.09	0.55

DISCUSSION

The purpose of this study was to develop a valid, reliable, and responsive evaluative PRO measure to assess longitudinal change in running ability after an RRI for clinical practice and research applications. After sustaining an RRI, runners monitor symptoms and manage the dynamic psychological response during physiologic healing and the progressive physiologic adaptation to restore preinjury performance. The 9-item UWRI accounts for how runners assess running ability through 2 components, running progression and symptom surveillance. Running progression (items 6-8) involves assessing different aspects of running through weekly volume, long-run distance, and running pace, which are unique components in the load application framework proposed by Bertelsen and colleagues.[4] Symptom surveillance incorporates how runners monitor symptoms while running (item 4), assess training response (items 1, 3, and 5), and describe the psychological response (items 2 and 9). Though different levels of running performance and experience create disparity in what is considered normal running, runners monitor symptoms to make informed training decisions.

Elite and competitive runners make a significant investment in their running careers and described elaborate, individualized schema for symptom monitoring and profound psychological responses following injury. Nonelite and recreational runners described succinct symptom monitoring processes and greater willingness to abstain from running. Commonalities in symptom surveillance themes were illuminated during item reduction. Predictably, run-

ners monitor symptoms while running, but runners of all performance levels frequently monitor symptoms between running sessions to determine readiness for future training. Psychological responses are inherently individualized; however, frustration is a common emotion that runners experience during recovery. The progressive increase in an athlete's confidence following injury is known to play an important role in the successful return to sport, and runners consistently reported that their running ability was fully restored when they established the confidence to train without fear of reinjury.[25,59] The UWRI is a parsimonious, clinically relevant measure reflecting runners' self-assessment of running ability and how they subsequently make decisions during the RRI recovery process.[38]

The iterative development process incorporated the target population and produced a sample with injury and runner characteristics that are consistent with other RRI studies.[28,54,56] Including the target population enhanced the UWRI's content validity and enabled data triangulation by source and method to ensure that the UWRI items accurately and comprehensively represent the recovery of distance-running ability following injury.[55] Participants and expert reviewers concluded that the UWRI was composed of comprehensible items that are likely to measure the construct due to their detailed, running-specific nature, while acknowledging the psychological factors associated with recovery. The 5-item response structure was chosen because it was preferred by the target population and is less burdensome to respondents than visual analog scales.[20] Participant responses showed that the items differentiate levels of function when used in a clinical setting and are responsive to change over time. Clinical implementation is enhanced by the low administration burden, short completion time, and simple scoring.

During preliminary testing (step 4), the UWRI beta test-retest reliability was adversely affected by a hard-stop question that blocked access to items 4 through 9 when the participant had not run during the prior 7 days. Excellent test-retest reliability was observed following modifications to present all items during each administration. The error associated with a single score was less than 2 points in the total UWRI score. It was logical to create a single composite score because the 2 correlated subscales collectively reflect the typical process runners use to gauge their running ability during recovery, which is the primary intent of the UWRI. The internal consistency of the UWRI, and that of each component factor, was within the accepted range during all clinical evaluations, indicating that item correlation was present without item redundancy.[6]

Construct-related validity was evaluated by comparing the change observed in the UWRI score with measures commonly used to assess RRIs.[5,55] The UWRI demonstrated the ability to assess changes in self-reported running ability, and the change in the UWRI total score was correlated with the changes measured by the GROC. The association between

TABLE 5 Psychometric Properties of the University of Wisconsin Running Injury and Recovery Index*

	Step 4: Preliminary Evaluation	Step 5: Reliability Assessment	Step 6: Psychometric Assessment
Item analysis†	r = 0.36-0.81	r = 0.47-0.82	r = 0.43-0.71
Internal consistency	α = .75 (.28, .93)	α = .83 (.76, .89)	α = .82 (.80, .84)
Factor 1	α = .81 (.71, .93)	α = .88 (.81, .94)	α = .89 (.87, .91)
Factor 2	α = .72 (.56, .88)	α = .74 (.62, .85)	α = .78 (.74, .80)
Reproducibility			
ICC (agreement)	0.56 (0.16, 0.81)	0.93 (0.89, 0.96)	NA
SEM (agreement)	4.46	1.47	NA
Construct validity			
UWRI change – GROC	r = 0.75 (0.39, 0.91)	NA	r = 0.61 (0.4, 0.76)
Responsiveness			
UWRI: significant improvement‡	13.33 ± 4.93	NA	9.38 ± 5.78
UWRI: slight improvement‡	10.42 ± 10.31	NA	4.5 ± 7.55
Ceiling effect§	UWRI, 0%; LEFS, 25%	UWRI, 0%	UWRI, 0.22%; LEFS, 12.43%
Floor effect§	UWRI, 3.45%; LEFS, 0%	UWRI, 0%	UWRI, 0%: LEFS, 0%

Abbreviations: GROC, global rating of change; ICC, intraclass correlation coefficient; LEFS, Lower Extremity Functional Scale; NA, not available; SEM, standard error of measurement; UWRI, University of Wisconsin Running Injury and Recovery Index.
**Values in parentheses are 95% confidence interval.*
†Range of item response (total score).
‡Values are mean ± SD. Significant improvement, GROC of +4 or +5; slight improvement, GROC of +2 or +3.
§Ceiling and floor effects are calculated as the percent of respondents achieving the maximum and minimum composite scores, respectively.

the UWRI and LEFS varied throughout the steps of this study, and the LEFS may not have the capacity to respond to functional improvements, because the initial evaluation scores were skewed toward the maximum potential score. The observed moderate associations were anticipated because the UWRI was designed to be a more critical assessment of running function than other PRO measures with greater temporal stability than the GROC.[18,53]

The UWRI was responsive to change, because individuals reporting clinically meaningful improvements on the GROC scale also reported greater change in UWRI scores than those reporting slight or no change.[50,55] Heightened responsiveness is further supported by the absence of UWRI ceiling or floor effects. The LEFS demonstrated a ceiling effect during step 4 and approached this threshold in the large data set used in step 6.[55] Responsiveness of the LEFS is further questioned because the majority of respondents, at the initial evaluation, had a score that was too high to achieve the 9-point minimum detectable change.[5] The iterative process used to develop the UWRI enhanced the specificity to running and produced a PRO measure capable of measuring changes in running ability regardless of body region or type of injury.

The strengths of this study include the repeated integration of the target population during item development, psychometric assessment of the UWRI using participants from the target population, and the involvement of multiple stakeholders in the development process. The sample population included a diverse representation of age, experience, running volume, speed, and injury. There are several limitations of this study, including participants lost to follow-up during prospective steps and the use of a retrospective design to conduct a psychometric assessment with clinical data. Further psychometric evaluation is warranted to evaluate the construct validity in combination with measures of running ability and other PRO measures. Clinically useful assessments of responsiveness, including minimum detectable change and minimal clinically important difference, may be determined in future studies. Future studies should also provide a better understanding of how the UWRI measures change in running ability for different performance abilities or injury types.

CONCLUSION

The UWRI is a reliable evaluative measure assessing running ability following an RRI, with minimal administrative burden. Content and construct-related validity assessments indicate that the UWRI is a more exacting evaluation of running ability than are other PRO measures currently used in this population. Continued psychometric evaluation in prospective clinical studies is warranted, because this novel measure of running ability is responsive to patient-perceived functional change. ◉

KEY POINTS

FINDINGS: The University of Wisconsin Running Injury and Recovery Index (UWRI) is a novel, running-specific patient-reported outcome measure reflecting how runners assess their own running ability while recovering from a running-related injury. The UWRI is a reliable assessment that can measure the change in running ability during an episode of care.

IMPLICATIONS: The UWRI systematically assesses postinjury distance-running ability with very little burden on providers or patients.

CAUTION: Construct-related validity and responsiveness require additional prospective, psychometric assessment conducted in combination with clinical care.

ACKNOWLEDGMENTS: *The authors thank Grace Trabulsi, DPT for her assistance during data collection and analysis.*

REFERENCES

1. Ardern CL, Taylor NF, Feller JA, Webster KE. A systematic review of the psychological factors associated with returning to sport following injury. *Br J Sports Med*. 2013;47:1120-1126. https://doi.org/10.1136/bjsports-2012-091203
2. Barton CJ. Managing RISK when treating the injured runner with running retraining, load management and exercise therapy. *Phys Ther Sport*. 2018;29:79-83. https://doi.org/10.1016/j.ptsp.2017.10.002
3. Barton CJ, Bonanno DR, Carr J, et al. Running retraining to treat lower limb injuries: a mixed-methods study of current evidence synthesised with expert opinion. *Br J Sports Med*. 2016;50:513-526. https://doi.org/10.1136/bjsports-2015-095278
4. Bertelsen ML, Hulme A, Petersen J, et al. A framework for the etiology of running-related injuries. *Scand J Med Sci Sports*. 2017;27:1170-1180. https://doi.org/10.1111/sms.12883
5. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. *Phys Ther*. 1999;79:371-383. https://doi.org/10.1093/ptj/79.4.371
6. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ*. 1997;314:572. https://doi.org/10.1136/bmj.314.7080.572
7. Carcia CR, Martin RL, Houck J, Wukich DK. Achilles pain, stiffness, and muscle power deficits: Achilles tendinitis. *J Orthop Sports Phys Ther*. 2010;40:A1-A26. https://doi.org/10.2519/jospt.2010.0305
8. Chatman AB, Hyams SP, Neel JM, et al. The Patient-Specific Functional Scale: measurement properties in patients with knee dysfunction. *Phys Ther*. 1997;77:820-829. https://doi.org/10.1093/ptj/77.8.820
9. 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. https://doi.org/10.1097/01.brs.0000164099.92112.29
10. Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007;7:541-546. https://doi.org/10.1016/j.spinee.2007.01.008
11. Damsted C, Parner ET, Sørensen H, Malisoux L, Nielsen RO. Design of ProjectRun21: a 14-week prospective cohort study of the influence of running experience and running pace on running-related injury in half-marathoners. *Inj Epidemiol*. 2017;4:30. https://doi.org/10.1186/s40621-017-0124-9
12. Dolak KL, Silkman C, Medina McKeon J, Hosey RG, Lattermann C, Uhl TL. Hip strengthening prior to functional exercises reduces pain sooner than quadriceps strengthening in females with patellofemoral pain syndrome: a randomized clinical trial. *J Orthop Sports Phys*

Ther. 2011;41:560-570. https://doi.org/10.2519/jospt.2011.3499

13. Dragulescu AA, Arendt C. xlsx: read, write, format Excel 2007 and Excel 97/2000/XP/2003 files. Available at: https://cran.r-project.org/web/packages/xlsx/index.html. Accessed August 30, 2019.
14. Fletcher TD. psychometric: applied psychometric theory. Available at: https://cran.r-project.org/web/packages/psychometric/index.html. Accessed August 30, 2019.
15. Forsdyke D, Smith A, Jones M, Gledhill A. Psychosocial factors associated with outcomes of sports injury rehabilitation in competitive athletes: a mixed studies systematic review. *Br J Sports Med*. 2016;50:537-544. https://doi.org/10.1136/bjsports-2015-094850
16. Gallagher J, Needleman I, Ashley P, Sanchez RG, Lumsden R. Self-reported outcome measures of the impact of injury and illness on athlete performance: a systematic review. *Sports Med*. 2017;47:1335-1348. https://doi.org/10.1007/s40279-016-0651-5
17. Gamer M, Lemon J, Fellows I, Singh P. irr: various coefficients of interrater reliability and agreement. Available at: https://cran.r-project.org/web/packages/irr/index.html. Accessed August 30, 2019.
18. Garrison C, Cook C. Clinimetrics corner: the Global Rating of Change Score (GRoC) poorly correlates with functional measures and is not temporally stable. *J Man Manip Ther*. 2012;20:178-181. https://doi.org/10.1179/1066981712Z.00000000022
19. Griggs CL, Schneider JC, Kazis LE, Ryan CM. Patient-reported outcome measures: a stethoscope for the patient history. *Ann Surg*. 2017;265:1066-1067. https://doi.org/10.1097/SLA.0000000000002165
20. Guyatt GH, Townsend M, Berman LB, Keller JL. A comparison of Likert and visual analogue scales for measuring change in function. *J Chronic Dis*. 1987;40:1129-1133. https://doi.org/10.1016/0021-9681(87)90080-4
21. Hernandez-Sanchez S, Hidalgo MD, Gomez A. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *Br J Sports Med*. 2014;48:453-457. https://doi.org/10.1136/bjsports-2012-091163
22. Hespanhol Junior LC, Huisstede BM, Smits DW, et al. The NLstart2run study: economic burden of running-related injuries in novice runners participating in a novice running program. *J Sci Med Sport*. 2016;19:800-804. https://doi.org/10.1016/j.jsams.2015.12.004
23. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10:407-415. https://doi.org/10.1016/0197-2456(89)90005-6
24. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Phys Ther*. 2009;89:125-135. https://doi.org/10.2522/ptj.20080234
25. Johnston LH, Carroll D. The context of emotional responses to athletic injury: a qualitative analysis. *J Sport Rehabil*. 1998;7:206-220. https://doi.org/10.1123/jsr.7.3.206
26. Juniper EF, Guyatt GH, Streiner DL, King DR. Clinical impact versus factor analysis for quality of life questionnaire construction. *J Clin Epidemiol*. 1997;50:233-238. https://doi.org/10.1016/S0895-4356(96)00377-0
27. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther*. 2009;17:163-170. https://doi.org/10.1179/jmt.2009.17.3.163
28. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis*. 1985;38:27-36. https://doi.org/10.1016/0021-9681(85)90005-0
29. Kluitenberg B, van Middelkoop M, Diercks RL, et al. The NLstart2run study: health effects of a running promotion program in novice runners, design of a prospective cohort study. *BMC Public Health*. 2013;13:685. https://doi.org/10.1186/1471-2458-13-685
30. Kluitenberg B, van Middelkoop M, Smits DW, et al. The NLstart2run study: incidence and risk factors of running-related injuries in novice runners. *Scand J Med Sci Sports*. 2015;25:e515-e523. https://doi.org/10.1111/sms.12346
31. Kluitenberg B, van Middelkoop M, Verhagen E, et al. The impact of injury definition on injury surveillance in novice runners. *J Sci Med Sport*. 2016;19:470-475. https://doi.org/10.1016/j.jsams.2015.07.003
32. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-163. https://doi.org/10.1016/j.jcm.2016.02.012
33. Lissitz RW, Green SB. Effect of the number of scale points on reliability: a Monte Carlo approach. *J Appl Psychol*. 1975;60:10-13. https://doi.org/10.1037/h0076268
34. Louw A, Diener I, Landers MR, Puentedura EJ. Preoperative pain neuroscience education for lumbar radiculopathy: a multicenter randomized controlled trial with 1-year follow-up. *Spine (Phila Pa 1976)*. 2014;39:1449-1457. https://doi.org/10.1097/BRS.0000000000000444
35. MacDermid JC, Silbernagel KG. Outcome evaluation in tendinopathy: foundations of assessment and a summary of selected measures. *J Orthop Sports Phys Ther*. 2015;45:950-964. https://doi.org/10.2519/jospt.2015.6054
36. Mallows A, Littlewood C, Malliaras P. Measuring patient-reported outcomes (PROs/PROMs) in people with Achilles tendinopathy: how useful is the VISA-A? *Br J Sports Med*. 2018;52:1221. https://doi.org/10.1136/bjsports-2017-097531
37. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int*. 2005;26:968-983. https://doi.org/10.1177/107110070502601113
38. Marx RG, Bombardier C, Hogg-Johnson S, Wright JG. Clinimetric and psychometric strategies for development of a health measurement scale. *J Clin Epidemiol*. 1999;52:105-111. https://doi.org/10.1016/S0895-4356(98)00148-6
39. McPoil TG, Martin RL, Cornwall MW, Wukich DK, Irrgang JJ, Godges JJ. Heel pain—plantar fasciitis. *J Orthop Sports Phys Ther*. 2008;38:A1-A18. https://doi.org/10.2519/jospt.2008.0302
40. Mohtadi NG, Griffin DR, Pedersen ME, et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). *Arthroscopy*. 2012;28:595-610.e1. https://doi.org/10.1016/j.arthro.2012.03.013
41. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19:539-549. https://doi.org/10.1007/s11136-010-9606-8
42. Nielsen RO, Buist I, Sørensen H, Lind M, Rasmussen S. Training errors and running related injuries: a systematic review. *Int J Sports Phys Ther*. 2012;7:58-75.
43. Pipkin A, Kotecki K, Hetzel S, Heiderscheit B. Reliability of a qualitative video analysis for running. *J Orthop Sports Phys Ther*. 2016;46:556-561. https://doi.org/10.2519/jospt.2016.6280
44. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall; 2008.
45. Psychological issues related to illness and injury in athletes and the team physician: a consensus statement—2016 update. *Curr Sports Med Rep*. 2017;16:189-201. https://doi.org/10.1249/JSR.0000000000000359
46. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014.
47. Ramskov D, Nielsen RO, Sørensen H, Parner E, Lind M, Rasmussen S. The design of the Run Clever randomized trial: running volume, intensity and running-related injuries. *BMC Musculoskelet Disord*. 2016;17:177. https://doi.org/10.1186/s12891-016-1020-0
48. Rauh MJ, Koepsell TD, Rivara FP, Margherita AJ, Rice SG. Epidemiology of musculoskeletal injuries among high school cross-country runners. *Am J Epidemiol*. 2006;163:151-159. https://doi.org/10.1093/aje/kwj022
49. Revelle W. psych: procedures for psychological, psychometric, and personality research. Available at: https://cran.r-project.org/web/packages/psych/index.html. Accessed August 30, 2019.
50. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes.

J Clin Epidemiol. 2008;61:102-109. https://doi.org/10.1016/j.jclinepi.2007.03.012
51. Robinson JM, Cook JL, Purdam C, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. *Br J Sports Med*. 2001;35:335-341. https://doi.org/10.1136/bjsm.35.5.335
52. Ryan MB, Valiant GA, McDonald K, Taunton JE. The effect of three different levels of footwear stability on pain outcomes in women runners: a randomised control trial. *Br J Sports Med*. 2011;45:715-721. https://doi.org/10.1136/bjsm.2009.069849
53. Schmitt J, Abbott JH. Global ratings of change do not accurately reflect functional change over time in clinical practice. *J Orthop Sports Phys Ther*. 2015;45:106-111. https://doi.org/10.2519/jospt.2015.5247
54. Taunton JE, Ryan MB, Clement DB, McKenzie DC, Lloyd-Smith DR, Zumbo BD. A retrospective case-control analysis of 2002 running injuries. *Br J Sports Med*. 2002;36:95-101. https://doi.org/10.1136/bjsm.36.2.95
55. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42. https://doi.org/10.1016/j.jclinepi.2006.03.012
56. van Gent RN, Siem D, van Middelkoop M, van Os AG, Bierma-Zeinstra SM, Koes BW. Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review. *Br J Sports Med*. 2007;41:469-480; discussion 480. https://doi.org/10.1136/bjsm.2006.033548
57. Walker N, Thatcher J, Lavallee D. Psychological responses to injury in competitive sport: a critical review. *J R Soc Promot Health*. 2007;127:174-180. https://doi.org/10.1177/1466424007079494
58. Watson CJ, Propps M, Ratner J, Zeigler DL, Horton P, Smith SS. Reliability and responsiveness of the Lower Extremity Functional Scale and the Anterior Knee Pain Scale in patients with anterior knee pain. *J Orthop Sports Phys Ther*. 2005;35:136-146. https://doi.org/10.2519/jospt.2005.35.3.136
59. Webster KE, Feller JA, Lambros C. Development and preliminary validation of a scale to measure the psychological impact of returning to sport following anterior cruciate ligament reconstruction surgery. *Phys Ther Sport*. 2008;9:9-15. https://doi.org/10.1016/j.ptsp.2007.09.003

EARN CEUs With *JOSPT*'s Read for Credit Program

JOSPT's **Read for Credit (RFC)** program invites readers to study and analyze selected *JOSPT* articles and successfully complete online exams about them for continuing education credit. To participate in the program:

1. Go to **www.jospt.org** and click on **Read for Credit** in the top blue navigation bar that runs throughout the site.
2. Log in to read and study an article and to pay for the exam by credit card.
3. When ready, click **Take Exam** to answer the exam questions for that article.
4. Evaluate the RFC experience and receive a personalized certificate of continuing education credits.

The RFC program offers you 2 opportunities to pass the exam. You may review all of your answers—including your answers to the questions you missed. You receive **0.2 CEUs**, or 2 contact hours, for each exam passed.

JOSPT's website maintains a history of the exams you have taken and the credits and certificates you have been awarded in **My CEUs** and **Your Exam Activity**, located in the right rail of the Read for Credit page listing available exams.

APPENDIX

UNIVERSITY OF WISCONSIN RUNNING INJURY AND RECOVERY INDEX

Instructions: Consider your current running injury over the past 7 days when answering each question; check (☒) the appropriate box.

1. How does your running injury impact your ability to perform daily activities?	☐ No impact	☐ Slightly impact	☐ Moderately impact	☐ Significantly impact	☐ Unable to perform
2. How frustrated are you by your running injury?	☐ Not frustrated	☐ Mildly frustrated	☐ Moderately frustrated	☐ Significantly frustrated	☐ Extremely frustrated
3. How much recovery have you made from your running injury?	☐ Complete recovery	☐ Significant recovery	☐ Moderate recovery	☐ Minimal recovery	☐ No recovery
4. How much pain do you experience while running?	☐ No pain	☐ Minimal pain	☐ Moderate pain	☐ Significant pain	☐ Unable to run
5. How much pain do you experience during the 24 hours following a run?	☐ No pain	☐ Minimal pain	☐ Moderate pain	☐ Significant pain	☐ Unable to run
6. How has your weekly mileage or weekly running time changed as a result of your injury?	☐ Same or greater than before my injury	☐ Minimally reduced	☐ Moderately reduced	☐ Significantly reduced	☐ Unable to run
7. How has the distance of your longest weekly run changed as a result of your injury?	☐ Same or longer than before my injury	☐ Minimally reduced	☐ Moderately reduced	☐ Significantly reduced	☐ Unable to run
8. How has your running pace or speed changed as a result of your injury?	☐ Same or faster than before my injury	☐ Minimally reduced	☐ Moderately reduced	☐ Significantly reduced	☐ Unable to run
9. How does your injury affect your confidence to increase the duration or intensity of your running?	☐ Confident to increase my running	☐ If I increase, I might be fine	☐ Neutral	☐ If I increase, I might get worse	☐ I cannot increase my running

ANDREA POLLI, PT, MSc[1-4] • JO NIJS, PT, PhD[1,2,5] • KELLY ICKMANS, PT, PhD[1,2,5]
BRIGITTE VELKENIERS, MD, PhD[6] • LODE GODDERIS, MD, PhD[3,7]

Linking Lifestyle Factors to Complex Pain States: 3 Reasons Why Understanding Epigenetics May Improve the Delivery of Patient-Centered Care

J Orthop Sports Phys Ther 2019;49(10):683-687. doi:10.2519/jospt.2019.0612

Persistent pain is a substantial challenge for the individual and for health care providers and systems. Fundamental and clinical research has helped unravel mechanisms and factors—from biological to psychosocial to cultural—that contribute, facilitate, or otherwise influence the pain experience.[10] Our understanding of pain as a subjective, complex, and heterogeneous experience challenges the biomedical model and strongly supports embracing a more comprehensive, person-centered, biopsychosocial approach.[14,43] The link between the pain experience and underlying biological or neurophysiological mechanisms is not straightforward but influenced by many psychosocial and cultural factors, which are in turn influenced by one individual's biology.[2,16,31]

Biological mechanisms are important to understanding pain. Some people may be more predisposed and more vulnerable to pain.[10] However, genetic predisposition (ie, a mutation of a gene or a group of genes) can only explain around 30% of the risk of developing persistent pain.[21] This clearly represents a gap in our capacity to explain pain. One important discovery of the past decades is that vulnerability to many chronic illnesses, including chronic pain, is not solely dictated by our genes, but rather by the interaction between our genes and environmental and lifestyle factors.[24,26] Early-life stress, physical activity, diet, therapeutic drugs, and exposure to toxic elements (tobacco, solvents, etc) are all capable of changing gene functioning.[18,36,40] This is relevant to orthopaedic and sports physical therapists, as treatments for persistent pain often aim to modify people's lifestyle.

People with persistent pain may also be vulnerable to other health conditions,

[1]Pain in Motion International Research Group, Brussels, Belgium. [2]Department of Physiotherapy, Human Physiology and Anatomy, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels, Belgium. [3]Center for Environment and Health, Department of Public Health and Primary Care, Katholieke Universiteit Leuven, Leuven, Belgium. [4]Research Foundation - Flanders, Brussels, Belgium. [5]Department of Physical Medicine and Physiotherapy, University Hospital Brussels, Brussels, Belgium. [6]Department of Internal Medicine and Endocrinology, University Hospital Brussels, Brussels, Belgium. [7]External Service for Prevention and Protection at Work, IDEWE, Heverlee, Belgium. No funding or sponsorship was received for the present paper. Andrea Polli is a PhD research fellow funded by the Research Foundation - Flanders. Dr Nijs holds the chair titled "Exercise Immunology and Chronic Fatigue in Health and Disease," funded by the Berekuyl Academy (Hierden, the Netherlands). The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Andrea Polli, Vrije Universiteit Brussel, PAIN Department, Brussels Health Campus, Laarbeeklaan 103, 1090 Jette, Brussels, Belgium. E-mail: andrea.polli@vub.be

such as inflammatory and metabolic disorders, cardiovascular disease, and cancer,[19,41] and have higher rates of anxiety, depressive symptoms, insomnia, and flu-like symptoms.[38] Taken together, these findings suggest that some underlying mechanisms might be shared by different conditions and could help explain such overall vulnerability. Research supports a role for altered immune system responses, central nervous system hyperexcitability, and stress-related responses in predicting chronic pain.[9,34] The same mechanisms have been associated with inflammatory and cardiovascular diseases, diabetes, and depression.[23]

Underlying mechanisms are of clear importance, and lifestyle and environmental factors play a fundamental role. It is in this context that the field of epigenetics, which aims to explain complex gene-environment interactions at the molecular level (**FIGURE**), has emerged. In the following sections, we describe what epigenetics is, how it is already making an impact in other medical fields, and why it might be important for orthopaedic and sports physical therapists to consider its impact on outcomes and management.

What Is Epigenetics?

Epigenetic mechanisms influence how a gene works (eg, how it encodes for functional proteins) without changing the DNA itself.[20] Every cell has the same genetic code—the same DNA sequence. However, different groups of cells have diverse characteristics and exert very different functions. This occurs because of highly regulated epigenetic mechanisms, which are able to selectively silence some genes while allowing others to express.[20] Consider the genome as a piano and the full set of keys as DNA; epigenetics determines which keys are played.

When reading this Viewpoint, one should remember that epigenetic mechanisms affect the degree to which DNA is accessed by transcription factors.[36] For example, acetylation of histones (a group of proteins wrapped in DNA) changes DNA arrangement and facilitates binding of transcription factors to DNA.[20] Then, transcription can start, eventually leading to protein expression. In contrast, methylation of DNA hinders binding of transcription factors, inhibiting protein expression.[20]

Micro ribonucleic acid (microRNA) interference might also influence gene expression. Technically, microRNA interference is not an epigenetic mechanism, as it does not influence DNA structure or accessibility.[37] However, microRNAs are key regulators of the expression of thousands of genes and proteins and have been proposed as promising biomarkers for a number of diseases (eg, rheumatoid arthritis, diabetes, neurological and psychiatric disorders, and cancer).[37]

Important Discoveries in Epigenetics

Cancer development and progression are strongly influenced by epigenetic mechanisms.[7] In particular, DNA in cancer cells is often hypomethylated in genes fostering tumor growth, and hypermethylated in genes encoding for tumor suppression factors.[7] New drugs are currently being tested in preclinical and phase 1 clinical trials to reverse epigenetic changes that might promote cancer progression.[15]

Drugs can induce broad epigenetic changes: long-term use of opioids increased

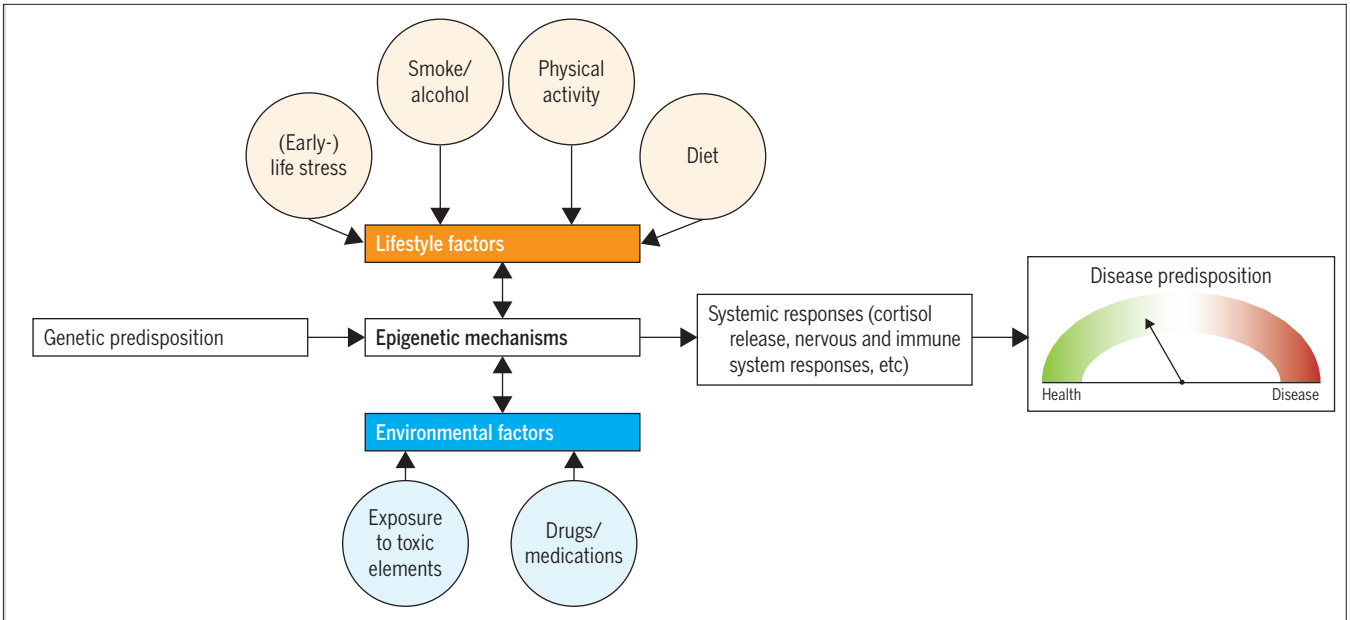


FIGURE. Schematic representation of the interactions between epigenetic mechanisms and lifestyle and environmental factors in the development of diseases. Environmental and lifestyle factors influence, via epigenetic mechanisms, biological responses and system functions. Importantly, this is often a dynamic and dual interaction, as lifestyle and environmental factors influence epigenetic processes, an epigenetic mechanism can influence the way our systems respond to environmental and lifestyle exposure. These interactions in turn facilitate either health or predisposition to certain chronic illnesses.

global DNA methylation, which was in turn associated with chronic pain symptoms.[27] DNA methylation of genes encoding for opiate receptors has been associated with the degree of clinical response to exogenous opioids,[27] suggesting that the molecular action of a drug is also influenced by an individual's epigenetic profile. This is important, as most people with persistent pain receive some medications for their symptoms, and opioids are commonly prescribed.[12]

Stressful events can induce dramatic changes at the molecular level and influence gene expression. Early-life stress or psychological traumas (eg, low maternal care during infancy) determine broad changes in DNA methylation in chromosome 18 and specific changes in DNA methylation of stress-related genes in animals.[25] Early-life stressful experience primed the stress system and determined an enhanced stress response and corticosteroid release at subsequent stress exposure during adulthood.[25] This suggests a dual and bidirectional relationship between environmental factors (eg, a stressful event) and a biological response (eg, future stress responses). In human studies, DNA methylation in stress-related genes is different in adults who experienced child abuse than in those who did not, increasing the chance of developing psychiatric disorders or depressive symptoms.[25]

The field of epigenetics has yet to have an impact on the fields of orthopaedic and sports physical therapy. However, the field is rapidly growing, and promising evidence has been published in other fields. Such evidence should not be overlooked but rather used to improve our understanding of biology and, consequently, our interventions.

Three Reasons Why Epigenetics Might Be Relevant to Understanding and Managing Pain

Improved Understanding of Underlying Mechanisms Would Improve the Clinician's Understanding of Pain and Promote Mechanism-Based Clinical Reasoning[9] Let us take the example of fibromyalgia. Initial evidence from independent research found the DNA of people with fibromyalgia to be differently methylated in over 900 different genes.[4,32] These genes encode for neurotrophins associated with neuronal sensitization such as the brain-derived neurotrophic factor, enzymes like histone deacetylases (HDACs), antioxidative products, immune and autonomic nervous system cells, and enzymes implicated in DNA repair and neuronal survival. Although these are small studies, some findings have been replicated[4,32] and suggest that greater understanding of the biological processes implicated in cellular responses to stress, immune system responses, and central nervous system abnormalities is needed to explain the pathophysiology of fibromyalgia.

These results are also supported by larger, longitudinal studies in people with chronic widespread pain.[28] Among 1708 people followed for 3 years, some regions in the DNA were differentially methylated in people with chronic widespread pain. These regions are related to genes known to regulate inflammatory cytokines (tumor necrosis factor-alpha and interleukin-17) and protein kinases. Targeting the correct underlying mechanisms is crucial to achieve successful analgesia.[45]

Epigenetics adds an additional layer to the clinician's understanding, providing a set of new biological mechanisms that explore complex processes and interactions at their root—where there is a direct influence on gene expression.

Epigenetics May Improve Subgrouping or Diagnosis in Complex Clinical Presentations[3] Diagnostic criteria for complex pain syndromes exist, but are almost always clinical,[3,44] a challenge for diagnosis.[42] A clearer biological characterization would likely facilitate diagnosis and guide treatments. Let us return to the example of fibromyalgia: the DNA methylation profile might identify patient subgroups in fibromyalgia.[4] Cluster analysis based on differentially methylated positions identified 3 different groups: 1 group included only people with fibromyalgia, 1 included only controls, and 1 included a mixed group of cases and controls. The 3 groups had different cortical excitability.[33] Expression of microRNAs in the cerebrospinal fluid can differentiate people with fibromyalgia from controls.[29] Another example: microRNA profiling in complex regional pain syndrome allowed reliable identification, based on inflammatory markers, of 1 subgroup—60% of the total sample—including only patients.[35]

Therefore, it might be possible to identify 1 or more subgroups of people with complex pain. Epigenetic marker assessment might also challenge clinical diagnostic criteria. Several investigations have attempted to find biological markers to improve the diagnosis of fibromyalgia and complex regional pain syndrome, without success.[3,44] Epigenetics might represent a step forward. If so, this would be of outstanding clinical relevance, and more studies are warranted. Many people with fibromyalgia struggle to achieve understanding and credibility of their condition.[39] We speculate that identifying 1 or more biomarkers might help patients to get the illness recognition they deserve.

Epigenetic Mechanisms Might Be Targets for Rehabilitation Interventions Reduced histone acetylation contributes to neuronal hyperexcitability and has been associated with hyperalgesia and allodynia in an animal model of neuropathic pain.[30] Hyperexcitability is successfully reversed by HDAC inhibitors,[30] a group of chemical compounds designed to restore normal histone acetylation. The efficacy of HDAC inhibitors is currently being tested in people with neurological disorders.[15]

The same nociceptive mechanism may be targeted with exercise. One bout of treadmill exercise reduces histone deacetylation in rats,[13] and a 60-minute cycling session may reduce HDACs in healthy people.[8] Exercise can reduce proinflammatory cytokines, such as tumor necrosis factor, and increase anti-inflammatory ones, such as interleukin-10, through epigenetic mechanisms.[22] Inflammation has

also been implicated in the pathogenesis of other chronic diseases (ie, diabetes) and cancer progression.[11] These findings might (1) explain why some lifestyle interventions are beneficial to prevent or improve some patients' conditions, and (2) suggest that exercise might work by promoting histone acetylation and regulating inflammatory markers.

Closing the Loop: Mechanism-Based Reasoning and Patient-Centered Care

We highlighted the complexity of persistent pain, determined by a diverse and ever-changing combination of biology (eg, central nervous system, immune system, and stress-related responses), psychology, and society. We acknowledged that important research suggests a need to embrace a patient-centered, biopsychosocial approach to improve outcomes.[14,43] We argue that only through in-depth understanding of complex mechanisms and by using mechanism-based reasoning can the clinician tailor interventions—the basic tenet of patient-centered care.[5]

Epigenetics is helping to unravel complex underlying mechanisms and might have at least 3 major clinical implications for orthopaedic and sports physical therapists. First, it promotes mechanism-based clinical reasoning by improved understanding of the pathophysiology of many health conditions and the underlying mechanisms of action of commonly used interventions, such as exercise and physical activity. Second, it might help patient subgrouping, allowing more targeted interventions. Finally, it might be used as a biomarker to monitor the effects of environmental factors and lifestyle interventions, such as physical activity, on health.

Several epigenetic biomarkers have been identified for diagnosis and prognosis in other conditions (eg, cancer), some of which have been commercialized.[6] In the next 10 to 15 years, as technology advances and costs reduce, biomarkers will likely find their place in the orthopaedic and sports arena. Discovering blood biomarkers is a priority in other fields of rehabilitation.[1] Epigenetic biomarkers have several advantages: they are more stable than other circulating biomarkers in fluids (serum, plasma, blood, etc),[17] but they are more influenced by environmental and lifestyle factors than are genetic markers and are thus more informative for prognosis and disease progression. For these reasons, we urge clinicians and clinical researchers to follow this rapidly growing area of research, as it might be soon contributing to patient assessment. ◉

REFERENCES

1. Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. *Neurorehabil Neural Repair*. 2017;31:864-876. https://doi.org/10.1177/1545968317732680
2. Brodal P. A neurobiologist's attempt to understand persistent pain. *Scand J Pain*. 2017;15:140-147. https://doi.org/10.1016/j.sjpain.2017.03.001
3. Bruehl S. An update on the pathophysiology of complex regional pain syndrome. *Anesthesiology*. 2010;113:713-725. https://doi.org/10.1097/ALN.0b013e3181e3db38
4. Ciampi de Andrade D, Maschietto M, Galhardoni R, et al. Epigenetics insights into chronic pain: DNA hypomethylation in fibromyalgia—a controlled pilot-study. *Pain*. 2017;158:1473-1480. https://doi.org/10.1097/j.pain.0000000000000932
5. Cloninger CR. Person-centered health promotion in chronic disease. *Int J Pers Cent Med*. 2013;3:5-12.
6. Costa-Pinheiro P, Montezuma D, Henrique R, Jerónimo C. Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics*. 2015;7:1003-1015. https://doi.org/10.2217/epi.15.56
7. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell*. 2012;150:12-27. https://doi.org/10.1016/j.cell.2012.06.013
8. Denham J, Marques FZ, O'Brien BJ, Charchar FJ. Exercise: putting action into our epigenome. *Sports Med*. 2014;44:189-209. https://doi.org/10.1007/s40279-013-0114-1
9. Denk F, McMahon SB. Chronic pain: emerging evidence for the involvement of epigenetics. *Neuron*. 2012;73:435-444. https://doi.org/10.1016/j.neuron.2012.01.012
10. Denk F, McMahon SB. Neurobiological basis for pain vulnerability: why me? *Pain*. 2017;158 suppl 1:S108-S114. https://doi.org/10.1097/j.pain.0000000000000858
11. Dufresne S, Rébillard A, Muti P, Friedenreich CM, Brenner DR. A review of physical activity and circulating miRNA expression: implications in cancer risk and progression. *Cancer Epidemiol Biomarkers Prev*. 2018;27:11-24. https://doi.org/10.1158/1055-9965.EPI-16-0969
12. Dunn KM, Saunders KW, Rutter CM, et al. Opioid prescriptions for chronic pain and overdose: a cohort study. *Ann Intern Med*. 2010;152:85-92. https://doi.org/10.7326/0003-4819-152-2-201001190-00006
13. Elsner VR, Lovatel GA, Bertoldi K, et al. Effect of different exercise protocols on histone acetyltransferases and histone deacetylases activities in rat hippocampus. *Neuroscience*. 2011;192:580-587. https://doi.org/10.1016/j.neuroscience.2011.06.066
14. Engel GL. The need for a new medical model: a challenge for biomedicine. *Science*. 1977;196:129-136. https://doi.org/10.1126/science.847460
15. Falkenberg KJ, Johnstone RW. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov*. 2014;13:673-691. https://doi.org/10.1038/nrd4360
16. Feder A, Nestler EJ, Charney DS. Psychobiology and molecular genetics of resilience. *Nat Rev Neurosci*. 2009;10:446-457. https://doi.org/10.1038/nrn2649
17. Glinge C, Clauss S, Boddum K, et al. Stability of circulating blood-based microRNAs – pre-analytic methodological considerations. *PLoS One*. 2017;12:e0167969. https://doi.org/10.1371/journal.pone.0167969
18. Godderis L, De Raedt K, Tabish AM, et al. Epigenetic changes in lymphocytes of solvent-exposed individuals. *Epigenomics*. 2012;4:269-277. https://doi.org/10.2217/epi.12.23
19. Gore M, Sadosky A, Stacey BR, Tai KS, Leslie D. The burden of chronic low back pain: clinical comorbidities, treatment patterns, and health care costs in usual care settings. *Spine (Phila Pa 1976)*. 2012;37:E668-E677. https://doi.org/10.1097/BRS.0b013e318241e5de
20. Henikoff S, Grosveld F. *Epigenetics & Chromatin* celebrates its first anniversary. *Epigenetics Chromatin*. 2009;2:13. https://doi.org/10.1186/1756-8935-2-13
21. Hocking LJ, Generation S, Morris AD, Dominiczak AF, Porteous DJ, Smith BH. Heritability of chronic pain in 2195 extended families. *Eur J Pain*. 2012;16:1053-1063. https://doi.org/10.1002/j.1532-2149.2011.00095.x
22. Horsburgh S, Robson-Ansley P, Adams R, Smith C. Exercise and inflammation-related epigenetic modifications: focus on DNA methylation. *Exerc Immunol Rev*. 2015;21:26-41.
23. Johnson AK, Zhang Z, Clayton SC, et al. The roles of sensitization and neuroplasticity in the long-term regulation of blood pressure and hypertension. *Am J Physiol Regul Integr Comp Physiol*. 2015;309:R1309-R1325. https://doi.org/10.1152/ajpregu.00037.2015
24. Kerr JI, Burri A. Genetic and epigenetic epide-

miology of chronic widespread pain. *J Pain Res*. 2017;10:2021-2029. https://doi.org/10.2147/JPR.S143869
25. Klengel T, Pape J, Binder EB, Mehta D. The role of DNA methylation in stress-related psychiatric disorders. *Neuropharmacology*. 2014;80:115-132. https://doi.org/10.1016/j.neuropharm.2014.01.013
26. Ladd-Acosta C, Fallin MD. The role of epigenetics in genetic and environmental epidemiology. *Epigenomics*. 2016;8:271-283. https://doi.org/10.2217/epi.15.102
27. Lirk P, Fiegl H, Weber NC, Hollmann MW. Epigenetics in the perioperative period. *Br J Pharmacol*. 2015;172:2748-2755. https://doi.org/10.1111/bph.12865
28. Livshits G, Malkin I, Freidin MB, et al. Genome-wide methylation analysis of a large population sample shows neurological pathways involvement in chronic widespread musculoskeletal pain. *Pain*. 2017;158:1053-1062. https://doi.org/10.1097/j.pain.0000000000000880
29. Masotti A, Baldassarre A, Guzzo MP, Iannuccelli C, Barbato C, Di Franco M. Circulating microRNA profiles as liquid biopsies for the characterization and diagnosis of fibromyalgia syndrome. *Mol Neurobiol*. 2017;54:7129-7136. https://doi.org/10.1007/s12035-016-0235-2
30. Matsushita Y, Araki K, Omotuyi O, Mukae T, Ueda H. HDAC inhibitors restore C-fibre sensitivity in experimental neuropathic pain model. *Br J Pharmacol*. 2013;170:991-998. https://doi.org/10.1111/bph.12366
31. Meeus M, Nijs J. Central sensitization: a biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. *Clin Rheumatol*. 2007;26:465-473. https://doi.org/10.1007/s10067-006-0433-9
32. Menzies V, Lyon DE, Archer KJ, et al. Epigenetic alterations and an increased frequency of micronuclei in women with fibromyalgia. *Nurs Res Pract*. 2013;2013:795784. https://doi.org/10.1155/2013/795784
33. Mhalla A, de Andrade DC, Baudic S, Perrot S, Bouhassira D. Alteration of cortical excitability in patients with fibromyalgia. *Pain*. 2010;149:495-500. https://doi.org/10.1016/j.pain.2010.03.009
34. Nijs J, Loggia ML, Polli A, et al. Sleep disturbances and severe stress as glial activators: key targets for treating central sensitization in chronic pain patients? *Expert Opin Ther Targets*. 2017;21:817-826. https://doi.org/10.1080/14728222.2017.1353603
35. Orlova IA, Alexander GM, Qureshi RA, et al. MicroRNA modulation in complex regional pain syndrome. *J Transl Med*. 2011;9:195. https://doi.org/10.1186/1479-5876-9-195
36. Polli A, Ickmans K, Godderis L, Nijs J. When environment meets genetics: a clinical review of the epigenetics of pain, psychological factors, and physical activity. *Arch Phys Med Rehabil*. 2019;100:1153-1161. https://doi.org/10.1016/j.apmr.2018.09.118
37. Pritchard CC, Cheng HH, Tewari M. MicroRNA profiling: approaches and considerations. *Nat Rev Genet*. 2012;13:358-369. https://doi.org/10.1038/nrg3198
38. Schiff M, Jaffe J, Freundlich B, Madsen P. New autoinjector technology for the delivery of subcutaneous methotrexate in the treatment of rheumatoid arthritis. *Expert Rev Med Devices*. 2014;11:447-455. https://doi.org/10.1586/17434440.2014.929492
39. Söderberg S, Lundman B, Norberg A. Struggling for dignity: the meaning of women's experiences of living with fibromyalgia. *Qual Health Res*. 1999;9:575-587. https://doi.org/10.1177/104973299129122090
40. Tabish AM, Poels K, Hoet P, Godderis L. Epigenetic factors in cancer risk: effect of chemical carcinogens on global DNA methylation pattern in human TK6 cells. *PLoS One*. 2012;7:e34674. https://doi.org/10.1371/journal.pone.0034674
41. Vaegter HB, Støten M, Silseth SL, et al. Cause-specific mortality of patients with severe chronic pain referred to a multidisciplinary pain clinic: a cohort register-linkage study. *Scand J Pain*. 2019;19:93-99. https://doi.org/10.1515/sjpain-2018-0094
42. Walitt B, Katz RS, Bergman MJ, Wolfe F. Three-quarters of persons in the US population reporting a clinical diagnosis of fibromyalgia do not satisfy fibromyalgia criteria: the 2012 National Health Interview Survey. *PLoS One*. 2016;11:e0157235. https://doi.org/10.1371/journal.pone.0157235
43. Wijma AJ, van Wilgen CP, Meeus M, Nijs J. Clinical biopsychosocial physiotherapy assessment of patients with chronic pain: the first step in pain neuroscience education. *Physiother Theory Pract*. 2016;32:368-384. https://doi.org/10.1080/09593985.2016.1194651
44. Wolfe F, Häuser W. Fibromyalgia diagnosis and diagnostic criteria. *Ann Med*. 2011;43:495-502. https://doi.org/10.3109/07853890.2011.595734
45. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. *Pain*. 2011;152:S2-S15. https://doi.org/10.1016/j.pain.2010.09.030

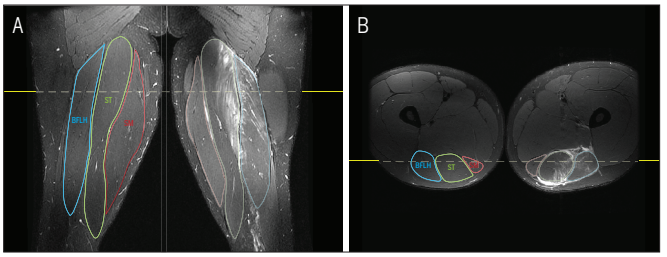


FIGURE 1. Magnetic resonance imaging 6 days after injury. (A) Coronal and (B) axial T2-weighted images of the thighs demonstrating increased T2 signal (increased brightness) in the left BFLH, surrounding a thickened proximal musculotendinous junction and tracking along the muscle's fascicles in a feathery pattern. Axial images reveal ST involvement. The yellow lines represent the locations of coronal and axial slices. Abbreviations: BFLH, biceps femoris long head; SM, semimembranosus; ST, semitendinosus.

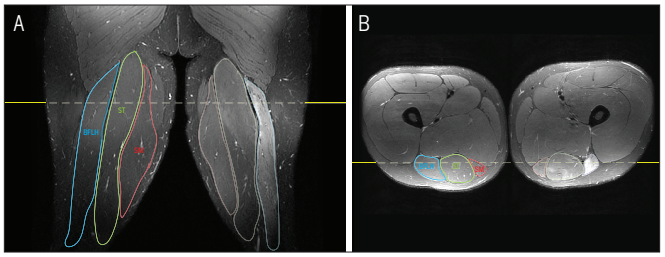


FIGURE 2. Magnetic resonance imaging 14 weeks after injury. (A) Coronal and (B) axial T2-weighted images of the thighs demonstrating a diffuse increase in signal (increased brightness) throughout the entirety of the left BFLH muscle and a portion of the ST, most consistent with subacute denervation edema. The yellow lines represent the locations of coronal and axial slices. Abbreviations: BFLH, biceps femoris long head; SM, semimembranosus; ST, semitendinosus.

Denervation Edema of Hamstring Muscles Following Acute Strain Injury

CHRISTA M. WILLE, PT, DPT, *Department of Orthopedics and Rehabilitation; Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI.*
KENNETH S. LEE, MD, *Department of Radiology, University of Wisconsin-Madison, Madison, WI.*
BRYAN C. HEIDERSCHEIT, PT, PhD, FAPTA, *Department of Orthopedics and Rehabilitation; Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI.*

A 20-year-old male American football player sustained a first-time, left hamstring strain injury (HSI) during competition. Magnetic resonance imaging indicated involvement of the biceps femoris long head (BFLH) and semitendinosus (ST) (**FIGURE 1**). Following 3 weeks of rehabilitation, the athlete was pain free, demonstrated full hamstring strength on clinical exam, and returned to competition without further incident.

Fourteen weeks after injury, magnetic resonance imaging was repeated as part of an ongoing study monitoring HSI recovery. Despite full athletic function, imaging revealed atrophy and increased signal intensity of the BFLH and a portion of the ST (**FIGURE 2**). Eccentric strength testing revealed that the involved limb was 15% weaker than the uninvolved limb, and an increase in strength in the uninvolved limb since preinjury testing. Due to lack of concomitant systemic symptoms and no personal history of inflammatory or autoimmune conditions, local tissue injury was the most likely differential diagnosis. Lack of subsequent trauma and diffuse edema evident throughout the BFLH, but absent in overlying subcutaneous tissues, indicated subacute denervation edema,[1,3] which commonly presents 2 to 4 weeks after denervation[1] and may be due to water shifting to extracellular spaces[2] as a result of decreased viability of local muscle tissue. The tibial branches of the sciatic nerve to the BFLH and a portion of the ST were likely damaged during the initial injury or by ischemic compression due to subsequent scarring. If innervation is not restored, atrophy with fatty infiltration can develop, indicating irreversible changes and a potential decrease in long-term muscle performance.[1] At 12 months post injury, the athlete reported no limitations and demonstrated a 10% increase in eccentric strength.

Because most individuals who sustain an HSI return to play within the time frame in which denervation edema may present, this condition may often go undetected. Continued monitoring of muscle performance, such as eccentric strength testing, is warranted after HSI to identify strength deficits. ◉ *J Orthop Sports Phys Ther 2019;49(10):761. doi:10.2519/jospt.2019.8598*

References
1. Fleckenstein JL, Watumull D, Conner KE, et al. Denervated human skeletal muscle: MR imaging evaluation. *Radiology*. 1993;187:213-218. https://doi.org/10.1148/radiology.187.1.8451416
2. West GA, Haynor DR, Goodkin R, et al. Magnetic resonance imaging signal changes in denervated muscles after peripheral nerve injury. *Neurosurgery*. 1994;35:1077-1085; discussion 1085-1086. https://doi.org/10.1227/00006123-199412000-00010
3. Wolf M, Wolf C, Weber MA. [Neurogenic myopathies and imaging of muscle denervation]. *Radiologe*. 2017;57:1038-1051. https://doi.org/10.1007/s00117-017-0316-x
The authors acknowledge the sports medicine staff at the University of Wisconsin-Madison Division of Athletics for their commitment to the welfare of the student-athletes and contributions to the Badger Athletic Performance program. The research presented was supported by the National Basketball Association and GE Healthcare Orthopedics and Sports Medicine Collaboration (MYT-015, D223) and under National Institutes of Health award TL1TR002375. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.